EMERGING TECHNOLOGIES POWERING RARE AND NEGLECTED DISEASE DIAGNOSIS AND THERAPY DEVELOPMENT

EDITED BY: Zhichao Liu, Shraddha Thakkar, Qais Y. Hatim, Ruth Roberts and Tieliu Shi

PUBLISHED IN: Frontiers in Pharmacology, Frontiers in Pediatrics and Frontiers in Genetics

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-064-0
DOI 10.3389/978-2-88976-064-0

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

EMERGING TECHNOLOGIES POWERING RARE AND NEGLECTED DISEASE DIAGNOSIS AND THERAPY DEVELOPMENT

Topic Editors:
Zhichao Liu, National Center for Toxicological Research (FDA), United States
Shraddha Thakkar, United States Food and Drug Administration, United States
Qais Y. Hatim, United States Food and Drug Administration, United States
Ruth Roberts, ApconiX, United Kingdom
Tieliu Shi, East China Normal University, China

Citation: Liu, Z., Thakkar, S., Hatim, Q. Y., Roberts, R., Shi, T., eds. (2022). Emerging Technologies Powering Rare and Neglected Disease Diagnosis and Therapy Development. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-064-0

Table of Contents

frontiers | Frontiers in Pharmacology

EDITORIAL
published: 05 April 2022
doi: 10.3389/fphar.2022.877401

Editorial: Emerging Technologies Powering Rare and Neglected Disease Diagnosis and Theraphy Development

Zhichao Liu[1]*, *Qais Hatim*[2], *Shraddha Thakkar*[2], *Ruth Roberts*[3,4] and *Tieliu Shi*[5,6]

[1]Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR, United States, [2]Office of Translational Sciences, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, United States, [3]ApconiX, BioHub, Alderley Park, United Kingdom, [4]Department of Biosciences, University of Birmingham, Birmingham, United Kingdom, [5]The Center for Bioinformatics and Computational Biology, The Institute of Biomedical Sciences and School of Life Sciences, Shanghai, China, [6]School of Statistics, Key Laboratory of Advanced Theory and Application in Statistics and Data Science-MOE, East China Normal University, Shanghai, China

Keywords: emerging technologies, rare diseases, drug repurposing, NGS—next generation sequencing, natural history

Editorial on the Research Topic

Emerging Technologies Powering Rare and Neglected Disease Diagnosis and Therapy Development

Although individual rare disease only affects a very small proportion of populations (i.e., usually less than 1:1500 in the United States or 1:2,000 in Europe), approximately 4% of the total world population experiences rare diseases collectively (Delavan et al., 2018). Obstacles such as incomplete knowledge of natural history hinder the etiology and pathogenesis understanding of rare and neglected diseases. Emerging technologies, including Artificial Intelligence (AI), advanced bioengineering, and next-generation sequencing (NGS), provide unprecedented opportunities to accelerate rare and neglected disease diagnosis and treatment development (Liu et al., 2019). The last day of February every year (i.e., rare disease day) raises awareness of rare diseases and promotes better diagnosis and treatment for eliminating the suffering of rare disease patients. It is an excellent opportunity to highlight this research topic for some recent progress and promising approaches to improve the diagnosis and therapy development of rare diseases. There are 20 papers collected in this research topic, covering a wide spectrum of research in rare diseases, including (1) promising application of NGS technologies in rare disease diagnosis; (2) clinical benefit of NGS-based genetic testing in rare diseases; (3) Genetic variant interpretation; (4) standardization of bioinformatics pipelines in NGS data analysis; (5) rare disease treatment development; and (6) natural history studies for better understating of etiology and pathogenesis of rare diseases.

More than 80% of rare diseases are of genetic origin. NGS-based causal genetic variant detections are indispensable for uncovering the etiology and pathogenesis of rare diseases and potentially facilitating treatment development. Ma et al. conducted a comprehensive genetic analysis based on 28 Chinese families with tyrosinase-positive oculocutaneous albinism (OCA) to identify genetic variants related to *OCA2* gene using NGS and confirmed the findings with sanger sequence. The causal association between genetic variants and clinical outcomes was discussed. Yang et al. employed whole-genome sequencing (WGS) and RNAseq to identify genetic variants that are responsible for peripheral neuroblastic tumors (PNTs) oncogenesis in a trio sample, finding a novel germline compound heterozygous mutation of the BRCA2 gene associated with familial PNTs.

OPEN ACCESS

Edited and reviewed by:
José A. G. Agúndez,
University of Extremadura, Spain

***Correspondence:**
Zhichao Liu
Zhichao.Liu@fda.hhs.gov

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *16 February 2022*
Accepted: *01 March 2022*
Published: *05 April 2022*

Citation:
Liu Z, Hatim Q, Thakkar S, Roberts R and Shi T (2022) Editorial: Emerging Technologies Powering Rare and Neglected Disease Diagnosis and Theraphy Development. Front. Pharmacol. 13:877401. doi: 10.3389/fphar.2022.877401

NGS has changed the landscape of genetic detection, providing unprecedented speed and resolution for uncovering individual genetic makeup. Clinical adoption of NGS has been significantly improved to offer diagnostic benefits for rare disease patients. Li et al. showed the potential utility of NGS on cerebrospinal fluid (CSF) for the diagnosis of cerebral alveolar echinococcosis (AE) in a patient with negative biopsy results in conventional clinical diagnosis procedure. Given that the patient had repeat seizures and progressive headaches for only 6 months and repeated target biopsy on the masses in the lung and liver showed fibrous connective tissue without positive findings, physicians successfully utilized the NGS results to diagnose the AE disease. They provided a "fit-for-purpose" treatment regimen (i.e., albendazole therapy). Liu et al. reported a case of a 1-year-old boy with metabolic acidosis and hypokalemia, a small penis, growth retardation, and G-6PD deficiency, whose clinical symptoms were complex and seemingly uncorrelated. The authors employed exome sequencing to identify the genetic mutation based on the trio-sample of the boy's family and found some *de novo* missense mutation that could be pathologic. Prader–Willi syndrome (PWS) is a complex genetic syndrome caused by the loss of function of genes in 15q11-q13 that are subject to regulation by genomic imprinting and expressed from the paternal allele only. Zhang et al. successfully implemented the preimplantation genetic diagnosis (PGD) to exclude imprinting deficiency in preimplantation embryos before transferring them into the mother's uterus. The lessons learned from the process further facilitate clinical utility of the procedure.

Complex structural variants (SVs) contribute to a large proportion of human genomic variation and cause rare diseases. Although NGS technologies, especially long-read-based sequencers, have tremendously increased SV detection sensitivity, the inconsistent results yielded by different bioinformatics pipelines still exist. Therefore, more studies on the pros and cons of varying bioinformatics pipelines are urgently needed to standardize the calling practice. Guo et al. conducted a comprehensive comparative study of three popular calling algorithms (i.e., PBSV, Sniffles, and PBHoney) on the GIAB benchmark samples, suggesting that all three pipelines exhibited better performance outside the tandem repeat regions (TRRs) than that of within TRRs. Sun et al. investigated the diagnosis power of the clinical WES technology in Epileptic encephalopathies (EEs) with an in-house pipeline for SNV identification and CNVkit for CNV detection. They found that WES can identify both SNVs/Indels and CNVs in a single flow cell and achieve a high diagnosis rate, suggesting that WES may serve as a first-tier diagnosis tool with cost-effective benefit in EEs.

Genetic variant interpretation and counseling are critical steps to enhance the clinical adoption of NGS testing. Feng et al. presented a case of a girl with non-ketotic hyperglycinaemia (NKH) and associate her genetic variants with clinical manifestations such as spasticity and bilateral cavitating leukoencephalopathy and a deficiency of a respiratory chain enzyme to provide proper genetic counseling. Gong et al. presented the detailed clinical features, such as infantile spasms (IS) and possibly the epilepsy of infancy with migrating focal seizures (EIMFS), and associated *KCNT2* gene mutations in developmental and epileptic encephalopathies (DEEs) to exemplify the genotype-phenotype association analysis in the clinical setting. Alpers' syndrome is an early inceptive neurodegenerative disorder with a poor prognosis, characterized by developmental regression, intractable epilepsy, and hepatic dysfunction. Li et al. deciphered the electroencephalogram (EEG) characteristics and clinical phenotype of different genotypes of the Alpers' syndrome and provided clues for early diagnosis. Jiao et al. identified a *de novo* variant (*p.Gly180Trp*) in the *SCAMP5* gene from four unrelated patients with epilepsy and neurodevelopmental delay, elucidating their phenotypic similarity and the difference of onset time, etc. Wang et al. presented the relationship between the patients' clinical features and the genetic variants in these patients identified by WES with a comprehensive literature survey, clarifying the mutational characteristics of causative genes in Hereditary spherocytosis (HS).

Natural history studies play a critical role in understanding the etiology, range of manifestations, and progression of rare diseases. Liu et al. carried out a meta-analysis of 36 encephalopathy patients. They identified a novel *de novo* genetic variant (i.e., (*c.116G>A*, *g.22229G>A*, *p.S39N*) in the GTPase domain of *DNM1L*) and elucidated its association with mitochondrial-related clinical phenotypes. Xu et al. investigated the pathogenicity of the novel mtDNA variant m. 9396G > A/MT-CO3 (*p.E64K*) in a patient with mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS) with biochemical assays. They found a muscle biopsy which confirmed the remarkable CIV deficiency caused by the novel mutation, characterized by ragged red fibers and generalized COX non-reactive muscle fibers. Hong et al. used a rat model to investigate the role of *GABRD* gene methylation in the nucleus accumbens in heroin-seeking behavior and suggested a potential treatment option to treat heroin addiction and relapse. Yang et al. revealed the association between a sex chromosome mosaicism present in the oral epithelial cells or gonad tissues of patients with sex development disorder (DSD) and with XYY in blood using Fluorescence *in situ* hybridization (FISH), enhanced the understanding of the genetic basis of DSD in males.

There are over 7,000 rare diseases. However, only less than 700 treatment options are available. There are tremendous unmet needs for rare disease treatment development. Sanchez et al. reviewed the progress of Alagille Syndrome (ALGS) treatment development, emphasized the emerging technologies such as NGS and advanced bioengineering in the development of targeted therapies for ALGS. Zhou et al. proposed actionable clinical decisions for CAD deficiency with specific manifestation such as co-occurrence of drug-resistant epilepsy, mainly focal or with generalization. They emphasized the need to act immediately upon discovering such variants in treatable disease genes. Wang et al. compared the eight patients with 3-Hydroxyisobutyryl-CoA hydrolase (HIBCH) deficiency to summarize the common clinical features caused by HIBCH genetic mutation. The outlined common clinical features further guided the treatment regimen strategy development, resulted in the improved clinical outcome of HIBCH patients. Liu et al. utilized the gene fusion profiles of neuroblastoma (NB)

patients to stratify the patients into newly defined subgroups with increased survival curves. Furthermore, the authors also carried out the precision medicine-based drug repurposing to the refined high-risk patient subgroups to optimize the treatment regimens.

REFERENCES

Delavan, B., Roberts, R., Huang, R., Bao, W., Tong, W., and Liu, Z. (2018). Computational Drug Repositioning for Rare Diseases in the Era of Precision Medicine. *Drug Discov. Today* 23, 382–394. doi:10.1016/j.drudis.2017.10.009

Liu, Z., Zhu, L., Roberts, R., and Tong, W. (2019). Toward Clinical Implementation of Next-Generation Sequencing-Based Genetic Testing in Rare Diseases: Where Are We? *Trends Genet.* 35, 852–867. doi:10.1016/j.tig.2019.08.006

Author Disclaimer: This editorial reflects the views of the authors and does not necessarily reflect those of the Food and Drug Administration. Any mention of commercial products is for clarification only and is not intended as approval, endorsement, or recommendation.

Conflict of Interest: RR is co-founder and co-director of ApconiX, an integrated toxicology and ion channel company that provides expert advice on non-clinical aspects of drug discovery and drug development to academia, industry, and not-for-profit organizations.

AUTHOR CONTRIBUTIONS

ZL wrote the first draft of the editorial. QH, RR, ST, and TS revised the editorial.

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.

frontiers in Pharmacology

CASE REPORT
published: 07 December 2020
doi: 10.3389/fphar.2020.608737

Case Report: Rapid Treatment of Uridine-Responsive Epileptic Encephalopathy Caused by CAD Deficiency

Ling Zhou[1†], Jie Deng[1†], Sarah L. Stenton[2,3], Ji Zhou[1], Hua Li[1], Chunhong Chen[1], Holger Prokisch[1,2,3] and Fang Fang[1]**

[1]Department of Neurology, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China, [2]Institute of Human Genetics, Technische Universität München, München, Germany, [3]Institute of Neurogenomics, Helmholtz Zentrum München, München, Germany

OPEN ACCESS

Edited by:
Zhichao Liu,
National Center for Toxicological Research (FDA), United States

Reviewed by:
Phillip Lawrence Pearl,
Harvard Medical School, United States
Gaetan Lesca,
Université Claude Bernard Lyon 1, France

***Correspondence:**
Fang Fang
fangfang@bch.com.cn
Holger Prokisch
prokisch@helmholtz-muenchen.de

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: 21 September 2020
Accepted: 28 October 2020
Published: 07 December 2020

Citation:
Zhou L, Deng J, Stenton SL, Zhou J, Li H, Chen C, Prokisch H and Fang F (2020) Case Report: Rapid Treatment of Uridine-Responsive Epileptic Encephalopathy Caused by CAD Deficiency. Front. Pharmacol. 11:608737. doi: 10.3389/fphar.2020.608737

We present two unrelated Chinese patients with CAD deficiency manifesting with a triad of infantile-onset psychomotor developmental delay with regression, drug-refractory epilepsy, and anaemia with anisopoikilocytosis. Timely translation into uridine supplementation, within 2-months of disease onset, allowed us to stop conventional anti-epileptic drugs and led to dramatic improvement in the clinical symptoms, with prompt cessation of seizures, resolution of anaemia, developmental progress, and prevention of development of severe and non-reversible manifestations. The remarkable recovery and prevention of advanced disease with prompt treatment, highlights the need to act immediately upon genetic diagnosis of a treatable disease. This further reinforces CAD deficiency as a treatable neurometabolic disorder and emphasises the need for a biomarker or genetic new born screening for early identification.

Keywords: CAD deficiency, epilepsy, anaemia with anisopoikilocytosis, uridine, developmental delay

INTRODUCTION

A trifunctional protein (carbamoyl phosphate synthetase, asparatate transcarbamylase and dihydroorotase) encoded by CAD, enzymatically accounts for the first three steps of the six-step *de novo* pyrimidine synthesis (Chen et al., 1989). Dysfunctional CAD results in pyrimidine deficiency, and corresponds clinically to EIEE-50 (early infantile epileptic encephalopathy 50, OMIM# 114010). This deficiency can be compensated by an alternative mechanism, whereby pyrimidine is recycled from uridine, giving rise to the treatment option of uridine supplementation with antiepileptic effect attributed to the salvage pathway (Nyhan, 2005). Following the discovery of the first patient with biallelic CAD mutations in 2015, and with a more recent retrospective series published of 20 subjects in Rymen et al. (2020) Genetic Medicine 2020, there remains a lack of detailed depictions of the clinical picture of CAD deficiency prior to or post uridine treatment (Ng et al., 2015; Koch et al., 2017; Del Caño-Ochoa et al., 2020; Rymen et al., 2020; Zhou et al., 2020).

Here, we report two further unrelated Chinese patients with CAD deficiency presenting with a triad of infantile-onset psychomotor developmental delay with regression, drug-refractory epilepsy, and anaemia with anisopoikilocytosis. In addition to the phenotypes reported to date, we report optic nerve involvement with clinical signs raising suspicion of visual impairment. Diagnostic whole exome sequencing (WES) was initiated within 1-month of

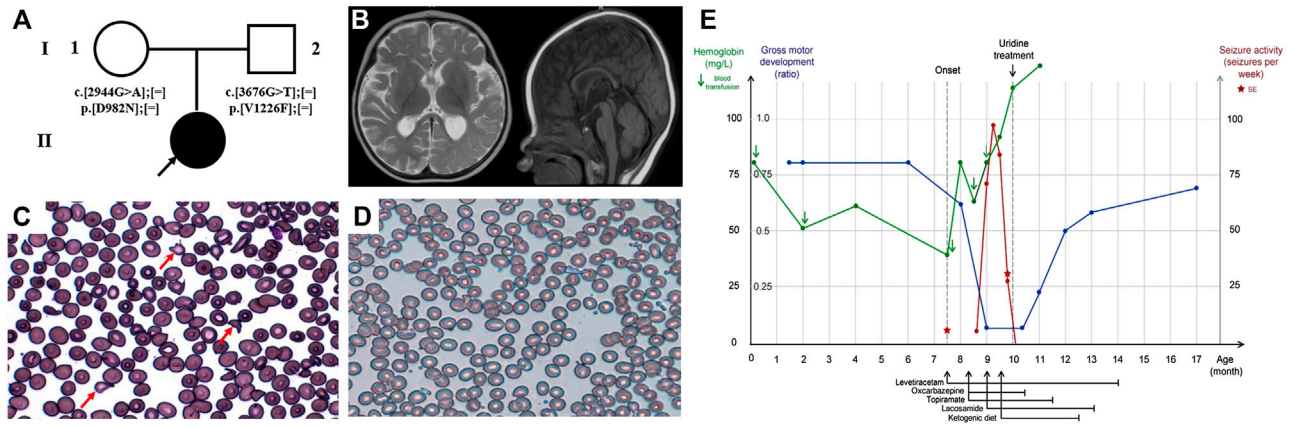

FIGURE 1 | Clinical features and uridine response of patient 1 with CAD deficiency. **(A)** Family pedigree of patient 1. **(B)** Brain MRI at 10-months. Peripheral blood smear **(C)** before treatment and **(D)** after treatment. **(E)** Clinical overview during the course of disease. Gross motor development represented as the observed development age according to the Nelson Gross Motor Development Milestones divided by the patient's biological age (Feigelman, 2016).

disease onset, and within 2-weeks of identifying the genetic diagnosis. This prompted immediate supplementation of uridine, achieving compelling improvement in the above clinical symptoms and prevention of the development of severe and non-reversible manifestations. This therapeutic success further supports the designation of CAD deficiency as a treatable epileptic encephalopathy and neurometabolic disorder, amongst a limited number of other notable exceptions (Dulac et al., 2014; van Karnebeek et al., 2014; Tarailo-Graovac et al., 2016).

CASE REPORT

Patient 1 is a girl currently 17-months of age. She is the first child of a healthy nonconsanguineous couple and was exclusively formula milk fed (**Figure 1A**). She received her first blood transfusion due to unexplained anaemia within one week of birth. Her neonatal period was otherwise uncomplicated. At the age of 7.5-months, she presented with recurrent afebrile seizures. The seizures were mostly of focal onset, with several episodes of status epilepticus, manifesting as eyes slanted to one side with ipsilateral or bilateral upper limb slightly shaking, lasting from several minutes to 1 h. She demonstrated global development delay prior to onset of the seizure attacks, and following seizure onset, she gradually lost head control, becoming hypotonic. Her brain MRI at 10-months showed mildly delayed myelination of the white matter, accompanied by mild atrophy-like change (**Figure 1B**). The interictal VEEG performed at 9-months recorded sharp and sharp-slow wave discharges on the left mid-to-posterior temporal lobe, while the ictal EEG documented 11 attack events (eyes turned right with slightly shaking of four limbs) originating from the left posteromedial temporal lobe (**Supplementary Figure S1A**). In addition, she presented with moderate to severe anaemia (haemoglobin 4–6.7, reference 11–16 g/dl; mean corpuscular volume 75.5–84.1, reference range 80–100 fl), necessitating five blood transfusions. The peripheral blood smear suggested abnormal erythrocytes of varying size and abnormal morphology (**Figure 1C**). Visual evoked potential (VEP) at 11-months suggested failure to evoke P100 wave, indicative of an optic nerve conduction block, with a normal examination of the fundus. Several anti-epileptic drugs (AEDs), including levetiracetam (LEV 50 mg/kg day), oxcarbazepine (OXC 50 mg/kg day), topiramate (TPM 4.5 mg/kg day), lacosamide (LCM 4.5 mg/kg day) and nitrazepam (NZP 0.25 mg/kg day), in addition to ketogenic diet therapy at a ratio of 2:1 were trialled for 1.5-months. These interventions, however, achieved poor efficacy. She achieved a seizure-free period for 1 month when taking LEV (20 mg/kg day), however her seizures subsequently recurred up to approximately 10 times per day. Two months after the first seizure, and just 2-weeks after identification of the CAD variants, she was administered oral uridine (100 mg/kg daily). Seizure cession was observed by day 2 of uridine administration. She stopped taking AEDs (OXC, TPM, KD, LCM, and LEV in order) approximately 2-months after commencing uridine and has remained seizure-free for 8 months follow-up to-date. Since treatment, a sequential interictal VEEG performed at 11-months of age recorded no epileptiform discharge, and the background activities were normal for the age (**Supplementary Figure S1B**). Re-examination of routine blood tests 1.5-months after commencing uridine supplementation demonstrated normalised haemoglobin as well as erythrocytes morphology, and the peripheral blood smear was almost normalised (**Figure 1D**). Re-examined VEP at 1-year of age demonstrated slightly decreased amplitude of P100 and prolonged latency. Moreover, she has made significant progress in development, and is now able to walk with assistance and speak several words, such as "mama," "papa," and "give me." Her clinical course is depicted in **Figure 1E**.

Patient 2 is a girl currently 17-months of age, born at term to a healthy nonconsanguineous family and fed by both breast and

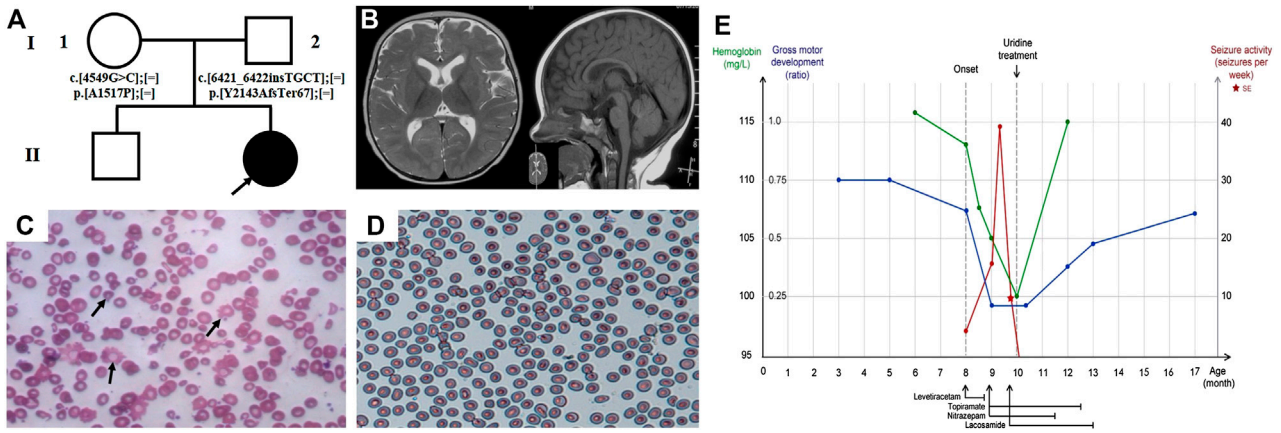

FIGURE 2 | Clinical features and uridine response of patient 2 with CAD deficiency. **(A)** Family pedigree of patient 2. **(B)** Brain MRI at 8.5-months. Peripheral blood smear **(C)** before treatment and **(D)** after treatment. **(E)** Clinical overview during the course of disease. Gross motor development represented as the observed development age according to the Nelson Gross Motor Development Milestones divided by the patient's biological age.

formula milk (**Figure 2A**). She presented with recurrent epileptic seizures at the age of 8-months. The seizures mainly presented focally with secondary generalisation. She experienced one episode of status epilepticus. When her seizures attack, her head and eyes tilted to the right, with salivating and lips turning blue, sometimes accompanied by ipsilateral or bilateral stiffness or waving of limbs, and usually lasting for one to 10 min. A mild developmental delay was noted prior to seizure onset, as she was able to raise her head at 3-months, roll over at 5-months, sit alone at 8-months, and she lacked visual fixation and following responses at 8-months. After seizure onset, she developmentally regressed and was unable to roll over or sit unsupported. Furthermore she developed hypotonia. The brain MRI performed at 8.5-months was unremarkable (**Figure 2B**). The ictal VEEG at 9-months recorded sharp (spike) slow wave discharges while monitoring a seizure attack (eyes turned right, salivating, with right limbs stiffness and shaking), and the interictal EEG documented sharp and spike slow wave in the bilateral occipitotemporal lobe (**Supplementary Figure S1C**). Her whole blood cell count indicated mild anaemia (haemoglobin 10–10.5, reference 11–16 g/dl; mean corpuscular volume 79.4–82.9, reference range 80–100 fl) and the peripheral blood smear showed abnormal erythrocytes of varying size and morphology, suggestive of anaemia with anisopoikilocytosis (**Figure 2C**). VEP performed at 8-months suggested an optic nerve conduction block, with severe prolongation of binocular P2 latency. She responded poorly to multiple AEDs, including LEV (10 mg/kg day, withdrew due to suspicious allergy), TPM (5.5 mg/kg day), LCM (5.5 mg/kg day), and NZP (0.15 mg/kg day). She achieved a short seizure free period for approximately 2-weeks at the first onset after receiving an intramuscular injection of phenobarbital, however seizures recurred up to almost 20 times a day thereafter. 2-months after the first seizure, and 6-days after identification of the CAD variants, supplementation with oral uridine (100 mg/kg daily) commenced. This led to the prompt cession of seizures by treatment day two. She stopped taking AEDs (LEV, NZP, TPM and LCM in order) approximately 4.5-months after commencing uridine, and has remained seizure-free during the 7.5-months follow-up period to-date. Moreover, she demonstrated dramatic improvement in her psychomotor development, as she can sit alone and has become more responsive and able to fix and follow objects. A sequential interictal VEEG at 1-year of age recorded no epileptiform discharge, except for few atypical sharp waves in the left posteromedial temporal lobe during sleep without epileptiform discharge (**Supplementary Figure S1D**). Routine blood tests and the peripheral blood smear both normalised (**Figure 2D**). A re-examined VEP at 13-months of age demonstrated normalized amplitude of P100 and latency. Moreover, she has made further significant progress in global development, and is now able to walk with assistance and speak one to two words, such as "mama" and "I want." Her clinical course is depicted in **Figure 2E**.

No adverse side effects of uridine administration were observed in either case. In-keeping with previous reports, despite CAD deficiency secondarily impairing glycosylation, urine purines and pyrimidines (including uracil, uric acid, hypoxanthine, and xanthine) and urine orotic acid testing prior to uridine supplementation were unremarkable, and urine purine testing showed only mild elevation of guanosine, deeming these biomarkers unsuitable in as biomarkers for the detection of CAD deficiency.

Genetically, compound heterozygous variants were identified in both patients in CAD (NM_004341) by trios-WES. Sanger sequencing was performed to confirm segregation by biparental transmission. In patient 1, the identified variants, c.3676G>T (p.V1226F) and c.2944G>A (p.D982N), were both interpreted as VUS (variant of unknown significance). As for patient 2, the identified variant, c.4549G>C (p.A1517P), was interpreted as likely pathogenic, while c.6421_6422insTGCT (p.Y2143AfsTer67) was predicted as VUS (**Figures 3A,B**). The

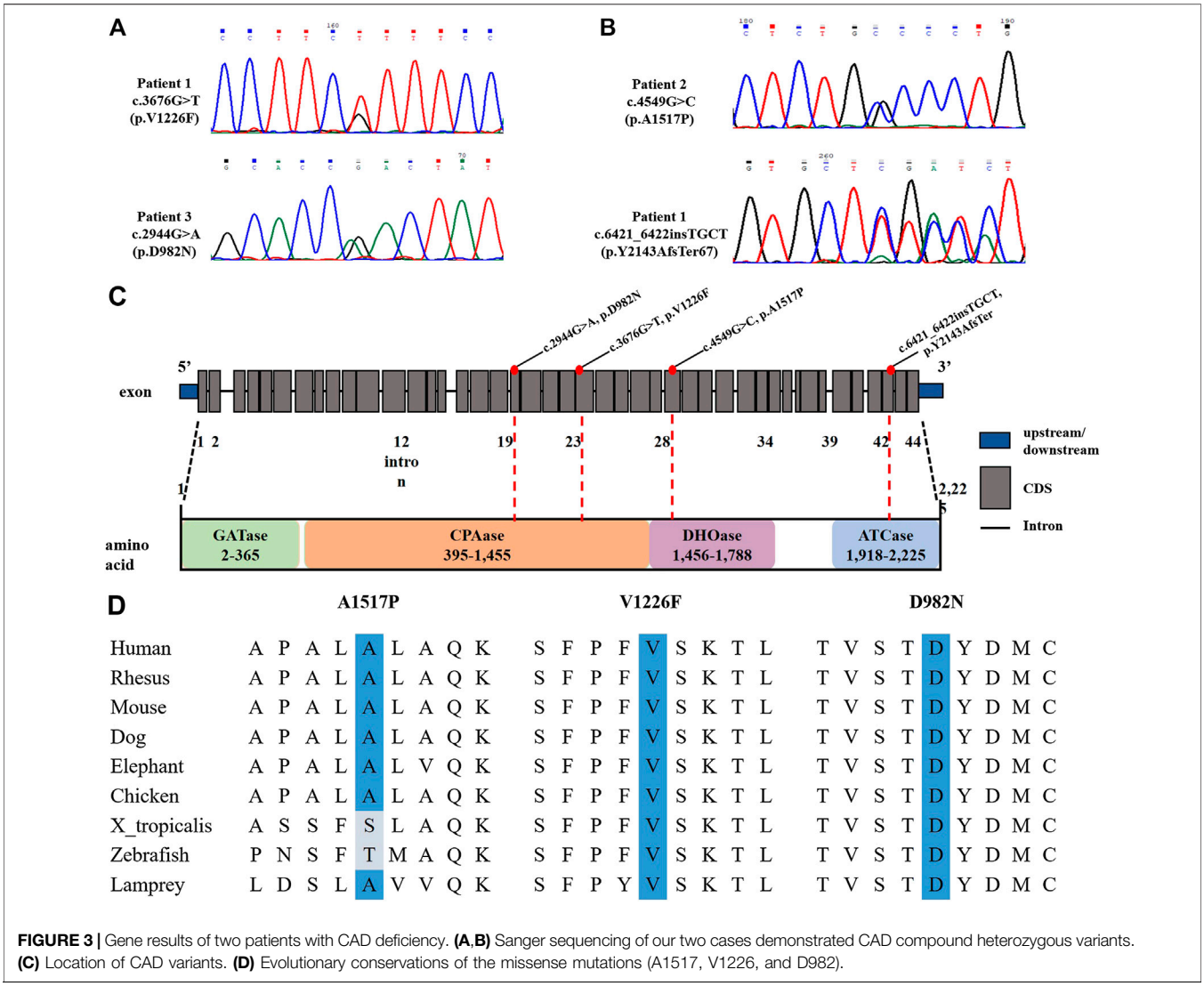

FIGURE 3 | Gene results of two patients with CAD deficiency. **(A,B)** Sanger sequencing of our two cases demonstrated CAD compound heterozygous variants. **(C)** Location of CAD variants. **(D)** Evolutionary conservations of the missense mutations (A1517, V1226, and D982).

CADD score, an in silico predictor of variant pathogenicity, was high in all missense variants (26.1, 30.0, and 24.4 for the p.V1226F, p.D982N, and A1517P, variants, respectively). The gnomAD allele frequency of the above variants were not listed, indicating the variants to be exceptionally rare. The pathogenicity of the variants was interpreted in accordance with the guidelines of the American College of Medical Genetics and Genomics (ACMG) (Richards et al., 2015). The human CAD gene consists of 44 exons and encodes four functional domains (GATase, CPAase, DHOase, and ATCase) that catalyse the first steps of *de novo* pyrimidine biosynthesis. The location of the novel variants in functional domains and their high-degree of conservation is depicted in (**Figure 3C**). According to sequence alignment, the amino acids A1517, V1226, and D982 are highly conserved across species, indicating evolutionary importance (**Figure 3D**). The mutation sites of the collectively described *CAD* mutations to date are distributed across all four functional domains of CAD, and include nonsense, splicing-site, and missense mutations, with no indication of specific phenotype-genotype correlation dependent on genomic position. They are all predicted to result in altered tertiary protein configuration and deteriorated protein function.

DISCUSSION

Here we recapitulate the clear clinical picture of CAD deficiency previously reported, with the co-occurrence of drug-resistant epilepsy, mostly focal or with generalization, developmental delay with regression, and anaemia with anisopoikilocytosis. Treatment with uridine supplementation was commenced despite the reported variants classification as VUS according to the ACMG criteria. Each of these phenotypes was responsive to supplementation, highlighting the need to act immediately upon discovery of such variants in treatable disease genes, where there is no known risk of side effects. Importantly, this targeted treatment allowed us to stop

conventional AEDs with known and significant side effect profiles.

Furthermore, due to recognition and diagnosis early in the natural course of the disease, we do not report the typical neuroradiological sequelae of progressive cerebral and cerebellar atrophy described in the advanced stage of disease (Koch et al., 2017). While of four previously described cases where uridine treatment was not implemented, two died in early childhood after succumbing to their neurodegeneration and in the two surviving patients, uridine treatment was applied with a delay of 15-months and 58-months from onset to treatment, respectively (Koch et al., 2017). Monitoring for future evolution of the brain MRI in our cases will require long-term follow-up to determine whether progressive cerebral and cerebellar atrophy will develop as expected in the late stage of disease. However, given their remarkable clinical improvement to-date, we believe prompt uridine supplementation may have prevented these complications.

Moreover, the clinical investigations suggest an additional phenotype in CAD deficiency, as visual impairment due to optic nerve conduction block was demonstrated by VEP in both patients. As with the typical triad of clinical phenotypes, this phenotype is ameliorated by uridine supplementation, as patients demonstrate improvement in pursuit of objects and in sequential VEP assessment following treatment. A patient recently reported by Ling Zhou et al. presented with reduced visual acuity and was diagnosed with strabismus at 2 years old, also here both were improved by uridine supplementation within approximately 1 year (Zhou et al., 2020).

As the natural course has shown to be lethal in early childhood without treatment, these cases highlight the dramatical efficiency and paramount importance of early CAD deficiency recognition and the need to act immediately, even without a (definite) genetic diagnosis. We do, however, appreciate the long-term evolution is not known, and will require a longer period of follow-up. To conclude, these case reports reinforce CAD deficiency as a treatable neurometabolic disorder and emphase the need for a biomarker or genetic new born screening for early identification.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

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

LZ, JD, and SS analysed the data and drafted the manuscript. JZ and HL collected and analysed the data. FF and HP critically revised and gave final approval for publication of the paper. Written consent for publication was obtained from the patients.

FUNDING

The work was supported by special project for capital health development (No. 2018-2-2096), by Prevention and Control of Major Chronic Non-Communicable Disease (No. 2016YFC1306203) and by the BMBF through the European network of mitochondrial disorders (GENOMIT, No. 10GM1603), by Cultivation Fund Project of the National Natural Science Foundation in Beijing Children's Hospital, Capital Medical University (GPQN201915), and Children's Medicine Research Project of Beijing Children's Hospital, Capital Medical University (YZQN202010).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2020.608737/full#supplementary-material

REFERENCES

Chen, K. C., Vannais, D. B., Jones, C., Patterson, D., and Davidson, J. N. (1989). Mapping of the gene encoding the multifunctional protein carrying out the first three steps of pyrimidine biosynthesis to human chromosome 2. *Hum. Genet.* 82, 40–44. doi:10.1007/BF00288269

Del Caño-Ochoa, F., Ng, B. G., Abedalthagafi, M., Almannai, M., Cohn, R. D., Costain, G., et al. (2020). Cell-based analysis of CAD variants identifies individuals likely to benefit from uridine therapy. *Genet. Med.* 22, 1598–1605. doi:10.1038/s41436-020-0833-2

Dulac, O., Plecko, B., Gataullina, S., and Wolf, N. I. (2014). Occasional seizures, epilepsy, and inborn errors of metabolism. *Lancet Neurol.* 13, 727–739. doi:10.1016/S1474-4422(14)70110-3

Feigelman, S. (2016). *Nelson textbook of pediatrics*. 20th Edn. Philadelphia, PA: Elsevier.

Koch, J., Mayr, J. A., Alhaddad, B., Rauscher, C., Bierau, J., Kovacs-Nagy, R., et al. (2017). CAD mutations and uridine-responsive epileptic encephalopathy. *Brain* 140, 279–286. doi:10.1093/brain/aww300

Ng, B. G., Wolfe, L. A., Ichikawa, M., Markello, T., He, M., Tifft, C. J., et al. (2015). Biallelic mutations in CAD, impair de novo pyrimidine biosynthesis and decrease glycosylation precursors. *Hum. Mol. Genet.* 24, 3050–3057. doi:10.1093/hmg/ddv057

Nyhan, W. L. (2005). Disorders of purine and pyrimidine metabolism. *Mol. Genet. Metabol.* 86, 25–33. doi:10.1016/j.ymgme.2005.07.027

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, 405–423. doi:10.1038/gim.2015.30

Rymen, D., Lindhout, M., Spanou, M., Ashrafzadeh, F., Benkel, I., Betzler, C., et al. (2020). Expanding the clinical and genetic spectrum of CAD deficiency: an epileptic encephalopathy treatable with uridine supplementation. *Genet. Med.* 22, 1589–1597. doi:10.1038/s41436-020-0933-z

Tarailo-Graovac, M., Shyr, C., Ross, C. J., Horvath, G. A., Salvarinova, R., Ye, X. C., et al. (2016). Exome sequencing and the management of

neurometabolic disorders. *N. Engl. J. Med.* 374, 2246–2255. doi:10.1056/NEJMoa1515792

van Karnebeek, C. D., Shevell, M., Zschocke, J., Moeschler, J. B., and Stockler, S. (2014). The metabolic evaluation of the child with an intellectual developmental disorder: diagnostic algorithm for identification of treatable causes and new digital resource. *Mol. Genet. Metabol.* 111, 428–438. doi:10.1016/j.ymgme.2014.01.011

Zhou, L., Han, X., Wang, T., and Wu, Y. (2020). A patient with CAD-deficiency responsive to uridine and literature review. *Front. Neurol.* 11, 64. doi:10.3389/fneur.2020.00064

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

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 18 December 2020
doi: 10.3389/fphar.2020.599191

Identification of an Identical *de Novo* SCAMP5 Missense Variant in Four Unrelated Patients With Seizures and Severe Neurodevelopmental Delay

Xianru Jiao[1], Manuela Morleo[2], Vincenzo Nigro[2,3], Annalaura Torella[3], Stefano D'Arrigo[4], Claudia Ciaccio[4], Chiara Pantaleoni[4], Pan Gong[1], Katheryn Grand[5], Pedro A. Sanchez-Lara[5], Joel Krier[6], Elizabeth Fieg[6], Andrew Stergachis[6], Xiaodong Wang[7] and Zhixian Yang[1]* on behalf of Telethon Undiagnosed Diseases Program*

[1]Department of Pediatrics, Peking University First Hospital, Beijing, China, [2]Telethon Institute of Genetics and Medicine, Naples, Italy, [3]Department of Precision Medicine, University of Campania "Luigi Vanvitelli", Naples, Italy, [4]Developmental Neurology Unit, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy, [5]Department of Pediatrics, Medical Genetics, Cedars-Sinai Medical Center, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States, [6]Brigham and Women's Hospital, Boston, MA, United States, [7]Cipher Gene Ltd., Beijing, China

OPEN ACCESS

Edited by:
Tieliu Shi,
East China Normal University, China

Reviewed by:
Juliet Taylor,
The University of Melbourne, Australia
Chin Moi Chow,
The University of Sydney, Australia

***Correspondence:**
Xiaodong Wang
xdwang@ciphergene.com
Zhixian Yang
zhixian.yang@163.com

Specialty section:
This article was submitted to Translational Pharmacology, a section of the journal Frontiers in Pharmacology

Received: 26 August 2020
Accepted: 11 November 2020
Published: 18 December 2020

Citation:
Jiao X, Morleo M, Nigro V, Torella A, D'Arrigo S, Ciaccio C, Pantaleoni C, Gong P, Grand K, Sanchez-Lara PA, Krier J, Fieg E, Stergachis A, Wang X and Yang Z (2020) Identification of an Identical de Novo SCAMP5 Missense Variant in Four Unrelated Patients With Seizures and Severe Neurodevelopmental Delay. Front. Pharmacol. 11:599191. doi: 10.3389/fphar.2020.599191

Objective: To establish and broaden the phenotypic spectrum of secretory carrier membrane protein (*SCAMP5)* associated with epilepsy and neurodevelopmental delay.

Methods: A Chinese patient was identified at the First Hospital of Peking University, and the three unrelated patients were recruited from two different countries (Italy and United States) through GeneMatcher. *SCAMP5* pathogenic variants were identified by whole exome sequencing; clinical data of the patients were retrospectively collected and analyzed.

Result: The onset age of seizures was ranged from 6 to 15 months. Patients had different types of seizures, including focal seizures, generalized tonic-clonic seizures and tonic seizure. One patient showed typical autism spectrum disorder (ASD) symptoms. Electroencephalogram (EEG) findings presented as focal or multifocal discharges, sometimes spreading to generalization. Brain magnetic resonance imaging (MRI) abnormalities were present in each patient. Severe intellectual disability and language and motor developmental disorders were found in our patients, with all patients having poor language development and were nonverbal at last follow-up. All but one of the patients could walk independently in childhood, but the ability to walk independently in one patient had deteriorated with age. All patients had abnormal neurological exam findings, mostly signs of extrapyramidal system involvement. Dysmorphic features were found in 2/4 patients, mainly in the face and trunk. All four unrelated patients were found to have the same heterozygous pathogenic *SCAMP5 de novo* variant (p. Gly180Trp).

Conclusion: Epilepsy, severe developmental delay, abnormal neurological exam findings, with or without ASD or variably dysmorphic features and were common in patients with *SCAMP5* variant. The onset time and type of seizure varied greatly. The EEG and brain MRI findings were not consistent, but diverse and nonspecific. The motor ability of patients with

heterozygous *SCAMP5* variant might have a regressive course; language development was more severely affected.

Keywords: epilepsy, Secretory carrier membrane protein 5, developmental delay, autism spectrum disorder, congenital deformity

INTRODUCTION

Secretory carrier membrane proteins (SCAMPs) are widely distributed integral membrane molecules implicated in regulating vesicular transport (Law et al., 2012). *SCAMP* 1–4 are expressed ubiquitously, whereas *SCAMP5* is known to be brain specific and is present in high levels in synaptic vesicles (SVs), and the function of SCAMP5 for vesicle transport is selective (Fernández-Chacón and Südhof, 2000; Park et al., 2018). The human *SCAMP5* maps to chromosome 15q. Previously, there have been cases of patients with a chromosomal microdeletion of 15q leading to a series of neurological diseases (Sharp et al., 2007; Klopocki et al., 2008; Castermans et al., 2010; Ahram et al., 2017; Huynh et al., 2018). In 2010, *SCAMP5* was studied as a candidate gene for autism spectrum disorder (ASD) (Castermans et al., 2010). However, no disease-associated human *SCAMP5* point mutations were reported until 2020, when Hubert et al. reported two unrelated patients with an identical heterozygous *SCAMP5* variant (p. Gly180Trp) who had ASD, intellectual disability, and seizures (Hubert et al., 2020). Studies in a *Drosophila melanogaster* model have shown the pathogenicity of that variant (Hubert et al., 2020). Zhang et al. subsequently identified a homozygous variant of *SCAMP5* (p. Arg91Trp) in Chinese siblings with pediatric epilepsy and juvenile Parkinson's disease (Zhang et al., 2020). *SCAMP5* variant knock-in mice showed typical early-onset epilepsy similar to that of humans (Zhang et al., 2020). The genetic pattern and specific phenotype of *SCAMP5* in human disease are unknown to date due to the limited numbers of reported cases.

Here, we present four unrelated patients with the identical *de novo SCAMP5* variant (p. Gly180Trp) from three countries, including detailed descriptions of initial clinical presentations and long-term follow-up (ranged from 1 to 33 years).

PATIENTS AND METHODS

Ethics Statement

This study was approved by the respective local Ethic Committees (Patient 1: Biomedical Research Ethical Committee of Peking University First Hospital, approval number: 2016-1135; Patient 2: local Ethical Committee of Federico II University of Naples; Patient 3: Cedars-Sinai Medical Center; Patient 4: Brigham and Women's Hospital). The individuals or their parents in this manuscript have been given the written informed consent to publish the case details.

Patients

Patient 1 was recruited from the First Hospital of Peking University. Additionally, three individuals (patients 2, 3, 4) with the identical *de novo* variant in *SCAMP5* were connected to this study through GeneMatcher (Sobreira et al., 2015). The results of electroencephalogram (EEG), magnetic resonance imaging (MRI), neurological examination findings, biochemical studies, plasma amino acids and urine organic acids test in four patients were respectively collected in First Hospital of Peking University, Telethon Undiagnosed Diseases Program, Cedars-Sinai Medical Center, and Brigham and Women's Hospital. Neurodevelopmental assessment was performed according to intelligence tests (Wechsler or Gesell intelligence scales) or clinical judgment and parents' questionnaires.

Genetic Analysis

DNA (3 μg) extracted from peripheral blood from probands and their parents was analyzed using whole-exome sequencing (WES) with standard protocol (Hubert et al., 2020). Sequence variants were checked with population databases gnomAD (http://gnomad.broadinstitute.org/) and evaluated using Polyphen2, SIFT, and Mutation Taster. The pathogenicity of variants were interpreted according to the American College of Medical Genetics (ACMG) guidelines (Richards et al., 2015). The variants were further confirmed by Sanger sequencing.

RESULTS

Clinical features of affected individuals with *SCAMP5* variants G180W were summarized in **Tables 1**, **2**.

Seizures, Electroencephalogram and Brain Magnetic Resonance Imaging Information

In total, four patients from three different countries were enrolled, including two males and two females. Among the four patients derived from four unrelated families, three patients were born after a normal pregnancy and uneventful delivery. Patient 4 was born 10 days post-term. Delivery was complicated by late decelerations during labor, occipital posterior presentation and a low forceps delivery, though Apgar scores were 9. All patients had seizures and the age of onset was ranged from 6 to 15 months. Patients had different types of seizures, including generalized tonic-clonic seizures (GTCS, 3/4), focal seizures (2/4), and tonic seizure (1/4). As for patient 1, the first seizures occurred at the age of 1 year and 3 months. Four days later, the child experienced two further episodes in the same day. After that, no convulsion was observed up to the last follow-up (2 years and 3 months) and he was not treated with antiepileptic drugs (AEDs). The seizures of patient 2 and patient 3 were controlled by AEDs at the age of 12 and 22 months, respectively. As for patient 4, from the first episode to the age of 22, recurrent epileptic seizures were consistently experienced, although treated with AEDs. However, after a hysterectomy to treat menorrhagia (around the

TABLE 1 | Demographic and clinical manifestations of affected individuals with *SCAMP5* variant G180W

	Patient 1	Patient 2	Patient 3	Patient 4	Hubert et al. patient 1	Hubert et al. patient 2
Origin	Chinese	Italian	American	American	Caucasian	Maghrebian
Current age/Sex	2 y 3 m/M	8 y 4 m/M	2 y 6 m/F	32 y/F	10 y/M	8 y 6 m/M
Gestation	At term	At term	At term	10 days post-term	At term	37 weeks
Birth history	Unremarkable	Unremarkable	Unremarkable	Late decelerations during labor, a low forceps delivery	Unremarkable	Intrauterine growth retardation
Age at seizure onset	15 m	11 m	12 m	6 m	30 m	33 m
Seizure type	FS	TS, GTCS	GTCS	FS, GTCS	Absences and atonic seizures, MS, GTCS	n.a.
Evolution	Not treated	Under control at 12 months	Under control at 22 months	Under control at about 22 years	Under control at 7 years	Not treated
Other clinical feature	No	Motor stereotypes, rare aggressiveness outburst	No	Overweight (BMI = 29)	Attention deficit, hyperactivity, stereotypies, autistic features	Aggressiveness, hyperactivity, stereotyped swinging Movements
Psychomotor development	Severe delay	Severe delay	Severe delay	Severe delay	Severe delay	Severe delay
Walking	1 y 5 m	3 y 10 m	No	4 y	3 y	2 y 6 m
Speech	No	No	No	No	No	No

FS, focal seizure; TS, tonic seizure; GTCS, generalized tonic-clonic seizures; MS, myoclonic seizure; n.a., not available; m, month; y, year.

age of 22), she was seizure free for over 10 years, and her epilepsy continues to be managed with valproic acid. Up to the last follow-up, all patients showed seizure free. In addition to seizures, patient 2 sometimes showed motor stereotypes and rare aggressiveness outburst.

EEG findings were obtained from three patients. Patient 1: initial EEG examination showed frontal-temporal and midline regions discharges, sometimes spreading to generalization and mild diffuse background slowing. At the last follow-up (2 years and 3 months), the EEG had no significant change (**Figure 1**). Patient 2: several EEGs demonstrated the presence of multifocal bilateral epileptic anomalies. At the last follow-up (8 years and 4 months), EEG presented diffuse and discontinuous epileptic activity, and short series of focal frontal-central left anomalies. Patient 3: EEG at the age of 12 months showed bilateral frontal-temporal discharges and mild diffuse background slowing. At the last follow-up (2 years and 6 months), EEG showed only slowing in the background without discharges. Brain MRI abnormalities were present in each patient. Brain MRI of patients 1 and 3 at 1 year old showed cyst in left choroid fissure and periventricular, respectively. Brain MRI of patient 2 demonstrated temporal mesial bilateral hypotrophy, especially affecting the hippocampus, aspecific white matter anomalies in peritrigonal and periventricular regions, diffuse supratentorial sulci enlargement, posterior corpus callosum thinning, and bilateral signal alterations in globus pallidus, substantia nigra, and dentate nucleus (**Figure 2**, at the age of 5). Brain MRI of patient 4 at ages 7, 13, and 31 have demonstrated consistent findings including local atrophy and FLAIR white matter hyperintensities (**Table 2**).

Neurodevelopment

All patients showed severe intellectual disability and developmental disorders, including motor and language development. Among them, three patients showed delayed motor development but could eventually walk independently at the last follow up, although with gait unsteadiness, and only one (patient 3) never reach autonomous walk. The motor development of patient 4 underwent a retrograde process. She crawled at 2 years and walked with a walker at 4 years, and required ankle foot orthosis braces from a very young age. She attained the ability to walk independently though at approximately age 30. She began to feel uncomfortable standing by herself and needed to hold onto something while standing or walking at all times and is now heavily reliant on a wheelchair (at the age of 32). At the last follow up, all patients were still backward in language development. They could understand simple instructions and speak a few words occasionally at the last follow up, which could not satisfy the normal communication.

Neurological Examinations and Dysmorphic Features

Neurological exam of all patients revealed drooling, hypotonia, limbs hyperreflexia, proximal limbs hyperkinesia, dystonic postures, mild tremor and gait abnormalities.

TABLE 2 | EEG, brain MRI, neurological examinations, and dysmorphic features of affected individuals with *SCAMP5* variant G180W

	Patient 1	Patient 2	Patient 3	Patient 4	Hubert et al. patient 1	Hubert et al. patient 2
Current age/ Sex	2 y 3 m/M	8 y 4 m/M	2 y 6 m/F	32 y/F	10 y/M	8 y 6 m/M
Interictal EEG	Frontal-temporal region discharges, sometimes spreading to generalization and mild diffuse background slowing	Multifocal bilateral discharges	Discharges in bilateral frontal-temporal and mild diffuse background slowing	n.a.	Frontal spikes and spike waves during sleep	Global slowing of background activity during drowsiness and sleep
Brain MRI	Age 1: Left choroid fissure cyst	Age 5: Temporal mesial bilateral hypotrophy, especially affecting the hippocampus, aspecific white matter anomalies in peritrigonal and periventricular regions, diffuse supratentorial sulci enlargement, posterior corpus callosum thinning, and bilateral signal alterations in globus pallidus, substantia nigra, and dentate nucleus.	Age 1: Periventricular cyst	Age 7: Thin corpus callosum, decreased white matter and mild dilation of lateral ventricles, subarachnoid spaces, and cerebellar fissures; age 13: Stable appearance of atrophic findings from age 7. New bilateral foci of FLAIR hyperintensity in periatrial white matter and focally in posterior limbs of internal capsules, likely new due to technical imaging differences; age 31: Slight atrophy of periventricular horns and thinning of corpus callosum.	Diffuse hyperintensity of the white matter, thin corpus callosum, mesial temporal sclerosis	Diffuse hyperintensity of the white matter, enlarged ventricules, mesial temporal sclerosis
Neurological exam	Dystonic postures while walking	Drooling, global hypotonia with limbs hyperreflexia, proximal limbs hyperkinesia, dystonic postures while walking	Hypotonia	Mild tremor, gait abnormalities	Ataxic gait, with trunk hypotonia, absent deep tendon reflexes, choreic delicate movements	Ataxic gait, global hypotonia
Facial dysmorphism	No	Progressive microcephaly, mildly narrow, depressed nasal bridge with broader tip, flat philtrum, large mouth with upper lip eversion, and protruding ears	No	Hypotelorism, bulbous nasal tip, downturned mouth, mild prognathic, mild anterior hairline elevation	n.a.	n.a.
Other dysmorphism	No	Scoliosis	No	Scoliosis	Varus equus	Varus equus

EEG, electroencephalogram; MRI, magnetic resonance imaging; n.a., not available; m, month; y, year.

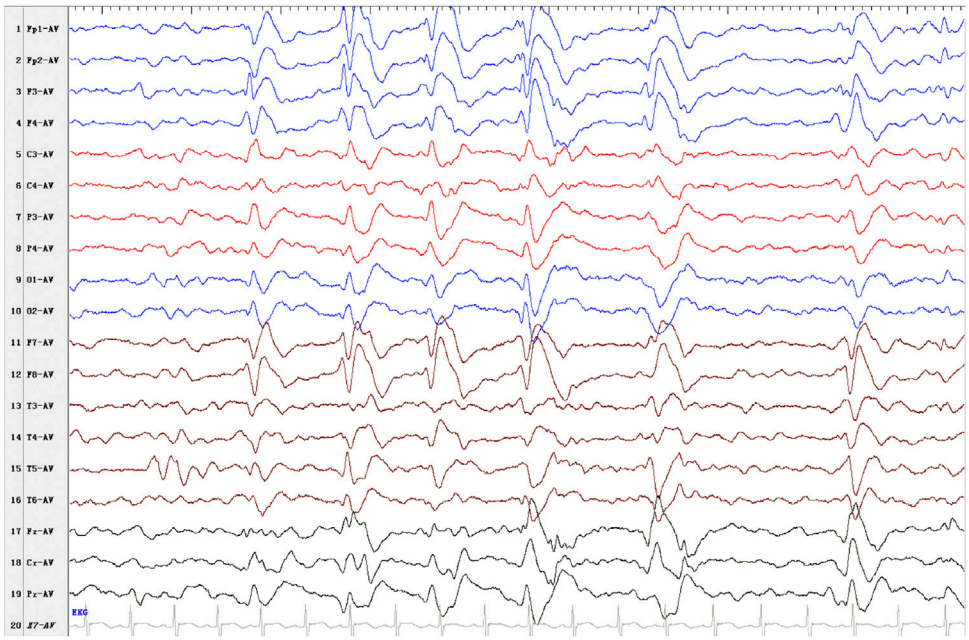

FIGURE 1 | Electroencephalogram (EEG) outcome of patient 1. Interictal EEG presented as discharges in frontal-temporal and midline regions, sometimes spreading to generalization at the last follow-up (2 years and 3 months).

Dysmorphic features were found in patient 2 and patient 4, mainly in the face and trunk. Facial deformities were varied, including progressive microcephaly, mildly narrow, depressed nasal bridge with broader tip, bulbous nasal tip, flat philtrum, large mouth with upper lip eversion, downturned mouth, mild prognathic, protruding ears, hypotelorism, and mild anterior hairline elevation. The other deformity was mainly scoliosis.

Secretory Carrier Membrane Protein *5* Variant

All four patients underwent exome sequencing and the same *de novo SCAMP5* (c.538G>T p.Gly180Trp) heterozygous variant was identified. The variant was predicted to be "probably damaging," "protein function affected" and "disease causing" to the protein structure by Polyphen-2, SIFT and Mutation Taster, respectively. The variant was absent in gnomAD population database, and classified as likely pathogenic according to ACMG guideline. The identical *SCAMP5* variant was reported in two patients by Hubert et al. (2020) with neurodevelopmental delay, ASD and epilepsy.

DISCUSSION

In 2020, Hubert et al. (2020) reported two unrelated patients with the identical heterozygous *de novo SCAMP5* variant (G180W) which was also identified in our cohort. Then, Zhang et al. (2020) identified a homozygous variant of *SCAMP5* (R91W) in Chinese siblings. To date, *SCAMP5* point mutations have been identified in four patients representing two different inheritance modes, and their pathogenicity has been confirmed by functional experiments (Hubert et al., 2020; Zhang et al., 2020). Using WES, we identified the identical *de novo* heterozygous *SCAMP5* variant (G180W) in four unrelated patients with a consistent phenotype, including epilepsy, severe developmental delay, abnormal neurological exams, and with or without ASD or variably dysmorphic features.

In 2010, a study identified *SCAMP5* as a candidate gene for ASD through detecting a *de novo* chromosomal translocation in ASD patients (Castermans et al., 2010). Both of the two patients reported by Hubert et al. exhibited characteristics of ASD, mainly in the form of stereotypical behaviors (Hubert et al., 2020). In our cohort, only patient 2 showed typical ASD symptoms, including motor stereotypes and rare aggressiveness outburst. Other patients have yet to find or obtain detailed information. In five cases with the novel 15q24 microdeletion syndrome that includes *SCAMP5*, ASD also had not been recorded (Sharp et al., 2007; Klopocki et al., 2008). Thus, ASD does not seem to be a universal feature of *SCAMP5* variants, even at the same mutation site.

Four patients here presented primarily with neurological abnormalities and began with epilepsy. The onset age of seizures ranged from 6 to 15 months, and three patients could be controlled with AEDs. In particular, patient 1 had only three episodes without AEDs therapy. Four cases reported internationally showed similar epileptic phenotypes (Hubert et al., 2020; Zhang et al., 2020). Seizures of two patients who had the same locus as ours occurred at 30 and 33 months, respectively (Hubert et al., 2020). Especially, in one patient, seizure was controlled at the age of seven, and the other did not require treatment (Hubert et al., 2020). The other two

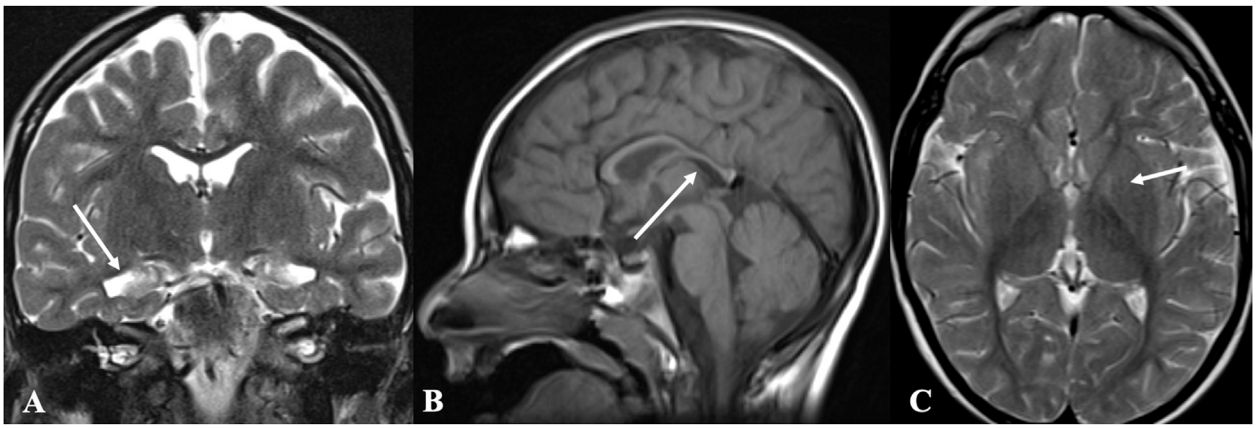

FIGURE 2 | Brain magnetic resonance imaging (MRI) of patient 2. Brain MRI demonstrated temporal mesial bilateral hypotrophy (arrow in **A**, coronal T2w), especially affecting the hippocampus, aspecific white matter anomalies in peritrigonal and periventricular regions, diffuse supratentorial sulci enlargement, posterior corpus callosum thinning (arrow in **B**, sagittal T1w), and bilateral signal alterations in globus pallidus (arrow in **C**, axial T2w), substantia nigra, and dentate nucleus.

reported by Zhang et al. carrying homozygous variants both had seizures at 5 months, but the progression was unknown (Zhang et al., 2020). Between the reported and newly identified cases, there were various forms of seizures, including focal seizures, GTCS, tonic seizure, myoclonic seizures, absences and atonic seizures, and GTCS form was the most common. Thus, seizure was the common phenotype of patients with *SCAMP5* variants, but the onset time and forms varied greatly. Seizures of a small number of patients occurred only a few times and did not require medical intervention, and most of patients could be controlled by AEDs. In 2014, Zheng et al. reported increased susceptibility to heat-induced seizures in adult Drosophila with *SCAMP* deficiencies (Zheng et al., 2014). However, the susceptibility to heat-induced of epilepsy was not present in our cohort nor in previously reported patients (Hubert et al., 2020; Zhang et al., 2020). This might be because the genome of Drosophila contained only a single *SCAMP*, which differed from that of humans, so the phenotypes did not overlap completely. In our cohort, EEG findings presented as focal or multifocal discharges, sometimes spreading to generalization. Abnormalities of brain MRI were also present in each patient, and those were varied without specificity. Whereas, in previous literature, similar EEG and brain MRI patterns were thought to be present in patients with *SCAMP5* variants (Hubert et al., 2020). Thus, the EEG and MRI findings of patients with *SCAMP5* variants were not consistent, but diverse and nonspecific.

Severe intellectual disability and developmental disorders, including motor and language, were found in six patients, namely four patients here and two patients reported by Hubert et al. (2020) with same variant, whereas this was not seen in affected sibling with homozygous variants. Furthermore, studies in a *Drosophila melanogaster* model showed that the heterozygous variant resulted in developmental and locomotor defects that were absent in homozygous knock-in mutant mice (Hubert et al., 2020; Zhang et al., 2020). Therefore, different mutation site or inheritance patterns of *SCAMP5* may have different effects on neurodevelopment. For motor development, one of them in our cohort experienced a progressive gait deterioration. A similar disease course occurred in the patient reported by Hubert et al. (2020), whose condition gradually worsened and developed an ataxic gait with trunk hypotonia at the age of seven. By knocking out the *SCAMP* in Drosophila, an accelerated age-dependent decline in climbing ability were found in adults (Zheng et al., 2014). Therefore, the motor ability of patients with *SCAMP5* variants might be regressive with age. However, deterioration of motor ability was not observed in most patients. Thus, long-mid term follow-up of additional patients with *SCAMP5* variant was necessary to arrive at firm conclusions. All six patients with *SCAMP5* heterozygous variant had poor language development and were still unable to speak, either as reported in the literature or in our patients (Hubert et al., 2020). We can therefore conclude that language development is more severely affected than motor development.

Abnormal manifestations of neurological exam, especially in the extrapyramidal nervous system, are present in our cohort, especially in the two relatively older patients (ages 8 and 32), whereas the two younger patients only showed mild neuromotor abnormalities (at about the age of two). Acquired movement disorder features also have been described in some patients with *SCAMP5* heterozygous variant (at about the age of 8 and 10) (Hubert et al., 2020). Moreover, the two patients carrying *SCAMP5* homozygous variants also suffered from Parkinson's disease at the age of 18 and 16, respectively (Zhang et al., 2020). The above symptoms suggest that the variants in *SCAMP5* might have an effect on neuromotor system function, especially in the extrapyramidal nervous system, and are age-related. Besides, some dysmorphic features could be seen in half of the patients in our cohort. Dysmorphic limbs features were also noted in two patients reported by Hubert et al. (2020) such as varus equus. Whereas, the two cases of homozygous *SCAMP5* mutation reported by Zhang et al. (2020) did not showed relevant

abnormal descriptions. Thus, dysmorphic characteristics might be specific phenotypes to *SCAMP5* heterozygous variant, or even specifically to G180W.

In this study, all four patients carried the previously identified *de novo* heterozygous *SCAMP5* variants (G180W) (Hubert et al., 2020). Previously, both heterozygous and homozygous *SCAMP5* variants had been associated with human disease (Hubert et al., 2020; Zhang et al., 2020). In 2020, Zhang et al. (2020) reported homozygous R91W variants in siblings with early-onset epilepsy and Parkinson's disease, while the parents with heterozygous R91W variants were normal. And that, the two patients did not show signs of ASD, intellectual disability, or other growth and developmental disorders (Zhang et al., 2020). It therefore appears that heterozygous and homozygous variants of *SCAMP5* are associated with distinct clinical phenotypes and inheritance patterns, and molecular characterization supports such a distinction. The two variants were located in different functional domains of the *SCAMP5*, with G180W located in the C-terminal tail and R91W located in the E peptide. Variants in different functional domains of *SCAMP5* might have different effects on gene functions. Therefore, further studies are needed to clarify the genetics and phenotype of *SCAMP5* in human.

It can be found from the description that *SCAMP5* is associated with a broad phenotype spectrum. *SCAMP5* is known to be brain specific and to be involved in vesicle transport (Fernández-Chacón and Südhof, 2000; Park et al., 2018). It may be due to its wide range of functions, leading to a wide range of clinical phenotypes. Studies have shown that the mutation of *SCAMP5* selectively activated excitatory synaptic signal transduction, leading to excitation-inhibition imbalance and recurrent seizures (Zhang et al., 2020). But the corresponding causes of other phenotypes need to be confirmed and linked with more research. Moreover, the phenotype spectrum and mutation sites also require more cases to be substantiated to determine whether the phenotype is associated with the mutation site or influenced by multiple factors.

CONCLUSION

We identified the same *de novo* heterozygous *SCAMP5* variant in four unrelated patients with a consistent phenotype, including epilepsy, severe developmental delay, abnormal neurological exam findings, with or without ASD or dysmorphic features. The age of onset and type of seizures varied greatly. The EEG and brain MRI findings were not typical, but diverse and nonspecific. Developmental milestones of *SCAMP5* mutated patients show a more severe involvement of language skills, as language is generally absent, but patients may experience a regression of gait abilities. Dysmorphisms are present in some patients but not in all, even if carrying the same *SCAMP5* variant. Finally, variants of *SCAMP5* in different domains might result in different genetic and phenotypic patterns.

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 authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The studies involving human participants were reviewed and approved by the Ethical Committee of Peking University First Hospital, local Ethical Committee of Federico II University of Naples, Cedars-Sinai Medical Center, and Brigham and Women's Hospital. Written informed consent to participate in this study was provided by the legal guardians (parents) of the patient. Written informed consent was obtained from the legal guardians (parents) of the patient for the publication of any potentially identifiable images or data included in this article. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

XW and ZY conceptualized and designed the study, coordinated the study overall, and revised the manuscript; XJ co-designed the study, drafted the initial manuscript, and revised the manuscript; MM, VN, AT, SD'A, CC, CP, PG, KG, PAS-L, JK, EF, and AS helped to collect and summarize data and revised the manuscript. Telethon Undiagnosed Diseases Program contributed to the recruitment and whole exome sequencing (WES) processes of a patient described in the manuscript. All authors approve of the final revision of the article.

FUNDING

This work was supported by National Nature Science Foundation of China (81771393), Beijing Municipal Science and Technology Commission (Z171100001017125), Beijing Natural Science Foundation (7202210), Capital's Funds for Health Improvement and Research (2020-2-4077), Telethon Foundation (Telethon Undiagnosed Diseases Program, TUDP), Fondazione Pierfranco e Luisa Mariani (CM22).

ACKNOWLEDGMENTS

We thank the patients and their families for participating.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2020.599191/full#supplementary-material.

REFERENCES

Ahram, D. F., Al-Sarraj, Y., Taha, R. Z., Elhag, S. F., Al-Shaban, F. A., El-Shanti, H., et al. (2017). A chromosomal microdeletion of 15q in a female patient with epilepsy, ID, and autism spectrum disorder: a case report. *Clin. Case Rep.* 5 (6), 1013–1017. doi:10.1002/ccr3.945

Castermans, D., Volders, K., Crepel, A., Backx, L., De Vos, R., Freson, K., et al. (2010). SCAMP5, NBEA and AMISYN: three candidate genes for autism involved in secretion of large dense-core vesicles. *Hum. Mol. Genet.* 19 (7), 1368–1378. doi:10.1093/hmg/ddq013

Fernández-Chacón, R. and Südhof, T. C. (2000). Novel SCAMPs lacking NPF repeats: ubiquitous and synaptic vesicle-specific forms implicate SCAMPs in multiple membrane-trafficking functions. *J. Neurosci.* 20 (21), 7941–7950. doi:10.1523/jneurosci.20-21-07941.2000

Hubert, L., Cannata Serio, M., Villoing-Gaudé, L., Boddaert, N., Kaminska, A., Rio, M., et al. (2020). *De novo* SCAMP5 mutation causes a neurodevelopmental disorder with autistic features and seizures. *J. Med. Genet.* 57 (2), 138–144. doi:10.1136/jmedgenet-2018-105927

Huynh, M.-T., Lambert, A.-S., Tosca, L., Petit, F., Philippe, C., Parisot, F., et al. (2018). 15q24.1 BP4-BP1 microdeletion unmasking paternally inherited functional polymorphisms combined with distal 15q24.2q24.3 duplication in a patient with epilepsy, psychomotor delay, overweight, ventricular arrhythmia. *Eur. J. Med. Genet.* 61 (8), 459–464. doi:10.1016/j.ejmg.2018.03.005

Klopocki, E., Graul-Neumann, L. M., Grieben, U., Tönnies, H., Ropers, H.-H., Horn, D., et al. (2008). A further case of the recurrent 15q24 microdeletion syndrome, detected by array CGH. *Eur. J. Pediatr.* 167 (8), 903–908. doi:10.1007/s00431-007-0616-7

Law, A. H. Y., Chow, C.-M., and Jiang, L. (2012). Secretory carrier membrane proteins. *Protoplasma* 249, 269–283. doi:10.1007/s00709-011-0295-0

Park, D., Lee, U., Cho, E., Zhao, H., Kim, J. A., Lee, B. J., et al. (2018). Impairment of release site clearance within the active zone by reduced SCAMP5 expression causes short-term depression of synaptic release. *Cell Rep.* 22 (12), 3339–3350. doi:10.1016/j.celrep.2018.02.088

Richards, S., Aziz, N., Aziz, N., Bale, S., Bick, D., Das, S., 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), 405–423. doi:10.1038/gim.2015.30

Sharp, A. J., Selzer, R. R., Veltman, J. A., Gimelli, S., Gimelli, G., Striano, P., et al. (2007). Characterization of a recurrent 15q24 microdeletion syndrome. *Hum. Mol. Genet.* 16 (5), 567–572. doi:10.1093/hmg/ddm016

Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). Gene matcher: a matching tool for connecting investigators with an interest in the same gene. *Hum. Mutat.* 36 (10), 928–930. doi:10.1002/humu.22844

Zhang, D., Yuan, C., Liu, M., Zhou, X., Ge, S., Wang, X., et al. (2020). Deficiency of SCAMP5 leads to pediatric epilepsy and dysregulation of neurotransmitter release in the brain. *Hum. Genet.* 139 (4), 545–555. doi:10.1007/s00439-020-02123-9

Zheng, J. C., Tham, C. T., Keatings, K., Fan, S., Yen-Chun Liou, S., Numata, Y., et al. (2014). Secretory carrier membrane protein (SCAMP) deficiency influences behavior of adult flies. *Front. Cell Dev. Biol.* 2, 64. doi:10.3389/fcell.2014.00064

Conflict of Interest: Author XW was employed by the company Cipher Gene Ltd.

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

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 21 January 2021
doi: 10.3389/fphar.2020.612200

Role of GABRD Gene Methylation in the Nucleus Accumbens in Heroin-Seeking Behavior in Rats

*Qingxiao Hong[1,2], Wenjin Xu[1,2], Zi Lin[1], Jing Liu[1], Weisheng Chen[1,2], Huaqiang Zhu[1,2], Miaojun Lai[1,2], Dingding Zhuang[1,2], Zemin Xu[1,2], Dan Fu[1,2], Wenhua Zhou[1,2] and Huifen Liu[1,2]**

[1]Laboratory of Behavioral Neuroscience, Ningbo Kangning Hospital, Ningbo Institute of Microcirculation and Henbane, School of Medicine, Ningbo University, Ningbo, China, [2]Key Laboratory of Addiction Research of Zhejiang Province, Ningbo, China

Epigenetic modifications such as DNA methylation play important roles in regulating gene expression and may mediate neuroplasticity and lead to drug-induced aberrant behaviors. Although several brain regions and neurobiological mechanisms have been suggested to be involved in these processes, there is remarkably little known about the effects of DNA methylation on heroin-seeking behavior. Using a Sprague-Dawley rat model, we show that heroin self-administration resulted in gamma-aminobutyric acid type A receptor subunit delta (GABRD) gene hypomethylation, which was associated with transcriptional upregulation of GABRD in the nucleus accumbens (NAc). Systemic L-methionine (MET) administration significantly strengthened the reinstatement of heroin-seeking behavior induced by heroin priming, whereas intra-NAc injections of the DNA methyltransferase (DNMT) inhibitor 5-aza-2′-deoxycytidine (5-Aza-dC) had the opposite effect on heroin-seeking. Meanwhile, 5-Aza-dC treatment decreased DNA methylation and upregulated the expression of GABRD in the NAc, whereas MET had the opposite effect. Our results also reveal that 5-Aza-dC might alter the methylation landscape of the GABRD gene by directly repressing DNMT1 and DNMT3A expression. Furthermore, reinstatement of heroin-seeking behavior was significantly inhibited by directly overexpressing GABRD and remarkably reinforced by GABRD gene silencing in the NAc. Collectively, these results suggest that targeting the GABRD gene and its methylation might represent a novel pharmacological strategy for treating heroin addiction and relapse.

Keywords: heroin, nucleus accumbens, GABRD, methylation, reinstatement

OPEN ACCESS

Edited by:
Tieliu Shi,
East China Normal University, China

Reviewed by:
Tao Jiang,
University of Illinois at Urbana-Champaign, United States
Francesca Felicia Caputi,
University of Bologna, Italy

****Correspondence:***
Huifen Liu
liuhuifen91@163.com

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *30 September 2020*
Accepted: *27 November 2020*
Published: *21 January 2021*

Citation:
Hong Q, Xu W, Lin Z, Liu J, Chen W, Zhu H, Lai M, Zhuang D, Xu Z, Fu D, Zhou W and Liu H (2021) Role of GABRD Gene Methylation in the Nucleus Accumbens in Heroin-Seeking Behavior in Rats. Front. Pharmacol. 11:612200. doi: 10.3389/fphar.2020.612200

INTRODUCTION

Heroin addiction is a chronic and relapsing brain disorder characterized by compulsive drug-seeking and the negative emotional state of withdrawal. Relapse to drug-seeking after abstinence has long been a challenge in the treatment of heroin addiction. Heroin addiction is characterized by specific behavioral alterations, indicating long-lasting alterations in gene and protein expression within specific reward-related brain regions. Accumulating evidence has demonstrated that epigenetic mechanisms such as DNA methylation regulate drug-induced gene expression profiles and enduring behavioral phenotypes (Nestler, 2014; Kalda and Zharkovsky, 2015). Recent data suggest that altered DNA methylation may indicate the changed gene expression

programs in response to experience (Meaney and Szyf, 2005; Klengel et al., 2013) and regulate the synaptic plasticity as well as be implicated in memory formation (LaPlant et al., 2010; Zovkic et al., 2013). Epigenetic modifications may also be a critical molecular mechanism that indicated the enduring addiction-related changes in brain plasticity (Robison and Nestler, 2011; Godino et al., 2015; Massart et al., 2015).

DNA methyltransferases (DNMTs) are responsible for adding methyl groups to cytosine-guanine dinucleotides (CpGs) in the genome (Suzuki and Bird, 2008). In mammals, the main DNMTs include DNMT1 which is responsible for DNA methylation maintenance and another two "*de novo*" methyltransferases that establish new methylation patterns (DNMT3A and DNMT3B) (Ribas et al., 2017). CpG islands methylation recruits co-repressor complexes to interfere with the binding of transcription factor to target DNA sequences (Jaenisch and Bird, 2003). Considerable available evidence demonstrates that drug-induced changes in gene expression are modulated by DNA methylation. Tian et al. (Tian et al., 2012) found that rewarding effects induced by cocaine could be significantly attenuated by reversing the global DNA hypomethylation. Besides, the hippocampus of juvenile rats exposed *in utero* to cocaine shows altered patterns of global DNA methylation and changed gene transcription, for example, genes encoding G-protein coupled receptor 73 (GPR73), polo-like kinase 2 (PLK2) and protein-tyrosine phosphatase non-receptor type 5 (PTPN5) (Novikova et al., 2008). Changes in gene expression occurs after cocaine self-administration, and the changes is shown to be correlate with increased expression of the methyl CpG-binding protein MeCP2 (Host et al., 2011). Additionally, methamphetamine administration may alter DNA methylation patterns through regulation of DNMT1 mRNA levels (Numachi et al., 2007). Interestingly, a recent clinical study showed that methamphetamine-induced changes in long interspersed element-1 methylation are associated with methamphetamine-induced paranoia and therefore may partially explain the pathophysiology of this type of psychosis (Kalayasiri et al., 2019). Previous preclinical study reported that κ1 opioid receptor (OPRK1) promoter methylation was significantly correlated with the length and frequency of drug use in Chinese male heroin addicts (Ji et al., 2018). Xu et al. (2016) demonstrated that methylation level of one CpG sites in brain derived neurotrophic factor (BDNF) promoter was significantly associated with addictive phenotypes of heroin-dependent individuals, including tension-anxiety, anger-hostility, fatigue-inertia, and depression-dejection. However, there are relatively few reports on animal model relevant to heroin addiction, little is known about the role of DNA methylation regulatory events in mediating the lasting effects of heroin-seeking behavior.

γ-aminobutyric acid (GABA) A receptors (GABAARs) are the crucial inhibitory neurotransmitter receptors in the central nervous system. As a class of transmembrane ligand-gated chloride channels, GABAARs consist of five subunits from the 19 known different subunit isoforms, α1–6, β1–3, γ1–3, δ, ε, θ, π, and ρ1–3 (Whiting, 2003; Olsen and Sieghart, 2009). The channels are usually formed by two α subunits, two β subunits, and one other subunit, and the resulting compositional differences are fundamentally responsible for the different functional and pharmacological properties of the channels (Brickley and Mody, 2012). Given that GABAARs are considered to be key targets for alcohol, sleep-promoting drugs and anaesthetics, their importance in regulating schizophrenia, epilepsy, and drug addiction has been highlighted. Previous studies reported that the expression of some GABAAR subunits, such as GABA type A receptor subunit (GABR) beta2 and GABR alpha2, was regulated by DNA methylation in mood disorders (Wang et al., 2016; Gatta et al., 2017; Ye et al., 2019). Besides, recent data suggest that DNA methylation is functionally correlated with the persistent cocaine-craving and may be partly negatively correlated with gene expression changes, for example, that of the GABR delta (GABRD) gene. Furthermore, a previous study showed that mRNA and protein levels of the GABRD were negatively modulated by promoter methylation in the cerebellum of individuals with alcohol use disorder (Gatta et al., 2017). However, no study has reported a correlation between GABRD methylation and heroin addiction.

In the present study, we combine global whole-genome and candidate gene approaches to investigate the effect of DNA methylation on heroin-seeking behavior by analysing DNA methylation profiles and the correlated gene expression changes in the NAc of rats. Our research also explores whether DNA methylation regulates GABRD gene transcription in the NAc after repeated heroin treatment, and whether the pharmacological enhancement or inhibition of DNMTs alters heroin self-administration and the reinstatement of heroin-seeking behavior in rats. Finally, we investigate whether overexpression or down-regulation of GABRD in the NAc of rats significantly alters the associated cues-induced or heroin priming reinstatement after extinction of heroin self-administration.

RESULTS

Global DNA Methylation

We performed whole-genome methylation sequencing to investigate changes in methylation profiles after heroin self-administration (PRJNA673675, SRA records will be accessible with the following link after the indicated release date 2024–10–01, https://www.ncbi.nlm.nih.gov/sra/PRJNA673675). On average, more than 331,106,942 sequence reads were obtained per sample, the differentially methylated regions (DMRs) were identified between the three groups. The methylKit and eDMR R package was used for comprehensive DMRs analysis, the eDMR can optimize regional methylation analysis based on bimodal normal distribution and weighted cost function. A comparative analysis revealed that there were 38,177 DMRs between the heroin self-administration and the yoked-saline group, and 23,129 DMRs existed between the yoked-heroin and the yoked-saline group. As shown in **Figure 1A**, the pink [Heroin Self-administration vs. Yoked-saline group (Control)], green (Yoked-heroin vs. Yoked-saline group) and blue (Heroin

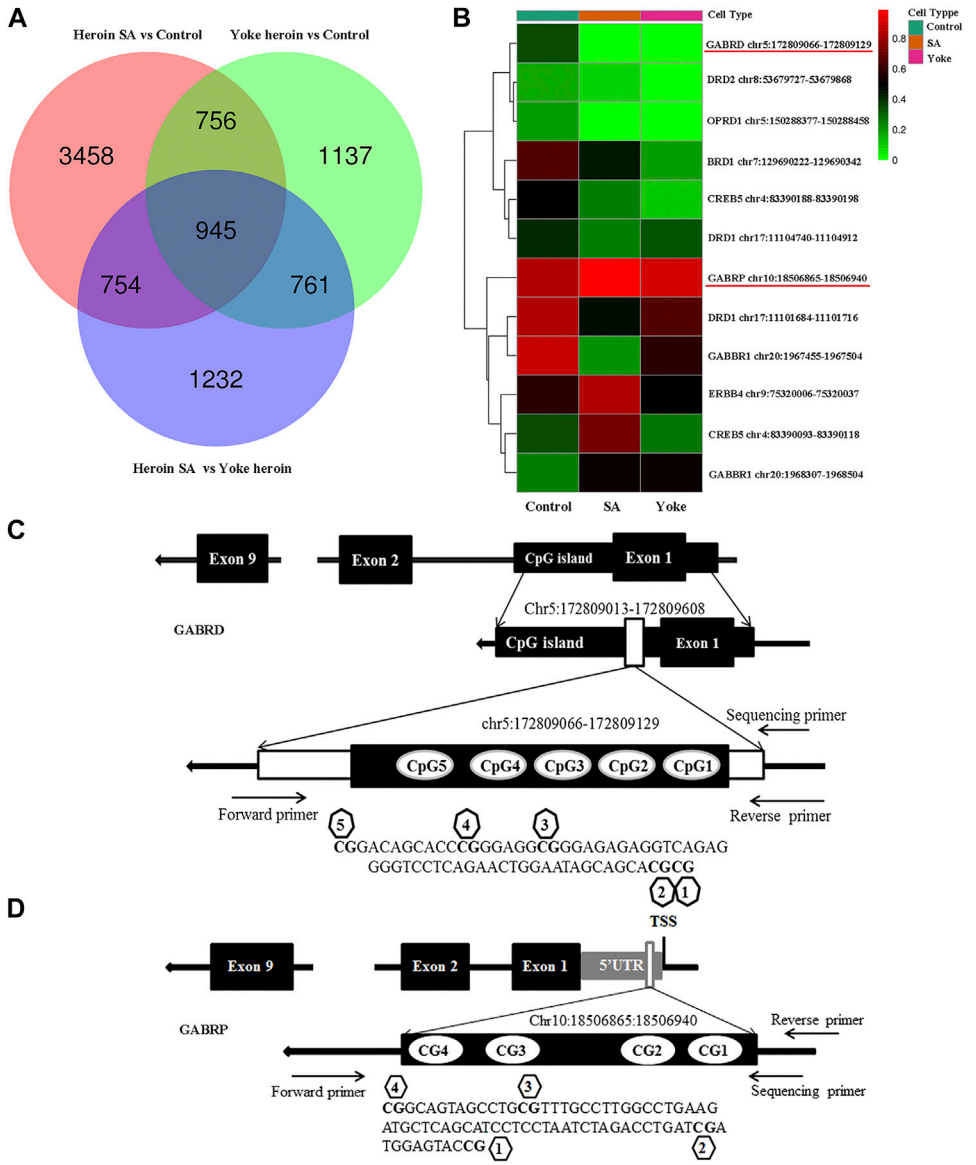

FIGURE 1 | DNA methylation of heroin self-administration rats and the locations of target sequences. **(A)** Venn diagram of promoter-DMRs related genes in the paired comparison of the three groups of rats (n = 3 per group). SA: self-administration. **(B)** A hierarchical clustering map of NAc gene methylation for yoked-saline (control), heroin self-administration, and yoked-heroin rats (red, high methylation; green, low methylation). **(C)** The target sequence of GABRD is located in a CpG island which contains 5 CpG sites. **(D)** The target sequence of GABRP is located in the 5′UTR which contains four CpG sites. SA: Self-administration, Control: yoked-saline group.

Self-administration vs. Yoked-heroin group) areas represent the genes that differentially methylated in promoter regions in the two corresponding groups respectively. And their area of overlap represents the core. In total, the co-existing 1,701 (756 + 945) genes were significantly differentially methylated in their promoter regions in the yoked-saline group compared against the heroin self-administration and the yoked-heroin group, respectively. The core value 945 in the middle represent the DMRs co-existing in the three sets. Some focused hypermethylated and hypomethylated genes are shown in **Figure 1B**. Of these, the methylation level of the GABRD gene was significantly reduced in the heroin self-administration and yoked-heroin groups compared with that in the yoked-saline group (DMR: chr5:172809066–172809129, located within CpG islands 1 kb downstream of the transcription start site (TSS) (**Figure 1C**) ($p < 0.05$, $p < 0.05$). Besides, methylation level of the GABRP gene (DMR: chr10:18506865–18506940, located within the 5′UTR) was shown to be increased in the heroin-self-administration compared with that in the yoked-saline group ($p < 0.05$) (**Figure 1D**).

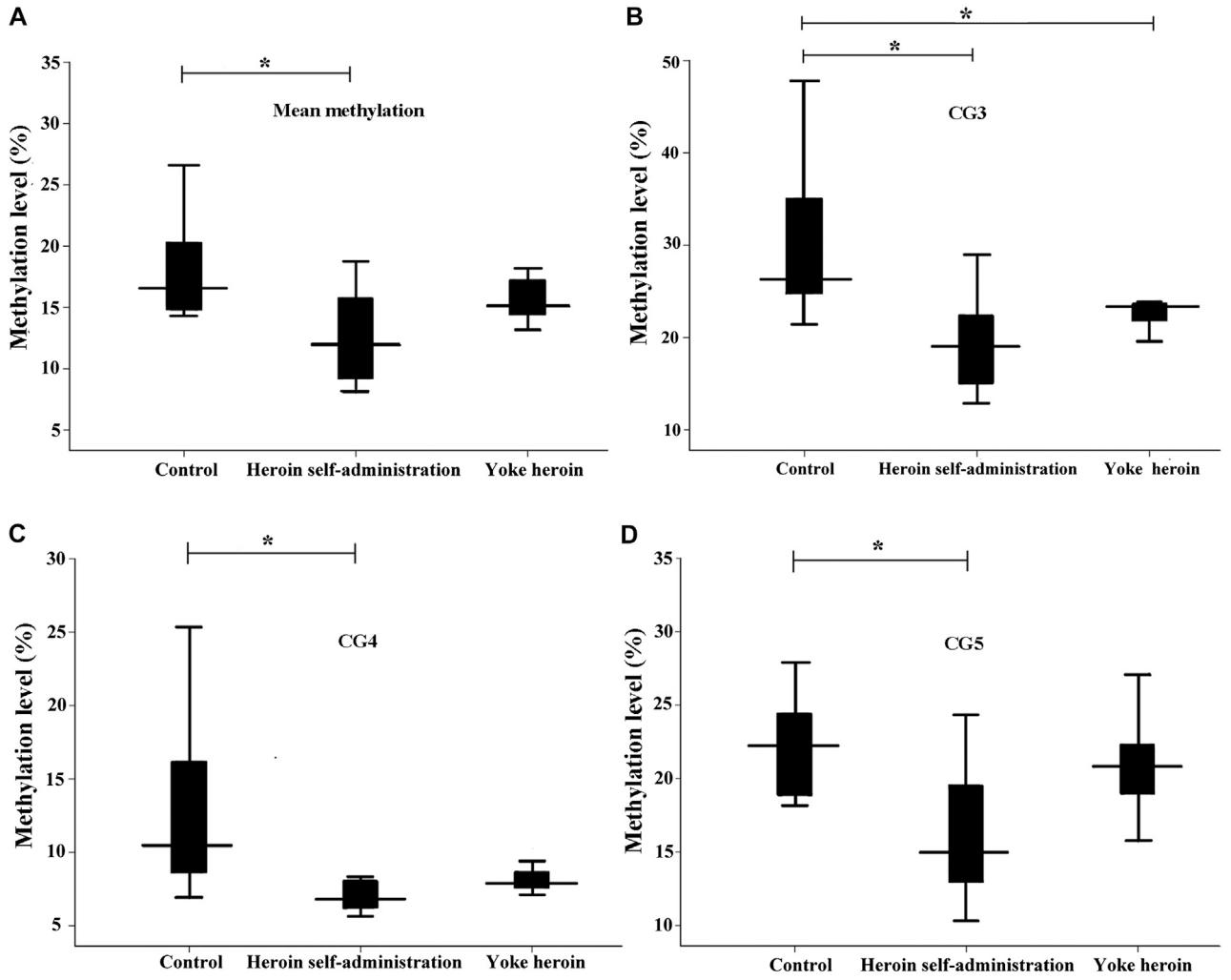

FIGURE 2 | Comparisons of GABRD methylation levels among three groups (n = 10 per group); Heroin self-administration group, which received heroin self-administration for 14 days, and iii) yoked-heroin group, which received amounts of heroin equal to those received self-administration over the same course of time (Yoke heroin). The methylation levels were calculated and the statistical differences evaluated by one-way ANOVA followed by Student–Newman–Keuls post hoc comparisons. Data were presented as means ± SDs. *p < 0.05.

Confirmation of the Changes in GABRD and GABRP Gene Methylation in Heroin-Administration Rats

To further confirm the results, pyrosequencing was performed. Significant correlations among five CpG sites in GABRD were observed (r > 0.50, p < 0.05). Therefore, the average percentage value for 1–5 CpG sites were used for subsequent analysis. The average GABRD methylation levels were significantly reduced in heroin self-administration group as compared to those in the yoked-saline group [$F_{(2,26)}$ = 5.91, p = 0.008; **Figure 2A**]. Further analysis of different CpG cites showed that methylation levels at CG3, CG4, and CG5 sites were decreased in the heroin self-administration group compared with those in the yoked-saline group [CpG3: $F_{(2,26)}$ = 7.82, p = 0.002, CpG4: $F_{(2,26)}$ = 4.95, p = 0.015, CpG5: $F_{(2,26)}$ = 6.18, p = 0.006, **Figures 2B–D**]. Additionally, the CG3 site exhibited reduced methylation in yoked-heroin group compared with that in the yoked-saline group (p = 0.042). However, the CG1 and CG2 sites exhibited no changes [CpG1: $F_{(2,26)}$ = 2.71, p = 0.085, CpG2: $F_{(2,26)}$ = 5.91, p = 0.058].

The sequence of the GABRP promoter that contains four CpG sites is shown in **Figure 1D**. Significant correlations between CG1 and CG2 were observed (r = 0.63, p = 1.08E−4); besides, there were significant correlation between CG3 and CG4 (r = 0.92, p = 1.62E−13). In total, the average methylation levels of GABRP were obviously increased in heroin self-administration and yoked-heroin groups compared with that in the yoked-saline group [$F_{(2,26)}$ = 16.77, p = 1.45E−5, **Figure 3A**]. Meanwhile, the methylation level at the CG3 site was significantly increased in yoked-heroin group compared with those in yoked-saline group and heroin self-administration groups [$F_{(2,29)}$ = 38.80, p = 6.34E−9, **Figure 3C**]. Moreover, there were statistically significant differences in CG3 + CG4 sites and CG4 sites

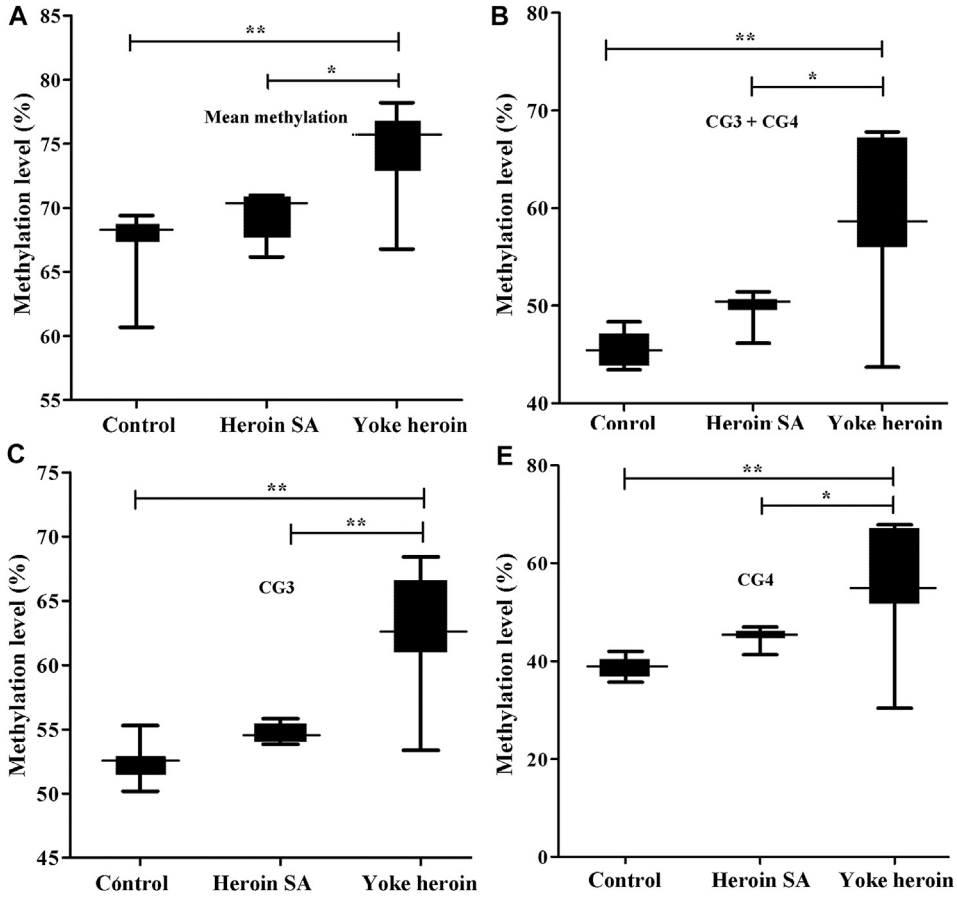

FIGURE 3 | Comparisons of GABRP methylation levels among the three groups (n = 10 per group). SA: Self-administration. The methylation levels were calculated and the statistical differences evaluated by one-way ANOVA followed by Student–Newman–Keuls post hoc comparisons. Data were presented as means ± SDs. *p < 0.05; **p < 0.001.

among the three groups [$F_{(2,29)}$ = 23.41, p = 8.87E−7, **Figure 3B**; $F_{(2,29)}$ = 16.62, p = 1.55E−5, **Figure 3D**]. Significant GABRP hypermethylation in heroin-self-administration and yoked-heroin groups, compared with that in the yoked-saline group were detected.

The relative mRNA expression of GABRD and GABRP were determined by RT-qPCR; one-way ANOVA showed that GABRD expression was significantly different in the three groups [$F_{(2,14)}$ = 4.86, p = 0.025]. The multiple comparisons illustrated that GABRD mRNA expression in the heroin self-administration group increased compared with that in the yoked-saline and yoked-heroin groups (p = 0.031 and p = 0.012, respectively **Figure 4A**). Whereas GABRP mRNA expression remained unchanged in three groups [$F_{(2,14)}$ = 0.056, p = 0.946]. These results revealed a significant negative correlation between GABRD gene methylation and mRNA expression in the NAc.

Detection of Enhancer Activities of the GABRD Fragments

Results of the double luciferase reporter gene assay showed that the fluorescence ratios among the pGL3-promoter, pGL3-promoter-GABRD, and Me-pGL3-promoter-GABRD groups were significantly different [$F_{(2,6)}$ = 1.97E3, p = 3.51E−9]. The post hoc comparisons showed fluorescence ratios in pGL3-promoter-GABRD which contains the GABRD segment that included five CpG sites were significantly higher than that in basal reporter plasmids (pGL3-promoter vector). These results revealed that the candidate fragment of GABRD had somehow enhancer-like activity (p < 0.05, **Figure 4B**). We then investigated the fluorescence ratios in the unmethylated (Pro-GABRD) or methylated forms of GABRD segment (Me-Pro-GABRD), the result demonstrated that fluorescence activity was significantly reduced in Me-Pro-GABRD compared with that in Pro-GABRD, these findings provided evidence that the methylated form of GABRD can weaken the gene activity significantly (p < 0.001, **Figure 4B**).

The Effects of MET on the Acquisition of Heroin Self-Administration and Heroin Reinstatement

Rats was injected with MET (s.c.) at 0.5 h prior to heroin self-administration training from the 10th day of training for 12

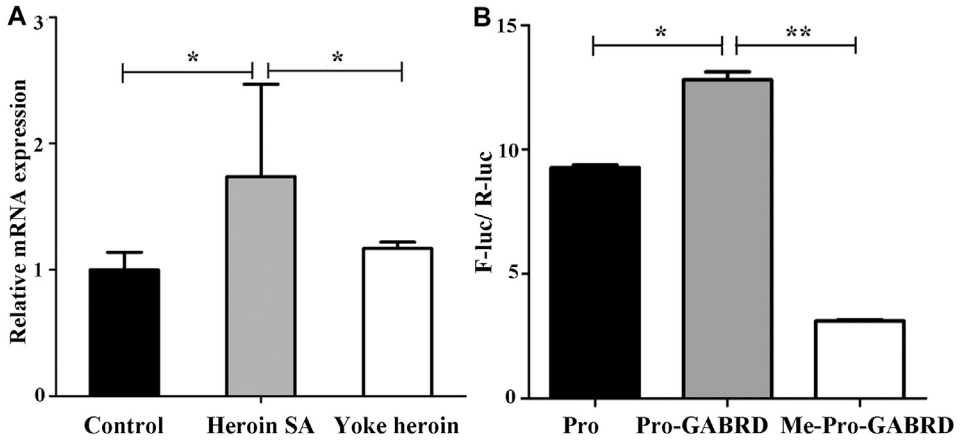

FIGURE 4 | GABRD mRNA expression level and enhancer-like activity of target sequence. **(A)** Relative mRNA expression of GABRD in the three groups (Control, $n = 6$; Heroin SA, $n = 6$; Yoke heroin, $n = 5$). **(B)** Dual luciferase reporter gene assay for GABRD ($n = 3$ per group). Control: Yoked-saline group; SA: Self-administration; Pro: pGL3-promoter vector; Me: Methylation; F-Luc: Firefly luciferase; R-Luc: Renilla luciferase. The relative mRNA expression of GABRD and the fluorescence ratios among three groups were compared using one-way ANOVA followed by Student–Newman–Keuls post hoc comparisons. The Data were presented as means ± SDs. $^{*}p < 0.05$; $^{**}p < 0.001$.

consecutive days. The two-way repeated-measures ANOVA indicated that the number of infusions and active nose-poke responses during heroin self-administration training after MET treatment showed no significant statistical differences [$F_{(1,9)}$ = 4.175, p = 0.071] [$F_{(1,9)}$ = 2.176, p = 0.157, **Figures 5A–C**]. Notably, MET-treated rats had significantly higher number of active nose-poke responses during the reinstatement induced by heroin priming compared with saline-treated rats [$F_{(1,27)}$ = 4.890, p = 0.036, **Figure 6B**] and showed an increasing trend in number of active nose-poke responses during the reinstatement induced by heroin cues; however, these differences were not statistically significant [$F_{(1,29)}$ = 0.507, p = 0.485, **Figure 6A**].

The Effects of MET on Regulation of GABRD Gene Methylation and mRNA Expression

Pyrosequencing was used to determine the effects of methyl supplementation on GABRD methylation. **Figure 6C** showed the methylation state of the five GABRD CpGs in the NAc. Significant correlations among these five CpG sites were observed ($r > 0.87$, $p < 0.001$). Two-way ANOVA revealed a significant main effect for MET [$F_{(1,\ 17)}$ = 5.41, p = 0.036] and interaction (MET and heroin) [$F_{(1,\ 17)}$ = 6.54, p = 0.023] on GABRD gene methylation; however, no effect of heroin on methylation was observed [$F_{(1,\ 17)} = 0.574$, $p = 0.461$]. Thereafter, post hoc multiple comparisons were performed, and the results showed that the GABRD gene exhibited an overall increase in DNA methylation in the Her + MET group, and there were significant differences between Her + MET vs. Her + Sal ($p = 0.003$), Her + MET vs. Con + MET (p = 0.034), and Her + MET vs. Con + Sal (p = 0.047) groups (**Figure 6C**). Additionally, there were significant effects of MET at CpG2 [$F_{(1,\ 17)} = 4.87$, $p = 0.045$], CpG3 [$F_{(1,\ 17)} = 5.83$, p = 0.030], and CpG4 [$F_{(1,\ 17)}$ = 6.54, p = 0.023]. Meanwhile, significant interaction of MET treatment and heroin effects at CpG3 [$F_{(1,\ 17)}$ = 6.87, p = 0.02] and CpG5 [$F_{(1,\ 17)}$ = 6.41, p = 0.024] were detected. Conversely, contrary to the methylation levels, mRNA expression level in the Her + MET group was obviously lower than that in the Her + Sal group [$F_{2,\ 28)}$ = 3.686, p = 0.038, **Figure 6D**].

The Effects of 5-Aza-dC on Heroin Priming Reinstatement

Given the ability of MET to reinforce the reinstatement of heroin-seeking induced by heroin priming, we hypothesized that microinjection with the methyltransferase inhibitor 5-Aza-dC might attenuate the reinstatement of heroin-seeking. Thus, we treated rats with intra-NAc microinjection of 5-Aza-dC or aCSF in the last 1, 3, 5 days of extinction training, followed by heroin priming reinstatement testing. The two-way ANOVA analysis revealed significant main effects of 5-Aza-dC [$F_{(1,\ 19)}$ = 14.402, p = 0.001], heroin [$F_{(1,19)}$ = 28.475, p = 5.46E−5], and interaction (heroin and 5-Aza-dC) [$F_{(1,19)}$ = 16.367, p = 8.40 E−4] on active responses during heroin reinstatement. Furthermore, the post hoc analysis showed that heroin priming reinstatement was prevented by pre-treatment with 5-Aza-dC (Her + AZA vs. Her + aCSF, p = 1.47 E−4, **Figure 7A**). Meanwhile, the active responses of heroin reinstatement obviously increased in the Her + aCSF group compared with that in the Con + aCSF group (p = 9.06 E−6) (**Figure 7A**).

The Effects of 5-Aza-dC on Methylation and the Expression of GABRD and DNMTs in the NAc

Pyrosequencing was used to investigate the effect of 5-Aza-dC on DNA methylation. The results revealed significant correlations among the five GABRD CpG sites ($r > 0.743$, $p < 0.001$). Two-way ANOVA showed a significant main effect of 5-Aza-dC [$F_{(1,\ 21)}$ = 21.83, p = 1.89E−4] and heroin [$F_{(1,\ 21)}$ = 12.99, p = 0.002] on the average methylation levels of these five CpG sites. Meanwhile,

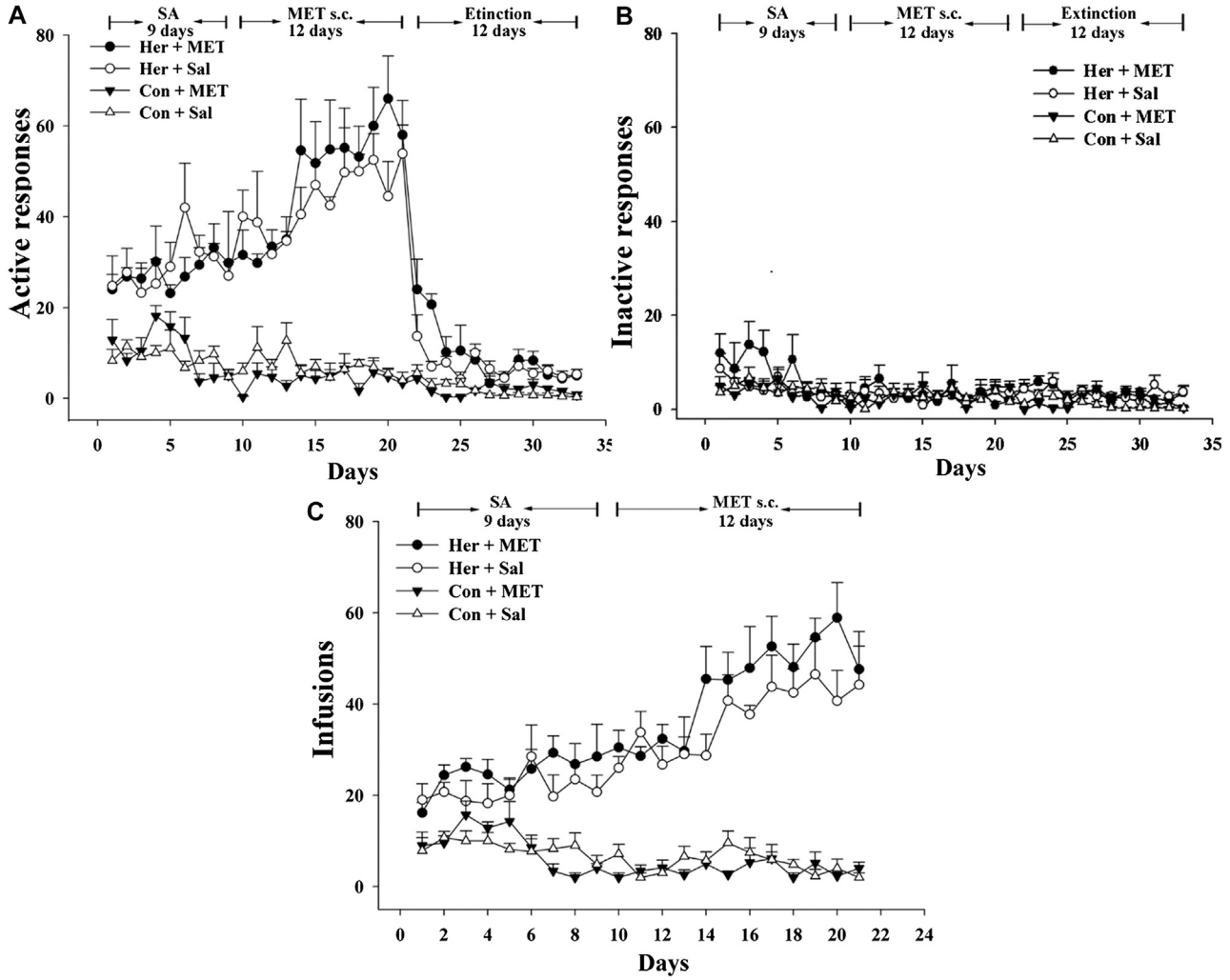

FIGURE 5 | Effect of systemic injection of MET on heroin self-administration (n = 5 per group). **(A)** Active, and **(B)** inactive nose-poke responses during training. **(C)** The number of infusions per session. Her + MET: heroin self-administration and MET-treated rats, Her + Sal: heroin self-administration and saline-treated rats, Con + MET: saline self-administration and MET-treated rats, Con + Sal: saline self-administration and saline-treated rats. The active, inactive nose-poke responses and infusions were measured by two-way repeated-measures ANOVA with treatment group and time as factors. Data were expressed as means ± SDs. SA: self-administration; s.c.: subcutaneous injection; MET: L-methionine; Sal: Saline.

there was significant effect of 5-Aza-dC at the CpG1 [$F_{(1,\ 21)}$ = 10.95, p = 0.004], CpG2 [$F_{(1,\ 21)}$ = 8.55, p = 0.009), CpG3 [$F_{(1,\ 21)}$ = 30.60, p = 2.98E−5], CpG4 [$F_{(1,\ 21)}$ = 12.71, p = 0.002], and CpG5 [$F_{(1,\ 21)}$ = 33.42, p = 1.77E−5] sites. Simultaneously, significant effects of heroin at CpG1 [$F_{(1,\ 21)}$ = 11.64, p = 0.003], CpG2 [$F_{(1,\ 21)}$ = 13.97, p = 0.002], CpG3 [$F_{(1,\ 21)}$ = 16.59, p = 0.001], CpG4 [$F_{(1,\ 21)}$ = 4.64, p = 0.045], and CpG5 [$F_{(1,\ 21)}$ = 10.35, p = 0.005] were detected. The significant effect of interaction was only indicated at CpG3 [$F_{(1,\ 21)}$ = 10.00, p = 0.005]. Furthermore, the multiple comparisons revealed that average methylation levels of the GABRD gene were significantly decreased after treatment with 5-Aza-dC (Her + AZA vs. Her + aCSF, p = 0.038, Con + AZA vs. Con + aCSF, p = 4.64E−4] (**Figure 7B**). Additionally, western blotting showed that treatment with 5-Aza-dC significantly altered DNMT1 and DNMT3A expression in NAc, but DNMT3B remained unaffected (**Figures 8A–D**). Two-way ANOVA revealed that DNMT1 and DNMT3A were down-regulated after 5-Aza-dC treatment in the NAc [$F_{(1,15)}$ = 6.10, p = 0.029 and $F_{(1,15)}$ = 9.79, p = 0.009] for main effect of 5-Aza-dC, respectively; $F_{(1,15)}$ = 5.50, p = 0.037 and $F_{(1,15)}$ = 5.40, p = 0.038 for main effect of heroin self-administration, respectively. However, no interaction effect 5-Aza-dC and heroin self-administration was observed [$F_{(1,15)}$ = 1.65, p = 0.223 and $F_{(1,15)}$ = 0.218, p = 0.649, respectively], but DNMT3B expression remained unaffected [$F_{(1,15)}$ = 0.002, p = 0.963 for 5-Aza-dC effect; $F_{(1,15)}$ = 0.456, p = 0.512 for heroin effect; $F_{(1,15)}$ = 0.735, p = 0.408 for interaction effect]. Post hoc multiple comparisons showed that DNMT1 and DNMT3A levels decreased in the heroin self-administration group after treatment with 5-Aza-dC in the NAc (Her + AZA vs. Her + aCSF, p = 0.021, p = 0.026, respectively, **Figures 8B,C**). However, there were decreasing trends of DNMT1 and DNMT3A protein expression levels after pretreatment with 5-

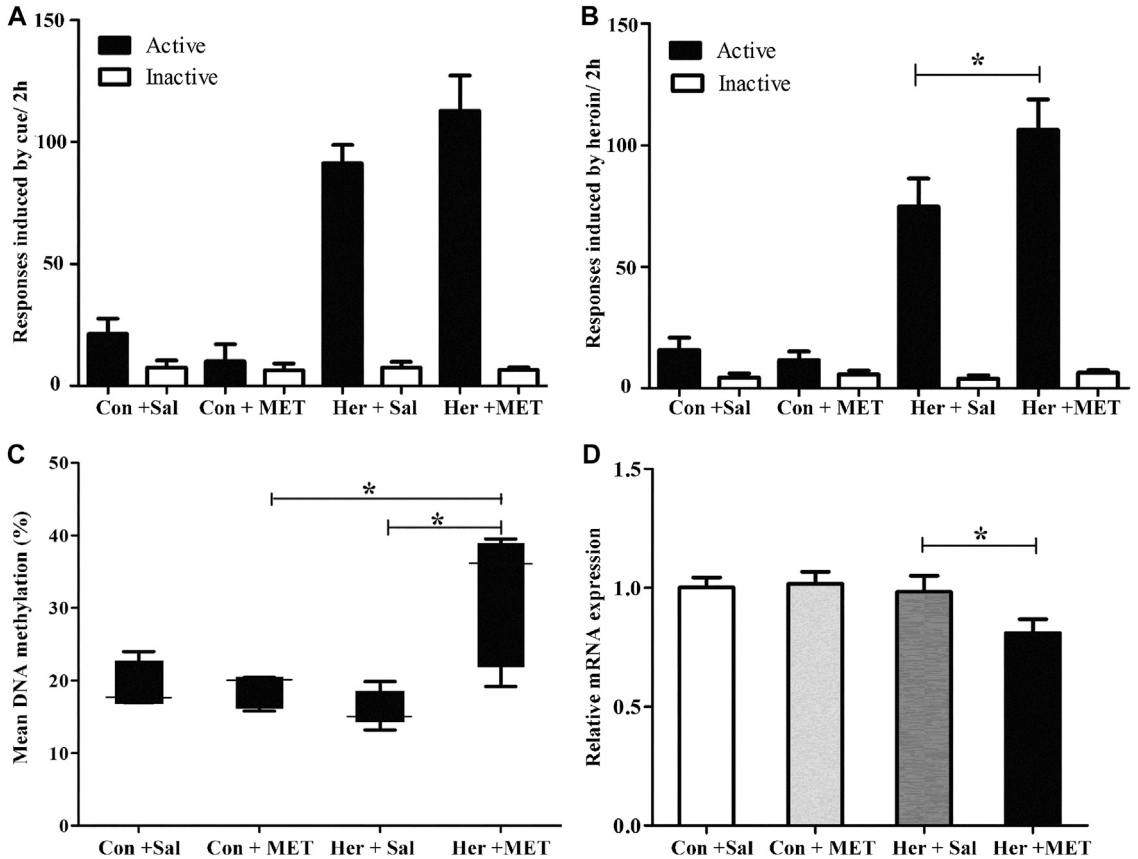

FIGURE 6 | Effect of systemic injection of MET on the reinstatement of heroin-seeking behavior and GABRD gene methylation and mRNA expression levels. **(A)** Systemic injection of MET had no effect on reinstatement of heroin-seeking behavior induced by heroin priming. **(B)** Systemic injection of MET enhanced reinstatement of heroin-seeking behavior induced by cues. **(C)** Systemic injection of MET increased the gene methylation of GABRD in heroin self-administration group. **(D)** Systemic injection of MET decreased the mRNA expression level of GABRD in heroin self-administration group. Con: saline self-administration; Her: Heroin; MET: L-methionine; Sal: Saline. The active and inactive nose-poke responses were measured by two-way repeated-measures ANOVA with treatment group and time as factors. Two-way ANOVA was used to analyze methylation level or mRNA expression of GABRD, with MET treatment and heroin self-administration as the independent variables. Data were expressed as means ± SDs. *$p < 0.05$.

Aza-dC in control group, although the differences were not statistically significant (Con + AZA vs. Con + aCSF, $p = 0.418$, $p = 0.084$, respectively, **Figures 8B,C**). Correspondingly, the level of the GABRD in the NAc was significantly up-regulated after intra-NAc microinjection of 5-Aza-dC (**Figures 9A,B**, Heroin + AZA vs. Heroin + aCSF: $p = 0.018$, Con + AZA vs. Con + aCSF: $p = 0.010$).

Effect of MET and 5-Aza-dC Treatment on Locomotion Activity

As shown in **Figure 9C**, statistical analysis revealed no significant effects of either systemic injection of MET or intra-NAc microinjection of 5-Aza-dC on the locomotor activities of rats compared with the corresponding vehicle control [MET vs. $\text{Saline}_{(s.c.)}$: 1,428.71 ± 151.36 vs. 1,585.29 ± 147.07, $p = 0.42$; 5-Aza-dC vs. $\text{aCSF}_{(intra\text{-}NAc)}$: 1,640.86 ± 204.93 vs. 1,456.29 ± 173.56, $p = 0.28$].

Effect of GABRD Overexpression or RNAi on Heroin Reinstatement

We investigated whether GABRD overexpression or RNAi in the NAc of rats would cause changes in heroin reinstatement. Our results showed that over-expression of GABRD in the NAc had no effect on cues-induced reinstatement of heroin-seeking behavior [$F_{(1,14)} = 0.083$, $p = 0.78$, **Figure 10A**], but the heroin priming reinstatement was significantly inhibited [$F_{(1,14)} = 8.83$, $p = 0.01$, **Figure 10B**]. At the same time, western blotting showed that the level of GABRD expression in the NAc significantly increased ($p = 0.038$, **Figures 10C,D**).

Our results also suggest that GABRD RNAi in the NAc could significantly enhance the cues-induced or heroin priming reinstatement [$F_{(1,16)} = 4.86$, $p = 0.04$; $F_{(1,16)} = 14.33$, $p = 0.002$, **Figures 11A,B**]. Correspondingly, the level of GABRD expression significantly decreased after microinjected with GABRD RNAi AAV vector ($p = 0.031$, **Figures 11C,D**).

DISCUSSION

In the present study, we revealed that heroin self-administration induced global DNA methylation changes in the NAc. The DMRs of some GABAARs were identified, which are responsible for

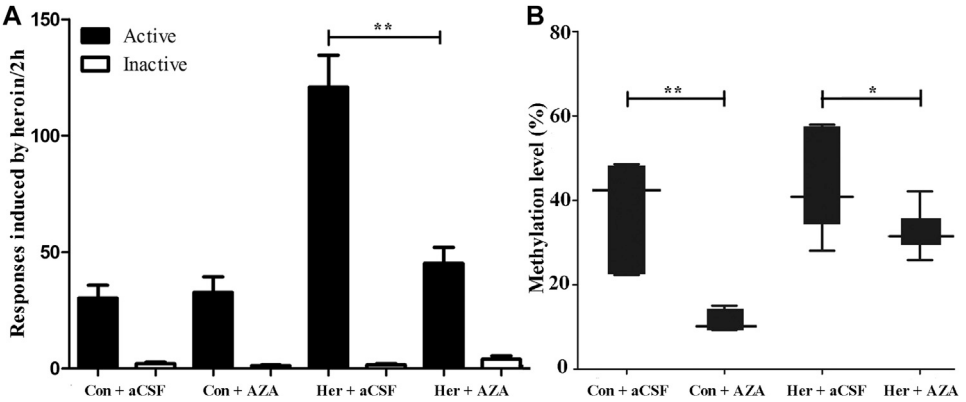

FIGURE 7 | Effect of intra-NAc microinjection of 5-Aza-dC on the reinstatement of heroin-seeking and GABRD gene methylation (n = 5 per group). **(A)** Intra-NAc microinjection of 5-Aza-dC abolished reinstatement of heroin-seeking behavior induced by heroin priming. **(B)** Intra-NAc microinjection of 5-Aza-dC inhibited the DNA methylation of GABRD. Con: saline self-administration; Her: Heroin; aCSF: artificial cerebrospinal fluid solution; AZA: 5-Aza-dC, 5-Aza-2-deoxycytidine. The active and inactive nose-poke responses were measured by two-way repeated-measures ANOVA with treatment group and time as factors. Two-way ANOVA followed by Student–Newman–Keuls post hoc comparisons was used to analyze methylation level of GABRD, with AZA treatment and heroin self-administration as the independent variables. Data were expressed as means ± SDs. $^{*}p < 0.05$; $^{**}p < 0.001$.

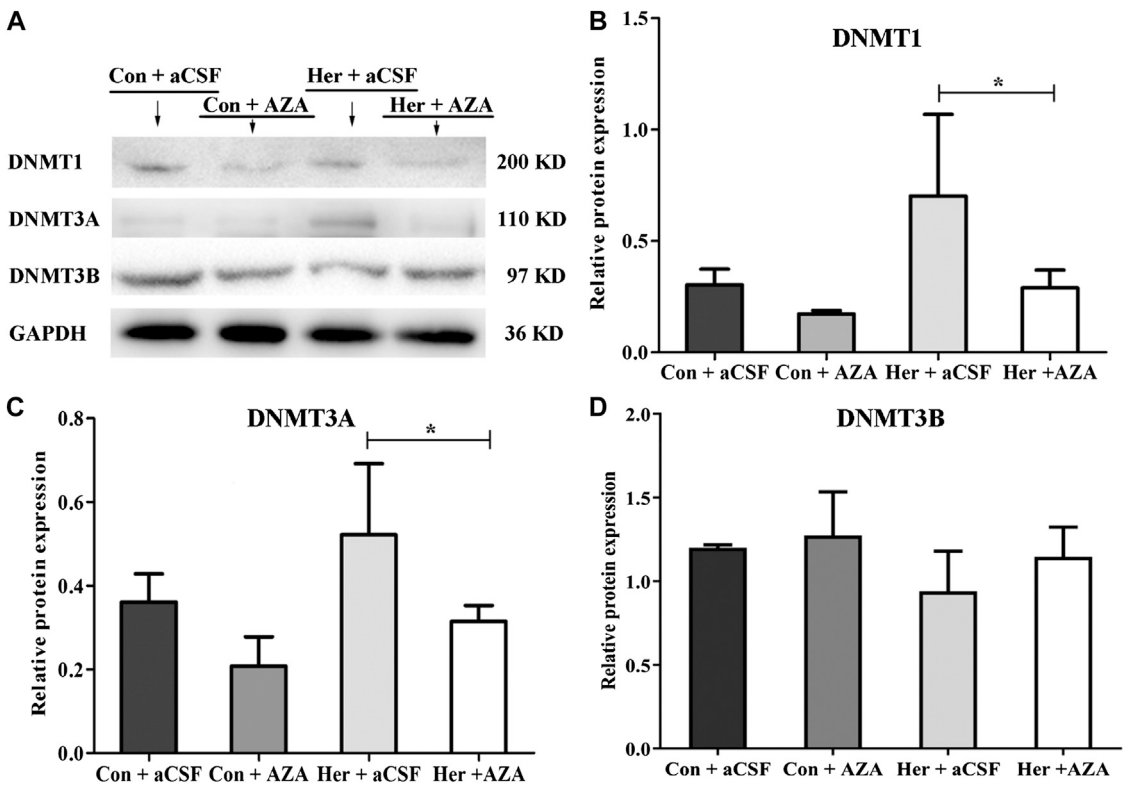

FIGURE 8 | Effect of 5-Aza-dC treatment on DNMT protein expression (n = 5 per group). **(A)** Effect of intra-NAc microinjection of 5-Aza-dC on the protein expression levels of DNMTs. **(B)** Intra-NAc microinjection of 5-Aza-dC abolished DNMT1 expression in the heroin-self-administration group. **(C)** Intra-NAc microinjection of 5-Aza-dC abolished DNMT3A expression in the heroin-self-administration group. **(D)** Intra-NAc microinjection of 5-Aza-dC had no effect on DNMT3B expression in four groups. Con: saline self-administration; Her: Heroin; aCSF: artificial cerebrospinal fluid solution; AZA: 5-Aza-dC, 5-Aza-2-deoxycytidine. The groups were compared using two-way ANOVA followed by Student–Newman–Keuls post hoc test. Data were expressed as means ± SDs. $^{*}p < 0.05$.

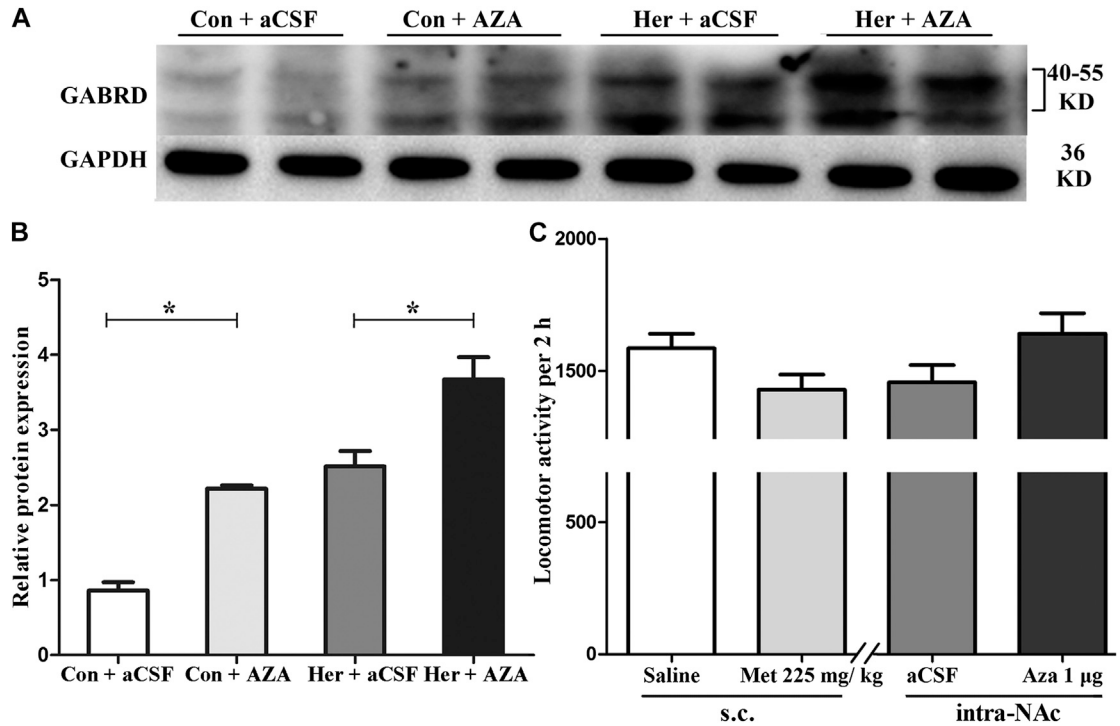

FIGURE 9 | Effect of 5-Aza-dC on the protein expression level of GABRD and effect of MET or 5-Aza-dC treatment on locomotor activity **(A)** and **(B)** Intra-NAc microinjection of 5-Aza-dC increased the protein expression level of GABRD (n = 5 per group). **(C)** MET or 5-Aza-dC had no effect on locomotion activity (n = 7 per group). Con: saline self-administration; Her: Heroin; aCSF: artificial cerebrospinal fluid solution; AZA: 5-Aza-dC, 5-Aza-2-deoxycytidine; MET: L-methionine; s.c.: subcutaneous injection. The relative protein expression level of GABRD in four groups were compared using two-way ANOVA followed by Student–Newman–Keuls post hoc test. Effect of MET and 5-Aza-dC treatment on locomotion activity were measured by Student t-test. Data were expressed as means ± SDs. *$p < 0.05$.

mediating the inhibitory effects of GABA. Subsequent validation by pyrosequencing also verified that DNA methylation at the GABRD CpG island significantly decreased and that GABRP promoter methylation significantly increased in the NAc of heroin self-administration rats. Besides, we demonstrated that MET, a methyl donor, had no effect on heroin self-administration, but it could reinforce the reinstatement of heroin-seeking induced by heroin priming while significantly increase DNA methylation levels and decrease mRNA expression of GABRD in the Her + MET group compared with that in the Her + Sal group. Additionally, 5-Aza-dC, a DNA methyltransferase inhibitor, was found to significantly attenuate the reinstatement induced by heroin priming, decrease GABRD methylation levels, and up-regulate its protein expression. With regard to the mechanisms by which the methylation landscape of GABRD was altered by 5-Aza-Dc in the NAc after reinstatement from heroin self-administration, our research found that 5-Aza-dC might alter the methylation landscape of GABRD through directly repressing DNMT1 and DNMT3A expression. Furthermore, 5-Aza-dC and MET had no significant effect on locomotor activity. Since pretreatment with 5-Aza-dC or MET failed to change the locomotor activity in heroin self-administration rats, the heroin induced reinstatement-related changes after intra-NAc injection of 5-Aza-dC or subcutaneous injection with MET might not be caused by impairment of coordinated motor ability or locomotor hypersensitization. Finally, we selectively overexpressed or silenced the GABRD gene in the NAc of rats to detect the effect of GABRD on heroin-seeking behavior. The results demonstrated that over-expression of GABRD in the NAc significantly reduced the reinstatement induced by heroin priming. On the contrary, GABRD RNAi led to an increase in cues-induced or heroin priming reinstatement.

Several genetic and environmental factors are known to contribute to the development of heroin addiction (Randesi et al., 2016). GABA inhibitory neurotransmitter receptors, especially GABAARs, are considered to play a crucial role in multiple drug-elicited neurobehavioral responses and cognitive impairment and to mediate the rewarding effects of drug abuse (Azevedo and Mammis, 2018; Engin et al., 2018; Mitchell et al., 2018). Previous studies have revealed that the NAc GABAergic system is an important mediator of heroin self-administration and that the GABRD receptors primarily control the excitability of the baseline neuronal network through shunting and tonic inhibition (Chuang and Reddy, 2018). Besides, GABRD may be essential for the cognitive and behavioral effects of consumed alcohol, such as behavioral tolerance, anxiolysis, and rewarding and reinforcing effects (Nie et al., 2011). In addition, mutations in GABRD are thought to be related to the pathology of epilepsy (Hernandez and Macdonald, 2019). However, the methylation of GABRD in heroin addiction is rarely mentioned.

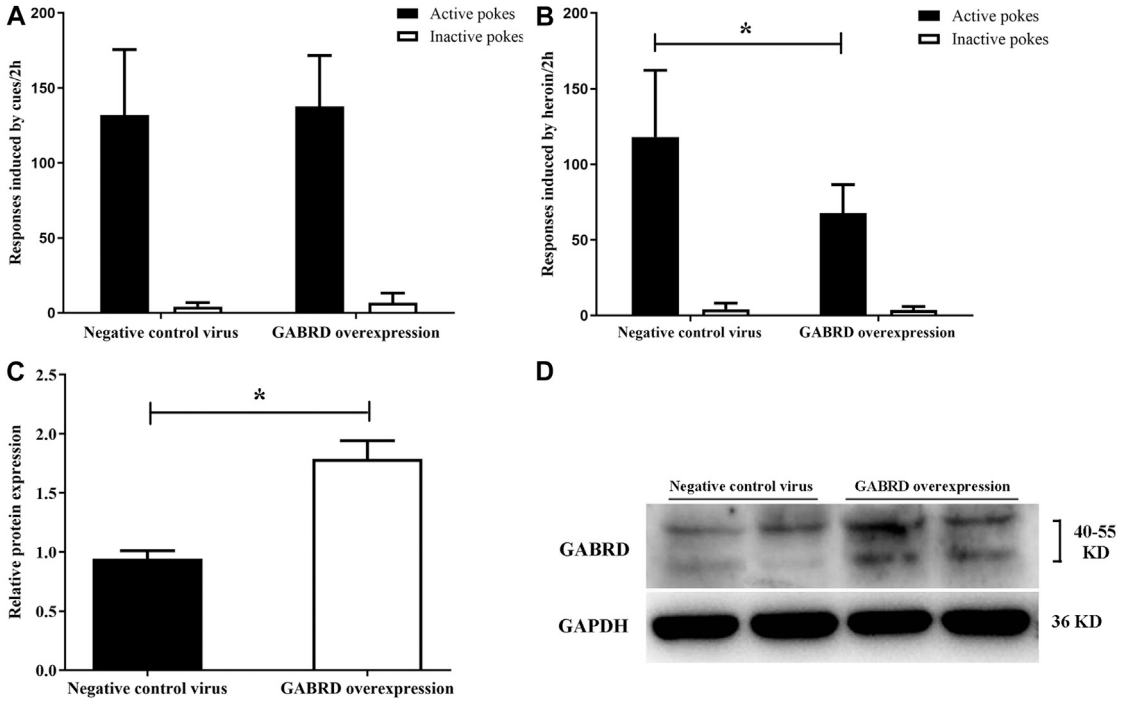

FIGURE 10 | Effects of overexpression of GABRD protein on the reinstatement (n = 8 per group). **(A)** Overexpression of GABRD had no effect on cues-induced reinstatement. **(B)** Overexpression of GABRD abolished the heroin priming-induced heroin-seeking behavior. **(C)** Grayscale analysis of protein expression level of GABRD after the overexpression of GABRD. **(D)** Western blot analysis of relative protein expression level of GABRD. Two groups were compared using Student t-test. Data were expressed as means ± SDs. $^*p < 0.05$.

In this study, rats were trained to self-administer heroin, and GABRD methylation was found to be significantly decreased after heroin self-administration training. Besides, the DNA methylation level of GABRD increased in the Her + MET group compared with that in the Her + Sal group after systemic injection of MET. On the contrary, intra-NAc microinjection of 5-Aza-dC significantly decreased GABRD gene methylation. Furthermore, gene methylation alterations in GABRD were negatively correlated with the level of GABRD expression; this result may be explained by the fact that the candidate fragment of GABRD has some enhancer-like activity, and consequently the methylated fragment of GABRD led to reduced expression. These data presented here further validate that changes in GABRD expression in the NAc affect the characteristics of heroin addiction in a DNA methylation-dependent manner.

These data were also generally consistent with findings showing that GABRD methylation in the NAc significantly decreased after cocaine self-administration (Massart et al., 2015). Research on cancer had also indicated that methylation in the c.g., 13916816 CpG site of GABRD was negatively correlated with mRNA expression and was related to the overall survival status of adult isocitrate dehydrogenase wild-type diffuse low-grade glioma patients (Zhang et al., 2019). However, there had been few reports regarding the correlation between GABRP methylation and drug addiction. In the present study, our results showed that GABRP methylation was elevated in the NAc of heroin self-administration rats; however, mRNA expression remained unchanged during heroin addiction. Our results were partly consistent with those of some former studies—for example, Zhao et al. (2013) showed that CpG methylation of the GABRP promoter in alcohol dependence (AD) patients was higher than that in their AD-discordant siblings; moreover, this was accompanied by functional GABRP downregulation. In the present study, the most important finding was that the DNA methylation status of the GABRD or GABRP genes, especially the GABRD gene, might serve as biomarkers for evaluating the severity of addiction or addiction-like behavior, and it can be inferred that transcriptional regulation that relies on epigenetic mechanism may contribute to the effects of heroin addiction. Previous studies had revealed that missense mutation in human GABRD receptor significantly reduced the surface expression of GABA-A receptors, reduced the GABA current and altered the channel gating frequency, which resulted in an impaired inhibitory neurotransmission. Moreover, the abundance or distribution of GABA-A receptors could affect the drug response (Chuang and Reddy, 2018). Reversible DNA cytosine methylation is a signal of chromatin condensation (Jordan et al., 2007). Thus, we speculated that changing level of chromatin condensation of GABRD in coding regions induced by methylation process might affect the GABA-A receptors features such as GABA sensitivity or ligand-gated chloride-ion channels gating frequency.

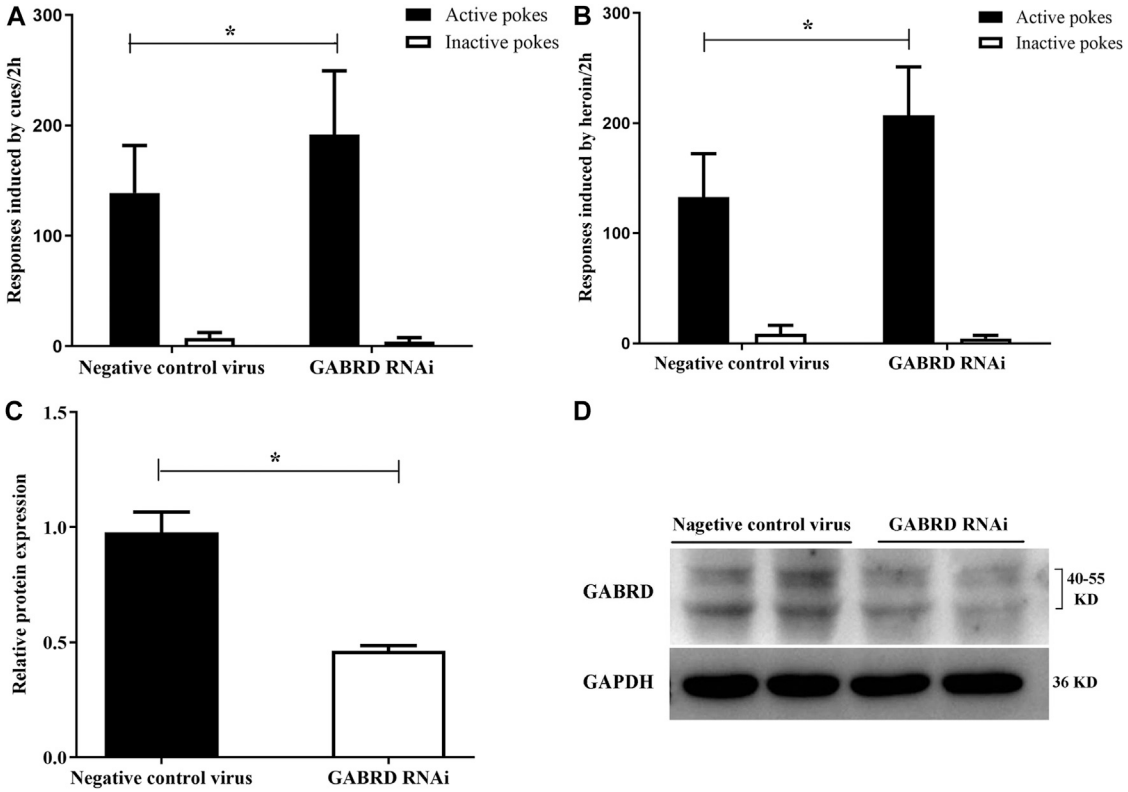

FIGURE 11 | Effects of down-regulation of GABRD on the reinstatement (n = 8 per group). **(A)** Down-regulation of GABRD reinforced the cues-induced reinstatement. **(B)** Knockdown of GABRD enhanced the heroin priming-induced heroin-seeking behavior. **(C)** Grayscale analysis of protein expression level of GABRD. **(D)** Western blot analysis of relative protein expression level of GABRD. Two groups were compared using Student t-test. Data were expressed as means ± SDs. *p < 0.05.

Recently, converging evidence showing global and site-specific changes in DNA methylation of other genes was also observed in the context of drug addiction (Lax and Szyf, 2018). For instance, a prior study noted that peripheral lymphocyte DNA from heroin addicts of Caucasian origin exhibited hypermethylation at the promoter region of the opioid receptor Mu 1 gene (Nielsen et al., 2012). Besides, long-term parental methamphetamine exposure in mice influenced offspring behavior and hippocampal DNA methylation at 70 hypermethylated and 39 undermethylated loci (Itzhak et al., 2015). In addition, several immediate early genes, such as activity regulated cytoskeletal-associated protein gene, FBJ osteosarcoma oncogene, nuclear receptor subfamily 4 group A member 1 gene, and early growth response 2, were confirmed to be differentially methylated in the frontal cortex and hippocampus of chronic methamphetamine-treatment mice (Cheng et al., 2015). Moreover, methamphetamine induced Tet methylcytosine dioxygenase gene-dependent DNA hypomethylation (Jayanthi et al., 2018). Furthermore, Zhu et al. (2017b) noted that chronic morphine administration induced hypermethylation of the glucocorticoid receptor (GR) promoter and then downregulated the expression of hippocampal GR. Research also suggested that chronic, heavy alcohol consumption led to changes in DNA methylation patterns.

Interestingly, in the present study, we found that pretreatment with MET significantly increased the reinstatement induced by heroin priming, but had no obvious effects on cues-induced reinstatement. Results of a previous study were partially consistent with ours and showed that MET attenuated cocaine-primed reinstatement had no effect on cues-induced cocaine reinstatement (Wright et al., 2015). This could be attributable to the fact that reinstatement induced by cues is more complicated than drug primed reinstatement. For example, Ambroggi et al. (2008) showed that the basolateral amygdala (BLA) and its projections to the NAc interacted to promote the reward-seeking behavioral response and that cues-evoked excitation of NAc neurons depends on BLA input. Therefore, firing activities and response of NAc neurons to MET is not enough to promote cues-induced heroin-seeking behavior. On the contrary, Tian et al. (2012) reported that MET could significantly weaken the rewarding effects elicited by cocaine. Besides, MET was proved to attenuate cocaine-primed reinstatement (Wright et al., 2015) and cause a robust decrease in cocaine conditioned place preference (CPP) (LaPlant et al., 2010; Tian et al., 2012). Importantly, our study illustrated that 5-Aza-dC injection into the NAc significantly inhibited the heroin-seeking behavior induced by heroin priming, which was consistent with the effects of 5-Aza-dC on addictive behaviors with other abused substances. 5-Aza-dC injection into

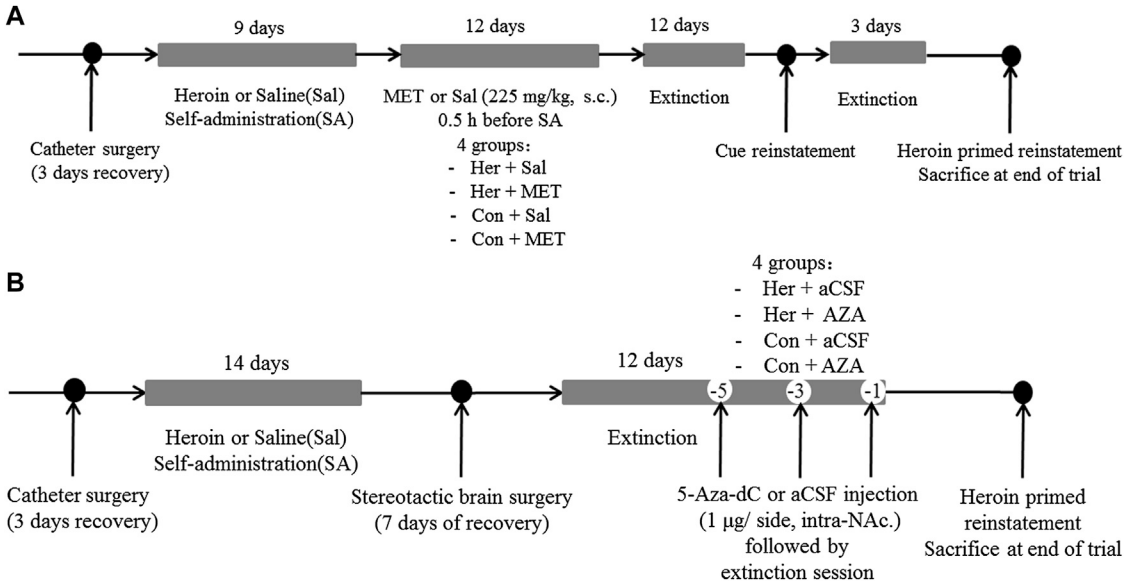

FIGURE 12 | Experimental design. **(A)** Experiment grouping and timeline for MET treatment. Rats were trained for heroin self-administration and subsequently injected with MET (225 mg/kg, s.c.) or saline at 0.5 h prior to heroin self-administration training for 12 consecutive days from the 10th day of training. This was followed by 12 days of extinction training and cues-induced or heroin priming reinstatement.i) heroin self-administration and MET-treated rats (Her + MET, $n = 5$), ii) heroin self-administration and saline-treated rats (Her + Sal, $n = 5$), iii) saline self-administration and MET-treated rats (Con + MET, $n = 5$), and iv) saline self-administration and saline-treated rats (Con + Sal, $n = 5$). **(B)** Experiment grouping and timeline for 5-Aza-dC treatment. Rats underwent bilateral NAc microinjection of 5-Aza-dC or aCSF (1 μl/side, 1 mg/ml, intra-NAc) in the last 1, 3, and 5 days of the extinction training followed by 2 h of heroin priming reinstatement. i) heroin self-administration and 5-Aza-dC treated rats (Her + AZA, $n = 5$), ii) heroin self-administration and aCSF-treated rats (Her + aCSF, $n = 5$), iii) saline self-administration and 5-Aza-dC treated rats (Con + AZA, $n = 5$), and iv) saline self-administration and aCSF-treated rats (Con + aCSF, $n = 5$). MET: L-methionine; s.c.: subcutaneous injection; aCSF: artificial cerebrospinal fluid solution; AZA: 5-Aza-dC, 5-Aza-2-deoxycytidine.

the medial prefrontal cortex (mPFC) significantly reduced alcohol consumption and alcohol preference, prevented excessive alcohol use in rats (Qiao et al., 2017). Furthermore, 5-Aza-dC injection into hippocampus CA1 area restrained acquisition of cocaine in mice but had no impact on cocaine-induced CPP, and injection into prelimbic cortex blocked cocaine-induced CPP but had no effect on acquisition of cocaine (Han et al., 2010). Moreover, 5-Aza-dC injection into the hippocampal CA1 in rats significantly weakened the consolidation and acquisition of morphine-induced CPP in rats (Zhang et al., 2014), disrupted the reconsolidation of morphine-associated withdrawal memory when injection of 5-Aza-dC into the agranular insular and BLA in rats (Liu et al., 2016). However, other studies has shown the different effects of DNA methylation inhibitors, for example, 5-Aza-dC injection into the prelimbic subregion of mPFC significantly potentiated the retrieval of morphine-induced CPP (Zhang et al., 2014), and intracerebroventricular 5-Aza-dC administration facilitated chronic intermittent ethanol-induced ethanol drinking (Qiang et al., 2014). Moreover, 5-Aza-dC injection into the cerebral ventricles of rats enhanced the reinforcing properties of cocaine (Fonteneau et al., 2017). The discrepancy between our results and these studies could be attributed to many factors, such as the the type of drugs, experimental paradigms, dosages, time points of MET and 5-Aza-dC administration, and the specific brain regions tested. As critical nodes within the mesocorticolimbic circuitry, the NAc plays a central role in reinforcement, motivation and drug seeking, the prelimbic subregion of mPFC is involved in extinction memory, the BLA mediates associative learning for both reward and emotions, and the hippocampus is essential for contextual learning memory (Ma et al., 2019). The neurobiological meanings of the brain regions involved in the plasticity underlying the effect of 5-Aza-dC should be noted. Undoubtedly, these findings revealed that 5-Aza-dC might interfere with the reconsolidation of addiction memory through different brain regions, which suggested a potential therapeutic target for drug addiction.

Another important finding in the present study was that decreased GABRD methylation and enhanced GABRD expression in the NAc after reinstatement from heroin self-administration induced by 5-Aza-dC treatment might partially contribute to decreased DNMT1 and DNMT3A expression. Previous studies had indicated that DNMT3A and DNMT3B in the medial prefrontal cortex were upregulated after alcohol exposure and that this upregulation could be reversed by 5-Aza-dc treatment (Qiao et al., 2017). Besides, 5-Aza-dC injected into left ventricle inhibited nicotine-induced an increase in DNMT3A and global DNA methylation (Ke et al., 2017). It should be note that DNMT1 and DNMT3A were more abundant than DNMT3B in adult brain. Furthermore, the ability to form DNMT-DNA adducts in DNMT1 was similar with DNMT3. Thus, it seemed that 5-Aza-dC treatment prefers to decrease DNMT1 and DNMT3A in the NAc when rats exposed to heroin. Further research should be undertaken to fully explore the role of DNMT isoforms in heroin reward and heroin seeking behavior.

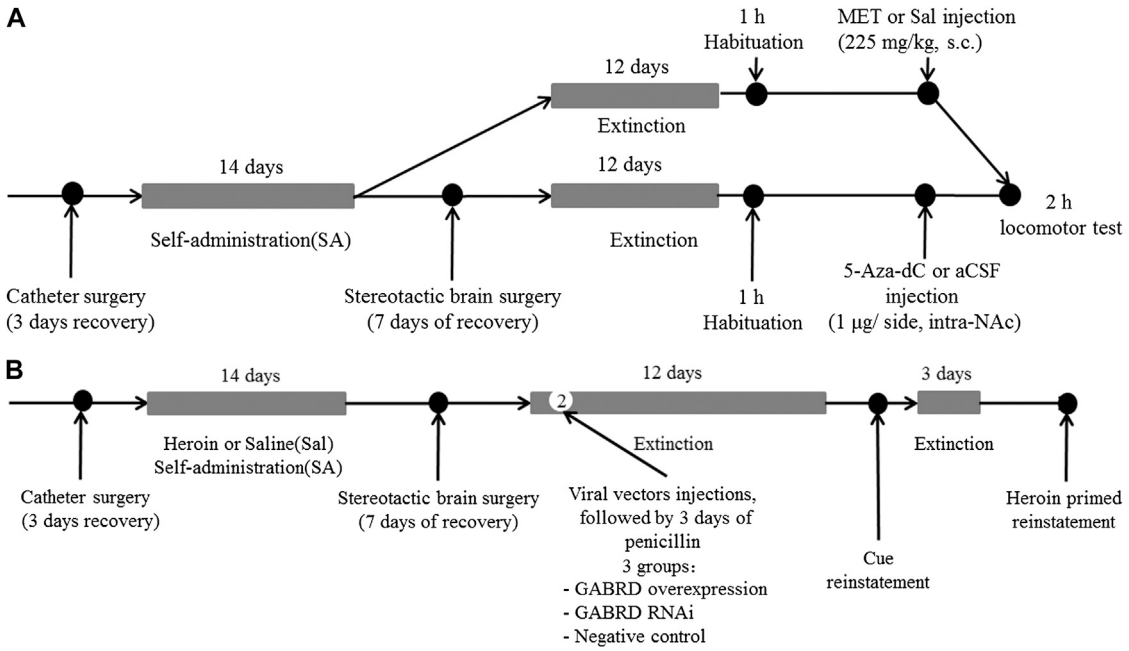

FIGURE 13 | Experimental design **(A)** Timeline for locomotor test (n = 7 per group). After extinction training from heroin self-administration, rats were habituated to the AccuScan chamber for 1 h and then injected with MET (225 mg/kg, s.c.), 5-Aza-dC (1 μg/side, intra-NAc), or vehicle (s.c., saline; intra-NAc, aCSF). Horizontal locomotor activities were recorded for 2 h. **(B)** Timeline for microinjection of recombinant AAV vectors targeting GABRD. On the second day of the extinction test, AAV vectors for GABRD overexpression, siRNA, or negative control were microinjected into the NAc of rats (n = 8 per group). After 12 days of extinction testing, a 2 h cues-induced or heroin priming reinstatement test was performed. i) overexpression of GABRD (GABRD overexpression), ii) knock-down of GABRD (GABRD RNAi), and iii) negative virus control group (Negative control). MET: L-methionine; s.c.: subcutaneous injection; aCSF: artificial cerebrospinal fluid solution; 5-Aza-dC: 5-Aza-2-deoxycytidine.

Notably, our results revealed that overexpression of GABRD could inhibit reinstatement of heroin-seeking behavior, while GABRD knockdown could promote heroin-seeking behavior. Coincidentally, Peng et al. (2004) revealed that down-regulation of nonsynaptic GABRD receptor levels in both interneurons and principal cells could lead to increased seizure susceptibility in the hippocampal formation in a temporal lobe epilepsy model. Nie et al. (2011) found that knockdown of GABRD receptors in the medial shell region of the NAc affected alcohol intake. Collectively, our data suggested that expression levels of GABRD in the NAc might regulate the responses to drug-seeking behavior; thus, drugs targeting GABRD could emerge as potential therapeutics for drug addiction.

In summary, the present results demonstrated that GABRD might be a potential biomarker and drug target for heroin addiction and relapse. Furthermore, epigenetic regulation through DNA methylation of GABRD may be the main regulatory mechanism underlying heroin addiction and responsible for the neuroadaptations induced by heroin. Besides, it was found that 5-Aza-dC administration might be a potential treatment for heroin relapse.

EXPERIMENTAL PROCEDURE

Animals

Adult male Sprague-Dawley rats weighing 250–300 g were purchased from the Experimental Animal Center of Zhejiang Province (Hangzhou, China). Rats were housed in a temperature- and humidity-controlled room with food and water freely available except when specified, and were maintained under a reversed 12-h light/dark cycle (lights on at 19:00 h, off at 07:00 h). Rat was weighed and handled daily for one week prior to surgery. All animal handling procedures were approved by the Laboratory of Behavioral Neuroscience, Animal Care and Use Committee of the Ningbo Institute of Microcirculation and Henbane (Ningbo, China) and conducted strictly in accordance with the National Institutes of Health (NIH) guidelines for the Care and Use of Laboratory Animals (NIH Publications No. 80–23, revised 1978).

Drugs

Heroin (diacetylmorphine HCl) was derived from the National Institute of Forensic Science (Beijing, China). The dose of heroin used for the self-administration experiment was chosen according to previous reports (Lai et al., 2014; Sun et al., 2015; Zhu et al., 2017a). The methyl donor L-methionine (MET) and the DNMT inhibitor 5′-aza-2′-deoxycytidine (5-Aza-dC) were purchased from Tocris Bioscience (MO, USA). 25 mg MET was dissolved in 1 ml sterile water and 0.2% dimethyl sulfoxide (DMSO, Sigma-Aldrich) and administered at a dose of 225 mg/kg body weight by subcutaneous injection (s.c.). For intra-NAc injections, 1 mg of 5-Aza-dC was dissolved in 1 ml of sterile water and 0.2% DMSO and microinjected in the bilateral NAc of each rat (1 μg/side).

Plasmids and Transient Transfections

GABRD cDNA was obtained by PCR amplification (forward primer:GGAGGTAGTGGAATGGATCCCGCC ACCATGGACG TTCTGGGCTGG CTGC, reverse primer:TCACCATGGTG GCGGGATCC ATGGTATACGCCGCCCAGTAG, product size: 1,393 bp) and ligated into the AAV vector GV467 after BamH1 restriction enzyme digestion using the ClonExpress TM II One Step Cloning Kit (Vazyme Biotech, Nanjing, China). The reaction was performed in a total volume of 10 μl containing 1 μl of PCR product and 2.5 μl of vector DNA. Subsequently, 10 μl of ligation mix was transformed into 100 μl TOP10 chemically competent cells (Invitrogen, Carlsbad, CA, USA). An appropriate amount of bacterial solution was spread on an Luria–Bertani (LB) agar plate containing 10 μg/ml ampicillin, and subsequently individual colonies were incubated in a shaking incubator at 30°C for 12–16 h. Clones that contained an insertion of the correct size were subsequently transferred to 10 ml of LB liquid medium containing the corresponding antibiotics and incubated overnight at 37°C. The plasmid was extracted using the Endo-Free midi Plasmid Kit (Omega, Norcross, GE, USA) according to the manufacturer's instructions and finally stably transfected into HEK293T cells in 6-well plates. Cellular green fluorescent protein was observed using a Olympus BX51 Upright Fluorescence Microscope (Olympus, Tokyo, Japan) at 24 h after transfection; a fluorescence rate of over 80% indicated appropriate GABRD overexpression. Cells were harvested 48 h later and stored at −80°C until further use.

The recombinant vector CV272 carrying double-stranded DNA oligo containing the interfering sequence of *GABRD* (CTCGAGAAGGTATATTGCTGTTGACAGTGAGCGAGACGT GAGGAACGCCATTGTCTAGTGAAGCCACAGATGTAGACA ATGGCGTTCCTCACGTCCTGCCTACTGCCTCGCAATTG, siRNA target sequence: GGACGTGAGGAACGCCATTGT) was transferred to the competent cells of the yeast strain GS115. Monoclonal colonies were identified by PCR, and positive colonies were then sequenced. The plasmid was extracted using the Qiagen Plasmid Plus Maxi Kit (QIAGEN, Hilden, Germany) according to the manufacturer's suggested protocols.

Surgery

Rats were catheterized in the right jugular vein as described previously (Lai et al., 2013). Briefly, rats were anesthetized using sodium pentobarbital (50 mg/kg, Serva) administered by intraperitoneal injection (i.p.), and atropine sulfate (0.3 mg/kg, s.c.) was injected before surgery. A silicon catheter (length 3.5 cm, inner diameter 0.5 mm, outer diameter, 0.94 mm) was plugged inthe right external jugular vein, and the other end of the catheter (10 cm, PE20) was passed subcutaneously to the incision of the back, where it exited into a custom-made fluid connector fixed to a jacket. The catheters were irrigated daily with 0.2 ml each of heparinized saline (100 IU) and cefazolin (0.1 g/ml). After 3 days of recovery, the self-administration procedures were initiated.

For stereotaxic injection, rats were anesthetized using sodium pentobarbital (50 mg/kg, i.p.) and installed in a stereotaxic frame (Stoelting, USA). 5-Aza-dC (1 μl per injection at a rate of 0.5 μl/min) was bilaterally injected into the NAc at the following sites according to the rat brain atlas of Paxinos and Watson (1998) (Paxinos and Watson, 1998): anterior–posterior, −1.4 mm from bregma; medial–lateral, ±1.5 mm from midline; dorsal–ventral, −(7.0–7.5) mm from bregma. The needle was maintained in place for an additional 3 min after each injection to ensure the sufficient diffusion of the solution, and then was slowly withdrawn. A Microinjection pump (MD-1001, Bioanalytical System Inc., IN) was used for all injections into the NAc. Finally, the rats were injected intramuscularly with penicillin (15,000 U daily) for at least 3 days to prevent infection.

Heroin Self-Administration

Firstly, in order to explore the potential changes of DNA methylation in heroin self-administration rats in the NAc, animal behavioral training was initiated after recovery as previously described (Hong et al., 2020). In brief, the rats were trained to self-administer heroin under a fixed ratio paradigm for 4 h daily in operant chambers which equipped with two nose pokes: an active nose-poke hole and an inactive one (ENV-114 M, Med Associates, St. Albans, VT). Each response in the active hole was immediately reinforced with an infusion of heroin (0.5 mg/kg) via the pump over 4 s (Lou et al., 2014), at the same time was accompanied by the noise of the infusion pump and illumination of the stimulus light located above the active nose poke for 20 s. The number of responses during this 20 s period was recorded, but there was no delivery of heroin. Responding on the inactive nose-poke produced no programmed consequences. Daily training sessions were conducted over 14 consecutive days. The criterion of stable heroin self-administration was defined as less than 10% of the variability in the number of active nose poking responses that rats touched on the last 3 days (Lai et al., 2013).

Yoked–Heroin Self-Administration and Saline Procedure

Both yoked-heroin self-administration (Yoke Heroin) and yoked-saline groups (Control) were tested simultaneously with the heroin self-administration group (Heroin self-administration) in different conditions. The yoked-heroin or yoked-saline rat was paired with the heroin self-administration rat. Yoked-heroin rats acquired intravenous infusions of heroin at the same rate, dose and number as rats in the heroin self-administration group. The yoked-saline group received intravenous infusions of saline at the same rate and number as the heroin self-administration group. Nose-pokes by the yoked rats were recorded but had no scheduled consequences.

Extinction and Reinstatement of Heroin Self-Administration

After heroin self-administration training sessions, the animals underwent daily 2 h extinction sessions for 12 days. During the extinction phase, the rats were replaced in the operant chambers for 2 h without conditioned cues and heroin, and responses in the active hole were recorded, but there is no programmed

consequences regardless of pressing the active or inactive pokes. Rats pressing the active nose-poke for less than 10% of the average active nose-poke responses during maintenance was determined to reach the extinction criterion.

Reinstatement tests were conducted 24 h after the last extinction training. During cues-induced reinstatement, rats were exposed to the same light and tone cues as those used during the period of heroin self-administration; active and inactive responses were recorded, but heroin was not available. During the heroin priming-induced reinstatement test, the rats were first given a single injection of heroin (0.25 mg/kg, s.c.) at 30 min before the reinstatement test session and then were exposed to the conditioned light and tone cues associated with heroin infusion for 5 s. The subsequent active nose-poke response led to the presentation of the conditioned stimulus, but heroin infusion was not available. Nose-pokes during the reinstatement were accumulated over 2 h.

Locomotor Activity Assessment

After extinction session, acrylic locomotor monitoring cages (AccuScan Instruments, Inc., Columbus, OH) were used to measure the horizontal locomotor activities of rats, each of which contained 16 photocell beams measuring the horizontal distance traveled. After 1 h of habituation to the chamber, beam breaks were continuously recorded once every 10 min for 2 h using the VersaMax/Digiscan System Software (AccuScan Instruments, Inc.).

Experimental Grouping and Treatments Received

Experiment 1 During the heroin self-administration training, rats were divided into three groups: i) yoked-saline group, which received saline injections (Control, $n = 10$) for 14 days, ii) heroin self-administration group ($n = 10$), which received heroin self-administration for 14 days, and iii) yoked-heroin group, which received amounts of heroin equal to those received self-administration over the same course of time (Yoke heroin, $n = 10$).

Experiment 2 For examining the effect of MET on heroin self-administration and cues-induced or heroin priming reinstatement of heroin-seeking, rats were divided into four groups: i) heroin self-administration and MET-treated rats (Her + MET, $n = 5$), ii) heroin self-administration and saline-treated rats (Her + Sal, $n = 5$), iii) saline self-administration and MET-treated rats (Con + MET, $n = 5$), and iv) saline self-administration and saline-treated rats (Con + Sal, $n = 5$) (**Figure 12A**). Rats were injected with MET (225 mg/kg, s.c.) or saline at 0.5 h prior to heroin self-administration training for 12 consecutive days from the 10th day of training; the dose was chosen based on a previous study (Wright et al., 2015). The self-administration behaviors of rats during this training were recorded. Then rats underwent the extinction training for 12 consecutive days, followed by reinstatement tests to assess cues-induced or heroin priming seeking behavior.

Experiment 3 To confirm the effect of 5-Aza-dC on DNA methylation and heroin priming reinstatement of heroin-seeking, rats were divided into four groups: i) heroin self-administration and 5-Aza-dC treated rats (Her + AZA, $n = 5$), ii) heroin self-administration and artificial cerebrospinal fluid (aCSF)-treated rats (Her + aCSF, $n = 5$), iii) saline self-administration and 5-Aza-dC treated rats (Con + AZA, $n = 5$), and iv) saline self-administration and aCSF-treated rats (Con + aCSF, $n = 5$) (**Figure 12B**). Rats underwent bilateral NAc microinjection of 5-Aza-dC or aCSF in the last 1, 3, and 5 days of the extinction training, each followed by a 2 h extinction session. For intra-NAc treatments, 5-Aza-dC (1 μg/side, 1 mg/mL) or aCSF was infused through an injection cannula placed within the guide cannula. The doses of 5-Aza-dC were defined based on a previous study (Qiao et al., 2017).

Experiment 4 To investigate the effect of MET or 5-Aza-Dc treatment on locomotor activity, rats that had undergone extinction training from heroin self-administration were habituated to the AccuScan chamber for 1 h and then injected with MET (225 mg/kg, s.c., $n = 7$), 5-Aza-dC (1 μg/side, intra-NAc, $n = 7$), or vehicle (s.c., saline; intra-NAc, aCSF, $n = 7$) and replaced back into the AccuScan chamber (**Figure 13A**). Horizontal locomotor activities were recorded for 2 h.

Experiment 5 To investigate whether changes in GABRD expression would affect changes in heroin reinstatement of heroin-seeking behavior, rats were randomly divided into three groups after heroin self-administration training: i) overexpression of GABRD (GABRD overexpression, $n = 8$), ii) knock-down of GABRD (GABRD RNAi, $n = 8$), and iii) negative virus control group (Negative control virus, $n = 8$). On the second day of the extinction test, AAV vectors for GABRD overexpression, RNAi, or negative control (1 μl each) were microinjected into the NAc of rats using a 5 μl syringe needle for bilaterally infusion at a rate of 0.5 μl/min. The needle was maintained in place for an additional 3 min after each injection to ensure the sufficient diffusion of the solution into the brain tissue. After 12 days of extinction testing, a 2 h cues-induced or heroin priming reinstatement test was performed (**Figure 13B**).

Tissue Processing

Once all testing was completed, rats were anesthetized using sodium pentobarbital (50 mg/kg, intramuscular) and rapidly decapitated. Brain tissues were dissected from the bregma using a rat brain matrix [2.2–0.8 mm according to coordinates from Paxinos and Watson (1998) rat brain atlas] (1.4 ± 0.1 mm thickness), and serial 1 mm brain sections were cut (Hong et al., 2020). The NAc were harvested bilaterally using a stainless-steel cannula (1.5 mm in diameter). Tissues were stored at −80°C until further use.

Illumina HiSeq

Genomic DNA was extracted from the NAc using the QIAamp DNA Mini kit (Qiagen, Hilden, Germany). DNA was sheared and then purified using the QIAquick PCR purification kit, and libraries were constructed using the Illumina library preparation kit. Thereafter, the QIAquick PCR purification kit was used to purify the libraries, and the EpiTect bisulfite kit (Qiagen, Hilden, Germany) was suitable for bisulfite

TABLE 1 | Primers for pyrosequencing and RT-qPCR.

Reaction	Gene	Primer	Sequence (5′–3′)
Pyrosequencing	GABRD	Forward	Biotin-GTGGTTAAGGGTAAGAATGAAGAGA
		Reverse	CCCATTAAATCCACCCAATTTTACTT
		Sequencing	CCACCCAATTTTACTTT
Pyrosequencing	GABRP	Forward	Biotin-ATTTATGGATGATGGGGTGAG
		Reverse	CCAAAAAACCAAACCTCTAAAACATA
		Sequencing	ACACTAACACACAACAATT
RT-qPCR	GABRD	Forward	AAGTCTGCCTGGTTCCATGATGTG
		Reverse	CTGTGGAGGTGATGCGGATGC
RT-qPCR	GABRP	Forward	TTCAATGTGGAGGTCAGCAGAAGC
		Reverse	GAGATGCTTGCGATGTCCAGAGTC
RT-qPCR	GAPDH	Forward	GACATGCCGCCTGGAGAAAC
		Reverse	AGCCCAGGATGCCCTTTAGT

conversion. Bisulfite converted libraries were amplified using PfuTurbo HotStart DNA polymerase (Agilent Technologies, CA, USA), size selected (200–300 bp), quantified using qPCR, and run through the Illumina HiSeq 2000 platform as previously described (Dehghanizadeh et al., 2018). The downstream analysis of the obtained sequencing data mainly involved four aspects of the programs/sequencing reads: Fastqc Control, Sequencing Depth Statistics, aligned to the genome, and differentially methylated region (DMR) or site analysis.

Double Luciferase Reporter Assay

Double luciferase reporter assay was carried out to verify the enhancer activity of the target sequence and examine the effect of methylation on its activity. Reporter gene vectors (pGL3-promoter) carrying GABRD DNA fragments (GABRD: chr5: 172809066–172809129) were constructed as follows. The pGL3-promoter was digested with NheI/XhoI, and the recombinant plasmids were transformed into TOP10 competent cells. The correct clones were subjected to sequencing, and the plasmid was extracted using TIANprep Mini Plasmid Kit (Tiangen Biotech, Beijing, China) after verification by sequencing. The constructed plasmids were then methylated using the CpG methyltransferase M.SssI Kit (Thermo, MA, USA) to establish the methylated groups. Thereafter, the EpiJET DNA Methylation Analysis Kit was used to perform the methylation test. HEK293T cells were cultured in Dulbecco's modified Eagle medium (Gibco, Invitrogen, Paisley, United Kingdom) with 10% fetal bovine serum (Gibco, Invitrogen, Paisley, United Kingdom) and 1% penicillin/streptomycin (Gibco, Invitrogen, Paisley, United Kingdom). HEK293T cells were stably co-transfected with recombinant plasmid or methylated plasmids using the Lipofectamine 2000 Transfection Reagent (Invitrogen, Carlsbad, USA) after the cells reached 70% confluence. The reporter assay was carried out using the Dual-Luciferase Reporter Assay system (Promega, Madison City, WI, USA) two days after co-transfection according to the manufacturer's protocol. The recombinant plasmids were divided into 3 groups: i) pGL3-promoter (Pro), ii) pGL3-promoter-GABRD (Pro-GABRD), iii) Me-pGL3-promoter-GABRD (Me-Pro-GABRD).

Bisulfite Sequencing

GABRD and GABR pi (GABRP) PCR and sequencing primers were designed using the PyroMark Assay Design 2.0 software. The sequences of the forward primers are shown in **Table 1**. All primers were synthesized by Sangon Biotechnology (Shanghai, China). Genomic DNA (500 ng) was bisulfite treated using the EZ DNA Methylation Gold™ kit (Zymo Research, CA, USA). Bisulfite-converted DNA were amplified as follows: 95°C for 7 min, followed by 40 cycles (95°C for 30 s, 55°C for 30 s, 72°C for 1 min), and 72°C for 5 min. Genes were amplified using Zymo Taq Premix (Zymo Research, CA, USA). The 20 μL PCR reaction was setup as follows: 10 μl Zymo Taq Premix, 6.5 μl DNase/RNase-free water, 1 μl primer pairs (20 pM), and 2.5 μL DNA. The PCR products were run on an Agilent Bioanalyser 2100 to check the specificity of primers according to the Agilent DNA 1000 kit manual (Aglient, CA, USA). A new gene run assay was designed using the PyroMark Q48 software, imported into a U disk, and then loaded into PyroMark Q48 Autoprep. Cleaning and injector priming programs were run to check the status of the instrument. Thereafter, the new run was initiated from the main menu, and reagents were pipetted into the designated injector according to the volumes shown on the instrument touchscreen. Cartridge lids were closed and locked after adding the required reagents. The PCR product (10 μl, about 250 ng template) and 3 μl magnetic beads were loaded into the correct wells of the PyroMark Q48 Disc. The disc was locked correctly into the instrument, a new absorber strip was inserted, and then the run was initiated. Finally, the methylation status of each CpG site was analyzed using the Pyro Q CpG software.

Real-Time RT-qPCR

Total RNA was extracted and reverse-transcripted into single-strand cDNA according to the method described previously (Hong et al., 2020). RT-PCR was used to analyze the mRNA levels of GABRD and GABRP with glyceraldehyde 3-phosphate dehydrogenase (GAPDH) for normalization; the nucleotide sequences of the primers are shown in **Table 1**. The PCR was performed in a LightCycler® 480 system (Roche Applied Science), and the $2^{-\Delta Cq}$ method was used to analyze the relative changes in mRNA expression levels.

Western Blot

For protein analysis, NAc samples were homogenized in 300 μl ice-cold lysis buffer (SDS: PMSF = 100: 1) and then centrifuged at 12,000 rpm for 20 min at 4°C. Supernatants were collected, and the total protein concentration was measured using the bicinchoninic acid protein assay reagent kit (Thermo Fisher Scientific, Pittsburgh, PA, USA). The protein solutions were mixed with loading buffer, boiled at 100°C for 5 min, and then separated by 10% SDS-polyacrylamide gel electrophoresis and electroblotted onto a polyvinylidene difluoride membrane (Millipore, Massachusetts, USA). Then the membrane was blocked with 5% skimmed milk for 1 h at 37°C. Subsequently, the membrane was incubated overnight at 4 °C with the appropriate antibodies—anti-GABRD (Sigma, SAB5200052), anti-GAPDH (Proteintech, 60004-1-Ig), anti-DNMT1 (Abclonal, A16729), anti-DNMT3A (Abclonal, A2065), and anti-DNMT3B (Abcam, ab2851). Thereafter, membranes were washed three times with PBST (1,000 mL 1×PBS and 1 mL Tween-20) (10 min each) and incubated with secondary antibody (Goat anti-mouse IgG (H + L) – HRP, 1:10,000, Jackson ImmunoResearch) for 2 h at 37 °C and then washed five times with PBST (5 min each). Protein bands were detected using a chemiluminescence detection system (Bio-Rad, Hercules, CA, USA). Protein expression intensities were expressed as ratios relative to that of GAPDH based on gray-scale analysis.

Statistical Analyses

Statistical analysis was performed using SPSS 16.0 (SPSS, Inc., Chicago, IL, USA) and GraphPad Prism 5.0 (GraphPad Software, Inc.). All data are expressed as the means ± standard deviations (SDs); $p < 0.05$ was considered indicative of a statistically significant difference. Gene methylation levels were analyzed using Pearson's regression analysis and one-way or two-way analysis of variance (ANOVA) followed by Student–Newman–Keuls post hoc comparisons. One-way and two-way ANOVA was used to analyze mRNA or protein expression, with treatment and heroin self-administration as the independent variables. Effects of MET or 5-Aza-dC on the nose-poke responses and infusions during heroin self-administration and reinstatement were analyzed using two-way repeated-measures ANOVA with treatment group and time as factors. Effect of MET and 5-Aza-dC treatment on locomotion activity, and Effect of GABRD overexpression or RNAi on protein expression level of GABRD were compared using Student *t*-test.

DATA AVAILABILITY STATEMENT

Our Illumina sequencing data to SRA database had been successfully processed and will be released on 2024-10-01, Accession to cite for these SRA data: PRJNA67367. Our SRA records will be accessible with the following link after the indicated release date: https://www.ncbi.nlm.nih.gov/sra/PRJNA673675.

ETHICS STATEMENT

The animal study was reviewed and approved by Laboratory of Behavioral Neuroscience, Animal Care and Use Committee of the Ningbo Institute of Microcirculation and Henbane (Ningbo, China).

AUTHOR CONTRIBUTIONS

HL designed the study; QH, WZ, and HL wrote the manuscript; WX, ZL, JL, and WC performed the experiments; HZ, ML, DZ, ZX, and DF analyzed the data.

FUNDING

This work was supported by the National Key R&D Program of China (Grant No. 2017YFC1310400), The Natural Science Foundation of Zhejiang (Grant No. LY18H090008), The Natural Science Foundation of China (Grant No. 81671321), and The Ningbo Science and Technology Project (Grant Nos. 2015C110026, 2017A04, 2019A610293, 2019C50076).

REFERENCES

Ambroggi, F., Ishikawa, A., Fields, H. L., and Nicola, S. M. (2008). Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron* 59 (4), 648–661. doi:10.1016/j.neuron.2008.07.004

Azevedo, C. A., and Mammis, A. (2018). Neuromodulation therapies for alcohol addiction: a literature review. *Neuromodulation* 21 (2), 144–148. doi:10.1111/ner.12548

Brickley, S. G., and Mody, I. (2012). Extrasynaptic GABA(A) receptors: their function in the CNS and implications for disease. *Neuron* 73 (1), 23–34. doi:10.1016/j.neuron.2011.12.012

Cheng, M. C., Hsu, S. H., and Chen, C. H. (2015). Chronic methamphetamine treatment reduces the expression of synaptic plasticity genes and changes their DNA methylation status in the mouse brain. *Brain Res.* 1629, 126–134. doi:10.1016/j.brainres.2015.10.021

Chuang, S. H., and Reddy, D. S. (2018). Genetic and molecular regulation of extrasynaptic GABA-A receptors in the brain: therapeutic insights for epilepsy. *J. Pharmacol. Exp. Therapeut.* 364 (2), 180–197. doi:10.1124/jpet.117.244673

Dehghanizadeh, S., Khoddami, V., Mosbruger, T. L., Hammoud, S. S., Edes, K., Berry, T. S., et al. (2018). Active BRAF-V600E is the key player in generation of a sessile serrated polyp-specific DNA methylation profile. *PLoS One* 13 (3), e0192499. doi:10.1371/journal.pone.0192499

Engin, E., Benham, R. S., and Rudolph, U. (2018). An emerging circuit pharmacology of GABAA receptors. *Trends Pharmacol. Sci.* 39 (8), 710–732. doi:10.1016/j.tips.2018.04.003

Fonteneau, M., Filliol, D., Anglard, P., Befort, K., Romieu, P., and Zwiller, J. (2017). Inhibition of DNA methyltransferases regulates cocaine self-administration by rats: a genome-wide DNA methylation study. *Gene Brain Behav.* 16 (3), 313–327. doi:10.1111/gbb.12354

Gatta, E., Auta, J., Gavin, D. P., Bhaumik, D. K., Grayson, D. R., Pandey, S. C., et al. (2017). Emerging role of one-carbon metabolism and DNA methylation enrichment on delta-containing GABAA receptor expression in the cerebellum of subjects with alcohol use disorders (AUD). *Int. J. Neuropsychopharmacol.* 20 (12), 1013–1026. doi:10.1093/ijnp/pyx075

Godino, A., Jayanthi, S., and Cadet, J. L. (2015). Epigenetic landscape of amphetamine and methamphetamine addiction in rodents. *Epigenetics* 10 (7), 574–580. doi:10.1080/15592294.2015.1055441

Han, J., Li, Y., Wang, D., Wei, C., Yang, X., and Sui, N. (2010). Effect of 5-aza-2-deoxycytidine microinjecting into hippocampus and prelimbic cortex on acquisition and retrieval of cocaine-induced place preference in C57BL/6 mice. *Eur. J. Pharmacol.* 642 (1–3), 93–98. doi:10.1016/j.ejphar.2010.05.050

Hernandez, C. C., and Macdonald, R. L. (2019). A structural look at GABAA receptor mutations linked to epilepsy syndromes. *Brain Res.* 1714, 234–247. doi:10.1016/j.brainres.2019.03.004

Hong, Q., Liu, J., Lin, Z., Zhuang, D., Xu, W., Xu, Z., et al. (2020). Histone 3 lysine 9 acetylation of BRG1 in the medial prefrontal cortex is associated with heroin selfadministration in rats. *Mol. Med. Rep.* 21 (1), 405–412. doi:10.3892/mmr.2019.10845

Host, L., Dietrich, J. B., Carouge, D., Aunis, D., and Zwiller, J. (2011). Cocaine self-administration alters the expression of chromatin-remodelling proteins; modulation by histone deacetylase inhibition. *J. Psychopharmacol.* 25 (2), 222–229. doi:10.1177/0269881109348173

Itzhak, Y., Ergui, I., and Young, J. I. (2015). Long-term parental methamphetamine exposure of mice influences behavior and hippocampal DNA methylation of the offspring. *Mol. Psychiatr.* 20 (2), 232–239. doi:10.1038/mp.2014.7

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* 33 (Suppl. 1), 245–254. doi:10.1038/ng1089

Jayanthi, S., Gonzalez, B., McCoy, M. T., Ladenheim, B., Bisagno, V., and Cadet, J. L. (2018). Methamphetamine induces TET1- and TET3-dependent DNA hydroxymethylation of crh and avp genes in the rat nucleus accumbens. *Mol. Neurobiol.* 55 (6), 5154–5166. doi:10.1007/s12035-017-0750-9

Ji, H., Liu, G., Xu, X., Liu, H., Xu, L., Hu, H., et al. (2018). Hypermethylation of the κ1 opioid receptor promoter in Chinese heroin and methamphetamine addicts. *Exp. Ther. Med.* 16 (3), 2392–2398. doi:10.3892/etm.2018.6514

Jordan, C., Li, H., Kwan, H. C., and Francke, U. (2007). Cerebellar gene expression profiles of mouse models for Rett syndrome reveal novel MeCP2 targets. *BMC Med. Genet.* 8, 36. doi:10.1186/1471-2350-8-36

Kalayasiri, R., Kraijak, K., Maes, M., and Mutirangura, A. (2019). Methamphetamine (MA) use induces specific changes in LINE-1 partial methylation patterns, which are associated with MA-induced paranoia: a multivariate and neuronal network study. *Mol. Neurobiol.* 56 (6), 4258–4272. doi:10.1007/s12035-018-1371-7

Kalda, A., and Zharkovsky, A. (2015). Epigenetic mechanisms of psychostimulant-induced addiction. *Int. Rev. Neurobiol.* 120, 85–105. doi:10.1016/bs.irn.2015.02.010

Ke, J., Dong, N., Wang, L., Li, Y., Dasgupta, C., Zhang, L., et al. (2017). Role of DNA methylation in perinatal nicotine-induced development of heart ischemia-sensitive phenotype in rat offspring. *Oncotarget* 8 (44), 76865–76880. doi:10.18632/oncotarget.20172

Klengel, T., Mehta, D., Anacker, C., Rex-Haffner, M., Pruessner, J. C., Pariante, C. M., et al. (2013). Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nat. Neurosci.* 16 (1), 33–41. doi:10.1038/nn.3275

Lai, M., Chen, W., Zhu, H., Zhou, X., Liu, H., Zhang, F., et al. (2013). Low dose risperidone attenuates cue-induced but not heroin-induced reinstatement of heroin seeking in an animal model of relapse. *Int. J. Neuropsychopharmacol.* 16 (7), 1569–1575. doi:10.1017/S1461145712001563

Lai, M., Zhu, H., Sun, A., Zhuang, D., Fu, D., Chen, W., et al. (2014). The phosphodiesterase-4 inhibitor rolipram attenuates heroin-seeking behavior induced by cues or heroin priming in rats. *Int. J. Neuropsychopharmacol.* 17 (9), 1397–1407. doi:10.1017/S1461145714000595

LaPlant, Q., Vialou, V., Covington, H. E., 3rd, Dumitriu, D., Feng, J., Warren, B. L., et al. (2010). Dnmt3a regulates emotional behavior and spine plasticity in the nucleus accumbens. *Nat. Neurosci.* 13 (9), 1137–1143. doi:10.1038/nn.2619

Lax, E., and Szyf, M. (2018). The role of DNA methylation in drug addiction: implications for diagnostic and therapeutics. *Prog. Mol. Biol. Transl. Sci.* 157, 93–104. doi:10.1016/bs.pmbts.2018.01.003

Liu, P., Zhang, J., Li, M., and Sui, N. (2016). Distinctive roles of 5-aza-2′-deoxycytidine in anterior agranular insular and basolateral amygdala in reconsolidation of aversive memory associated with morphine in rats. *Front. Behav. Neurosci.* 10, 50. doi:10.3389/fnbeh.2016.00050

Lou, Z. Z., Chen, L. H., Liu, H. F., Ruan, L. M., and Zhou, W. H. (2014). Blockade of mGluR5 in the nucleus accumbens shell but not core attenuates heroin seeking behavior in rats. *Acta Pharmacol. Sin.* 35 (12), 1485–1492. doi:10.1038/aps.2014.93

Ma, B., Mei, D., Wang, F., Liu, Y., and Zhou, W. (2019). Cognitive enhancers as a treatment for heroin relapse and addiction. *Pharmacol. Res.* 141, 378–383. doi:10.1016/j.phrs.2019.01.025

Massart, R., Barnea, R., Dikshtein, Y., Suderman, M., Meir, O., Hallett, M., et al. (2015). Role of DNA methylation in the nucleus accumbens in incubation of cocaine craving. *J. Neurosci.* 35 (21), 8042–8058. doi:10.1523/JNEUROSCI.3053-14.2015

Meaney, M. J., and Szyf, M. (2005). Maternal care as a model for experience-dependent chromatin plasticity?. *Trends Neurosci.* 28 (9), 456–463. doi:10.1016/j.tins.2005.07.006

Mitchell, S. J., Maguire, E. P., Cunningham, L., Gunn, B. G., Linke, M., Zechner, U., et al. (2018). Early-life adversity selectively impairs alpha2-GABAA receptor expression in the mouse nucleus accumbens and influences the behavioral effects of cocaine. *Neuropharmacology* 141, 98–112. doi:10.1016/j.neuropharm.2018.08.021

Nestler, E. J. (2014). Epigenetic mechanisms of drug addiction. *Neuropharmacology* 76 Pt B, 259–268. doi:10.1016/j.neuropharm.2013.04.004

Nie, H., Rewal, M., Gill, T. M., Ron, D., and Janak, P. H. (2011). Extrasynaptic delta-containing GABAA receptors in the nucleus accumbens dorsomedial shell contribute to alcohol intake. *Proc. Natl. Acad. Sci. U. S. A.* 108 (11), 4459–4464. doi:10.1073/pnas.1016156108

Nielsen, D. A., Utrankar, A., Reyes, J. A., Simons, D. D., and Kosten, T. R. (2012). Epigenetics of drug abuse: predisposition or response. *Pharmacogenomics* 13 (10), 1149–1160. doi:10.2217/pgs.12.94

Novikova, S. I., He, F., Bai, J., Cutrufello, N. J., Lidow, M. S., and Undieh, A. S. (2008). Maternal cocaine administration in mice alters DNA methylation and gene expression in hippocampal neurons of neonatal and prepubertal offspring. *PLoS One* 3 (4), e1919. doi:10.1371/journal.pone.0001919

Numachi, Y., Shen, H., Yoshida, S., Fujiyama, K., Toda, S., Matsuoka, H., et al. (2007). Methamphetamine alters expression of DNA methyltransferase 1 mRNA in rat brain. *Neurosci. Lett.* 414 (3), 213–217. doi:10.1016/j.neulet.2006.12.052

Olsen, R. W., and Sieghart, W. (2009). GABA A receptors: subtypes provide diversity of function and pharmacology. *Neuropharmacology* 56 (1), 141–148. doi:10.1016/j.neuropharm.2008.07.045

Paxinos, G., and Watson, C. (1998). The rat brain in stereotaxic coordinates. San Diego: Academic Press.

Peng, Z., Huang, C. S., Stell, B. M., Mody, I., and Houser, C. R. (2004). Altered expression of the delta subunit of the GABAA receptor in a mouse model of temporal lobe epilepsy. *J. Neurosci.* 24 (39), 8629–8639. doi:10.1523/JNEUROSCI.2877-04.2004

Qiang, M., Li, J. G., Denny, A. D., Yao, J. M., Lieu, M., Zhang, K., et al. (2014). Epigenetic mechanisms are involved in the regulation of ethanol consumption in mice. *Int. J. Neuropsychopharmacol.* 18 (2) . doi:10.1093/ijnp/pyu072

Qiao, X., Yin, F., Ji, Y., Li, Y., Yan, P., and Lai, J. (2017). 5-Aza-2′-deoxycytidine in the medial prefrontal cortex regulates alcohol-related behavior and Ntf3-TrkC expression in rats. *PLoS One* 12 (6), e0179469. doi:10.1371/journal.pone.0179469

Randesi, M., van den Brink, W., Levran, O., Blanken, P., Butelman, E. R., Yuferov, V., et al. (2016). Variants of opioid system genes are associated with non-dependent opioid use and heroin dependence. *Drug Alcohol Depend.* 168, 164–169. doi:10.1016/j.drugalcdep.2016.08.634

Ribas, L., Vanezis, K., Imues, M. A., and Piferrer, F. (2017). Treatment with a DNA methyltransferase inhibitor feminizes zebrafish and induces long-term expression changes in the gonads. *Epigenet. Chromatin.* 10 (1), 59. doi:10.1186/s13072-017-0168-7

Robison, A. J., and Nestler, E. J. (2011). Transcriptional and epigenetic mechanisms of addiction. *Nat. Rev. Neurosci.* 12 (11), 623–637. doi:10.1038/nrn3111

Sun, A., Zhuang, D., Zhu, H., Lai, M., Chen, W., Liu, H., et al. (2015). Decrease of phosphorylated CREB and ERK in nucleus accumbens is associated with the incubation of heroin seeking induced by cues after withdrawal. *Neurosci. Lett.* 591, 166–170. doi:10.1016/j.neulet.2015.02.048

Suzuki, M. M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* 9 (6), 465–476. doi:10.1038/nrg2341

Tian, W., Zhao, M., Li, M., Song, T., Zhang, M., Quan, L., et al. (2012). Reversal of cocaine-conditioned place preference through methyl supplementation in mice:

altering global DNA methylation in the prefrontal cortex. *PLoS One* 7 (3), e33435. doi:10.1371/journal.pone.0033435

Wang, L., Jiang, W., Lin, Q., Zhang, Y., and Zhao, C. (2016). DNA methylation regulates gabrb2 mRNA expression: developmental variations and disruptions in L-methionine-induced zebrafish with schizophrenia-like symptoms. *Gene Brain Behav.* 15 (8), 702–710. doi:10.1111/gbb.12315

Whiting, P. J. (2003). The GABAA receptor gene family: new opportunities for drug development. *Curr. Opin. Drug Discov. Dev* 6 (5), 648–657. doi:10.1046/j.1471-4159.2002.01098.x

Wright, K. N., Hollis, F., Duclot, F., Dossat, A. M., Strong, C. E., Francis, T. C., et al. (2015). Methyl supplementation attenuates cocaine-seeking behaviors and cocaine-induced c-Fos activation in a DNA methylation-dependent manner. *J. Neurosci.* 35 (23), 8948–8958. doi:10.1523/JNEUROSCI.5227-14.2015

Xu, X., Ji, H., Liu, G., Wang, Q., Liu, H., Shen, W., et al. (2016). A significant association between BDNF promoter methylation and the risk of drug addiction. *Gene* 584 (1), 54–59. doi:10.1016/j.gene.2016.03.010

Ye, Q., Trivedi, M., Zhang, Y., Bohlke, M., Alsulimani, H., Chang, J., et al. (2019). Brain iron loading impairs DNA methylation and alters GABAergic function in mice. *Faseb. J.* 33 (2), 2460–2471. doi:10.1096/fj.201801116RR

Zhang, H., Zhang, L., Tang, Y., Wang, C., Chen, Y., Shu, J., et al. (2019). Systemic screening identifies GABRD, a subunit gene of GABAA receptor as a prognostic marker in adult IDH wild-type diffuse low-grade glioma. *Biomed. Pharmacother.* 118, 109215. doi:10.1016/j.biopha.2019.109215

Zhang, J. J., Han, J., and Sui, N. (2014). Okadaic acid blocks the effects of 5-aza-2-deoxycytidine on consolidation, acquisition and retrieval of morphine-induced place preference in rats. *Neuropharmacology* 86, 282–293. doi:10.1016/j.neuropharm.2014.08.005

Zhao, R., Zhang, R., Li, W., Liao, Y., Tang, J., Miao, Q., et al. (2013). Genome-wide DNA methylation patterns in discordant sib pairs with alcohol dependence. *Asia Pac. Psychiatr.* 5 (1), 39–50. doi:10.1111/appy.12010

Zhu, H., Lai, M., Chen, W., Mei, D., Zhang, F., Liu, H., et al. (2017a). N-acetylaspartylglutamate inhibits heroin self-administration and heroin-seeking behaviors induced by cue or priming in rats. *Neurosci. Bull.* 33 (4), 396–404. doi:10.1007/s12264-017-0140-3

Zhu, J., Zhu, F., Zhao, N., Mu, X., Li, P., Wang, W., et al. (2017b). Methylation of glucocorticoid receptor gene promoter modulates morphine dependence and accompanied hypothalamus-pituitary-adrenal axis dysfunction. *J. Neurosci. Res.* 95 (7), 1459–1473. doi:10.1002/jnr.23913

Zovkic, I. B., Guzman-Karlsson, M. C., and Sweatt, J. D. (2013). Epigenetic regulation of memory formation and maintenance. *Learn. Mem.* 20 (2), 61–74. doi:10.1101/lm.026575.112

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

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 08 March 2021
doi: 10.3389/fphar.2021.605803

Cinical, Metabolic, and Genetic Analysis and Follow-Up of Eight Patients With *HIBCH* Mutations Presenting With Leigh/Leigh-Like Syndrome

Junling Wang[1], *Zhimei Liu*[1], *Manting Xu*[1], *Xiaodi Han*[1], *Changhong Ren*[1], *Xinying Yang*[1], *Chunhua Zhang*[2] *and Fang Fang*[1]*

[1]*Department of Neurology, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China,* [2]*Department of Research, Development of MILS International, Ishikawa, Japan*

OPEN ACCESS

Edited by:
Ruth Roberts, ApconiX, United Kingdom

Reviewed by:
Charlotte L. Alston, Wellcome Trust Center for Mitochondrial Research (WT), United Kingdom
Cheryl D. Cropp, Samford University, United States

***Correspondence:**
Fang Fang
fangfang@bch.com.cn

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *16 September 2020*
Accepted: *22 January 2021*
Published: *08 March 2021*

Citation:
Wang J, Liu Z, Xu M, Han X, Ren C, Yang X, Zhang C and Fang F (2021) Cinical, Metabolic, and Genetic Analysis and Follow-Up of Eight Patients With HIBCH *Mutations Presenting With Leigh/Leigh-Like Syndrome. Front. Pharmacol. 12:605803. doi: 10.3389/fphar.2021.605803*

3-Hydroxyisobutyryl-CoA hydrolase (*HIBCH*, NM_014362.3) gene mutation can cause HIBCH deficiency, leading to Leigh/Leigh-like disease. To date, few case series have investigated the relationship between metabolites and clinical phenotypes or the effects of treatment, although 34 patients with *HIBCH* mutations from 27 families have been reported. The purpose of this study was to analyze the phenotypic spectrum, follow-up results, metabolites, and genotypes of patients with HIBCH deficiency presenting with Leigh/Leigh-like syndrome and explore specific metabolites related to disease diagnosis and prognosis through retrospective and longitudinal studies. Applying next-generation sequencing, we identified eight patients with *HIBCH* mutations from our cohort of 181 cases of genetically diagnosed Leigh/Leigh-like syndrome. Six novel *HIBCH* mutations were identified: c.977T>G [p.Leu326Arg], c.1036G>T [p.Val346Phe], c.750+1G>A, c.810-2A>C, c.469C>T [p.Arg157*], and c.236delC [p.Pro79Leufs*5]. The Newcastle Pediatric Mitochondrial Disease Scale (NPMDS) was employed to assess disease progression and clinical outcomes. The non-invasive approach of metabolite analysis showed that levels of some were associated with clinical phenotype severity. Five (5/7) patients presented with elevated C4-OH in dried blood spots, and the level was probably correlated with the NPMDS scores during the peak disease phase. 2,3-Dihydroxy-2-methylbutyrate in urine was elevated in six (6/7) patients and elevated S-(2-caboxypropyl) cysteamine in urine was found in three patients (3/3). The median age at initial presentation was 13 months (8–18 months), and the median follow-up was 2.3 years (range 1.3–7.2 years). We summarized and compared with all reported patients with *HIBCH* mutations. The most prominent clinical manifestations were developmental regression/delay, hypotonia, encephalopathy, and feeding difficulties. We administered drug and dietary treatment. During follow-up, five patients responded positively to treatment with a significant decrease in NPMDS scores. Our research is the largest case series of patients with *HIBCH* mutations.

Keywords: HIBCH gene, Leigh/Leigh-like syndrome, C4-OH, 2,3-dihydroxy-2-methylbutyrate (23DH2MB), mitochondrial disorders, children

INTRODUCTION

Leigh syndrome (LS), as the most common presentation of mitochondrial disorders in children, is a devastating neurodegenerative disease accompanied by symmetrical lesions in basal ganglia, brainstem, and with genetic and clinical heterogeneity. It sometimes presents in patients with normal laboratory findings, atypical neuroimaging or symptoms, but highly suggestive of LS, which is known as Leigh-like syndrome (Baertling et al., 2014). It is related to over 95 disease-associated genes and still in expanding scenery (Lake et al., 2016; Rahman and Rahman, 2018; Baldo and Vilarinho, 2020). LS has an estimated incidence of 1:40,000 newborns (Rahman et al., 1996). There is no research on the incidence of LS in Chinese. Currently, there is no standard treatment for LS and it is very challenging to performing large-scale clinical research, due to the variable phenotypes, genetic heterogeneity and early death.

The mitochondrial nuclear gene *HIBCH* (cytogenetic location: 2q32.2; GenBank accession number: NM_004092.3; OMIM* 610690), encodes the 3-hydroxyisobutyryl-CoA hydrolase (HIBCH) protein. It is mainly responsible for converting 3-hydroxyisobutryl-CoA to 3-hydroxyisobutyric acid in valine catabolism, as well as 3-hydroxypropanoyl-CoA to 3-hydroxypropionic acid, a hypothetical minor pathway for propionic acid catabolism, which feeds into the valine catabolic pathway (**Figure 1A**) (Shimomura et al., 1994; Peters et al., 2014). *HIBCH* mutations were first identified to cause HIBCH deficiency (HIBCHD, OMIM #250620) in 2007 (Loupatty et al., 2007), and it was reported that it could lead to Leigh/Leigh-like disease due to secondary oxidative phosphorylation (OXPHOS) defects (**Figure 1B**) (Ferdinandusse et al., 2013; Yamada et al., 2014; Soler-Alfonso et al., 2015; Stiles et al., 2015).

The most comprehensive case series of *HIBCH* mutations was a report of five Turkish cases from two unrelated families (Schottmann et al., 2016). To date, just 31 *HIBCH* gene mutations have been reported, with 34 patients on a broad phenotypic spectrum ranging from early death (Brown et al., 1982; D'Gama et al., 2020) to just movement disorder with survival into adulthood, although the estimated incidence of HIBCH deficiency was 1 in ~130,000 individuals (Stiles et al., 2015). The current understanding and detection of *HIBCH* mutations are insufficient. There is a lack of systematic research on the clinical, genotype, and biochemical characteristics of patients with *HIBCH* mutations. In the early stage, detection of HIBCH enzymatic activity and respiratory chain enzymes in patients'muscle or skin tissues can help clinicians make a diagnosis. However, enzymatic tests of biopsied tissue is an invasive method and sometimes yields inconsistent results (Loupatty et al., 2007; Reuter et al., 2014), limiting their clinical application. At present, genetic testing for disease diagnosis is more widespread. We propose that non-invasive metabolites should be investigated to assist early diagnosis and timely treatment.

Herein, we analyzed the genetic, metabolic, and clinical features and follow-up findings of 8 Chinese patients with *HIBCH* gene mutations identified by next-generation sequencing (NGS) combined with metabolic examination and who presented with Leigh/Leigh-like Syndrome.

METHODS

Patients

Eight patients with identified *HIBCH* mutations who were from our large cohort of case series with 181 cases of genetically

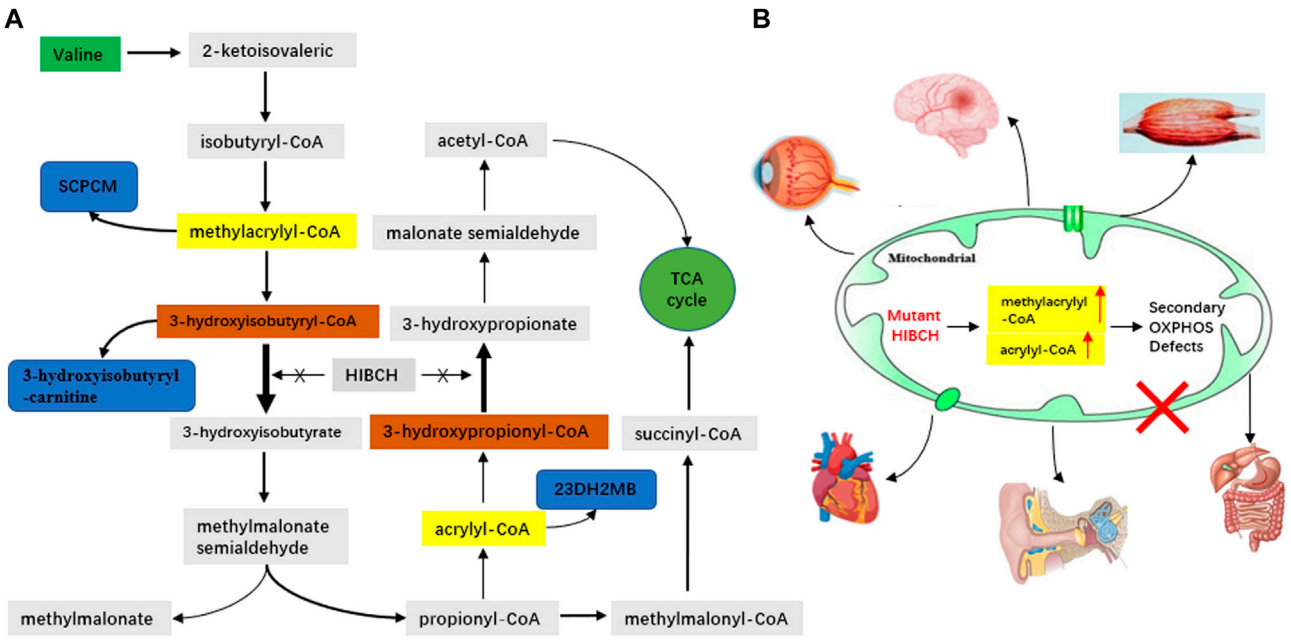

FIGURE 1 | Schematic representation of the valine metabolic pathway involved in HIBCH **(A)**, and mechanism of HIBCH deficiency with the resulting clinical phenotypes **(B)**. SCPCM, S-(2-caboxypropyl)cysteamine; 23DH2MB, 2,3-dihydroxy-2-methylbutyrate.

diagnosed Leigh/Leigh-like syndrome were recruited to participate in this study from October 2012 to December 2019.

Genetic Analyses

Due to different times of recruitment, we applied different NGS approaches: targeted panel sequencing (Patients 1 and 2) (Fang et al., 2017) and whole-exome sequencing (WES) (Patients 3, 4, 5, 6, 7, and 8). We performed Sanger sequencing to validate the identified *HIBCH* mutations and test the parental origin of available family members.

Briefly, genomic DNA was extracted from peripheral blood using TIANamp Blood DNA Kit (Tiangen, Beijing, China). Then 200 ng high-quality genomic DNA were utilized for library preparation. Genomic DNA was enzymatic sheared, end repaired, phosphorylation of the 5′ prime ends, a-tailing of the 3′ends and ligated to sequencing adapters, then PCR amplified following standard library preparation SOP, by using KAPA Hyper Plus Kits (Kapa Biosystems, Wilmington, MA, United States). The post-PCR libraries were captured by SureSelectXT Human All Exon V6 (Agilent, United States) or clinical exome analysis (6110 genes) (Agilent, United States), respectively. The final enriched libraries were sequenced with 2 × 150 bp on Illumina sequencers (Illumina, San Diego, CA, United States). The average sequencing depth for each sample was 149.3X (Patient 1), 174.9X (Patient 2), 148.7X (Patients 3 and 4), 133.9X (Patient 5), 130.5X (Patient 6), 132.02X (Patient 7), and 130.2X (Patient 8). Sequence data were aligned to the human genome reference (UCSC Genome Browser build hg19) using Burrows-Wheeler Aligner (BWA) (Li, 2012). Variant filtering and annotation was carried out with ANNOVAR software (Wang et al., 2010). Variants were screened as follows: 1) Preference to the variants related to the diseases, small INDEL, canonical splice sites, and nonsense variants. 2) Minor allele frequency (MAF) in normal populations <5% (except for known MAF ≥5% pathogenicity). 3) Preference to variants in the Human Gene Mutation Database (HGMD), ClinVar. 4) Preference to variants in the Online Mendelian Inheritance in Man database. Pathogenic variants were defined according to Standards and guidelines for the interpretation of sequence variants published by the American College of Medical Genetics (ACMG) in 2015 with Human Genome Variation Society nomenclature (Richards et al., 2015).

Additionally, The deletion at 2q32.2 including *HIBCH* gene was firstly identified by WES as the loss of heterogeneity in the Patient 8. Droplet-based digital PCR (ddPCR) was performed in the patient and his father to verify this deletion. Using QX200™ Droplet digital™ PCR systems, we performed the processes of sample dispersion, amplification and Quantification in turn. The number of droplets was obtained in the droplet reader based on the fluorescence amplitude (Hindson et al., 2011). And the data was analyzed by the software QuantaSoft™ analysis Pro 1.0.596.

Metabolite Measurements

Amino acids and acylcarnitines in dried blood spots (DBS) were analyzed by electrospray tandem mass spectrometry (LC-MS/MS) for C4-OH (composed of 3-hydroxyisobutyryl-carnitine, 3-hydroxy-butyryl-carnitine, and malonylcarnitine [C3-DC]). Samples were prepared as follows: a 3-mm diameter disc was punched in a microtiter plate, 100 μl internal standard solution was added, and the plate was covered. After gentle shaking at 45°C for 45 min, the solvent after elution was transferred into a new microtiter plate and analyzed immediately. An LCMS-8040 (Shimadzu, Japan) in positive mode was utilized for acylcarnitines and amino acids analysis. For MS/MS data acquisition, the peak including C4-OH (248.2>85) was observed with multiple reaction monitoring.

Seven patients' urine metabolites, including 2,3-dihydroxy-2-methylbutyrate (23DH2MB), were analyzed by standard gas chromatography-mass spectrometry (GC-MS). Briefly, 100 μl urine was incubated with 40 units urease at 37°C for 15 min. After adding an internal standard (20 μg of n-heptadecanoic acid), the sample was centrifuged to deproteinize by adding 1,000 μl of ethanol, then the supernatant was evaporated. The residue was completely dried under a nitrogen stream for 3 min and derivatized with 100 μl of N, O-bis(trimetnylsilily) trifluorcetamide and 10 μl of trimethylchlorosilane at 90°C for 40 min (Zhang et al., 2000). A GCMS-QP2010 Ultra GC-MS system (Shimadzu, Kyoto, Japan) was used with an ultra Alloy capillary column (30 × 0.25 mm internal diameter with 0.25-μm film thickness, Frontier Lab, Fukushima, Japan). The temperature was programmed to increase from 60 to 350°C at 17°C/min. Next, 2 μl of derivated sample was injected in split mode. The mass chromatographic quantitation of urine metabolites including 23DH2MB was based on the fragment ion peaks area compared with the corresponding urine creatinine fragment ion area ratio.

Through retrospective re-analysis, quantitative urine screening for S-(2-caboxypropyl) cysteamine (SCPCM) was measured by LCMS/MS in urine of cases. This project was not routinely performed in the metabolic laboratory. The process included sample and label pretreatment. The SCPCM concentration in urine of 140 control children was determined by the detection and verification of an SCPCM standard, and the reference value was obtained using the SCPCM standard concentration curve.

Data Collection and Follow-Up

At the first visit, clinical data were collected retrospectively, consisting of demographic features, early developmental milestones before first symptoms, age at presentation, initial present, precipitating causes, metabolite findings, and radiographic data.

During the follow-up period, all eight patients were assessed every 6 months to 1 year at our neurology clinic. Each follow-up evaluated motor, cognition and linguistics functions; muscle strength and tone; feeding situation; growth index; recurrence; metabolite analyses; and response to treatment.

Follow-up MRI results were classified into four levels: progression, stable, evolution, and regression (Bonfante et al., 2016). The Newcastle Pediatric Mitochondrial Disease Scale (NPMDS) was adopted to assess disease progression (Phoenix et al., 2006), and each patient was scored with the NPMDS before medication (peak period) and after treatment (the last follow-up). NPMDS scores were correlated with the degree of disease burden.

TABLE 1 | *HIBCH* mutations in eight patients.

Patient	Exon	Nucleotide variation[a]	Amino acid variation[b]	Reported/ novel	Parental origin	AF	Predication of pathogenicity			
							Mutation taster (score)	SIFT (score)	CADD (score)	PROVEAN (score)
1	12	c.977T>G	p.Leu326Arg	Novel	ND	—	Disease causing (1)	Damaging (0.00)	28.1	Damaging (–5.12)
	13	c.1027C>G	p.His343Asp	Reported	Maternal	4.067e-06	Disease causing (1)	Tolerable (0.089)	22.9	Damaging (–3.88)
2	7	c.452C>T	p.Ser151Leu	Reported	Maternal	—	Disease causing (1)	Damaging (0.0)	34.0	Damaging (–5.79)
	7	c.469C>T	p.Arg157*	Novel	Paternal	1.22e-05	Disease causing (1)	NA	43.0	NA
3		c.750+1G>A	NA	Novel	Maternal	—	NA	NA	34.0	NA
	13	c.1036G>T	p.Val346Phe	Novel	ND	—	Disease causing (1)	Damaging (0.00)	35.0	Damaging (–4.03)
4		c.750+1G>A	NA	Novel	Maternal	—	NA	NA	34.0	NA
	13	c.1036G>T	p.Val346Phe	Novel	ND	—	Disease causing (1)	Damaging (0.00)	35.0	Damaging (–4.03)
5	13	c.1027C>G	p.His343Asp	Reported	Maternal	4.067e-06	Disease causing (1)	Tolerable (0.089)	22.9	Damaging (–3.88)
		868 kb deletion including *HIBCH*		Novel	Paternal	NA	NA	NA	NA	NA
6	13	c.1027C>G	p.His343Asp	Reported	Paternal	4.067e-06	Disease causing (1)	Tolerable (0.089)	22.9	Damaging (–3.88)
		c.439-2A>G	NA	Reported	Maternal	—	NA	NA	34.0	NA
7	4	c.236delC	p.Pro79Leufs*5	Novel	Paternal	—	Disease causing (1)	NA	NA	NA
	13	c.1027C>G	p.His343Asp	Reported	Maternal	4.067e-06	Disease causing (1)	Tolerable (0.089)	22.9	Damaging (–3.88)
8		c.810-2A>C	NA	Novel	Maternal	—	NA	NA	34.0	NA
	9	c.743A>G	p.His248Arg	Reported	Paternal	8.86e-06	Disease causing (1)	Damaging (0.03)	23.1	Damaging (–3.37)

[a]*Transcript, NM_014362.3; Allele Frequency; —, Unrecorded; ND, not detection; NA, not available.*
[b]*NP_055177.2.*

Higher and lower scores indicate a more severe phenotype and clinical improvement, respectively.

Review 34 Reported Cases With *HIBCH* Mutations

We reviewed the data of 34 published patients, of which 4 cases had insufficient clinical data (Charng et al., 2016; 336; Ceyhan-Birsoy, et al., 2019; Jou et al., 2019; Hu et al., 2020). The data of onset and current age, clinical symptoms, metabolites, neuroimaging, genotype, and natural history were analyzed to summarizes the similarities and differences between our patients and other reported cases, both with the *HIBCH* mutations.

Analysis of Other 173 Patients With Leigh/Leigh-Like Syndrome

We grouped confirmed genes of 181 patients with Leigh/Leigh-like syndrome according to pathomechanism including OXPHOS complexes (I, II, IV, V), mitochondrial DNA-related protein synthesis, mitochondrial homeostasis, substrate, mitochondrial cofactors, and inhibitors. The age of onset, predisposing factors, initial symptoms, main symptoms, brain imaging, laboratory examinations, and clinical outcomes of 173 patients were investigated to analyze similarities and differences between subjects with *HIBCH* gene mutations and other gene mutations associated with Leigh/Leigh-like syndrome.

RESULTS

HIBCH Genetic Analyses

Ten mutations, including six novel mutations, were identified by NGS of the *HIBCH* gene with compound heterozygous states (**Table 1**), and according to the ACMG criteria, the ten variants were classified as likely pathogenic or pathogenic (**Supplementary Table S1**). No other candidate genes that could fully explain clinical symptoms were found. They were confirmed by Sanger sequencing. The paternal blood samples of Patient 1 and two siblings (Patients 3 and 4) were not available. The identified mutations were distributed on coding exons 4, 7, 9, 12, and 13 (**Figure 2**). Two novel missense mutations, c.977T>G (p.Leu326Arg) and c.1036G>T (p.Val346Phe), were absent from the Exome Aggregation Consortium (ExAC), 1,000 Genomes, gnomAD_genome, and ESP6500 databases, but conserved amino acid residues in multiple species and several in silico tools predicted these mutations as disease causing. Two novel splicing mutations (c.750+1G>A and c.810-2A>C) were predicted as likely pathogenic. One novel nonsense mutation c.469C>T (p.Arg157*) can disrupt protein function by changing the stop codon and was also absent from population databases

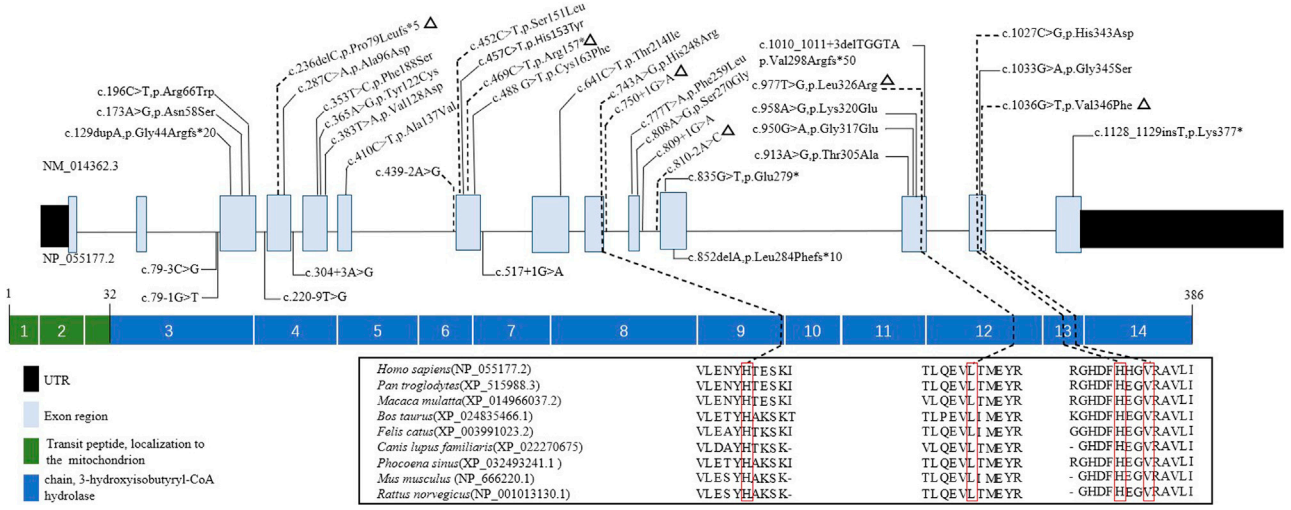

FIGURE 2 | Schematic representation the *HIBCH* gene structure showing disease-associated mutation locations and conservation of mutant amino acid residues with known protein domains. Notes: dashed line, mutations in our study; triangles, our novel mutations including missense, splicing, frameshift and nonsense mutations; solid line, mutations reported in the literature.

(gnomAD_genome, ExAC, 1,000 Genomes, and ESP6500). Another novel mutation c.236delC, p.Pro79Leufs*5, was predicted to be loss of function due to a premature truncation. The HGMD includes c.1027C>G (p.His343Asp) and c.439-2A>G, which have been reported in the literature (Zhu et al., 2015; Yang et al., 2018; Xu et al., 2019). The former was identified in four patients of this cohort, and from our NGS data there were no region of homology of the variant site. c.743A>G (p.His248Arg) and c.452C>T (p.Ser151Leu) are included in Clinvar, and both have conserved residues predicted to be damaging and disease causing. In addition, Sanger sequencing and droplet digital PCR identified maternal c.1027C>G (p.His343Asp) and a paternal 868-kb heterozygous deletion from 2q32.2 including the *HIBCH* gene in Patient 5 homozygous for the c.1027C>G (p.His343Asp) *HIBCH* variant. Our identified 6 novel mutation and 31 other reported mutation sites are summarized in **Figure 2**.

Metabolite Results and Laboratory Investigations

There were 5 (5/8) patients with DBS showing elevated C4-OH during either the peak or recovery phase. Six (6/7) had increased 23DH2MB levels in urine (**Figure 3**), including four cases in the peak stage and two in the recovery stage. Through retrospective analysis, we identified an increase in urine SCPCM levels in three cases. The data are shown in **Table 2**. There were no significant abnormalities of amino acid levels in DBS. In addition, three patients had mild elevation of lactic acid in blood, but no cases had elevated lactate in cerebrospinal fluid.

Additionally, we found that the NPMDS scores during the peak phase may be correlated with the value of C4-OH and SCPCM. We also observed that patients with lower C4-OH during the peak phase were more likely to have a positive response to treatment.

Demographic and Clinical Features

In this case series, eight patients (four male, four female) were from seven non-consanguineous Chinese families, and Patients three and four were siblings. The clinical features are summarized in **Table 3**. Two patients were ethnic minorities (Patient 6 Mongolian and Patient 8 Hmong). Seven were born at term with normal birth measurements, and the perinatal period was generally normal for all subjects. Early developmental milestones before disease onset were almost normal. All patients developed initial symptoms at ages ranging from 8 to 18 months at a median of 13 months. With regard to the initial symptoms, three cases developed encephalopathy, two cases had developmental regression, two cases showed developmental delay, and one case was dystonia. Six patients had precipitating causes, with five cases triggered by infection and one case by vaccination. Brain MRI scans of all cases at onset revealed T2 hyperintensity in the bilateral basal ganglia, and four cases also had brainstem lesions, similar to manifestation. In short, it was Leigh pattern imaging (**Figure 4**).

The median age of diagnosis was 2.0 years (range 1.0–5.7 years). During the disease course, the frequently prominent clinical features were hypotonia (8/8), developmental delay (8/8), developmental regression (7/8), encephalopathy (6/8), feeding difficulties (5/8), dystonia (4/8), ataxia (4/8), metabolic acidosis (4/8), strabismus (4/8), nystagmus (3/8), and seizures (3/8). Video electroencephalogram (VEEG) in two cases (Patients 3 and 4) with tonic seizures showed bilateral central and apical areas with sharp and slow wave emission that was more remarkable on the right side. The VEEG of Patient 8 with focal seizures presented as continuous release of slow waves and multiple spikes in the right frontal and temporal regions. Two cases (Patients 3 and 8) developed status epilepticus during acute encephalopathy. Three (Patients 2, 3, and 8) developed respiratory failure during acute encephalopathy, and all

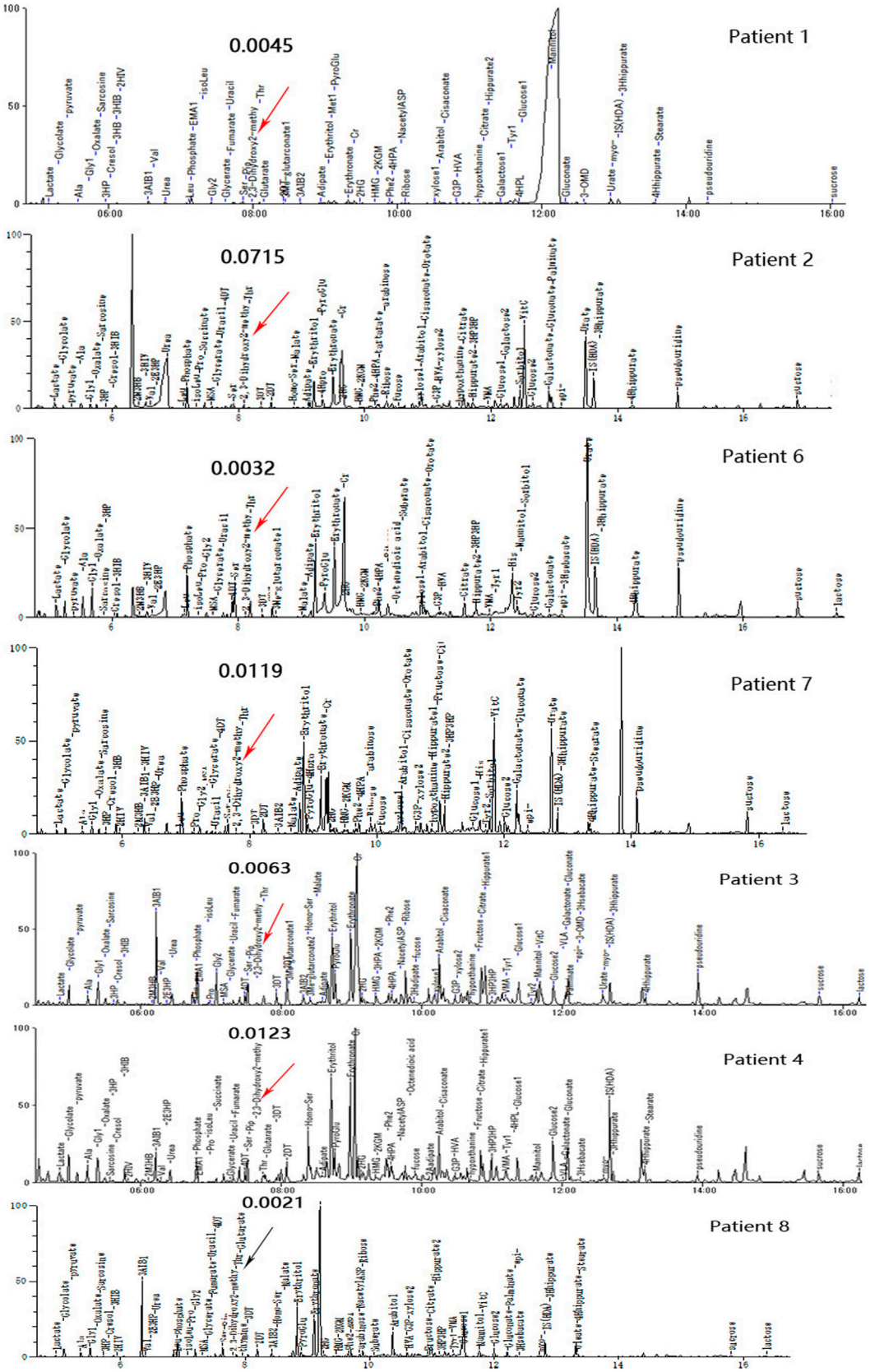

FIGURE 3 | Urine organic acid profiles of seven patients in different phases; red arrows indicate abnormally high 23DH2MB peaks, and the black arrow represents a normal 23DH2MB peak. Patients 1, 2, 6, and 7 with elevated 23DH2MB levels in the peak stage. Patients 3 and 4 with elevated 23DH2MB levels in the recovery stage. Patient 8 with a normal 23DH2MB level after adopting a valine-restricted diet.

TABLE 2 | Metabolite results of extensive investigations of eight patients with *HIBCH* mutations.

	C4-OH (0.00–0.26 μmol/L)			23DH2MB (0.0005–0.0029)		SCPCM (<0.624 μmol/mmol Cr)		Lactic acid	
	Peak phase	Recovery phase	Neonatal period	Peak phase	Recovery phase	Peak phase	Recovery phase	Blood (0.5–2.2 mmol/L)	CSF (1.0–2.78 mmol/L)
Patient 1	0.200	—	—	0.0045↑	—	—	—	1.51	1.46
Patient 2	1.664↑	—	—	0.0715↑	—	8.02↑	—	4.13↑	—
Patient 3	—	0.184	—	—	0.0063↑	—	—	1.60	—
Patient 4	0.579↑	0.183	—	—	0.0123↑	—	2.74↑	0.92	—
Patient 5	0.221	—	—	—	—	—	—	5.23↑	—
Patient 6	0.73↑	0.174	—	0.0032↑	—	2.45↑	—	3.23↑	—
Patient 7	0.485↑	—	—	0.0199↑	—	—	—	1.50	—
Patient 8	1.574↑	1.235↑	0.425↑	—	0.0021	—	—	1.2–2.1	1.50

↑, *elevated level;* —, *not detected; CSF, cerebrospinal fluid.*

required intubation and ventilatory support for ~1 week, of which one case (Patient 3) had a failed extubation and underwent a second intubation for ventilator-assisted breathing. Two cases (Patients 2 and 8) had two episodes of acute encephalopathy both induced by infection in 1 year. During acute encephalopathy period, two cases (Patient 2 and 4) were both associated with thyroid dysfunction (T3, T4 and TSH all decreased) and one case (Patient 6) with hepatic dysfunction (ALT 375 U/L, reference range 0–40 U/L), and about half a year after the peak period, these indicators gradually returned to normal.

Treatment

Once the patients diagnosed, we fully informed the parents about treatments and recommended pharmacologic therapy and adopting a valine-restricted diet. The drugs consisted of antioxidants and OXPHOS complex cofactors including L-carnitine (50 mg/kg/day), coenzyme Q10 (10 mg/kg/day), thiamine (10 mg/kg/day), and riboflavin (10 mg/kg/day), additionally including some symptomatic drugs, such as levetiracetam, or baclofen. For the dietary treatment, we worked with nutritionists to develop a plan to limit valine intake with protein intake of 1.0–1.5 g/kg/d, recommended energy intake of 80 kcal/kg/d, and no restriction on carbohydrates. It was recommended to use maple glycosuria formula milk powder and supplement leucine and isoleucine.

Two cases (Patients 1 and 5) only used drug treatments, 3 (Patients 6, 7, and 8) initiated drug and diet therapy, and 3 cases (Patients 2, 3, and 4) abandoned therapy. The NPMDS scores in the peak phase and at the last follow-up (recovery phase) were shown in **Figure 5.** The treated patients presented with significant decreased scores.

Follow-Up

The median duration between initial symptom onset and the last follow-up was 2.3 years (range 1.3–7.2 years). NPMDS scores during the peak phase and at the last evaluation, as well as MRI findings were showed in **Table 3**. The clinical conditions of three cases (Patients 1, 5, and 7) improved, while 2 (Patient 6 mild delay and Patient 8) were relatively stable; all five cases were in the treatment group. Conversely, all three cases (Patients 2, 3, and 4) in the non-treatment group progressed.

Patient 1 exhibited developmental regression following pneumonia at age 14 months but subsequently recovered. Unfortunately, at 4 years 5 months, she experienced acute encephalopathy due to a febrile infection, which again led to developmental regression, and she lost the skills of sitting and speaking. She received drug therapy thereafter and slowly recovered and gradually improved. The patient is currently 8 years 4 months old, with a normal growth index. She is independent and attends grade 3 at a normal elementary school but sometimes has trouble concentrating and wandering. She has some movement intolerance and walks with a slightly abnormal right lower limb posture. Despite having several fevers and other infections since initiating treatment, none have worsened her condition. Her last MRI scan showed significant improvement (**Figure 6**).

For Patient 5, roseola infantum at age 13 months caused her condition to deteriorate, presenting with no head control, no language, and unable to sit unassisted. Following drug therapy, she gradually recovered. Now she is 2 years 4 months old, can understand basic sentences, walk a dozen steps unsupported, and speak 2 and 3 words.

Patient 7 has received drug and diet treatments for 6 months and has improved. The therapy included drugs of L-carnitine 500 mg/day, coenzyme Q10 100 mg/day, thiamine 100 mg/day, riboflavin 100 mg/kg/day, and limit protein intake 12 g/d, no restriction on carbohydrates. At present she is 2 years 7 months, can walk without assistance but occasionally falls, can understand complex sentences, and pronounce monosyllables.

Patient 6 have received the standard therapy of drug and diet plus baclofen (1 mg/kg/day) since the onset of encephalopathy for nearly 1 year. Despite four bouts of pneumonia during this time, his condition did not deteriorate. He is 3 years and 2 months old, with dystonia as the main manifestation. He can walk for a few steps with support, understand general language, and pronounce sequential words.

Patient 8 was initiated on drug therapy plus levetiracetam (30 mg/kg/day) since onset at age 11 months with partial responses. However, 10 months later he experienced a second encephalopathy episode followed by frequent seizures and status epilepticus. Since being placed on a valine-restricted diet, he has shown a dramatic response.

TABLE 3 | Clinical features and follow-ups of eight patients with *HIBCH* mutation.

	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6	Patient 7	Patient 8
Gender	Female	Female	Male	Male	Female	Male	Female	Male
Ethnic origin	Han	Han	Han	Han	Han	Mongolian	Han	Hmong
Perinatal problems	—	—	—	—	DFM	Jaundice	—	—
Age at onset	1 years 2 months	10 months	1 year 6 months	8 months	1 year 1 months	1 year 8 months	1 year 1 months	11 months
Initial presentation	DR	Encephalopathy	DD	Encephalopathy	DR	PD	DD	Encephalopathy
Precipitating causes	Pneumonia	Diarrhea	—	Febrile illness	ES	HAV	—	Influenza
Age at diagnosis	4 years 7 months	1 year	5 years 8 months	4 years	1 year 2 months	2 years 2 months	1 year 10 months	1 year 2 months
Developmental regression/delay	+/+	+/+	+/+	+/+	+/+	+/+	−/+	+/+
Hypotonia	+	+	+	+	+	+	+	+
Encephalopathy	+	+	+	+	−	+	−	+
Metabolic acidosis	−	+	−	+	−	+	−	+
Feeding difficulties	−	+	+	+	−	+	−	+
Dystonia	−	+	+	−	−	+	−	+
Ataxia	+	−	+	+	−	−	−	+
Strabismus	−	−	+	+	−	+	−	+
Nystagmus	−	−	+	+	−	−	−	+
Seizures	−	−	+	+	−	−	−	+
MRI-basal ganglia of T2 hyperintensity	+	+	+	+	+	+	+	+
MRS (Lactate peak)	−	ND	+	−	ND	−	ND	−
NMPDS score in the peak phase	35.4	45.8	36.2	44.8	38.1	45.6	34.2	48.8
Therapy	+[a]	−	−	−	+[a]	+[b]	+[b]	+[b]
Age of last assessment	8 years 4 months	4 years	8 years	6 years 4 months	2 years 4 months	3 years 2 months	2 years 7 months	2 years 3 months
Follow-up MRI	Evolution	NA	Progression	Progression	NA	Stable	Evolution	Progression
NMPDS score in the last assessment	5.2	61.3	53.1	48.4	19.9	32.1	15.8	38.7

DFM, decreased fetal movement; DR, developmental regression; DD, developmental delay; PD, paroxysmal dystonia; ES, exanthema subitum; HAV, hepatitis A vaccine; ND, not detection; NA, not available.

[a]*Therapy with drug [L-carnitine (50 mg/kg/day), coenzyme Q10 (10 mg/kg/day), thiamine (10 mg/kg/day), riboflavin (10 mg/kg/day) and symptomatic drugs].*

[b]*Therapy with drug plus valine-restricted diet; Progression, new lesions present and/or extension of previously visualized lesions; Stable, no change in T2 or DWI; Evolution, normalization of DWI with persistent T2 changes, or decreased size of the T2 signal changes as a result of encephalomalacia.*

Although he has been seizure-free for 5 months, there has been no improvement in development. He is 2 years 4 months old, with no head control, language, or eye contact.

The conditions of all three cases without long-term treatment (Patients 2, 3, and 4) deteriorated. They had poor response to treatment in the early stage, and all eventually stopped therapy. At the last follow-up, Patient 2 was 4 years old, with height 90 cm (<3rd percentile) and weight 8.5 kg (<3rd percentile). She has poor head control with no eye contact or language. Patient 3 is currently 8 years old. He has no head control and is unable to sit unsupported or actively grasp objects. He knows simple words but is nonverbal. Patient 4 is presently 6 years 4 months old; his clinical course and physical examination are similar to that of his older brother (Patient 3).

Six (6/8) patients underwent follow-up neuroimaging (**Table 3**). Brain MRI showed progression in three cases, two of which abandoned therapy, and 1 (Patient 8) case with therapy had lesion progression to the entire right cerebral hemisphere with evident brain volume loss. One case (Patient 3) had corpus callosum thinning and a lactic acid peak on magnetic resonance spectroscopy (MRS). Brain MRI improved in two cases and was relatively stable in one case, and they were all receiving therapy.

Analysis of 34 Reported and Our Cases With *HIBCH* Mutations

The patient's general condition, clinical, imaging, and metabolic characteristics were summarized in **Table 4**, and there were no significant statistical difference between our cases series and other reported cases. In our cohort, there were no deaths, probably due to the older age of onset, but 7 (22%) of the reported cases had been fatal. The main symptoms of total cases were developmental delay/regression, hypotonia, distonia, encephalopathy, and feeding difficulties. T2 hyperintensity in the bilateral basal ganglia occurs in an average of 95% of cases. Brain atrophy and brainstem lesions were also common neuroimaging abnormalities. All patients tested for SCPCM presented with elevated levels, suggesting a more specific metabolite of the disease, and more samples were needed for verification. In both groups, there was an almost equal proportion (63 vs. 65%) of cases with elevated C4-OH levels. However, the

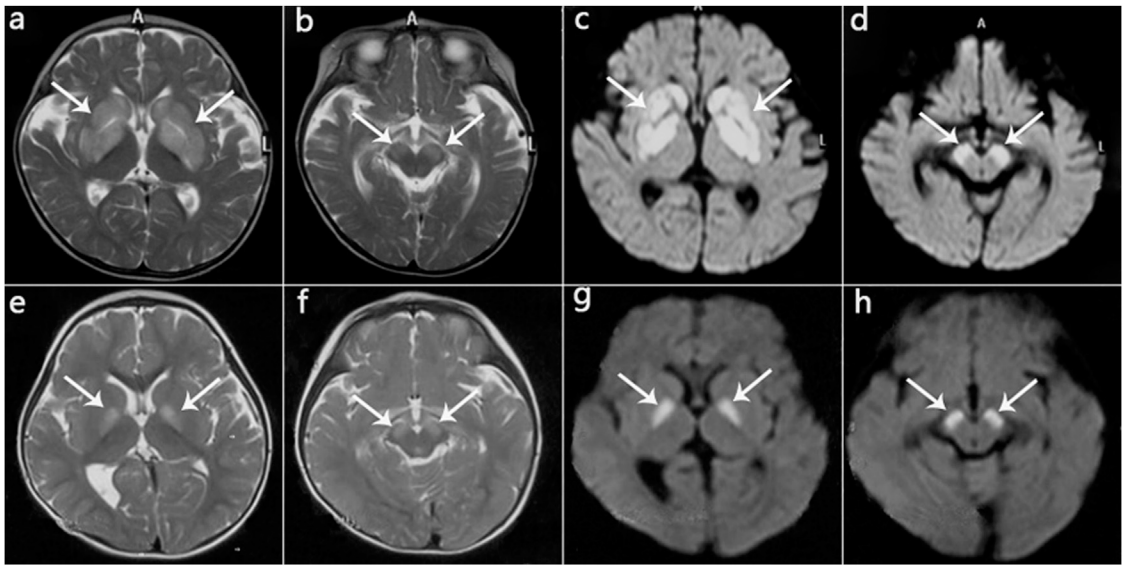

FIGURE 4 | MRI manifestations of 2 patients (Patient 2, Patient 5) during disease onset. **(a-d)**: Patient 2 at age 10 months; hyperintensity of the globus pallidus, putamen, caudate nucleus, and cerebral peduncle on T2WI **(a,b)** and DWI **(c,d)**. **(e-h)**: Patient 5 at age 13 months; hyperintensity on T2WI **(e,f)** and DWI **(g,h)** in the globus pallidus and cerebral peduncle.

percentage of patients with elevated 23DH2MB appeared to be higher in our cases, but the difference was not statistically significant due to the small sample size.

Analysis With 173 Patients With Leigh/Leigh-Like Syndrome

There were 43 other mutant genes identified in 173 cases with Leigh/Leigh-like syndrome (**Supplementary Figure S1**). We summarized the clinical, genetic, metabolic and imaging characteristics of the 173 cases. Due to baseline heterogeneity and large differences in sample size, we only performed statistical descriptions of the data (**Supplementary Table S2**). The symptoms of hypotonia, dystonia, encephalopathy, and feeding difficulties were more common in patients with *HIBCH* mutations. In metabolites, the elevated levels of blood C4-OH (62.5 vs. 2.3%) and urine 23DH2MB (85.7 vs. 5.8%) were significant in the *HIBCH* mutation group. Among the 173 patients group, 10 cases showed elevated 23DH2MB levels, including 8 with *ECHS1* mutation, 1 with *SURF1* mutation, and 1 with *MT-ND6* mutation, and the latter 2 patients showed slightly elevated 23DH2MB levels and normal C4-OH levels. This implied that elevated levels of C4-OH combined with 23DH2MB in patients with clinically diagnosed Leigh/Leigh-like Syndrome may serve as specific metabolites for *HIBCH* mutation patients.

DISCUSSION

We analyzed the clinical, metabolic, and genetic features of 8 patients with *HIBCH* mutations associated with Leigh/Leigh-like syndrome and reported their follow-up data providing additional natural history information. We also investigated the relationships between metabolite levels, disease severity, and clinical outcomes.

In our case series, earlier onset age was associated with a poorer prognosis. Infection or vaccination could also accelerate or exacerbate the onset or recurrence. There have been 34 patients from 27 unrelated families reported with *HIBCH* mutations, 4 of which have insufficient clinical data. Twenty-six (26/32) had initial symptoms within 2 years of age, including seven cases who died within an onset age of 1 year old (Brown et al., 1982; Ferdinandusse et al., 2013; Yamada et al., 2014; Peters et al., 2015; D'Gama et al., 2020). The significant clinical presentations were developmental delay/regression, hypotonia, encephalopathy, feeding difficulties and dystonia, and the last 4 symptoms were more specific in patients with Leigh/Leigh-like syndrome caused by *HIBCH* mutation in our cohort. Interestingly, seizures seemed to be related to phenotype severity, but it occurred infrequently with only 3 patients in this cohort. Among 10 (10/33) patients described elsewhere, most of had poor responses to treatment and worse prognoses (Loupatty et al., 2007; Ferdinandusse et al., 2013; Reuter et al., 2014; Candelo et al., 2019; Karimzadeh et al., 2019; D'Gama et al., 2020; Marti-Sanchez et al., 2020). The neuroimaging of all 8 patients showed symmetrical lesions in basal ganglia with/without brainstem. Lactate peaks on MRS (Soler-Alfonso et al., 2015; Candelo et al., 2019; D'Gama et al., 2020) and corpus callosum abnormalities (Brown et al., 1982; Peters et al., 2015; Tan et al., 2018; Candelo et al., 2019; D'Gama et al., 2020) are uncommon in cases with *HIBCH* mutation, but it may be associated with a more severe phenotype. Additionally, in our cohort there were 2 cases with thyroid dysfunction during the onset phase who had negative prognoses, and only 1 case were

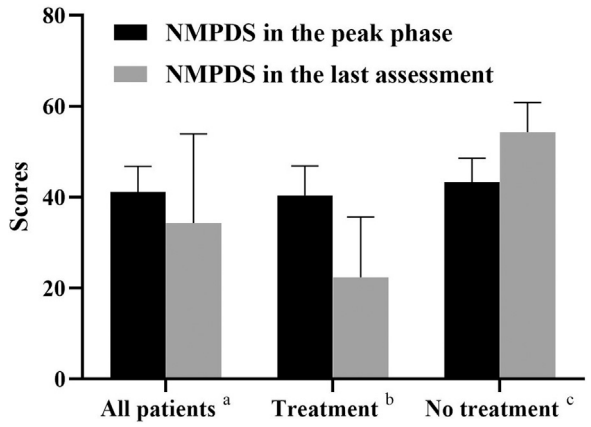

FIGURE 5 | NMPDS scores in the peak phase and last assessment. a, All the recruited patients (n = 8); b, Patients who received therapy (Patient 1,5,6,7,8; n = 5); c, Patients who gave up therapy (Patient 2,3,4; n = 3).

reported in the recent literature (Marti-Sanchez et al., 2020). Furthermore, one case had hepatic dysfunction as a rare presentation, and it was proposed that HIBCH might protect cells against methacrylyl-CoA toxicity (Ishigure et al., 2001). The follow-ups showed differences in the prognoses of these patients, which may be related to therapeutic interventions or their genotype.

It is difficult to correlate the genotypes and phenotypes of HIBCH deficiency. The variation sites reported to date are distributed in all exons except exons 1 and 2. Of these, exon 12 has the most mutation sites and cases, but phenotype severity varies (Ferdinandusse et al., 2013; Schottmann et al., 2016). Additionally, there have been no identified pathogenic mutations impacting the key residues of Glu120, Gly146, Glu169, or Asp177 (substrate binding). It was reported that truncating mutations tend to cause a more severe phenotype (Tan et al., 2018). One study observed a longer survival in cases with homozygous mutations located on the protein surface than in those with variants inside or near the catalytic site (Marti-Sanchez et al., 2020). Disease severity may be related to residual HIBCH enzyme activity (Yamada et al., 2014). The p.His343Asp mutation seemed to be associated with less severe clinical presentations in this cohort. To date, there has been no confirmed founder or hot spot mutation of the *HIBCH* gene. In our cases, c.1027C>G [p.His343Asp] was identified in half of patients, in addition to two of the four Chinese cases previously reported (Zhu et al., 2015; Xu et al., 2019), but it has not been described in other countries. Additionally, there is no region of homology from our NGS data, and it is at a extremely low frequency in population databases. We speculate that p.His343Asp may be a hot spot mutation in Chinese. The major limitation of our study is the absent detection of HIBCH enzyme activity in recruited patients. Instead, we assessed specific metabolites in blood and urine, so that patients could easily accept and assist in disease diagnosis.

HIBCH deficiency leads to the accumulation of 3-hydroxy-isobutyryl-CoA, but it was hypothesized that methacrylyl-CoA

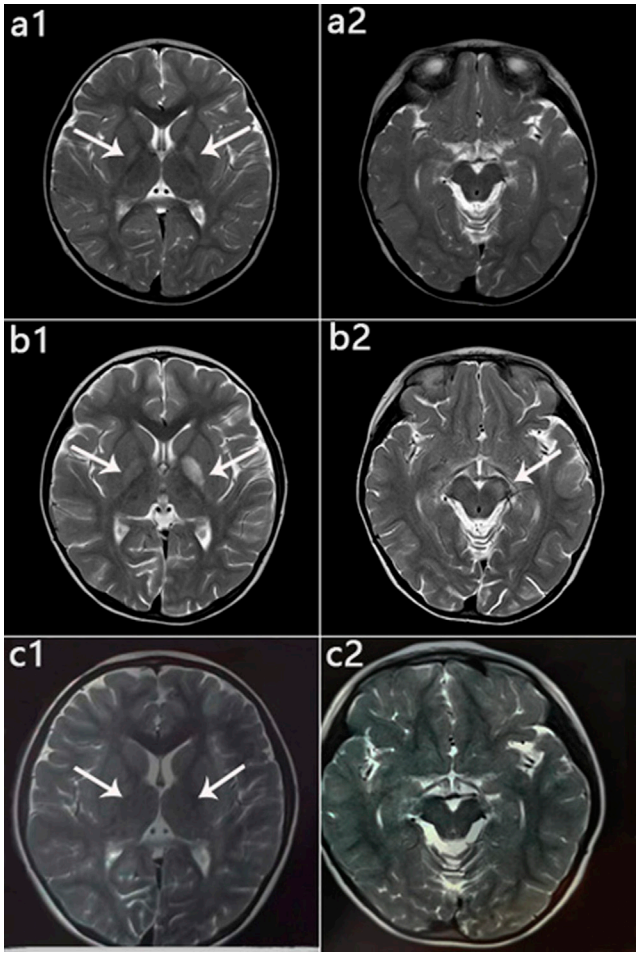

FIGURE 6 | Patient 1: MRI performance in the peak and recovery phases. **(a1,a2)**: Initial MRI obtained at age 2 years; hyperintensity on T2WI in the globus pallidus **(a1)**, normal cerebral peduncle **(a2)**. **(b1,b2)**: MRI performed at 4 years 5 months during the acute stage; hyperintensity on T2WI in the globus pallidus with swelling on the left side **(b1)** and in the left cerebral peduncle **(b2)**. **(c1,c2)**: MRI performed at 6 years 1 month during the recovery stage; only slightly hyperintensity on T2WI in globus pallidus **(c1)**, and Abnormal signal disappeared in the cerebral peduncle **(c2)**.

(**Figure 1**), as an upstream highly reactive intermediates, resulted in disease pathogenesis and neurotoxicity. The excessive generated methacrylyl-CoA can react with cysteamine to form SCPCM. 23DH2MB was derived from acryloyl-CoA, a compound homologous to methacrylyl-CoA, involving in valine metabolic pathway. In our cohort, elevated C4-OH and 23DH2MB levels are rare in Leigh/Leigh-like syndrome patients with other gene mutations. Our results indicate that C4-OH in the blood and 23DH2MB and SCPCM in the urine can be used as specific metabolites to assist the diagnosis and evaluate disease severity, although they are not completely specific.

The elevated C4-OH level is the direct result of the accumulation of 3-hydroxy-isobutyryl-CoA, according to acylcarnitine analyses in approximately 15 reported patients with confirmed *HIBCH* mutations. In this cohort, a higher C4-OH level in the peak phase was associated with a more

TABLE 4 | The clinical, neuroimaging and metabolic characteristics of our cases and the reported *HIBCH* mutation patients.

	Our cases	Reported cases	Total cases
General demographic features	N = 8	N = 32	N = 40
Sex	4 F: 4 M	17 F: 15 M	21 F: 19 M
Family history	2 (25%)	15 (47%)	17 (43%)
Age onset, median (range)	13, (8–20) months	4, (0–84) months	6, (0–84) months
Precipitating cause	6 (75%)	12 (38%)	18 (45%)
Current age, median (range)	4.0, (2.3–8.3) years	5.0, (0.25–43) years	5.0, (0.25–43) years
Death cases	0	7 (22%)	7 (18%)
Death age, median (range)	—	36, (0.9–96) months	36, (0.9–96) months
Main clinical features	N = 8	N = 30	N = 38
Developmental delay	5 (63%)	24 (80%)	29 (76%)
Developmental regression	7 (88%)	17 (57%)	24 (63%)
Hypotonia	8 (100%)	20 (67%)	28 (74%)
Encephalopathy	6 (75%)	13 (43%)	19 (50%)
Dystonia	4 (50%)	15 (50%)	19 (50%)
Feeding difficulties	5 (63%)	14 (47%)	19 (50%)
Ataxia	4 (50%)	11 (37%)	15 (40%)
Seizures	3 (38%)	10 (33%)	13 (34%)
Nystagmus	3 (38%)	11 (37%)	14 (37%)
Strabismus	4 (50%)	6 (20%)	10 (26%)
Optic atrophy	0	5 (17%)	5 (13%)
Brain imaging involvements	N = 8	N = 30	N = 38
Basal ganglia	8 (100%)	28 (93%)	36 (95%)
Brainstem	6 (75%)	15 (50%)	21 (55%)
White matter	3 (38%)	10 (33%)	13 (34%)
Cerebellum	3 (38%)	7 (23%)	10 (26%)
Brain atrophy	3 (38%)	18 (60%)	21 (55%)
Corpus callosum	1 (13%)	5 (17%)	6 (16%)
MRS lactate peak	1 (13%)	3 (10%)	4 (11%)
Metabolic studies			
Increased blood lactate level	3/8 (38%)	12/27 (44%)	15/35 (43%)
Elevated urine SCPCM	3/3 (100%)	2/2 (100%)	5/5 (100%)
Elevated blood C4-OH	5/8 (63%)	15/23 (65%)	20/31 (65%)
Elevated urine 23HD2MB	6/7 (86%)	2/8 (25%)	8/15 (53%)
Decreased HIBCH activity in skin fibroblasts	—	11/11 (100%)	11/11 (100%)

severe phenotype, indicating poor prognosis. However, in current clinical laboratories, typical LCMS/MS cannot distinguish between 3-hydroxyisobutyryl-carnitine and 3-hydroxybutyryl carnitine isomers. This requires comprehensive clinical analysis. Sometimes C4-OH was partially elevated in the asymptomatic neonatal period (Stiles et al., 2015) and was easily overlooked. For example, Patient 8 had a mild increase in C4-OH in the asymptomatic neonatal period, but it did not attract enough attention to make a timely diagnosis and initiate treatment.

Probably due to different stages and different severity of the disease, 23DH2MB in our cases showed a higher positive rate compared with reported cases. 23DH2MB levels did appear to change in response to treatment. After Patient 8 adopted a valine-restricted diet for 1 month, 23DH2MB fell to the reference value, while C4-OH had not yet returned to a normal level. The first measurement of 23DH2MB was in normal urine with trace amounts (Thompson et al., 1975), and also detected in patients with propionate disorders. In the valine pathway, it is originated from acryloyl-CoA, but the exact mechanism is not clear. 23DH2MB abnormalities were also identified in a few Leigh syndrome patients with *ECHS1* mutations (Peters et al., 2014; Yang and Yu, 2020). There was only 2 reported patient with an *HIBCH* mutation with elevated urinary 23DH2MB excretion (Ferdinandusse et al., 2013; Marti-Sanchez et al., 2020). We identified 6 patients with *HIBCH* mutations with elevated 23DH2MB levels.

SCPCM as a conjugate of the toxic metabolite methacrylyl-CoA, was firstly identified in HIBCH deficiency patients (Brown et al., 1982). It was an unusual amino acid in the urine of *HIBCH* mutation (Peters et al., 2015) and *ECHS1* mutation patients, and cannot be detected in routine metabolic screening. In our case series, urinary SCPCM level correlated with both disease severity and prognosis. To date, there have been 5 patients identified with *HIBCH* mutations and elevated SCPCM levels in urine, including 3 cases in our study. SCPCM seems to be more specific for disease diagnosis, but more samples are needed for verification.

There is currently no consensus on HIBCH deficiency treatment approaches. A low-valine, carbohydrate-rich diet may be effective (Loupatty et al., 2007; Soler-Alfonso et al., 2015). This treatment can avoid excessive ATP production from aberrant valine metabolism. Other possible drug therapies involve the administration of antioxidants and supplementation with vitamins/cofactors, often termed nutraceuticals. Fewer than 10 cases have reported treatment outcomes that include improvement (Loupatty et al., 2007; Soler-Alfonso et al., 2015), progression and even death (Yamada et al., 2014). It is difficult to correlate therapeutic effects with gene mutation sites and clinical manifestations, which may be related to residual HIBCH enzyme activity. Patients in this cohort

with the c.1027C>G [p.His343Asp] mutation seemed to responded well to treatment, similar to previous studies (Zhu et al., 2015; Xu et al., 2019).

In real clinical work, metabolites can be affected by the severity of disease and diet state, and false positive or negative results may occur. Genetic testing, especially targeted NGS (WES/panels), is the first-line auxiliary diagnostic approach. However, the targeted NGS testing is costly, and it can not identify possible pathogenic variants in deep intron regions and larger deletions/duplications. If the metabolites and clinical phenotypes are well defined and consistent, it may be recommended to directly perform the less costly approach of Sanger sequencing. Unfortunately, we did not perform enzymatic detection of HIBCH. We should combine clinical, metabolic and genetic analysis to diagnose the disease. Early diagnosis and treatment may delay disease progression of the disease. The main limitations of our study were the small sample size and the absence of formal clinical trials.

In conclusion, we identified 10 rare *HIBCH* gene variants in 8 patients with Leigh/Leigh-like syndrome, including 6 novel variants. The phenotype of Leigh/Leigh-like syndrome caused by *HIBCH* mutations was not significantly different from other gene mutations, except for more comment presentations with hypotonia, dystonia, and acute encephalopathy. The genotype-phenotype correlation requires further research. However, measurements of metabolites including C4-OH, 23DH2MB and SCPCM were relatively specific and also associated with disease severity, therapeutic effects, and possibly prognosis. The non-invasive tools of NGS and metabolite analyses should be considered first-line auxiliary diagnostic approaches. Patients with Leigh/Leigh-like syndrome with *HIBCH* mutations may have a positive prognosis, especially with early diagnosis and timely therapy.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Beijing Children's Hospital, Capital Medical University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

FF, JW, and ZL Conceived and designed the research. XH, CZ, CR, and XY collected the data and performed the study. JW, ZL, and MX analyzed the data. JW wrote the manuscript. FF revised the manuscript. All authors reviewed and approved the final version.

FUNDING

This work was supported by the Capital Health Research and Development Fund (2018-2-2096), and National Natural Science Foundation of China(81541115).

ACKNOWLEDGMENTS

We would like to thank the patients and the many clinicians and laboratory scientists who contributed to this study. We also acknowledge MILS (Beijing) Medical Laboratory.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.605803/full#supplementary-material.

REFERENCES

Baertling, F., Rodenburg, R. J., Schaper, J., Smeitink, J. A., Koopman, W. J., Mayatepek, E., et al. (2014). A guide to diagnosis and treatment of Leigh syndrome. *J. Neurol. Neurosurg. Psychiatry* 85 (3), 257–265. doi:10.1136/jnnp-2012-304426

Baldo, M. S., and Vilarinho, L. (2020). Molecular basis of Leigh syndrome: a current look. *Orphanet J. Rare Dis.* 15 (1), 31. doi:10.1186/s13023-020-1297-9

Bonfante, E., Koenig, M. K., Adejumo, R. B., Perinjelil, V., and Riascos, R. F. (2016). The neuroimaging of Leigh syndrome: case series and review of the literature. *Pediatr. Radiol.* 46, 443–451. doi:10.1007/s00247-015-3523-5

Brown, G. K., Hunt, S. M., Scholem, R., Fowler, K., Grimes, A., Mercer, J. F., et al. (1982). Beta-hydroxyisobutyryl coenzyme A deacylase deficiency: a defect in valine metabolism associated with physical malformations. *Pediatrics* 70 (4), 532–538. doi:10.1203/00006450-198212000-00014

Candelo, E., Cochard, L., Caicedo-Herrera, G., Granados, A. M., Gomez, J. F., Díaz-Ordoñez, L., et al. (2019). Syndromic progressive neurodegenerative disease of infancy caused by novel variants in HIBCH: report of two cases in Colombia. *Intractable Rare Dis. Res.* 8 (3), 187–193. doi:10.5582/irdr.2019.01014

Ceyhan-Birsoy, O., Murry, J. B., Machini, K., Lebo, M. S., Yu, T. W., Fayer, S., et al. (2019). Interpretation of genomic sequencing results in healthy and ill newborns: results from the BabySeq project. *Am. J. Hum. Genet.* 104 (1), 76–93. doi:10.1016/j.ajhg.2018.11.016

Charng, W., Karaca, E., Akdemir, Z. C., Gambin, T., Atik, M. M., Gu, S., et al. (2016). Exome sequencing in mostly consanguineous Arab families with neurologic disease provides a high potential molecular diagnosis rate. *BMC Med. Genomics* 9 (1), 42. doi:10.1186/s12920-016-0208-3

D'Gama, A. M., Brucker, W. J., Zhang, T., Gubbels, C. S., Ferdinandusse, S., Shi, J., et al. (2020). A phenotypically severe, biochemically "silent" case of HIBCH deficiency in a newborn diagnosed by rapid whole exome sequencing and enzymatic testing. *Am. J. Med. Genet.* 182 (4), 780–784. doi:10.1002/ajmg.a.61498

Fang, F., Liu, Z., Fang, H., Wu, J., Shen, D., Sun, S., et al. (2017). The clinical and genetic characteristics in children with mitochondrial disease in China. *Sci. China Life Sci.* 60, 746–757. doi:10.1007/s11427-017-9080-y

Ferdinandusse, S., Waterham, H. R., Heales, S. J., Brown, G. K., Hargreaves, I. P., Taanman, J. W., et al. (2013). HIBCH mutations can cause Leigh-like disease with combined deficiency of multiple mitochondrial respiratory chain enzymes and pyruvate dehydrogenase. *Orphanet J. Rare Dis.* 8, 188. doi:10.1186/1750-1172-8-188

Hindson, B. J., Ness, K. D., Masquelier, D. A., Belgrader, P., Heredia, N. J., Makarewicz, A. J., et al. (2011). High-throughput droplet digital PCR

system for absolute quantitation of DNA copy number. *Anal. Chem.* 83 (22), 8604–8610. doi:10.1021/ac202028g

Hu, C. P., Li, X. H., Zhao, L., Shi, Y. Y., Wu, B. B., Zhou, S. Z., et al. (2020). Clinical and molecular characterization of pediatric mitochondrial disorders in south of China. *Eur. J. Med. Genet.* 63, 103898. doi:10.1016/j.ejmg.2020.103898

Ishigure, K., Shimomura, Y., Murakami, T., Kaneko, T., Takeda, S., Inoue, S., et al. (2001). Human liver disease decreases methacrylyl-CoA hydratase and beta-hydroxyisobutyryl-CoA hydrolase activities in valine catabolism. *Clin. Chim. Acta* 312 (1–2), 115–121. doi:10.1016/s0009-8981(01)00597-6

Jou, C., Ortigoza-Escobar, J. D., O'Callaghan, M. M., Nascimento, A., Darling, A., Pias-Peleteiro, L., et al. (2019). Muscle involvement in a large cohort of pediatric patients with genetic diagnosis of mitochondrial disease. *J. Clin. Med.* 8 (1), 68. doi:10.3390/jcm8010068

Karimzadeh, P., Saberi, M., Sheidaee, K., Nourbakhsh, M., and Keramatipour, M. (2019). 3-Hydroxyisobutyryl-CoA hydrolase deficiency in an Iranian child with novel HIBCH compound heterozygous mutations. *Clin. Case Rep.* 7 (2), 375–380. doi:10.1002/ccr3.1998

Lake, N. J., Compton, A. G., Rahman, S., and Thorburn, D. R. (2016). Leigh syndrome: one disorder, more than 75 monogenic causes. *Ann. Neurol.* 79 (2), 190–203. doi:10.1002/ana.24551

Li, H. (2012). Exploring single-sample SNP and INDEL calling with whole-genome de novo assembly. *Bioinformatics* 28, 1838–1844. doi:10.1093/bioinformatics/bts280

Loupatty, F. J., Clayton, P. T., Ruiter, J. P., Ofman, R., Ijlst, L., Brown, G. K., et al. (2007). Mutations in the gene encoding 3-hydroxyisobutyryl-CoA hydrolase results in progressive infantile neurodegeneration. *Am. J. Hum. Genet.* 80 (1), 195–199. doi:10.1086/510725

Marti-Sanchez, L., Baide-Mairena, H., Marcé-Grau, A., Pons, R., Skouma, A., López-Laso, E., et al. (2020). Delineating the neurological phenotype in children with defects in the ECHS1 or HIBCH gene. *J. Inherit. Metab. Dis.* 1–14. doi:10.1002/jimd.12288

Peters, H., Buck, N., Wanders, R., Ruiter, J., Waterham, H., Koster, J., et al. (2014). ECHS1 mutations in Leigh disease: a new inborn error of metabolism affecting valine metabolism. *Brain* 137, 2903–2908. doi:10.1093/brain/awu216

Peters, H., Ferdinandusse, S., Ruiter, J. P., Wanders, R. J., Boneh, A., and Pitt, J. (2015). Metabolite studies in HIBCH and ECHS1 defects: implications for screening. *Mol. Genet. Metab.* 115 (4), 168–173. doi:10.1016/j.ymgme.2015.06.008

Phoenix, C., Schaefer, A. M., Elson, J. L., Morava, E., Bugiani, M., Uziel, G., et al. (2006). A scale to monitor progression and treatment of mitochondrial disease in children.*Neuromuscul. Disord.* 16 (12), 814–820. doi:10.1016/j.nmd.2006.08.006

Rahman, J., and Rahman, S. (2018). Mitochondrial medicine in the omics era. *Lancet* 391 (10139), 2560–2574. doi:10.1016/S0140-6736(18)30727-X

Rahman, S., Blok, R. B., Dahl, H. H., Danks, D. M., Kirby, D. M., Chow, C. W., et al. (1996). Leigh syndrome: clinical features and biochemical and DNA abnormalities. *Ann. Neurol.* 39 (3), 343–351. doi:10.1002/ana.410390311

Reuter, M. S., Sass, J. O., Leis, T., Köhler, J., Mayr, J. A., Feichtinger, R. G., et al. (2014). HIBCH deficiency in a patient with phenotypic characteristics of mitochondrial disorders. *Am. J. Med. Genet. A.* 164A (12), 3162–3169. doi:10.1002/ajmg.a.36766

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

Schottmann, G., Sarpong, A., Lorenz, C., Weinhold, N., Gill, E., Teschner, L., et al. (2016). A movement disorder with dystonia and ataxia caused by a mutation in the HIBCH gene. *Mov Disord.* 31 (11), 1733–1739. doi:10.1002/mds.26704

Shimomura, Y., Murakami, T., Fujitsuka, N., Nakai, N., Sato, Y., Sugiyama, S., et al. (1994). Purification and partial characterization of 3-hydroxyisobutyryl-coenzyme A hydrolase of rat liver. *J. Biol. Chem.* 269 (19), 14248–14253. doi:10.1016/s0021-9258(17)36781-9

Soler-Alfonso, C., Enns, G. M., Koenig, M. K., Saavedra, H., Bonfante-Mejia, E., and Northrup, H. (2015). Identification of HIBCH gene mutations causing autosomal recessive Leigh syndrome: a gene involved in valine metabolism. *Pediatr. Neurol.* 52 (3), 361–365. doi:10.1016/j.pediatrneurol.2014.10.023

Stiles, A. R., Ferdinandusse, S., Besse, A., Appadurai, V., Leydiker, K. B., Cambray-Forker, E. J., et al. (2015). Successful diagnosis of HIBCH deficiency from exome sequencing and positive retrospective analysis of newborn screening cards in two siblings presenting with Leigh's disease. *Mol. Genet. Metab.* 115 (4), 161–167. doi:10.1016/j.ymgme.2015.05.008

Tan, H., Chen, X., Lv, W., Linpeng, S., Liang, D., and Wu, L. (2018). Truncating mutations of HIBCH tend to cause severe phenotypes in cases with HIBCH deficiency: a case report and brief literature review. *J. Hum. Genet.* 63 (7), 851–855. doi:10.1038/s10038-018-0461-8

Thompson, J. A., Markey, S. P., and Fennessey, P. V. (1975). Gas-chromatographic/mass-spectrometric identification and quantitation of tetronic and deoxytetronic acids in urine from normal adults and neonates. *Clin. Chem.* 21 (13), 1892–1898. doi:10.1093/clinchem/21.13.1892

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38 (16), e164. doi:10.1093/nar/gkq603

Xu, Y., Zhang, J., Yu, K., Feng, F., Sun, X., Li, C., et al. (2019). A therapeutic regimen for 3-hydroxyisobutyryl-CoA hydrolase deficiency with exercise-induced dystonia. *Eur. J. Paediatr. Neurol.* 23 (5), 755–759. doi:10.1016/j.ejpn.2017.11.004

Yamada, K., Naiki, M., Hoshino, S., Kitaura, Y., Kondo, Y., Nomura, N., et al. (2014). Clinical and biochemical characterization of 3-hydroxyisobutyryl-CoA hydrolase (HIBCH) deficiency that causes Leigh-like disease and ketoacidosis. *Mol. Genet. Metab. Rep.* 1, 455–460. doi:10.1016/j.ymgmr.2014.10.003

Yang, H. Y., Wu, L. W., Deng, X. L., Yin, F., and Yang, L. F. (2018). Diagnosis and treatment of 3-hydroxyisobutyryl-CoA hydrolase deficiency: a case report and literature review. *Zhongguo Dang Dai Er Ke Za Zhi* 20 (8), 647–651. doi:10.7499/j.issn.1008-8830.2018.08.009

Yang, H., and Yu, D. (2020). Clinical, biochemical and metabolic characterization of patients with short-chain enoyl-CoA hydratase(ECHS1) deficiency: two case reports and the review of the literature. *BMC Pediatr.* 20 (1), 50. doi:10.1186/s12887-020-1947-z

Zhang, C., Xu, K., Dave, U. P., Wang, Y., and Matsumoto, I. (2000). Inborn errors of metabolism discovered in Asian department of pediatrics and mental retardation research center. *J. Chromatogr. B Biomed. Sci. Appl.* 746 (1), 41–49. doi:10.1016/s0378-4347(00)00087-6

Zhu, H., Bao, X., and Zhang, Y. (2015). 3-Hydroxy-isobutyryl-CoA hydrolase deficiency in a child with Leigh-like syndrome and literature review. *Zhonghua Er Ke Za Zhi.* 53 (8), 626–630. doi:10.3760/cma.j.issn.0578-1310.2015.08.017

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 12 April 2021
doi: 10.3389/fgene.2021.616693

Disorder of Sexual Development Males With XYY in Blood Have Exactly X/XY/XYY Mosaicism in Gonad Tissues

Yongjia Yang[1], Fang Chen[1], Zhenqing Luo[1], Yu Zheng[1], Jiayong Zheng[2], Yuyan Fu[1], Weijian Chen[3] and Haiyan Luo[4]**

[1] The Laboratory of Genetics and Metabolism, Hunan Children's Research Institute, Hunan Children's Hospital, University of South China, Changsha, China, [2] Key Laboratory of Obstetrics and Gynecology, Wenzhou People's Hospital, Wenzhou, China, [3] Department of Pathology, Hunan Children's Hospital, University of South China, Changsha, China, [4] Center for Diagnosis and Treatment of Rare Diseases, Hunan Children's Hospital, University of South China, Changsha, China

Y chromosome represents masculinization. The extra Y chromosome of XYY patients usually leads to over-masculinization phenotypes. The occurrence of several DSD cases with XYY in blood is controversial. Is XYY associated with disorder of sex development (DSD)? What is the mechanism behind DSD in males with XYY in blood? To this end, this study retrospectively analyzed blood-karyotype data of 4,437 DSD male children and karyotypes data of 6,259 newborn males as the control. Exome sequencing (ES) was performed to test whether the patients with DSD and with XYY in blood had other variants on known DSD-genes. Testicular biopsy was performed. Fluorescence *in situ* hybridization (FISH) was used to test whether a sex chromosome mosaicism was present in the oral epithelial cells or gonad tissue of patients with DSD and with XYY in blood. Among 4,437 DSD males who received cytogenetic evaluation, 14 patients with 47,XYY were identified. By contrast, five individuals among the 6,259 controls had 47,XYY. XYY in blood is more frequent among males with DSD than in other males (p = 0.004). The XYY karyotypes were confirmed again by GTG-banding in blood samples and by FISH performed on oral epithelial cells. ES on seven XYY DSD patients was successfully performed, but results did not identify any pathogenic variant on 55 known DSD genes. Gonad biopsy (n = 3) revealed testicular dysplasia and true hermaphroditism. FISH of gonad tissues (n = 3) showed that all of the samples had mosaic for X/XY/XYY. This study is the first to investigate the relationship between XYY in blood and DSD. The knowledge that XYY is in the blood and in oral cells have X/XY/XYY mosaicism in gonadal tissue is new for both researchers and clinicians who seek to understand the genetic basis of DSD males.

Keywords: XYY syndrome, DSD, gonad biopsy, FISH, exome sequencing, mosaicism

OPEN ACCESS

Edited by:
Shraddha Thakkar, United States Food and Drug Administration, United States

Reviewed by:
Fulya Taylan, Karolinska Institutet, Sweden
Ibtessam Ramzi Hussein, King Abdulaziz University, Saudi Arabia

****Correspondence:***
Haiyan Luo
luohaiyan0429@126.com
Weijian Chen
cwj_7@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 October 2020*
Accepted: *19 March 2021*
Published: *12 April 2021*

Citation:
Yang Y, Chen F, Luo Z, Zheng Y, Zheng J, Fu Y, Chen W and Luo H (2021) Disorder of Sexual Development Males With XYY in Blood Have Exactly X/XY/XYY Mosaicism in Gonad Tissues. Front. Genet. 12:616693. doi: 10.3389/fgene.2021.616693

INTRODUCTION

The diagnosis of chromosome diseases is highly dependent on GTG-binding of blood samples. The 47,XYY syndrome is a common sex chromosome aneuploidy that occurs in 1 out of 1,000 male births (Bardsley et al., 2013; Gao et al., 2014; Jo et al., 2016; Yao et al., 2019). Although XYY is diagnosed routinely through GTG-banding of peripheral blood, the phenotypes of XYY

TABLE 1 | Clinic information for 14 DSD males with XYY in blood, which observed from 4,437 DSD children.

ID	Age[a]	Karyotype in blood	Genital abnormalities	Other phenotypes
A56	0–5	47,XYY	Penile dysplasia and absent right testis	Intellectual disability
A57	0–5	47,XYY	Penile and testicle dysplasia	Intellectual disability and bilateral strabismus
A58	0–5	47,XYY	Penile dysplasia; penis adduction; and supra urethral dehiscence	Intellectual disability and short stature
A59	5–10	47,XYY	Penile dysplasia and penis adduction	–
A60	0–5	47,XYY	Penile dysplasia; penile flexion deformity; hypospadias; and testis dysplasia	–
A61	0–5	47,XYY	Hypospadias and scrotum division	Epilepsy
A62	5–10	47,XYY	Penile dysplasia and testis dysplasia	–
A63	5–10	47,XYY	Hypospadias and penile scrotal translocation deformity	–
A64	0–5	47,XYY	Hypospadias; penile flexion deformity; and testis dysplasia	–
A65	0–5	47,XYY	Penile dysplasia and hypospadias	Right ectopic kidney
A66	0–5	47,XYY	Hypospadias, testis dysplasia; penile dysplasia; and absent right testis	Female reproductive tissues observed
A67	0–5	47,XYY	Hypospadias; penile dysplasia; and penis adduction	–
A68	5–10	47,XYY	Penile dysplasia; penis adduction; and testis dysplasia	Intellectual Disability
A69	5–10	47,XYY	Hypospadias and penile dysplasia	–

[a] *Year range.*

patients may vary greatly, ranging from no phenotype and relatively few abnormalities to multi-systemic symptoms; for a specific symptom, the severity can vary among individuals (Kim et al., 2013; Bardsley et al., 2013).Some scholars believe that XYY is associated with a status of over-masculinization, because the existence of an extra Y chromosome and XYY individuals usually results in tall stature, impulsivity, and/or sex organ overdevelopment (macroorchidism and macropenis) (Bardsley et al., 2013; Jo et al., 2015). Several reports stated that some men institutionalized for antisocial behavior were found to have an increased frequency of the XYY karyotype, and males in prison with XYY had higher testosterone than healthy age-matched controls (Hook, 1973; Schiavi et al., 1984). For decades, the claim that XYY males tend to exhibit more aggressive, anti-social, and criminal behavior than XY males is controversial, but this hypothesis has never been substantiated (Lenroot et al., 2009).

Another more pointed argument was the opposite of over-masculinization; XYY has also been sporadically reported in connection with several cases of disorder of sex development (DSD) (Boczkowski, 1970; Grace and Campbell, 1978; Rivera et al., 1979; Terada et al., 1984; Okamoto et al., 1988; Diego Nuñez et al., 1992; Suzuki et al., 1999; Benasayag et al., 2001; Monastirli et al., 2005; Bardsley et al., 2013; Latrech et al., 2015). Such DSD phenotypes of XYY patients include micropenis, testicular dysplasia, true-hermaphrodite, and complete sex reversal (**Supplementary Table 1**) (Boczkowski, 1970; Grace and Campbell, 1978; Rivera et al., 1979; Terada et al., 1984; Okamoto et al., 1988; Diego Nuñez et al., 1992; Suzuki et al., 1999; Benasayag et al., 2001; Monastirli et al., 2005; Bardsley et al., 2013; Latrech et al., 2015). However, due to the fact that all previously reported DSD XYY patients are sporadic cases, and the DSD frequency is as high as 1/200 in a general male population, it is still unknown whether DSD is associated with XYY or just the coincidence of XYY and DSD. Given that more than half of patients with DSD can be traced to a pathogenic variant on one of 55 known genes related to DSD (Eggers et al., 2016; Wang et al., 2017), it is also possible that the DSD phenotypes are the consequence of an extra Y chromosome and a pathogenic variant on one of the known DSD genes.

MATERIALS AND METHODS

Ethics Statement

The study protocol was approved by the Academic Committee of Hunan Children's Hospital (Approval number: HCHLL58, Changsha City, Hunan Province, China). All participants or their parents provided written informed consent to partake in this study.

Patient Clinical and Karyotype Data

Identifying and diagnostic information and karyotype data were obtained from the records of all males diagnosed with DSD from Hunan Children's Hospital (Changsha City, China) from July 2010 to June 2018. DSD was evaluated and diagnosed by pediatric urologists. Karyotype data of neonatal umbilical cord blood from Wenzhou People's Hospital (Wenzhou City, China) were used as the control karyotype data. These data included a total of 10,086 karyotype data of the general population, of which 6,259 are males, obtained from May 2012 to April 2018.

Cytogenetic Analysis

Peripheral venous blood was collected in a vacutainer sodium heparin vial. Slides were prepared from phytohemagglutinin-stimulated peripheral lymphocyte cultures by using standard cytogenetic methods. Giemsa (GTG) banding at a 400-band level to a 550-band level was performed in accordance with the standard laboratory protocol. Two different cultures, corresponding to two different series of slides from each sample, were separately prepared and analyzed. At least 40 metaphases were analyzed for each individual. For the second round of GTG-banding evaluation, 100 metaphases were analyzed per patient.

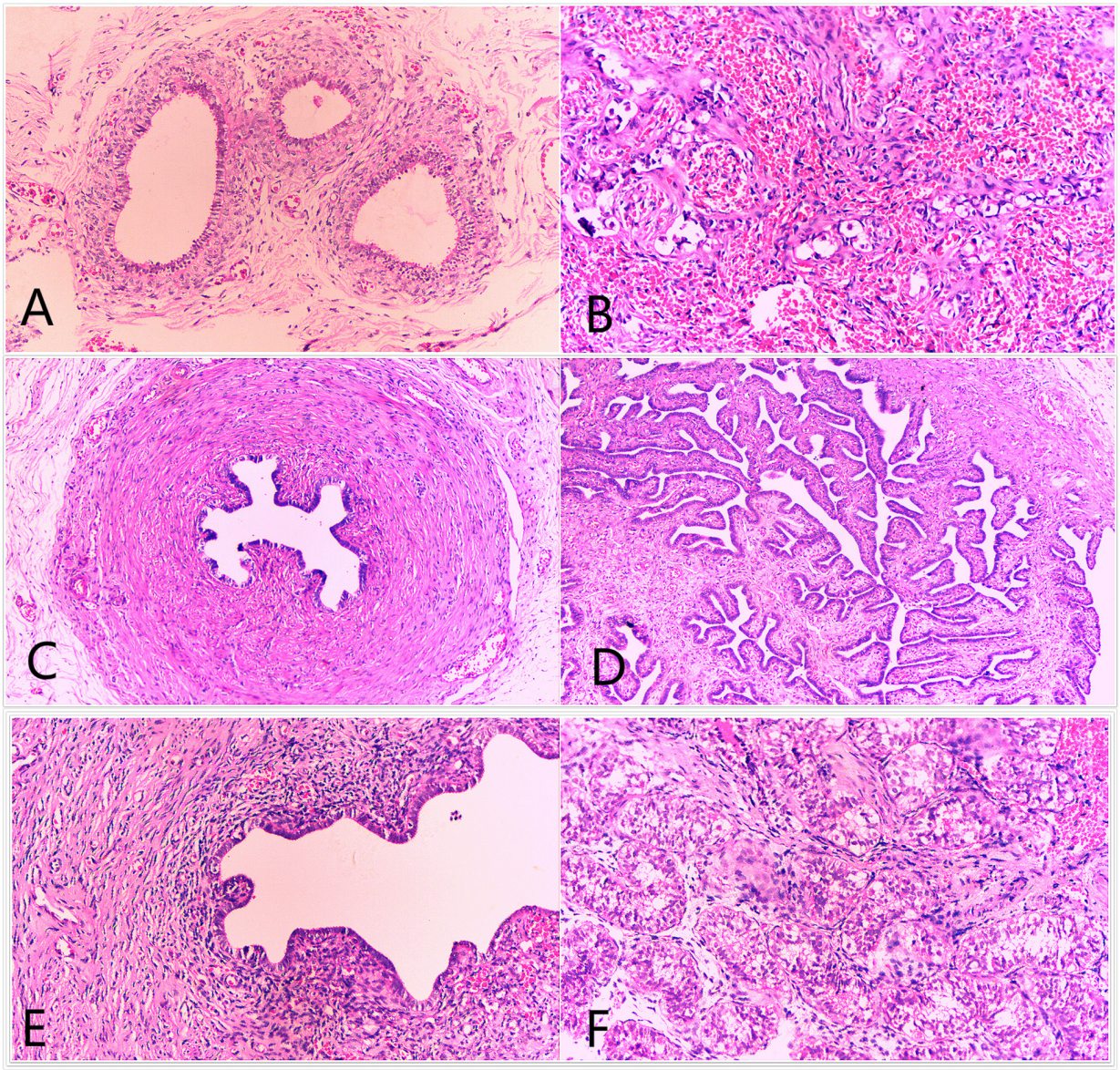

FIGURE 1 | H&E staining of the biopsy gonad tissues from a severest DSD patient with XYY in blood (A66). For A66, surgical exploration of scrotum and groin revealed **(A)** the structure of epididymis, **(B)** the structure of ovary (primodial follicles are visible), **(C)** the structure of vas deferens, **(D)** the structure of fallopian tubes, and **(E)** structure of uterus. For A66, right testis biopsy revealed the dysplasia of spermatogenic tubules **(F)**.

Exome Sequencing

Genomic DNA (200 ng) of each individual was sheared by Biorupter (Diagenode, Belgium) to acquire 150–200 bp fragments. The ends of DNA fragment were repaired and Illumina Adaptor was added (Fast Library Prep Kit, iGeneTech, Beijing, China). After constructing a sequencing library, the whole exons were captured with AIExome Enrichment Kit V1 (iGeneTech, Beijing, China) and sequenced on Illumina platform (Illumina, San Diego, CA, United States) with 150 base-paired end reads. Raw reads were filtered to remove low quality reads by using FastQC. Then, clean reads were mapped to the reference genome GRCh37 by using Bwa. After removing duplications, SNV and InDel were called and annotated by using GATK. 11.9 G bases were obtained for each sample. The average yield was ∼16.6 Gb with an error rate of $<0.1\%$. Furthermore, $>80\%$ bases had a Phred quality score of ≥ 30 (Q30).

Gonad Biopsy, Hematoxylin and Eosin Stain (H&E), and Fluorescence *in situ* Hybridization

Gonadal biopsy was performed for three DSD children with XYY. H&E was performed routinely for all available tissue specimens according to standard procedures. Fluorescence

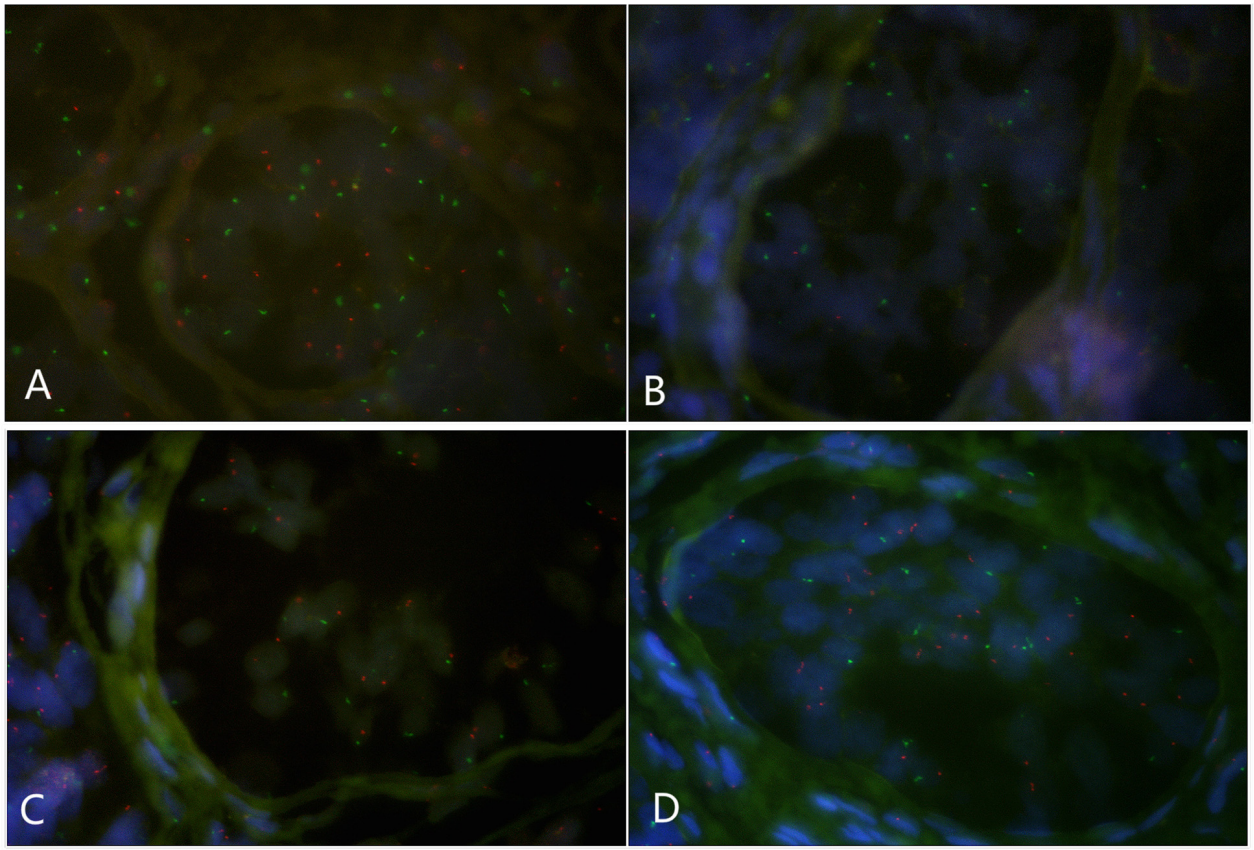

FIGURE 2 | Fluorescence *in situ* hybridization (FISH) on spermatogenic tubules (green spot: FISH centromeric signals of Chromosome X; red spot: signals of Chromosome Y). **(A)** A spermatogenic tubule from a normal testicle (NC1553), 23 green and 22 red spots in the tubule implicates XY; BCD: Spermatogenic tubules from A58 testes. In panel **(B)**, 17 green and 2 red signals detected implicates X/XY mosaicism. In panel **(C)**, 17 green and 19 red signals detected implicates XY. In panel **(D)**, 17 green and 34 red signals detected implicates XYY.

in situ hybridization (FISH) analyses were performed on oral epithelial cells and on gonad tissues. The centromere satellite probes Xp11.1-q11.1 Alpha Satellite DNA Vysis CEP X (DXZ1) Spectrum Green Probe and Yp11.1-q11.1 Alpha Satellite DNA Vysis CEP Y (DYZ3) Spectrum Orange Probe were purchased from Abbott-Vysis company (Abbott Park, IL, United States).

RESULTS

Disorder of Sex Development Males Have an High Frequency of XYY

From July 2010 to June 2018, 4,437 DSD males received genetic and clinical evaluation in Hunan Children's Hospital. Of these 4,437 DSD males, 3,885 were simplex with one phenotype or deformity and 552 were complex phenotypes with more than one phenotype. Among the 3,885 simplex DSD male children, 2,717 had hypospadias (accounting for 61.24% of overall DSD), 897 had penis dysplasia (20.22%), 107 had cryptorchidism (2.41%), and 93 had hypoplastic testis (2.10%), whereas 71 had other conditions (1.60%). The 552 complex DSD children presented the following: (i) two or more DSD phenotypes (n = 459) (10.34%); and (ii) DSD phenotype and other malformations, such as developmental delay, skeletal deformities, and others (n = 93) (2.10%).

Among 4,437 males with DSD, 136 (3.07%) were affected by cytogenetic abnormalities (**Supplementary Table 2**), including 47,XXY (n = 28), 45,X/46,XY (n = 21), 47,XYY (n = 14), 46,XX (n = 12), 46,XY/46,XX (n = 8), 46,XY,del(9p22) (n = 6), different chromosome translocation (n = 17), and others (n = 30). Therefore, XYY was one of major chromosome abnormalities in DSD males.

The clinical data for 14 male children with DSD and XYY are provided in **Table 1**. Blood endocrine test was performed for eight of the 14 patients with DSD and XYY (**Supplementary Table 3**), among which two patients exhibited pituitary prolactin above the normal (27.47 and 19.86; the normal range for children was 1–19 pg/ml).

The karyotype data of neonatal umbilical cord blood were used as the control for this study. In 6,259 male controls, a total of 51 individuals had abnormal karyotype (**Supplementary Table 4**), and among them, five individuals had XYY (phenotypes were unavailable). This result was consistent with the reported finding that the XYY frequency in the general male population was 1/1,000 (Bardsley et al., 2013; Jo et al., 2016).

TABLE 2 | Fluorescence *in situ* hybridization (FISH) signal counts per spermatogenic tubules for one normal control (NC1553) and three DSD children with XYY in blood (A58, A66, and A60).

NC 1553	Green	Red	Green/ red	Predicted karyotype	A58	Green	Red	Green/ red	Predicted karyotype	A66	Green	Red	Green/ red	Predicted karyotype	A60	Green	Red	Green/red	Predicted karyotype
Tub1	16	15	1.07	XY	Tub1	22	44	0.50	XYY	Tub1	22	4	5.5	X/XY	Tub1	13	0	–	X
Tub2	31	35	0.89	XY	Tub2	10	21	0.48	XYY	Tub2	27	10	2.7	X/XY	Tub2	25	4	6.25	X/XY
Tub3	29	28	1.04	XY	Tub3	14	28	0.50	XYY	Tub3	33	11	3	X/XY	Tub3	27	8	3.38	X/XY
Tub4	14	17	0.82	XY	Tub4	18	6	3	X/XY	Tub4	15	1	15	X/XY	Tub4	31	7	4.43	X/XY
Tub5	18	19	0.95	XY	Tub5	10	11	0.91	XY	Tub5	19	2	9.5	X/XY	Tub5	27	5	5.40	X/XY
Tub6	16	14	1.14	XY	Tub6	41	92	0.45	XYY	Tub6	37	19	1.95	X/XY	Tub6	17	3	5.67	X/XY
Tub7	25	25	1.00	XY	Tub7	27	50	0.54	XYY	Tub7	40	18	2.22	X/XY	Tub7	22	0	–	X
Tub8	14	14	1.00	XY	Tub8	26	48	0.54	XYY	Tub8	9	9	1.00	XY	Tub8	22	2	11.00	X/XY
Tub9	11	11	1.00	XY	Tub9	43	44	0.98	XY	Tub9	39	11	3.55	X/XY	Tub9	19	7	2.71	X/XY
Tub10	12	13	0.92	XY	Tub10	39	74	0.53	XYY	Tub10	19	2	9.50	X/XY	Tub10	33	2	16.50	X/XY
Tub11	12	11	1.09	XY	Tub11	20	20	1.00	XY	Tub11	16	3	5.33	X/XY	Tub11	27	9	3.00	X/XY
Tub12	17	16	1.06	XY	Tub12	15	16	0.94	XY	Tub12	29	0	–	X	Tub12	35	6	5.83	X/XY
Tub13	13	13	1.00	XY	Tub13	17	2	8.5	X/XY	Tub13	17	0	–	X	Tub13	16	4	4.00	X/XY
Tub14	19	22	0.86	XY	Tub14	28	51	0.55	XYY	Tub14	29	4	7.25	X/XY	Tub14	15	3	5.00	X/XY
Tub15	28	31	0.90	XY	Tub15	22	4	5.5	X/XY	Tub15	22	3	7.33	X/XY	Tub15	24	4	6.00	X/XY

Tub = spermatogenic tubule, Green = green spot signal of X centromere; Red = red spot signal of Y centromere.

Between-group comparison showed significant difference in the frequency of XYY in controls (5/6,259, 0.08%) and in the DSD group (14/4437, 0.32%) ($p = 0.004$). If only complex DSD was considered, then the frequency of XYY in complex DSD (14/552, 2.54%) was even higher than that in the control group (5/6259, 0.08%) ($p < 0.001$).

Repeating GTG Banding Assay on Blood and FISH on Oral Cells Did Not Support Sex Chromosome Aneuploid Mosaicism for Patients With DSD

To test if sex chromosome aneuploid mosaicism was present in the 14 patients with DSD (the first round of GTG banding reported as 47,XYY; 40 metaphase cells were analyzed per patient), we intended to re-visit all of them to replicate the GTG banding. However, only 10 patients were followed up, and peripheral venous blood samples were obtained from seven individuals (listed in **Supplementary Table 5**). The GTG banding of blood was performed again for these seven patients. For each one, 100 metaphase cells were analyzed. None of these seven patients with DSD presented with sex chromosome aneuploid mosaicism. All have true 47,XYY in blood.

We further collected oral epithelial cells from these seven patients with XYY and DSD and carried out FISH tests on them by using the centromere satellite probes of the Y and X chromosomes. The oral cells from all seven patients (oral epithelial cells) have two signals of the Y chromosome and one signal of the X chromosome (**Supplementary Figure 1**).

None of Seven DSD-XYY Patients Harbor Pathogenic Variant on 55 Known DSD Genes

DSD has 1/200 chance of occurrence among general males, and more than half of DSD males have pathogenic variants on known DSD genes (Eggers et al., 2016). The XYY frequency in a male population is about 1/1,000 (**Supplementary Table 4**) (Bardsley et al., 2013; Jo et al., 2016). It is possible for a patient with DSD to have XYY aneuploidy and a DSD-gene mutation. To test this hypothesis, exome sequencing (ES) was performed for seven patients with DSD and with XYY in blood (for those whose genomic DNA data were available). ES was performed, and an average of 10.51 Gb data were obtained within the target region of each sample. The quality statistics for ES are shown in **Supplementary Table 5**.

We focused on 55 known causative genes for male DSD (**Supplementary Table 6**) (Eggers et al., 2016). We used the following filtering steps: (i) rare variants (MAF < 0.005, gnomAD_Eas); (ii) variants absent in in-house controls (201 ES data of males without DSD; parts are published before (Yang et al., 2019; Zhu et al., 2019); and (iii) considering damaging variants (loss-of-function and damaging missense variants) (Yang et al., 2019). Only one variant remained, namely, chr17:77753158, C to T (CBX2,NM_032647:exon3:c.117-3C > T) on A61 (**Supplementary Figure 2A**). Results of literature search showed the existence of a girl with 46,XY who harbored compound heterozygous variants on CBX2 (Biason-Lauber et al., 2009).

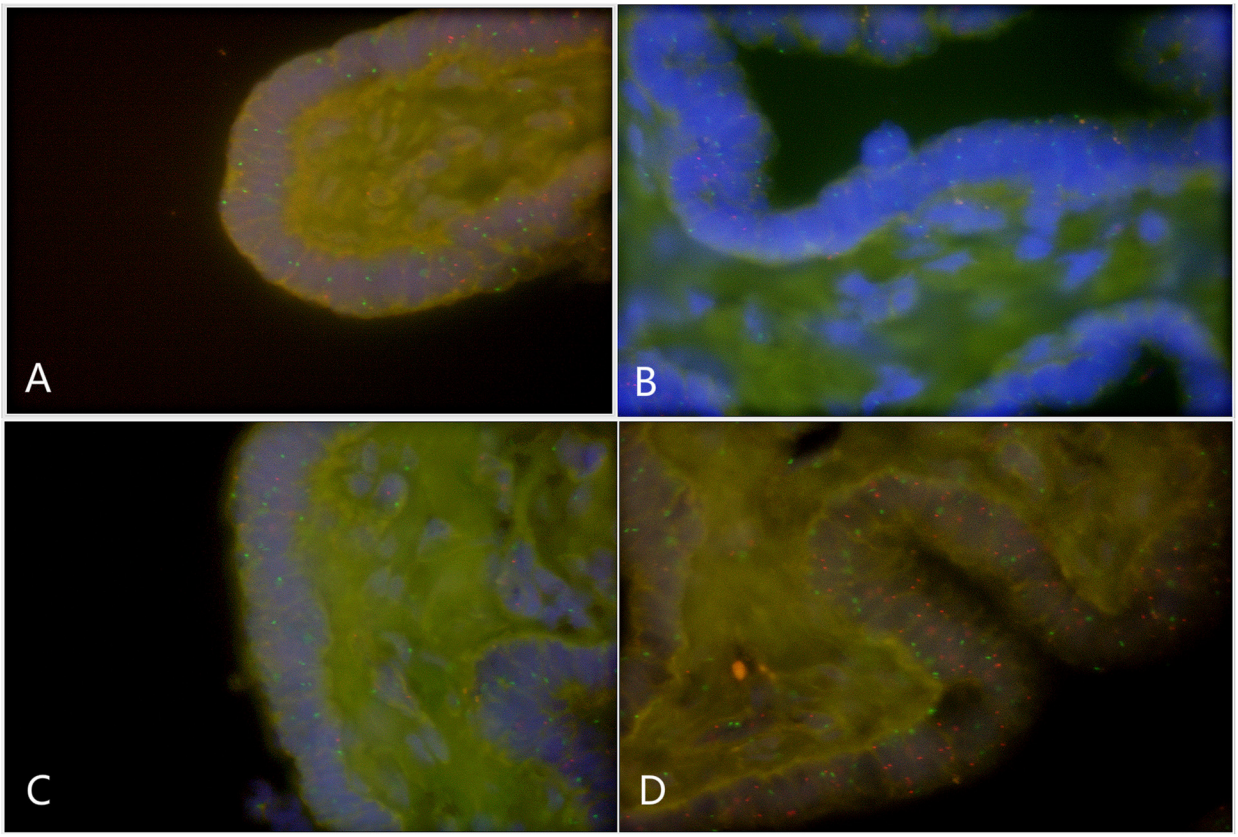

FIGURE 3 | Fluorescence *in situ* hybridization (FISH) on fallopian tubes tissue that originated from a DSD boy with XYY in blood (A66) (green spot: FISH centromeric signals of Chromosome X; red spot: signals of Chromosome Y). **(A)** X/XY signals; **(B)** X/XY signals; **(C)** X/XY signals; **(D)** XYY signals.

Read map of ES data for CBX2 was re-checked, and the whole CBX2 coding regions were covered (>30×) by ES for A61, but no other variant was identifiable (data not shown). We intended to test the pathogenicity for the CBX2 variant through segregation analysis of the family. Results showed that the CBX2-c.117-3C > T variant was shared by his unaffected grandfather (**Supplementary Figure 2**). Therefore, the significance of the variant CBX2-c.117-3C > T is unknown. Altogether, no direct evidence supported the assumption that XYY patients harbored a DSD gene pathogenic variant, which co-contributed to DSD.

Gonad Biopsy and H&E Staining

The pathogenic analysis of gonad tissue may give clues to elucidate the mechanism of the development of DSD in XYY males. Most patients with DSD and with XYY in blood were subjected to urological surgery for deformity correction. Three of them (or their guardians) agreed to undergo gonad biopsy. H&E staining of the biopsy tissue showed that 3/3 DSD-XYY patients exhibited dysplasia of spermatogenic tubules of testicles (two bilateral; one unilateral) (**Figure 1F** and **Supplementary Figure 3**). In one of the three (1/3) DSD-XYY patients (A66), only the right testicle was observed in the process of operation (data not shown).

Through laparoscopic, ureteroscopic, and cystoscopic examinations, fragmentary gonad tissues were found and excised. H&E staining of the biopsied tissues showed the structures of bisexual reproductive organ-tissues, including epididymis, ovary, vas deferens, fallopian tubes, and uterus (**Figures 1A–E**).

Fluorescence *in situ* Hybridization on Gonad Tissues Revealed X/XY/XYY Mosaicism

Paraffin embedded sections showed incomplete nucleus of the tissue cells. For the evaluation of chromosome aneuploidy, analyzing centromeric signal of FISH by using single cells as a unit on slides would be unreasonable. We then intended to evaluate the centromeric signal by using spermatogenic tubules as a unit. Indeed, in a normal testicle donated by a boy that died due to a traffic accident (NC1553), we found that the X or Y signal per tubule are basically equal (**Figure 2A**, **Table 2**, and **Supplementary Figure 4**). FISH results on spermatogenic tubules of three patients with DSD (A58,A66 and A60; their karyotypes are 47,XYY in blood) indicated that all of their testicular tissues have X/XY/XYY mosaicism, although the mosaic percentage varied (**Figures 2A–D**, **Table 2**, and **Supplementary Figure 4**).

For patient A66, FISH was performed for the fallopian tube tissue, and the results also indicated X/XY/XYY mosaicism (**Figure 3**).

DISCUSSION

Since 1963, at least 23 Patients with DSD with the karyotype of XYY in blood have been reported (**Supplementary Table 1**) (Boczkowski, 1970; Grace and Campbell, 1978; Rivera et al., 1979; Terada et al., 1984; Okamoto et al., 1988; Diego Nuñez et al., 1992; Suzuki et al., 1999; Benasayag et al., 2001; Monastirli et al., 2005; Bardsley et al., 2013; Latrech et al., 2015). Phenotypes for these patients with DSD are variable, e.g., female-external-genitalia, bisexual appearances, or hypospadias (**Supplementary Table 1**) (Boczkowski, 1970; Grace and Campbell, 1978; Rivera et al., 1979; Terada et al., 1984; Okamoto et al., 1988; Diego Nuñez et al., 1992; Suzuki et al., 1999; Benasayag et al., 2001; Monastirli et al., 2005; Bardsley et al., 2013; Latrech et al., 2015). Combining the well-characterized knowledge of XYY leading to over-masculinization (macroorchidism, macropenis, and tall stature) and information on DSD in XYY patients, scientists and clinical practitioners tend to assume that XYY leads to bipolarized sex determination.

However, because all DSD XYY patients in the literature are sporadically reported cases, and DSD frequency in general males is as high as 1/200 (the coincidence of XYY and a DSD? Latrech et al. (2015) 14), a clear evidence of XYY's association with DSD is lacking. To address this issue, we compared the XYY frequencies of 4,437 patients with DSD and 6,259 newborn-general males. Statistical results indicated a solid association between XYY in blood and male DSD ($p = 0.004$).

Given that the condition of more than 50% of patients with DSD can be explained by pathogenic variant on one of the 55 known DSD genes (Eggers et al., 2016), another possibility exists, i.e., the DSD phenotypes in XYY patients are due to the coincidence of XYY and a DSD gene mutation. To test this coincidence, we analyzed the coding regions of 55 known male DSD genes in seven patients with DSD and with XYY in blood by next generation sequencing. However, we did not identify any definite pathogenic variants on these 55 genes in patients with DSD with XYY in blood.

To further explore the mechanism underlying DSD in males with XYY in blood, gonad biopsy was performed on three patients. H&E staining of gonad tissues revealed the presence of dysplasia in spermatogenic tubules in all three patients and bisexual reproductive organ in one of the three patients. The FISH test results revealed the occurrence of X/XY/XYY mosaic in gonad tissues for all three patients with DSD and with XYY in blood.

One limitation of this study is that the genetic analysis (including but not limited to the gonad FISH test) did not involve adult infertile men with XYY in blood. Previously studies have shown that the majority of individuals with XYY in blood are fertile because of the loss of the extra Y before meiosis (Rives et al., 2005), and the remaining patients with XYY in blood are infertile (Suzuki et al., 1999; Borjian Boroujeni et al., 2017). However, the mechanism that explains why a number of XYY men are infertile is not fully understood. This limitation can be compensated by a recent study (Sciurano et al., 2019) carried out by Sciurano et al., who described a 35-year-old man with primary infertility. The man had hypotrophic testes (10 and 12 ml testicular volume; the normal volume is more than 15 ml) and inguinal hernia during childhood; Yq12 FISH detected X/XY/XYY mosaicism (50% XYY; 44%XY; 6%X, respectively) in Sertoli cells (Sciurano et al., 2019).

In conclusion, we identified X/XY/XYY mosaicism in gonad tissues and detected the lack of sex chromosome mosaicism on oral epithelial cells on patients with DSD and with XYY in blood. These findings explain the DSD phenotypes in patients with XYY in blood and also indicate the uncharacterized proliferation-mode of gonad cells. Further study is needed to explore such novel cell proliferative mode.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The name of the repository and accession number can be found below: National Center for Biotechnology Information (NCBI) ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/, VCV000996035.1

AUTHOR CONTRIBUTIONS

HL, YY, and WC designed the research, analyzed the experimental data, and wrote the manuscript. YY, FC, ZL, YZ, JZ, YF, WC, and HL performed the sample collection and the research. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (31501017 to YY) and the Hunan Health Commission Research Fund (B2019019 to YY).

ACKNOWLEDGMENTS

The authors are grateful to the patients for agreeing to undergo repeated examinations and sampling.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.616693/full#supplementary-material

REFERENCES

Bardsley, M. Z., Kowal, K., Levy, C., Gosek, A., Ayari, N., Tartaglia, N., et al. (2013). 47, XYY syndrome: clinical phenotype and timing of ascertainment. *J. Pediatr.* 163, 1085–1094. doi: 10.1016/j.jpeds.2013.05.037

Benasayag, S., Rittler, M., Nieto, F., Torres de Aguirre, N., Reyes, M., and Copelli, S. (2001). 47,XYY karyotype and normal SRY in a patient with a female phenotype. *J. Pediatr. Endocrinol. Metab.* 14, 797–801.

Biason-Lauber, A., Konrad, D., Meyer, M., De Beaufort, C., and Schoenle, E. J. (2009). Ovaries and female phenotype in a girl with 46,XY karyotype and mutations in the CBX2 gene. *Am. J. Hum. Genet.* 84, 658–663. doi: 10.1016/j.ajhg.2009.03.016

Boczkowski, K. (1970). XYY karyotype in eunuchoidal phenotypic female. *J. Clin. Endocrinol. Metab.* 30, 111–113. doi: 10.1210/jcem-30-1-111

Borjian Boroujeni, P., Sabbaghian, M., Vosough Dizaji, A., Zarei, M. S., Almadani, N., Mohammadpour, L. F., et al. (2017). Clinical aspects of infertile 47,XYY patients: a retrospective study. *Hum. Fertil.* 22, 88–93. doi: 10.1080/14647273.2017.1353143

Diego Nuñez, M. A., Prieto Veiga, J., Rey Sánchez, F., Salazar Veloz, J. F., De Manueles Jiménez, J., Santos Borbujo, J., et al. (1992). Clinical polymorphism of the XYY syndrome. *Ann. Esp. Pediatr.* 37, 140–144.

Eggers, S., Sadedin, S., van den Bergen, J. A., Robevska, G., Ohnesorg, T., Hewitt, J., et al. (2016). Disorders of sex development:insights from targeted gene sequencing of a large international patient cohort. *Genome Biol.* 17:243.

Gao, Y., Xie, B., and Liu, R. (2014). Delivering noninvasive prenatal testing in a clinical setting using semiconductor sequencing platform. *Sci. China Life Sci.* 57, 737–738. doi: 10.1007/s11427-014-4696-0

Grace, H. J., and Campbell, G. D. (1978). XYY karyotype, female phenotype and gonadal dysgenesis. A case report. *S. Afr. Med. J.* 54, 284–286.

Hook, E. B. (1973). Behavioral implications of the human XYY genotype. *Science* 179, 139–150. doi: 10.1126/science.179.4069.139

Jo, H. C., Lee, S. W., Jung, H. J., and Park, J. E. (2016). Esthesioneuroblastoma in a boy with 47,XYY karyotype. *Korean J. Pediatr.* 59(Suppl. 1), S92–S95.

Jo, W. H., Jung, M. K., Kim, K. E., Chae, H. W., Kim, D. H., Kwon, A. R., et al. (2015). XYY syndrome:a 13-year-old boy with tall stature. *Ann. Pediatr. Endocrinol. Metab.* 20, 170–173. doi: 10.6065/apem.2015.20.3.170

Kim, W., Khadilkar, A. C., Ko, E. Y., and Sabanegh, E. S. Jr. (2013). 47,XYY Syndrome and male infertility. *Rev. Urol.* 15, 188–196.

Latrech, H., Skikar, I., Gharbi Mel, H., Chraïbi, A., and Gaouzi, A. (2015). Disorder of sexual development and congenital heart defect in 47XYY: clinical disorder or coincidence? *Case Rep. Endocrinol.* 2015:802162.

Lenroot, R. K., Lee, N. R., and Giedd, J. N. (2009). Effects of sex chromosome aneuploidies on brain development: evidence from neuroimaging studies. *Dev. Disabil. Res. Rev.* 15, 318–327. doi: 10.1002/ddrr.86

Monastirli, A., Stephanou, G., Georgiou, S., Andrianopoulos, C., Pasmatzi, E., Chroni, E., et al. (2005). Short stature, type E brachydactyly, exostoses, gynecomastia, and cryptorchidism in a patient with 47,XYY/45,X/46,XY mosaicism. *Am. J. Med. Sci.* 329, 208–210. doi: 10.1097/00000441-200504000-00008

Okamoto, E., Yabumoto, H., Terakawa, T., Shima, H., Ikoma, F., Sakamoto, H., et al. (1988). XYY syndrome: report of three cases. *Hinyokika Kiyo* 34, 191–195.

Rivera, H., Hernandez, A., Martinez, R., Plascencia, L., Cuevas, A., and Cantu, J. M. (1979). Malformed genitalia in the 47,XYY genotype. *Ann. Genet.* 22, 225–227.

Rives, N., Milazzo, J. P., Miraux, L., North, M. O., Sibert, L., and Macé, B. (2005). From spermatocytes to spermatozoa in an infertile XYY male. *Int. J. Androl.* 28, 304–310. doi: 10.1111/j.1365-2605.2005.00540.x

Schiavi, R. C., Theilgaard, A., Owen, D. R., and White, D. (1984). Sex chromosome anomalies, hormones, and aggressivity. *Arch. Gen. Psychiatry* 41, 93–99. doi: 10.1001/archpsyc.1984.01790120097012

Sciurano, R. B., Rahn, I. M., González, A. B., Rey, V. G., Benavente, R., and Solari, A. J. (2019). Selective advantage of euploid spermatocytes I in an azoospermic 47,XYY man with gonadal mosaicism. *Hum. Reprod.* 34, 568–573. doi: 10.1093/humrep/dey387

Suzuki, Y., Sasagawa, I., Kaneko, T., Tateno, T., Iijima, Y., and Nakada, T. (1999). Bilateral cryptorchidism associated with 47,XYY karyotype. *Int. Urol. Nephrol.* 31, 709–713.

Terada, T., Yanagi, S., Nakada, T., and Katayama, T. (1984). A case of XYY syndrome with male infertility and retentio testis. *Hinyokika Kiyo* 30, 701–707.

Wang, Y., Gong, C., Wang, X., and Qin, M. (2017). AR mutations in 28 patients with androgen insensitivity syndrome (Prader grade 0-3). *Sci. China Life Sci.* 60, 700–706. doi: 10.1007/s11427-017-9084-9

Yang, Y., Zheng, Y., Li, W., Li, L., Tu, M., Zhao, L., et al. (2019). SMAD6 is frequently mutated in nonsyndromic radioulnar synostosis. *Genet. Med.* 21, 2577–2585. doi: 10.1038/s41436-019-0552-8

Yao, H., Gao, Y., Zhao, J., Xu, H., Luo, Y., Yuan, Y., et al. (2019). Genome-wide detection of additional fetal chromosomal abnormalities by cell-free DNA testing of 15,626 consecutive pregnant women. *Sci. China Life Sci.* 62, 215–224. doi: 10.1007/s11427-017-9344-7

Zhu, G., Zheng, Y., Liu, Y., Yan, A., Hu, Z., Yang, Y., et al. (2019). Identification and characterization of NF1 and non-NF1 congenital pseudarthrosis of the tibia based on germline NF1 variants: genetic and clinical analysis of 75 patients. *Orphanet. J. Rare Dis.* 14:221. doi: 10.1034/j.1399-0004.2000.570308.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.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 23 April 2021
doi: 10.3389/fphar.2021.608778

Unraveling Gene Fusions for Drug Repositioning in High-Risk Neuroblastoma

Zhichao Liu[1], Xi Chen[1], Ruth Roberts[2,3], Ruili Huang[4], Mike Mikailov[5] and Weida Tong[1]**

[1]Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR, United States, [2]ApconiX, BioHub at Alderley Park, Alderley Edge, United Kingdom, [3]University of Birmingham, Edgbaston, Birmingham, United Kingdom, [4]National Center for Advancing Translational Sciences, National Institutes of Health, Rockville, MD, United States, [5]Office of Science and Engineering Labs, Center for Devices and Radiological Health, US Food and Drug Administration, Silver Spring, MD, United States

OPEN ACCESS

Edited by:
Meng Yao Lu,
National Taiwan University Hospital, Taiwan

Reviewed by:
Henrik Green,
Linköping University, Sweden
Yen-Yi Liu,
China Medical University, Taiwan
Yen-Lin Liu,
Taipei Medical University, Taiwan

***Correspondence:**
Zhichao Liu
zhichao.liu@fda.hhs.gov
Weida Tong
Weida.tong@fda.hhs.gov

Specialty section:
This article was submitted to Translational Pharmacology, a section of the journal Frontiers in Pharmacology

Received: 21 September 2020
Accepted: 23 March 2021
Published: 23 April 2021

Citation:
Liu Z, Chen X, Roberts R, Huang R, Mikailov M and Tong W (2021) Unraveling Gene Fusions for Drug Repositioning in High-Risk Neuroblastoma. Front. Pharmacol. 12:608778. doi: 10.3389/fphar.2021.608778

High-risk neuroblastoma (NB) remains a significant therapeutic challenge facing current pediatric oncology patients. Structural variants such as gene fusions have shown an initial promise in enhancing mechanistic understanding of NB and improving survival rates. In this study, we performed a comprehensive *in silico* investigation on the translational ability of gene fusions for patient stratification and treatment development for high-risk NB patients. Specifically, three state-of-the-art gene fusion detection algorithms, including ChimeraScan, SOAPfuse, and TopHat-Fusion, were employed to identify the fusion transcripts in a RNA-seq data set of 498 neuroblastoma patients. Then, the 176 high-risk patients were further stratified into four different subgroups based on gene fusion profiles. Furthermore, Kaplan-Meier survival analysis was performed, and differentially expressed genes (DEGs) for the redefined high-risk group were extracted and functionally analyzed. Finally, repositioning candidates were enriched in each patient subgroup with drug transcriptomic profiles from the LINCS L1000 Connectivity Map. We found the number of identified gene fusions was increased from clinical the low-risk stage to the high-risk stage. Although the technical concordance of fusion detection algorithms was suboptimal, they have a similar biological relevance concerning perturbed pathways and regulated DEGs. The gene fusion profiles could be utilized to redefine high-risk patient subgroups with significant onset age of NB, which yielded the improved survival curves (Log-rank p value ≤ 0.05). Out of 48 enriched repositioning candidates, 45 (93.8%) have antitumor potency, and 24 (50%) were confirmed with either on-going clinical trials or literature reports. The gene fusion profiles have a discrimination power for redefining patient subgroups in high-risk NB and facilitate precision medicine-based drug repositioning implementation.

Keywords: neuroblastoma, structural variants, gene fusions, next-generation sequencing, drug repositioning, precision medicine 3

INTRODUCTION

Neuroblastoma (NB) is the most common and deadly pediatric malignancy, and the average age of patients is about 1–2˚years at diagnosis (Matthay et al., 2016; Fletcher et al., 2018). Approximately 70% of NB patients have a metastatic disease with a less than 30% event-free survival rate (Moroz et al., 2011). Several therapy and treatment options, such as immunotherapeutic strategies, local irradiation, autologous stem cell transplantation (ASCT) combined with chemotherapy, have improved the survival rate of NB patients. However, a substantial number of NB patients, particularly in the high-risk group, still suffer profound treatment-related morbidity (Semeraro et al., 2015). Therefore, advanced treatment options are still urgently needed to improve the survival rate while eliminating adverse events.

The molecular understanding of high-risk NB is mostly elusive. It becomes significant hurdles to advance NB prognosis and therapy development (Almstedt et al., 2020). Beyond the frequently detected gene alternations such as *ALK* activations (Mossé et al., 2008), *MYCN* amplification (Zeid et al., 2018), and LMO1 expression (Wang et al., 2011), advancement in sequencing technologies provided a more in-depth and width view of the molecular basis of NB (Pugh et al., 2013; Bueno et al., 2016; Boeva et al., 2017; Ma et al., 2018). More and more complex genetic events such as gene fusions have been identified and could differential high-risk NB patients (Rorie et al., 2004; Mertens et al., 2015; Peifer et al., 2015). For example, (Peifer et al., 2015) discovered recurrent genomic rearrangements of the telomerase reverse transcriptase gene (*TERT*), occurring only in high-risk NB patients. Furthermore, *TERT*-related fusions could be used to define a new patient subgroup in high-risk NB with adverse clinical outcomes. These promising findings trigger deeper thinking on how to translate these genetic findings into therapy development.

Gene fusion is a result of structural variants (SVs), including insertion, deletion, inversion, and translocation that joins the two separate transcripts. Commonality and diversity of gene fusions in cancer were discussed elsewhere (Lee et al., 2017; Xu et al., 2018; Gao et al., 2018; Hu et al., 2018; Picco et al., 2019). A lot of computational tools have been developed to detect fusion transcripts based on DNA/RNA sequencing (Kim and Salzberg, 2011; Jia et al., 2013; Beccuti et al., 2014; Davidson et al., 2015; Heyer et al., 2019). Some comparative studies have also been conducted to prioritize the fusion detection tools based on statistical measures such as precision and F-scores (Wang et al., 2013; Liu et al., 2016; Kumar et al., 2016; Zhang et al., 2016). The conclusion drawn from these comparative studies mainly suggested combined top performance callers generate consensus results for further experimental verification (Liu et al., 2016; Kumar et al., 2016). Fusion detection algorithms with different mathematical equations and hypothesis behind, it is still an open question on how to anchor and apply these algorithms based on biological relevance. Furthermore, two tumors rarely shared the same gene fusions due to tumor heterogeneity, which limited the discrimination power for stratification of cancer into informative subtypes based on the individual fused transcript. The question has been raised on whether the patient gene fusion profiles could be used to regroup the patients, which are predictive of clinical outcomes such as patient survival, therapy response, and tumor pathology.

Moreover, oncogenic gene fusion not only expands our understanding of tumor biology but provides possible therapeutic targets for treatment development. For example, Imatinib and ponatinib were approved US FDA to treat chronic myeloid leukemia (CML), which targets BCR-ABL1 fusion (Druker, 2008). Moreover, crizotinib/ceritinib inhibiting *ALK* fusion was approved to treat non–small cell lung cancer (Mertens et al., 2015). However, these approved drugs are mainly tyrosine kinase inhibitor (TKI) inhibitors, which suffers some severe adverse drug reactions (ADRs) such as cardiotoxicity and drug-induced liver injury (Liu et al., 2017; Xu et al., 2018). It is interesting to explore the probability of utilizing transcriptomic response, interplayed with gene fusions for a distinct patient subgroup for alternative treatment development while minimizing toxicity.

To explore the potential solutions for these unsolved questions, we conducted a comprehensive genomic analysis of 498 human NB cases (**Figure 1**). First, A landscape of gene fusions was presented by using three state-of-art fusion calling algorithms, including ChimeraScan (Iyer et al., 2011), SOAPfuse (Jia et al., 2013), and TopHat-Fusion (Kim and Salzberg, 2011). Then, high-risk NB patients were regrouped based on detected gene fusion profiles and evaluated by survival analysis. Next, differentially expressed genes (DEGs) of the redefined high-risk NB patient subgroups were extracted and functionally analyzed. Finally, Repositioning candidates for the redefined high-risk subgroups were enriched with a large-scale of transcriptomic profiles in LINCS L1000 Connectivity Map (CMap). The proposed framework provides a promising approach to translate novel genetic findings into therapy development.

MATERIALS AND METHODS

RNA-Seq Dataset of Neuroblastoma

The tumor samples of 498NB patients were enrolled from seven countries under the consent of respective clinical trials. The range of patients at diagnosis was from 0 to 295.5˚months. The patients were classified based on the International Neuroblastoma Staging System (INSS, https://www.cancer.org/cancer/neuroblastoma/detection-diagnosis-staging/staging.html) and the MYCN-amplified (MNA) was measured. The ratio of MNA patients and total patients in each stage: stage 1 (3/121), stage 2 (5/78), stage 3 (15/63), stage 4 (65/183), and stage 4 (4/53). Furthermore, 176 patients were classified as high-risk ones based on the Revised International Neuroblastoma Response Criteria. The clinical characteristics of 498 NB patients were listed in **Supplementary Table S1**.

The detailed sample preparation was described elsewhere (Oberthuer et al., 2006). Briefly, the patients' tumor samples were manually checked by a pathologist and ensure the sample contains at least 60% of tumor content. The total RNA was then isolated from 30 to 60 mg of snap-frozen tissue obtained before

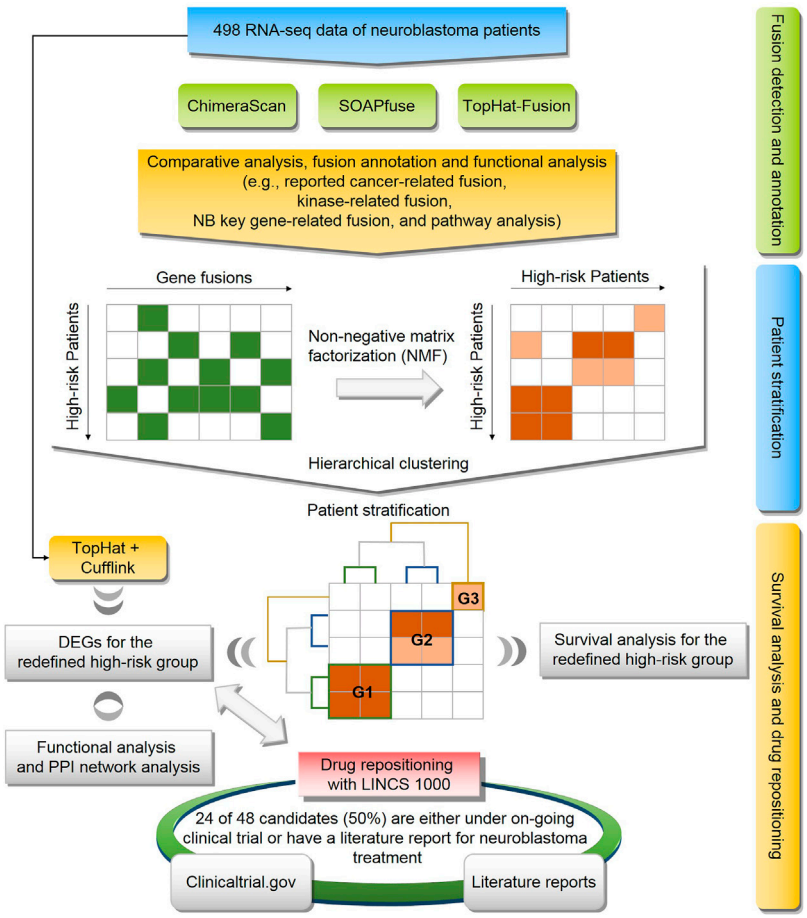

FIGURE 1 | Flowchart of study: the workflow consists of three components including **(1)** fusion detection by three algorithms (i.e., ChimeraScan, SOAPfuse, and TopHat-Fusion) and fusion annotation by publicly available database and pathway analysis; **(2)** The high-risk patients were regrouped based on gene fusion profiles using non-negative matrix factorization (NMF) and hierarchical cluster analysis (HCA); **(3)** Survival analysis and drug repositioning for the redefined patient subgroup.

cytotoxic treatment using the FastPrep FP120 cell disruptor (Qbiogene-Inc, Carlsbad, CA) and the TRIzol reagent (Invitrogen, Karlsruhe, Germany). Last, the RNA integrity was assessed, and the samples were selected with an RNA integrity number of more than 7.5. The RNA-seq data of 498 primary NB samples were generated in FDA Sequencing Quality Control (SEQC) phase I project. The raw RNA-seq data could be downloaded from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) database with series accession number GSE62564 (Zhang et al., 2015). The purified mRNA was extracted from total RNA using Dynabeads® mRNA Purification Kit (Invitrogen), and ERCC RNA spike-in was injected based on the user guide for RNA sequencing. Then, the non-stranded TruSeqs™ protocol was used to conduct Library preparation. Next, clusters were generated based on the TruSeq PE Cluster Kit v3 reagent preparation guide. Last. High-throughput shotgun sequencing was performed on the Illumina HiSeq 2000 platform with the paired-end 100-bp reads. A total of 30,753,066,000 reads were produced, which enabled high coverage of the entire genome spectrum of NB. The average reads per sample was 6.1790e+07 ± 9.5474e+06, and the reads were distributed in a range of 39270986 to 10442490 (see **Supplementary Figure S1**). More detailed information on sequencing data generation was described elsewhere (Zhang et al., 2015).

Detection of Transcript Expression and Fusion Transcripts

There are more than 20 state-of-art fusion transcript detection tools (Liu et al., 2016). In this study, we applied the three most cited tools, including TopHat-Fusion (Kim and Salzberg, 2011), ChimeraScan (Iyer et al., 2011), and SOAPfuse (Jia et al., 2013), to detect fusion transcripts from the RNA-seq data of 498 primary NB samples. In this study, we used the human reference genome sequence (hg19, downloaded from the UCSC Genome Browser: http://hgdownload.soe.ucsc.edu/downloads.html#human) to detect the transcript-level expression and fusion transcripts.

To generate a transcript-level expression, we used TopHat2 v2.1.0 (Kim et al., 2013) to align the raw reads to the UCSC human genome and quantified the expression levels of all the transcripts with FPKM (Fragments Per Kilobase of transcript per

Million fragments mapped) values using Cufflinks v2.2.1 (Trapnell et al., 2010). All the parameters in the pipelines of alignment and quantification were set as default. The expressions at transcription-level were further summarized to the annotated genes in the RefSeq database (https://www.ncbi.nlm.nih.gov/refseq/) by using the highest FPKM value in those of transcripts as the expression of a gene when there were multiple transcripts corresponded to one gene (Beccuti et al., 2014).

There are many parameters in different fusion transcript detection algorithms, which have a significant impact on calling performance. However, optimization of parameter setting for each algorithm with different datasets is far beyond the scope of our paper. Therefore, we mainly followed the default parameter setting in each algorithm.

TopHat-Fusion

For TopHat-Fusion, all raw reads in FASTQ files were first aligned to the reference genome by using the TopHat2 v2.1.0 (Kim et al., 2013). Then, the initially unmapped reads were split into small segments and remapped to the reference genome for identifying the initial fusion candidates by TopHat-Fusion. Lastly, the fusions, for which the number of fusion spanning reads was higher than five and the sum of the fusion spanning reads, and the supporting mate pairs were greater than 10, were kept as candidate fusion transcripts.

ChimeraScan

The ChimeraScan pipeline utilized Bowtie (version 1.1.2) to align the raw reads to the reference genome (Langmead and Salzberg, 2012). The subsequent procedures of nominating the candidate fusions, detecting the spanning reads, and filtering the false positives were conducted with the default settings.

SOAPfuse

For SOAPfuse (Li et al., 2009), the raw reads were mapped to the reference genome by using the SOAP2 (version 2.21) algorithm. The single-end and paired-end mapped reads were kept for identifying the candidate gene fusions. The unmapped reads were then aligned to the annotated transcripts (Ensemble release), and the mapped reads were retained. Finally, the unmapped reads in the second step were iteratively trimmed and realigned to the annotated transcripts until the length of the reads was less than 30 nucleotides. The reads still unmapped to the annotated transcripts were filtered out. All the aligned reads in the three steps were used to detect the gene fusions by seeking the span-reads. The maximum hits for each span-read, and the filter parameters for identifying the gene fusions, were set as default.

Gene Fusion Annotation

The detected fusion transcripts were annotated by the following strategies. First, the reported fusion transcripts were curated and combined based on three public resources, including the Catalogue of Somatic Mutations in *Cancer* (COSMIC, https://cancer.sanger.ac.uk/cosmic/download) (Forbes et al., 2008), Tumor Fusion Gene Data Portal (http://tumorfusions.org/) (Xu et al., 2018), and ChimerDB 3.0 (https://academic.oup.com/nar/article/45/D1/D784/2605708) (Lee et al., 2017). Second, a list of genes associated with neuroblastoma risk was curated by literature survey by querying against PubMed and other databases using key words "genes" and "neuroblastoma". Third, a file of human protein kinases was downloaded from UniProt (https://www.uniprot.org/docs/pkinfam). All the annotation data sets were listed in **Supplementary Table S2**.

Fusion Transcripts-Based Stratification and Survival Analysis

The 176 high-risk NB patients were stratified into new subgroups based on the detected fusion transcript profiles. First, the patient-fusion transcript profiles matrix was constructed based on fusion transcripts detected from each algorithm. Then, NMF was used to decompose the patient-fusion transcript profiles matrix (***F***: 176 patients × unique number of fusion transcripts) into two matrices. 1) patient subgroup assignment (***W***: 176 patients × *k* subgroups) 2) fusion transcript assignment (***H***: k subgroups × unique number of fusion transcripts). This procedure was repeated 500 times. Consequently, the patient-patient relationship matrix (176 patients × 176 patients) was generated, and each cell of the matrix represent the probability of any two patients assigned to the same patient subgroup. Subsequently, the Hierarchical clustering analysis was used to create the consensus assignment of patients into *k* subgroups. Finally, Kaplan-Meier survival analysis was conducted for the comparisons between the subgroups, and the *p* values were calculated using the Log-rank test. All the procedures were performed in R (version 3.4.1) with the packages *NMF v0.21.0*, *ggplot2 v2.2.1*, *survminer v0.4.2*, and *survival v2.4.3*.

Importance Fusions for Distigusing the Patients Subgroups

To further investigate the important fusions and classification performance of gene fusions for redefined patient groups, we employed the XGboost binary classifier (Chen and Guestrin, 2016). XGBoost (Extreme Gradient Boosting) is an ensemble machine learning algorithm for regression and classification problems based on the Gradient Boosting Decision Tree (GBDT), which has been widely applied in biomedical applications (Ji et al., 2019). Specifically, the 176 high-risk patients were stratified into patient subgroups, and we developed the XGboost classifier on the gene fusion profiles and their redefined patient groups. The important gene fusion profiles and performance metrics (i.e., The area under the receiving operating characteristic curve (AUC)) based on 100-run 5-fold cross-validation were calculated. The calculation was performed in R (version 3.4.1) with packages xgboost version 1.3.2. The detailed hyperparameters, including binary: logistic objective, max-depth 6, step size of each boosting step 50, were used.

Differentially Expressed Genes in Patient Subgroups

To identify the differentially expressed genes (DEGs), we separately compared the transcript profiles of the patients in *k*

subgroups with those of the patients in the control group, in which the survival days of the patients were longer than the median survival days among the stages 1 and 2. The DEGs were finally identified using the R packages *limma* and *edgeR* with an adjusted *p* value less than 0.05 as a cut-off value (Robinson and Smyth, 2008). The genes in the DEG list were ranked by their fold changes in descending order and the top/down 500 genes were extracted for further analysis. The functional analysis of extracted DEGs in each patient subgroup was conducted by using The Database for Annotation, Visualization, and Integrated Discovery (DAVID, https://david.ncifcrf.gov/) [44].

Enrichment of Repositioning Candidates

Drug-induced transcriptional profiles of NIH LINCS project (http://www.lincsproject.org/) (Corsello et al., 2017) were employed to enrich repositioning candidates specifically for each patient subgroup. The hypothesis behind genome-based repositioning is that if the drug signature is reversely correlated with the disease signature, the drug could be potentially used to treat the diseases. For LINCS data, LINCS L1000 characteristic direction signatures search engine (L1000CDS^2) was used to reversely compare the DEG in each patient subgroup to the drug transcriptional signatures in LINCS project (Duan et al., 2016). The L1000CDS^2 used data sets including LINCS L1000 level 3 normalized data and level 5 moderated Z-scores (MODZ), which were downloaded from lincscloud.org and GEO (GSE70138). There are a total of transcriptomic profiles generated from 98 cell lines.

Immune Cell Gene Signatures

Although neuroblastoma is typically considered to be an immunologically 'cold' tumor (Szanto et al., 2020), several studies have demonstrated the presence of tumor-infiltrating lymphocytes (e.g., T cells and NK cells), in human neuroblastoma tumors (Coughlin et al., 2006; Hishiki et al., 2018; Wienke et al., 2021). Therefore, we further investigated the immune-related gene expression in the refined patient subgroup. Immune cell gene expression data in mouse cell lines and tissues were extracted from the Immunological Genome Project (ImmGen) (Heng and Painter, 2008). The pre-processing and normalization of data were described previously (Painter et al., 2011; Kidd et al., 2015). Specifically, 304 differential state gene expression (e.g., fold change values) covering 11,153 mapped ortholog human Entrez gene ids were generated between two steady-state profiles from 221 unique immunological cell types (stored at https://github.com/iguana128/Gene-fusion_NB). In this study, we ranked order gene expression profile from high to low based on fold change values for each of 304 immune-related states. Then, the top/down 500 genes in each immunological state were selected as DEGs for further analysis.

Code Availability

The scripts of RNA-seq analysis, data analysis, and data curation were listed in the GitHub (https://github.com/iguana128/Gene-fusion_NB). Furthermore, the scripts for generating all the figures and tables were provided as well.

RESULTS

Overview of Gene Fusion Transcripts

Three algorithms, including ChimeraScan (Iyer et al., 2011), SOAPfuse (Jia et al., 2013), and TopHat-Fusion (Kim and Salzberg, 2011), were employed to harbor gene fusion transcripts among 498 human NB cases. A total of 9,153 (i.e., 9,153/498 = 18.37 per patient), 2,004 (i.e., 2,004/498 = 4.02 per patient), and 1,057 (i.e., 1,057/498 = 2.12 per patient) unique gene fusions were detected by ChimeraScan, SOAPfuse, and TopHat-Fusion, respectively (**Supplementary Table S3**). We found 47.07% (4,308/9,153), 59.28% (1,188/2,004), and 43.61% (461/1,057) of total detected fusions only specific to the individual patient for ChimeraScan, SOAPfuse, and TopHat-Fusion, respectively. ChimeraScan is more sensitive to detect more gene fusions than another two algorithms. There were only 221 gene fusion transcripts identified by at least two algorithms, which represents the divergence of fusion detection algorithms (**Figure 2A**). We further investigated the distribution of detected gene fusions across the patient subgroups defined by using the International Neuroblastoma Staging System (INSS). Specifically, the number of identified gene fusions were increased from the early stages (stage 1 ~ stage 3) to late stages (stage 4 and high-risk group), indicating the higher stage has more complex tumor compositions. For ChimeraScan and TopHat-Fusion, more across-chromosomal fusions were detected than that of inter-chromosomal fusions. However, SOAPfuse harbored more inter-chromosomal fusion transcripts (**Figure 2B**). Furthermore, the concordances among the three algorithms for individual patients also tended to be increased in patients in stage 4 and high-risk groups, although overall concordances were suboptimal (less than 5% overlapped ratio) (**Supplementary Figure S2**).

The detected gene fusions from the three algorithms were further annotated by the curated knowledge concerning reported cancer-related genes fusions, kinase protein family, and neuroblastoma related key genes. Similarly, patients from late stages (stage 4 and high-risk group) enriched more regarding reported cancer-related fusions, kinase-related fusions, and neuroblastoma key genes related fusions. The total number of annotated fusions was increased from TopHat-Fusion to SOAPfuse and ChimeraScan (**Figure 2C**). **Table 1** summarized the annotated fusions identified by the three algorithms. SOAPfuse detected more reported cancer-related fusions (90) than that of ChimeraScan (26) and TopHat-Fusion (10). *Axon guidance* pathway was enriched by the genes involving reported cancer-related gene fusions from both ChimeraScan and SOAPfuse. However, the enriched pathways (e.g., *Hippo signaling pathway* and *Oxytocin signaling pathway*) based on reported gene fusions identified by TopHat-Fusion were distinct. *MAPK signaling pathway* was enriched by kinase-related fusions from both ChimeraScan and TopHat-Fusion, although the involved gene fusions were entirely different. It indicated that fusion detection algorithms have a relatively higher similarity in the biological levels. It was interesting that some neuroblastoma related essential genes are more susceptible to form gene fusions. For example, ChimeraScan was more sensitive to *ALK* related

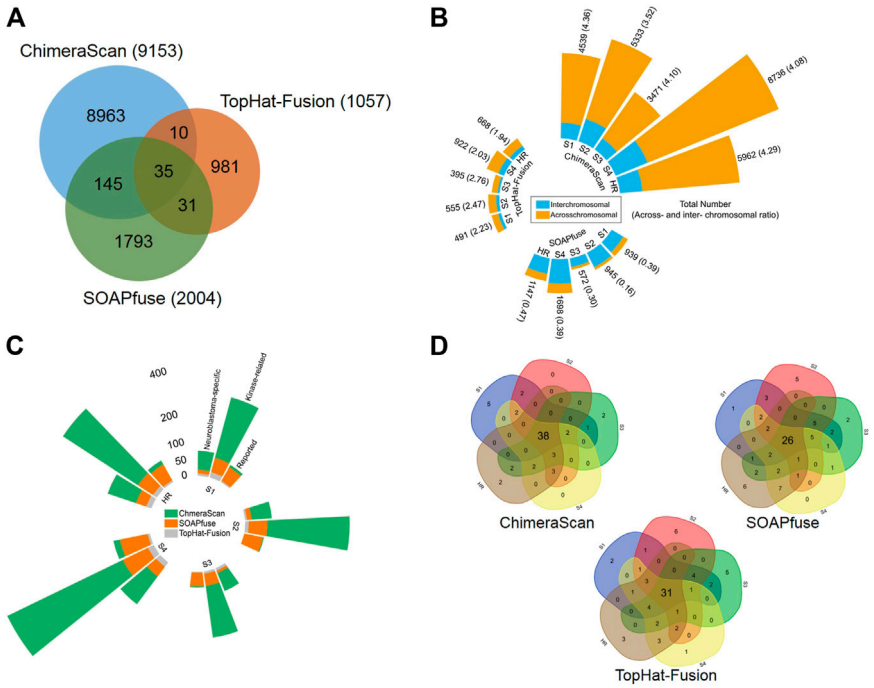

FIGURE 2 | Comparative analysis and annotation of identified gene fusions by three algorithms, including ChimeraScan, SOAPfuse, and TopHat-Fusion: **(A)** a Venn diagram of detected fusions among three algorithms; **(B)** Distribution of detected gene fusions in different INSS clinical stages and high-risk group. The yellow and blue color represents the inter- and across-chromosomal gene fusion, respectively; **(C)** The detected fusions in each clinical stage were further annotated by reported cancer-related gene fusions, kinase-related fusions, and neuroblastoma key gene-related gene fusions; **(D)** The high-frequency gene fusions for each algorithm were extracted across the different clinical stages.

fusions, while SOAPfuse and TopHat-Fusion detected more *MYCN* and *LOM1* related fusions. *DDX1* associated fusions were identified by all three algorithms.

We further extracted the top fifty high-frequent gene fusions across the patients in each stage for the three algorithms (**Figure 2D**). There were 38, 26, and 31 high-frequency gene fusions that appeared in all clinical stages for ChimeraScan, SOAPfuse, and TopHat-Fusion, respectively (**Supplementary Table S4**). However, as mentioned above, substantial low-frequency gene fusions are more patient-specific.

Fusion Transcripts-Based Stratification

To further investigate the discrimination power of gene fusion profiles, we stratified 176 high-risk NB patients into new subgroups based on the detected fusion transcripts in each of the three algorithms. Here, only fusions appeared more than one time in the high-risk group were employed. Consequently, the high-risk patient and gene fusion profiles matrices were constructed with dimensions 176 patients × 2,509 fusions, 176 patients × 446 fusions, and 176 patients × 351 fusions for ChimeraScan, SOAPfuse, and TopHat-Fusion, respectively. Then, 176 patients were divided into a predefined number of subtypes (k = 6, 6, and 3) by using average subtype assignment frequency from 500 runs of the NMF algorithm. The predefined subtype k was optimized by using cophenetic degree and sparseness parameters in NMF (**Supplementary Figure S3**). Then, patient subtype assignment matrices (176 patients × 176 patients) for each fusion detection method were further clustered by using hierarchical clustering analysis (HCA) (**Figure 3A**). The 176 patients were clustered into four subgroups for each fusion algorithm. The overall survival distribution of each subtype was illustrated in **Figure 3B**. It is illustrated that there existed one subgroup (number of patients = 99, 120, and 67) with significantly smaller overall survival time (median overall survival = 798, 883, and 735˚days) than another three groups (median overall survival = 1,242, 1,241, and 1,229˚days). It was found 39 patients were overlapped by the subgroups based on fusion profiles from ChimeraScan (39/99 = 39.4%), SOAPfuse (39/120 = 32.5%), and TopHat-Fusion (39/67 = 58.2%), respectively (**Supplementary Figure S4**). The Kaplan-Meier survival analysis showed significantly different survival times between the redefined patient subgroups with p values (0.014, 0.022, and 0.032) and hazard ratios (1.670, 1.638, and 1.617) for the three algorithms (**Figure 3C**; **Supplementary Table S5**). It was indicated that the fusion transcript profiles could be used to further distinguish the patients from the high-risk group with improved survival rates.

Table 2 listed the clinical characteristics of redefined high-risk patients. MYCN-amplified patients are more enriched in the refined high-risk patient group (73.7, 72.5, and 59.7% for ChimeraScan, SOAPfuse, and TopHat-Fusion, respectively). In contrast, the lower percentage of MYCN-amplified patients was classified into other subgroups (i.e., 24.7, 8.9, and 47.7% for ChimeraScan, SOAPfuse, and TopHat-Fusion, respectively). It

TABLE 1 | Annotated gene fusions in different fusion detection algorithms.

Fusion detection algorithms	Number of identified fusions	Representative fusions	Involved KEGG pathways	Involved genes	*p* values
Reported cancer-related gene fusions					
ChimeraScan	26	TRA@_TRA@; CNRIP1_PPP3R1; CAMTA2_SPAG7; C11orf48_INTS5 BMPR1B_PDLIM5; DDX17_DMC1; GCC2_RANBP2; GPR128_TFG; POR_RHBDD2; SIPA1L3_WDR62	hsa04360: Axon guidance	NCK1, MET, PPP3R1, PPP3CC	5.5E-3
			hsa04660: T cell receptor signaling pathway	NCK1, PPP3R1, PPP3CC	3.5E-2
SOAPfuse	90	CLSTN1_CTNNBIP1; CYB5R4_RIPPLY2; ACTN4_EIF3K; ANKDD1A_PLEKHO2; BAG2_ZNF451; EIF4E3_FOXP1; CNRIP1_PPP3R1; TBCEL_TECTA; ADIPOR1_CYB5R1; POR_RHBDD2	hsa04520: Adherens junction	ACTN4, MET, LMO7, CTNND1, IQGAP1	7.2E-3
			hsa04360: Axon guidance	ABLIM2, NCK1, MET, PPP3R1, PPP3CC	4.9E-2
TopHat-fusion	10	DDX1_NBAS; PLB1_PPP1CB; ALK_GALNT14; CCND1_ORAOV1; CNRIP1_PPP3R1; DNAJB4_FUBP1; EIF4E3_FOXP1; BAIAP2_TBCD; BMPR1B_PDLIM5; KDM4A_ST3GAL3	hsa04390: Hippo signaling pathway	CCND1, BMPR1B, PPP1CB	1.2E-2
			hsa04921: Oxytocin signaling pathway	CCND1, BMPR1B, PPP1CB	1.3E-2
Kinase-related gene fusions					
ChimeraScan	422	AK095450_FER; FLJ25037_KSR1; MAST2_TIMM23; PHKG2_TH; DAPK3_MYO9A; MXD4_PKN1; AF070581_PAK3; MAP3K15_SNORD10; STK38L_TRIM8; AK124179_PASK	hsa04010: MAPK signaling pathway	FGFR2, FGFR1, FGFR3, MAPKAPK5, MAP4K2, MAPKAPK3, MKNK2, MKNK1, AKT1, MAP3K6, MAP3K3, PAK1, AKT3, MAP2K5, AKT2, PRKCA, TAOK2, MAP2K2, NLK, MAP2K3, MAP2K4, TAOK3, NR4A1, PRKCG, MAPK11, MAPK10, STK4, FLNA, STK3, PRKCB, MAPK1, MAP4K4, MAPK12, MAPK13, NTRK1, MAPK14, PDGFRB, MAPK9, MAPK7, MAP3K12	2.7E-18
			hsa04012: ErbB signaling pathway	Prkca, ERBB3, MAP2K2, ERBB2, CAMK2G, MAP2K4, RPS6KB2, PRKCG, MAPK10, PRKCB, AKT1, MAPK1, PTK2, PAK3, GSK3B, MAPK9, PIK3CA, CAMK2B, PAK1, ABL1, CAMK2A, AKT3, AKT2	2.5E-15
SOAPfuse	93	LIMK2_RNF185; CHCHD2_PHKG1; PRKAA1_TTC33; FES_MAN2A2; BCKDK_KAT8; INSL3_JAK3; TAOK2_TMEM219; NTRK1_PEAR1; MAP2K5_SKOR1; ACVR2B_CNOT6	hsa04722: Neurotrophin signaling pathway	IRAK4, MAPK1, PDPK1, MAP2K1, RPS6KA2, CAMK2G, NTRK1, CAMK2B, MAPK7, AKT3, AKT2, MAP2K5	4.5E-8
			hsa04921: Oxytocin signaling pathway	MAPK1, MAP2K1, ROCK2, CAMK2G, PRKAA1, CAMK2B, EEF2, GNAS, MAPK7, SRC, PRKCB, CAMK1D, MAP2K5	8.7E-8
TopHat-fusion	29	IRAK3_RBMS1; STK24_STK24P1; LOC407835_MAP2K2; DAPK1_RPS29; BMX_HNRNPDL; C14orf166_MERTK; MAPK11_MAPK12; ENSG00000226049_TLK2; CDK4_TMEM132C; ALK_GALNT14	hsa04010: MAPK signaling pathway	PAK2, MAPK12, TAOK1, RPS6KA2, MAP2K2, MAP2K4, MAPK11	1.8E-4
			hsa04660: T cell receptor signaling pathway	PAK2, MAPK12, MAP2K2, MAPK11, CDK4	3.9E-4
Neuroblastoma key genes-related fusions					
ChimeraScan	138	CHD5_FOXI3; ALK_ANKS1A; ALK_USP11; CNOT3_DLK1; DLK1_FLJ00420; DLK1_KIAA0691; BC035411_DDX1; DDX1_SMA4; AX748330_RASSF7; PDGFA_SLC29A4	hsa04010: MAPK signaling pathway	PTPN7, BDNF, CACNG8, PDGFA, CACNG7, NTRK1, MAP2K4, PPP3R1, TP53, NR4A1, STK4, STK3	2.9E-5
			hsa04210: Apoptosis	TNFRSF10C, TNFRSF10B, NTRK1, PIK3CD, CASP8, TP53	3.1E-4
SOAPfuse	43	HOXC4_HOXC6; ENSG00000198353_HOXC6; LMO1_RIC3; NTRK1_PEAR1; KIF1B_PGD; DDX1_NBAS; CAMTA1_VAMP3; DDX1_MYCNOS; DDX1_MYCNUN; NPRL2_ZMYND10	hsa01130: Biosynthesis of antibiotics	ODC1, NME2, NME1-NME2, NME1, PGD	5.6E-3
			hsa00240: Pyrimidine metabolism	NME2, NME1-NME2, NME1	5.0E-2
TopHat-fusion	22	ANGPT2_MCPH1-AS1; FMO4_TOP1; EDARADD_ENO1; NME2_NME2P1; DDX1_NBAS; ALK_GALNT14; MYCN_NBAS; DDX1_MYCNUT; FOXR1_PAFAH1B2; HACE1_SCML4	hsa01130: Biosynthesis of antibiotics	ODC1, NME2, NME1-NME2, ENO1	5.1E-3
			hsa01100: Metabolic pathways	ODC1, NME2, NME1-NME2, PAFAH1B2, GALNT14, ENO1	4.6E-2

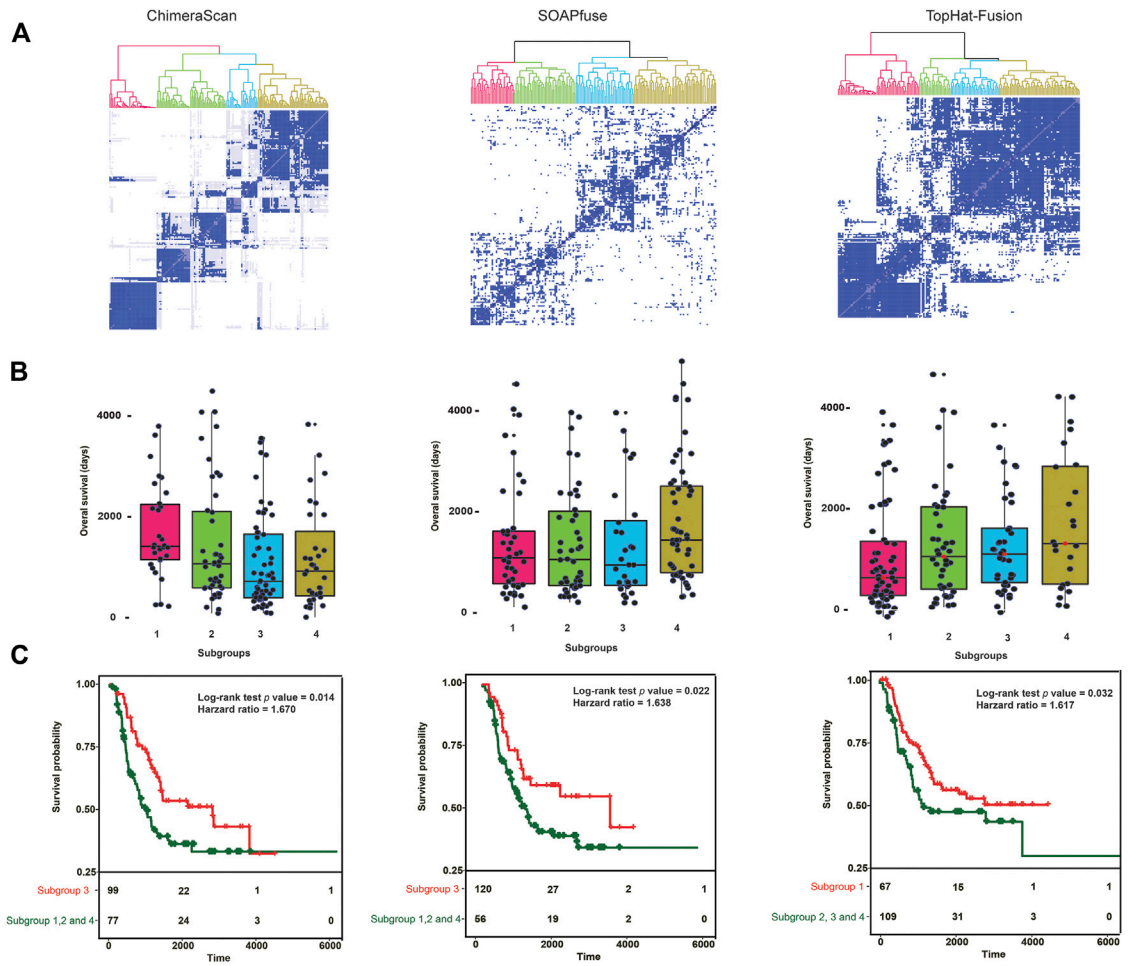

FIGURE 3 | Gene fusion profile-based patient stratification and survival analysis: **(A)** hierarchical cluster analysis (HCA) for redefining patient subgroups in high-risk NB patient groups for each fusion detection algorithm. First, patient assignment matrices were generated from average results of 500 runs with nonnegative matrix factorization (NMF) with patient gene fusion profiles. Then, the HCA was performed on the patient assignment matrices to redefine the patient groups. The color denotes the redefined patient groups; **(B)** The overall survival time (Days) of patients in the redefined patient subgroups. The subgroup with lowest median survival time was considered as the redefined high-risk group; **(C)** Kaplan-Meier survival analysis was conducted between the redefined high-risk subgroups and the combination of other subgroups. Log-Rank *p* value and the hazard ratio was calculated.

was indicated that the redefined patient group might be MYCN-amplified, especially for ChimeraScan and SOAPfuse. Furthermore, the age of diagnosis was more than 18°months for over 92% of refined high-risk patients and 96% of patients in other groups, highlighting the early diagnosis of high-risk NB patients is very challenging. Additionally, a higher percentage of males were assigned into the redefined high-risk patient group, e.g., 57.5 and 62.7% for ChimeraScan and TopHat-Fusion, respectively. However, a higher rate of females was classified into other subgroups. The detail clinical information for redeinfed high-risk patients were listed in **Supplementary Table S5**.

To further investigate the important gene fusion profiles that could distinguish the high-risk group patients, we developed an XGboost binary classifier. **Figure 4** illustrated the top 10 important features derived from XGboost classifiers for the three fusion detection tools. Furthermore, the average AUC values of 100-run 5-fold cross-validations were ranked with the following order: SOAPfuse (0.862 ± 0.031) > ChimeraScan (0.849 ± 0.018) > TopHat-Fusion (0.799 ± 0.013). The high AUC and small stand deviation among 100 5-fold cross-validation results indicated the reliable classification results could be obtained based on the gene fusion profiles derived from the three detection tools.

Differentially Expressed Genes Associated With Patient Subgroups

We next sought for the DEGs associated with the redefined high-risk subgroups and examined their underlying mechanism. The DEGs were generated by comparing the transcript profiles between the patients in the redefined high-risk subgroup, and patients with the overall survival days were longer than the median survival days in stages 1 and 2. Here, the

TABLE 2 | Clinical characteristics of redefined high-risk patients

Clinical characteristics	Group	Algorithm	Number	Percentage of total
MYCN status				
Normal	Redefined high-risk	ChimeraScan	26	27.3%
		SOAPfuse	33	27.5%
		TopHat-fusion	27	40.3%
	Others	ChimeraScan	57	74.0%
		SOAPfuse	50	89.3%
		TopHat-fusion	56	51.3%
Amplified	Redefined high-risk	ChimeraScan	73	73.7%
		SOAPfuse	87	72.5%
		TopHat-fusion	40	59.7%
	Others	ChimeraScan	19	24.7%
		SOAPfuse	5	8.9%
		TopHat-fusion	52	47.7%
N.A.	Redefined high-risk	ChimeraScan	0	0
		SOAPfuse	0	0
		TopHat-fusion	0	0
	Others	ChimeraScan	1	1.3%
		SOAPfuse	1	1.8%
		TopHat-fusion	1	1.0%
Age at diagnosis				
<18°months	Redefined high-risk	ChimeraScan	6	6.1%
		SOAPfuse	9	7.5%
		TopHat-fusion	5	7.5%
	Others	ChimeraScan	3	3.9%
		SOAPfuse	0	0%
		TopHat-fusion	4	3.7%
>18°months	Redefined high-risk	ChimeraScan	93	93.9%
		SOAPfuse	111	92.5%
		TopHat-fusion	62	92.5%
	Others	ChimeraScan	74	96.1%
		SOAPfuse	56	100%
		TopHat-fusion	105	96.3%
Sex				
Male	Redefined high-risk	ChimeraScan	57	57.5%
		SOAPfuse	52	43.3%
		TopHat-fusion	42	62.7%
	Others	ChimeraScan	54	70.1%
		SOAPfuse	43	76.8%
		TopHat-fusion	69	63.3%
Female	Redefined high-risk	ChimeraScan	42	42.5%
		SOAPfuse	68	56.7%
		TopHat-fusion	25	37.7%
	Others	ChimeraScan	23	29.9%
		SOAPfuse	13	23.2%
		TopHat-fusion	40	26.7%

top 500 up-and down-regulated genes were extracted (**Supplementary Table S6**). The overlapped DEGs associated with refined high-risk subgroups from the three algorithms occupied 65.3% of 1,000 DEGs for each fusion detection method (**Figure 5A**). The shared KEGG pathways enriched by using DEGs were illustrated in **Figure 5B**. Three KEGG pathways, including *Ribosome*, *Cell cycle*, and *DNA replication,* were enriched by all the three algorithms. The DEG from TopHat-Fusion enriched more immune-related pathways such as *Primary immunodeficiency* and *an Intestinal immune network for IgA production* (**Supplementary Table S7**). **Figure 5C** highlighted the top ten up- and down-regulated genes associated with refined high-risk subgroups. *MYCN*, *MYCNOS*, and *SLC30A3* were the most up-regulated genes, while *APOD*, *INSRR*, and *PIRT* were the most down-regulated genes across the three algorithms. Those genes have been reported to play a different regulation role in NB development (Huang and Weiss, 2013; Olsson et al., 2016). Notably, *MYCN* has been found in ~25% of high-risk NB patients and correlated with poor diagnosis (Huang and Weiss, 2013).

Another interesting finding here is the onset age of NB in the redefined high-risk NB patient group is significantly lower than another three redefined subgroups for ChimeraScan and SOAPfuse. It is indicated that the development of an adaptive immune system may play a role in neuroblastoma evolution (**Figure 6**). We further compared the DEG in a redefined high-risk group with the highly expressed gene signatures in

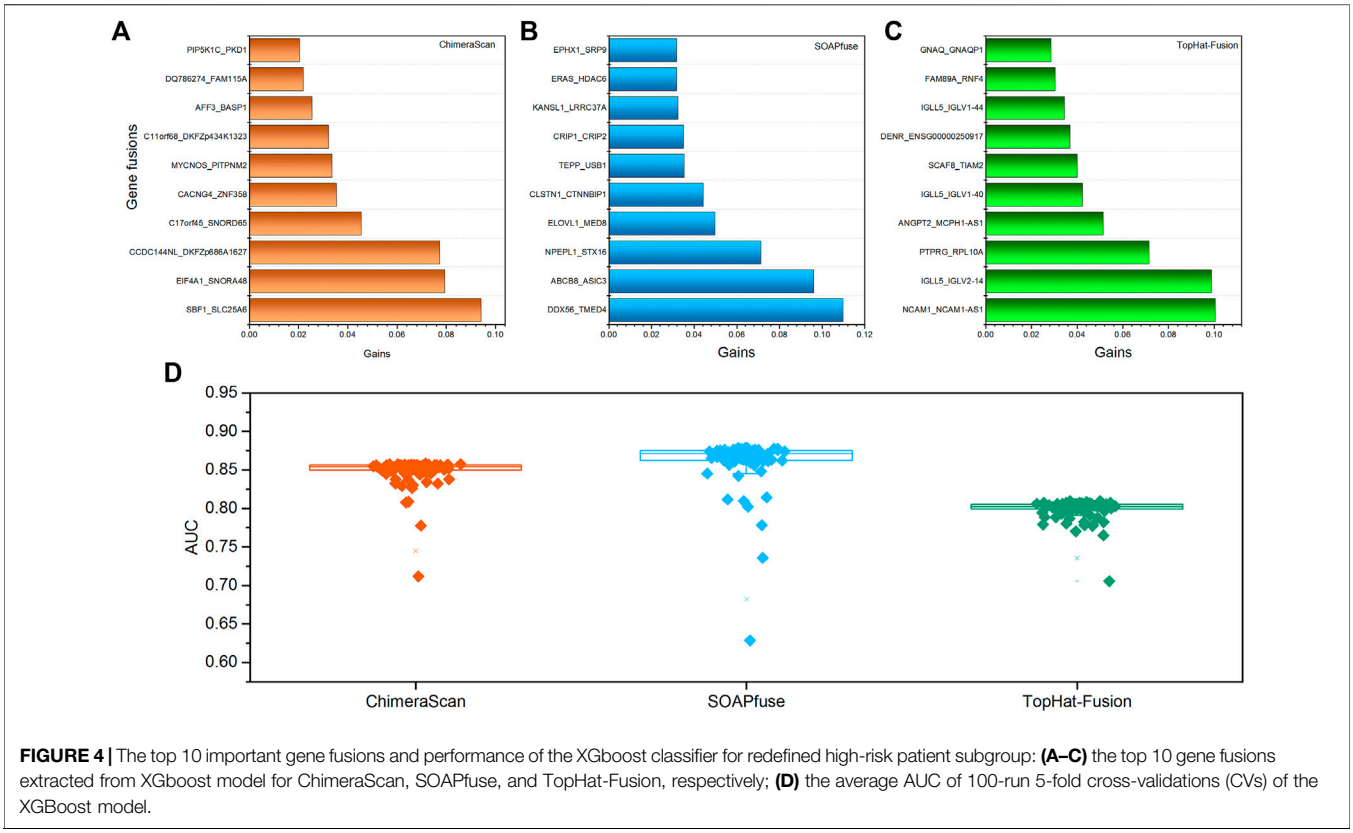

FIGURE 4 | The top 10 important gene fusions and performance of the XGboost classifier for redefined high-risk patient subgroup: **(A–C)** the top 10 gene fusions extracted from XGboost model for ChimeraScan, SOAPfuse, and TopHat-Fusion, respectively; **(D)** the average AUC of 100-run 5-fold cross-validations (CVs) of the XGBoost model.

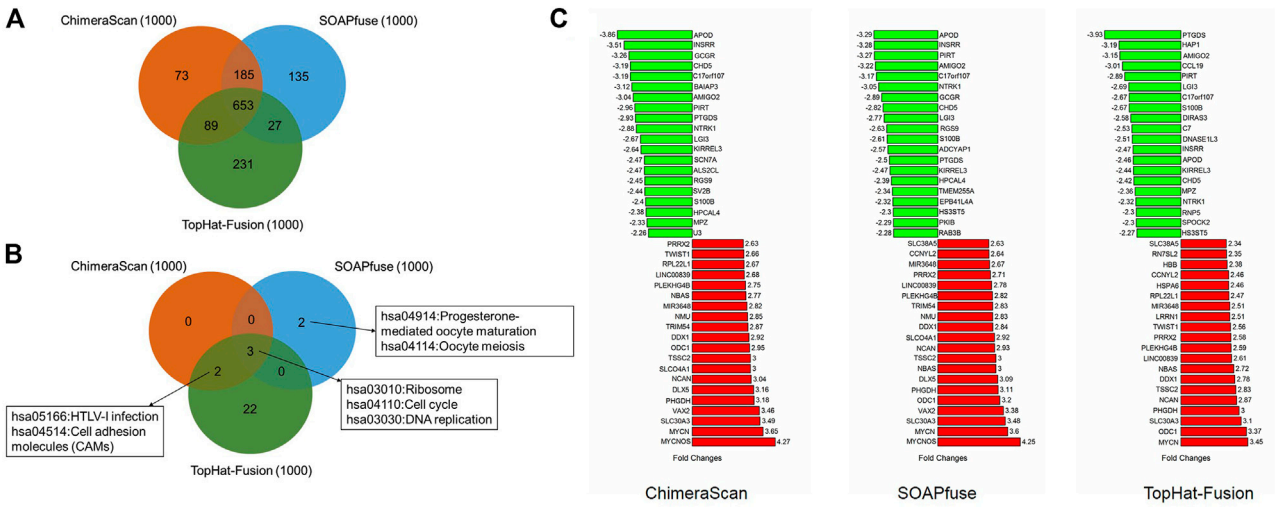

FIGURE 5 | The comparative analysis of the differential expressed genes (DEGs) and related KEGG pathways regulated by the redefined high-risk group: **(A)** a Venn diagram of DEGs regulated by the redefined high-risk group for the three fusion detection algorithms; **(B)** a Venn diagram of KEGG pathways enriched by DEGs; **(C)** the top 20 up- and down-regulated genes of DEGs derived from the three algorithms.

304 immune-related states obtained from the ImmGen (https://www.immgen.org/) (Heng and Painter, 2008). The similarities between DEG in the redefined high-risk subgroup and immune-related cell types were calculated and rank-ordered (**Supplementary Table S8**). The percentage of overlapped ranked immune-related cell types among the three fusion detection algorithms was illustrated in **Figure 7**. The enriched immune-related cell types were very similar among the three

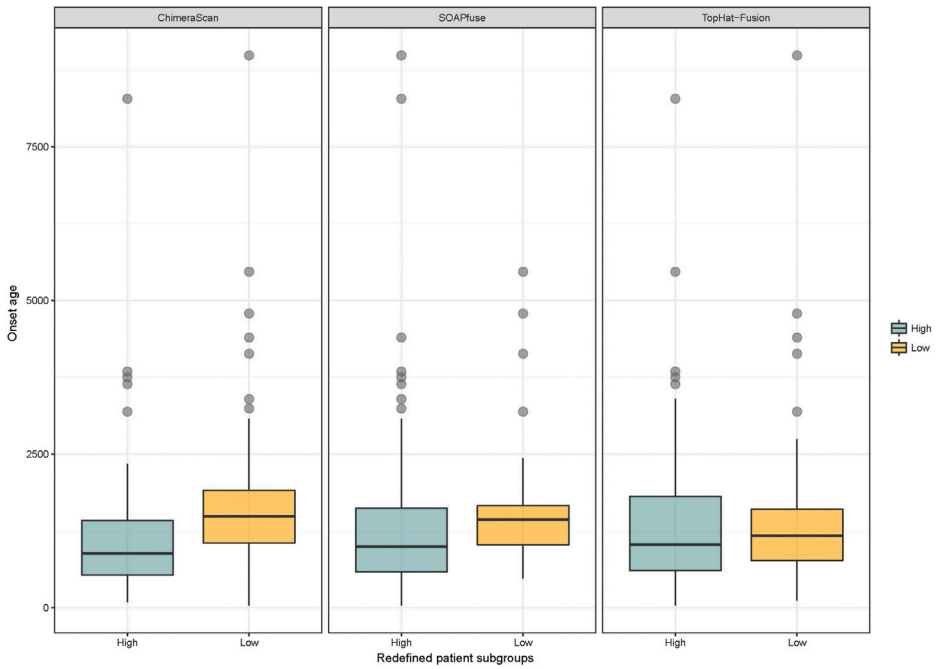

FIGURE 6 | Distribution of onset age in the redefined high-risk and low-risk NB patients: blue and yellow colors represent the redefined high-risk and low-risk patient subgroups, respectively. The student's t-test was used to generated *p* value.

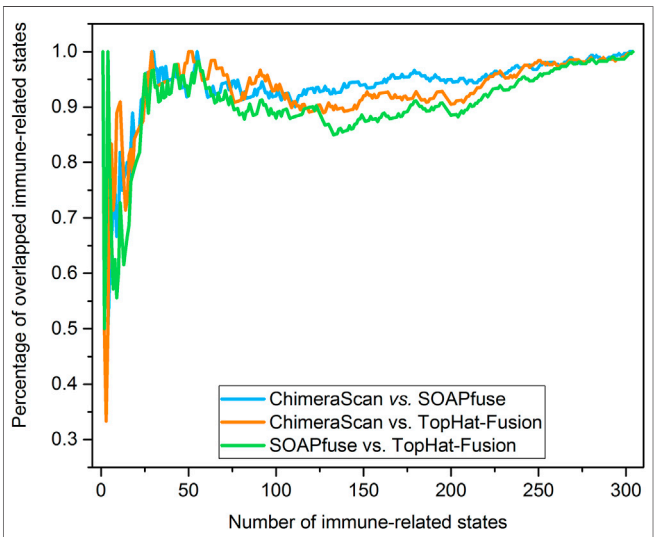

FIGURE 7 | Percentage of overlapped immune-related cell types: the immune-related cell types were enriched and rank-ordered from high to low by comparing the DEG regulated by the redefined high-risk subgroup for each fusion detection algorithm and gene signatures of immune-related cell type from ImmGen (https://www.immgen.org/). Then, the percentage of overlapped enriched immune-related was calculated between any two rank-ordered immune-related cell type lists. The top-ten enriched immune-related cell types were illustrated in the sub-table for each fusion detection algorithm.

algorithms. Furthermore, we listed the top ten enriched immune-related cell types from each method and found T8 cell types dominated (Orentas et al., 2012).

Repositioning Candidates for Redefined High-Risk Subgroup

We further explore the potential repositioning candidates that could be used to treat the patients in the redefined high-risk subgroup. The L1000 drug signatures were employed to compare the DEGs of the redefined high-risk group. LINCS L1000 characteristic direction signatures search engine (L1000CDS^2) was applied to enrich Repositioning candidates. If the drug signatures were reversely correlated with DEG in a refined high-risk group, the drug was considered as potential repositioning candidates. Consequently, we got the top 50 repositioning candidates for patients in the redefined high-risk group from each fusion detection algorithms. **Supplementary Table S9** listed the detailed information on repurposing candidates, including compound/time/dose/cell line enriched from L1000CDS^2. Interestingly, a total of 37 Repositioning candidates were overlapped by at least two refined high-risk groups from different fusion detection algorithms (**Supplementary Table S9**).

To verify the enriched repositioning candidates, we conducted a two-step analysis. **Table 3** compiles the candidates for the redefined high-risk groups defined by the three fusion detection algorithms. First, we queried clinical trial studies through http://clinicaltrial.gov (www.clinicaltrials.gov) to seek for the clinical evidence of repositioning candidates for the treatment of neuroblastoma. Notably, three repurposing candidates, including selumetinib, vemurafenib, and trametinib, were in clinical trials for potentially treating NB. Second, we further carried out a comprehensive literature survey through PubMed (https://www.ncbi.nlm.nih.gov/pubmed/) using the keywords

TABLE 3 | Summary information of Repositioning candidates for neuroblastoma patients in refined high-risk groups.

Repositioning candidates	Clinical phases	Mode of actions (MoAs)	Approved/Investigated therapeutic categories	Evidence	Confirmation sources
On-going clinical studies in clinicaltrial.gov					
selumetinib (BRD-K57080016)	Phase 2	Mitogen-activated protein kinase (MEK) inhibitor	Solid tumors such as neuroblastoma; non-Hodgkin lymphoma	Efficacy evaluation of selumetinib in treating patients with solid tumors such as recurrent neuroblastoma, non-Hodgkin lymphoma, or histiocytic disorders with MAPK pathway activation mutations that have spread to other places in the body and have come back or do not respond to treatment. (A pediatric MATCH treatment trial)	NCT03213691
vemurafenib (PLX4032, RG7204)	Phase 2	V600E mutated BRAF inhibitor	Melanomas; 8% of all solid tumors, including neuroblastoma, melanoma, colorectal, thyroid and other cancers	Efficacy evaluation of vemurafenib in treating patients with relapsed or refractory advanced solid tumors such as recurrent neuroblastoma, non-hodgkin lymphoma, or histiocytic disorders with BRAF V600 mutations (A pediatric MATCH treatment trial)	NCT03220035
trametinib	Phase 1	MEK inhibitor	Refractory solid tumors such as neuroblastoma; lymphomas; multiple myeloma	Next generation personalized neuroblastoma therapy by using trametinib	NCT02780128
Literature support from PubMed					
BRD-K68548958 (C646)	Investigational	Histone acetyltransferase p300 inhibitor	Prostate and lung cancers	Mouse *in vitro* (primary murine cortical neurons):C646 as a selective histone acetylation could regulate the expression of omega-3 polyunsaturated fatty acid, docosahexaenoic acid (DHA)-metabolizing enzyme Alox15 in neuroblastoma cells, which affect cognition and memory in brain development	Pmid: 29235036
BMS-536924	Investigational	Insulin-like growth factor-I receptor (IGF-IR)	Childhood sarcomas	Human *in vitro*: 28 sarcoma and neuroblastoma cell lines were screened for *in vitro* response to BMS-536924 to identify sensitive and resistant cell lines. Notably, Ewing's sarcoma, rhabdomyosarcoma, and neuroblastoma are more responsive to BMS-536924, suggesting these specific subtypes may represent potential targeted patient subpopulations for the IGF-IR inhibitor	Pmid: 19117999
mitoxantrone	Approved	DNA-reactive agent	Breast cancer; acute myeloid leukemia; non-Hodgkin's lymphoma	Patient cohort study: A distinct side population (SP) was found in neuroblastoma cells from 15 of 23 patients (65%). These cells also expressed high levels of ABCG2 and ABCA3 transporter genes and had a greater capacity to expel cytotoxic drugs, such as mitoxantrone, resulting in better survival	PMID:15381773
MK-2206	Phase 2	Allosteric AKT inhibitor	Colorectal cancer; breast cancer; other solid tumors	Human *In vitro*: Combination of an allosteric akt inhibitor MK-2206 with etoposide or rapamycin enhances the antitumor growth effect in high-risk neuroblastoma patients	Pmid: 22550167

(Continued on following page)

TABLE 3 | (*Continued*) Summary information of Repositioning candidates for neuroblastoma patients in refined high-risk groups.

Repositioning candidates	Clinical phases	Mode of actions (MoAs)	Approved/Investigated therapeutic categories	Evidence	Confirmation sources
BRD-K65814004	Investigational	NADH/NADPH oxidase inhibitor	Antibiotics	Human *in vitro*: Neuroblastoma cell death was attenuated by ROS-scavengers and was dose-dependently inhibited by the NADPH oxidase inhibitor diphenyleneiodonium chloride (DPI)	Pmid: 16260066
BRD-A36630025 (SN-38)	Approved	a topoisomerase I inhibitor	Colon cancer, and small cell lung cancer	Mouse *in vivo*: Nanoparticle delivery of an SN-38 conjugate is more effective than irinotecan in a mouse model for treating neuroblastoma	Pmid: 25684664
DL-PDMP	Investigational	Glucosyltransferase inhibitor	Lewis lung carcinoma cell metastasis	Mouse *in vitro* (murine neuroblastoma cells): DL-PDMP is a potent inhibitor of glucosylceramide synthase, resulting in inhibition of the synthesis and shedding of gangliosides, which may contribute to the observed bone marrow depression in neuroblastoma patients	Pmid: 9809988
saracatinib (AZD-0530)	Phase 2	Dual kinase inhibitor, with selective actions as a src inhibitor and a bcr-abl tyrosine-kinase inhibitor	Alzheimer's disease and schizophrenia	Human *in vitro* (neuroblastoma SKNSH cells): Iron depletion results in src kinase inhibition of saracatinib with associated cell cycle arrest in neuroblastoma cells	Pmid: 25825542
wortmannin	Phase 2	Covalent inhibitor of phosphoinositide 3-kinases (PI3Ks)	Recurrent glioblastoma	Human *in vitro* (neuroblastoma SKNSH cells): PI3K pathway inhibition down-regulates surviving expression and enhances TRAIL-mediated apoptosis in neuroblastomas. PI3K pathway may play a crucial role in neuroblastoma cell survival; therefore, treatment with inhibitors of PI3K such as LY294002 or wortmannin may provide potential novel therapeutic options	Pmid: 15065019
palbociclib	Approved	CDK4/6 inhibitor	ER-positive and HER2-negative breast cancer	Human *in vitro*: Selective inhibition of CDK4/6 using palbociclib may provide a new therapeutic option for treating neuroblastoma	Pmid: 26225123
naproxol (BRD-K34014345)	Approved	Nonsteroidal anti-inflammatory drugs (NSAIDs)	Inflammation	Human/rat *in vitro*: Protective effects of the former four non-steroidal anti-inflammatory drugs such as naproxol against apoptosis might be mainly due to their direct nitric oxide radical scavenging activities in neuronal cells	Pmid: 11259508 Pmid: 15975708
Nutlin-3	Phase 1	*cis*-imidazoline analogs	Retinoblastoma	Human *in vitro*: Amplification or overexpression of MYCN sensitizes neuroblastoma cell lines with wild-type p53 to MDM2-p53 antagonists such as Nutlin-3 and MI-63, which may therefore be particularly effective in treating high-risk MYCN-amplified neuroblastoma	Pmid: 21725357
AS605240	Investigational	Selective PI3K inhibitors	Diabetics; rheumatoid arthritis; pulmonary fibrosis; cancer	Human/mouse *in vitro*: Phosphoinositide 3-kinases (PI3K) in selected neuroblastoma tumors with the inhibitor AS605240, which has been shown to express low toxicity and relative specificity for the PI3K species γ	Pmid: 20224967

(Continued on following page)

TABLE 3 | (*Continued*) Summary information of Repositioning candidates for neuroblastoma patients in refined high-risk groups.

Repositioning candidates	Clinical phases	Mode of actions (MoAs)	Approved/Investigated therapeutic categories	Evidence	Confirmation sources
gossypol (BRD-K19295594)	Investigational	Natural phenol derived from the cotton plant	Contraceptive and antimalarial	Human *in vitro*: Mcl1 appears as a predominant pro-survival protein contributing to chemoresistance in neuroblastoma, and Mcl1 inactivation may represent a novel therapeutic strategy. Optimization of compounds with higher Mcl1 affinity, or combination with additional Mcl1 antagonists such as gossypol, may enhance the clinical utility	Pmid: 19556859
ixazomib (MLN2238)	Approved	Proteasome inhibitor	Multiple myeloma	Mouse *in vivo*: ixazomib not only inhibits neuroblastoma cell proliferation and induces apoptosis but also enhances dox-induced cytotoxicity in neuroblastoma cells, suggesting that combination therapy including ixazomib with traditional therapeutic agents such as dox is a viable strategy that may achieve better outcomes for NB patients	Pmid: 27687684
canertinib (CI - 1033)	Discontinued	Irreversible tyrosine-kinase inhibitor	Various of cancer types	Human *in vitro*: Non-EGFR ERBB family members such as canertinib contribute directly to neuroblastoma growth and survival, and pan-ERBB inhibition represents a potential therapeutic target for treating neuroblastoma	Pmid: 20564646
AZD8055	Phase 1/2	mTOR inhibitor	Recurrent gliomas	Human *in vitro* and mouse *in vivo*: AZD8055 can induce cell cycle arrest, autophagy and apoptosis. AZD8055 strong antitumor activity on neuroblastoma *in vitro* and *in vivo*, which may be further investigated for treatment in clinical trials for high risk NB	Pmid: 29499203
teniposide	Approved	Podophyllotoxin derivatives	Acute lymphocytic leukemia (ALL)	Children with neuroblastoma have a significantly higher incidence of acute reactions to teniposide than patients with other malignancies ($p = 0.008$), and that these reactions cannot be prevented by premedication with antiallergic drugs	Pmid: 3857970
vorinostat	Approved	Histone deacetylases (HDAC) inhibitor	Cutaneous T cell lymphoma (CTCL)	Mouse *in vitro:* Vorinostat created a permissive tumor microenvironment (TME) for tumor-directed mAb therapy by increasing macrophage effector cells expressing high levels of fc-receptors (FcR) and decreasing the number and function of myeloid-derived suppressor cells (MDSC) in high-risk neuroblastoma	Pmid: 27471639
NVP-TAE684	Investigational	ALK inhibitor	Lung cancer and others	Human/mouse *in vitro*: The transforming potential of the putative gain-of-function ALK mutations as well as their signaling potential and the ability of two ATP-competitive inhibitors, crizotinib (PF-02341,066) and NVP-TAE684, to abrogate the activity of ALK for neuroblastoma patients	Pmid: 21838707

(Continued on following page)

TABLE 3 | (*Continued*) Summary information of Repositioning candidates for neuroblastoma patients in refined high-risk groups.

Repositioning candidates	Clinical phases	Mode of actions (MoAs)	Approved/Investigated therapeutic categories	Evidence	Confirmation sources
torin-2	Investigational	Selective mTOR inhibitor	Various of cancer types	Torin-2 with potency against both mTOR and PI3K was more effective in promoting cytotoxicity when combined with crizotinib. Our findings should enable a more precise selection of molecularly targeted agents for patients with ALK-mutated neuroblastoma	Pmid: 25228590

('Repositioning candidate name' AND 'neuroblastoma') to convey the finding here. Consequently, we further found 21 drugs have literature citations to support their potential use for NB patients. Collectively, we found 50% of candidates (24 (3 + 21) out of 48 repositioning candidates) to either have on-going clinical trial or literature supports for NB treatment. In total, 45 out of 48 repurposing candidates (93.8%) have been reported with antitumor potency.

DISCUSSION

One of the critical aspects of precision medicine is to translate the novel genetic findings into therapy development. Structural variants (SVs) such as gene fusions have been identified in various tumor types, indicating some potential discriminative power to improve the patient survival curves. However, it was still elusive how to translate these novel genetic elements into therapy development. Several challenges become significant hurdles toward the practical application of novel genetic findings in a clinical setting. First, the inconsistency of gene fusions among the gene fusion detection algorithms still exists. Technically, several comprehensive comparison studies have been reported and suggested the consensus approach as a solution to improve reproducibility. However, the inconsistency among algorithms may be explained by biological relevance, which has not thoroughly investigated. Second, the discrimination power of gene fusion has only been assessed individually. The patient gene fusion profiles may provide a more robust solution to redefine the patient groups toward better survival rates. Finally, it is still a gap to utilize novel genetic findings into therapy development, although the possibility was always discussed elsewhere (Cully, 2015). We therefore aimed to explore the opportunity to apply patients' gene fusion profiles to stratify the high-risk neuroblastoma patients into subgroups for improving the survival rate and implementing the precision medicine-based drug repositioning.

Overall, three popular gene fusion detection algorithms were comprehensively assessed with RNA-seq data of 498 NB patients. The technical reproducibility among the three algorithms was suboptimal, which reflected the interior difference of mathematical proof behind the algorithms. Moreover, the sample heterogenicity is still a significant factor for fusion detection algorithm selection. The number of gene fusions was correlated with clinical stages from low-risk to high-risk. The sensitivity of fusion detection algorithms was increased from TopHat-Fusion to SOAPfuse and ChimeraScan. Notably, the biological relevance of detected gene fusions from the three algorithms shared substantial similarity regarding regulating pathways and DEGs.

More importantly, the gene fusion profiles derived from each algorithm have a discrimination power to redefine patient subgroups in the high-risk group with improved survival rates with Log-Rank p values less than 0.05. Furthermore, highly overlapped repositioning candidates (37 out of 48 candidates) could be enriched based on DEGs from different algorithms, and 50% of repositioning candidates (24 out of 48 repositioning candidates) could be verified by on-going clinical studies and literature reports.

The detected gene fusions were annotated by the current knowledge (**Table 1**). For example, the *Axon guidance* pathway was enriched by genes involved in reported cancer-related gene fusions based on ChimeraScan and SOAPfuse. Axon guidance plays a central role in controlling neuronal migration and neuronal survival (Chédotal et al., 2005). The expression change of proteins such as slits, semaphorins, and netrins involved in the *Axon guidance* pathways induce the pathological changes in neural circuits which predisposed to neurological disorder in adult and NB in children (Van Battum et al., 2015). Another interesting finding here is some kinase-related fusions could perturb immune-related pathways such as the T cell receptor signaling pathway. It may help investigate the immune cell types regulated by gene fusion profiles in the redefined patient subgroups (**Figures 5**, **6**). The abnormalities of immune systems in children with NB have been observed. However, the underlying mechanism is not fully understood. We found the onset age of NB patients in the redefined patient subgroup was significantly smaller than the others. Furthermore, some specific immune cell types could be enriched by all the three algorithms, which may provide more biological hints for better understanding the interplay between immune systems and pathogenesis of NB.

Several repurposing candidates hold a promise for further investigation on clinical usage for treating NB. Over 93% of enriched repositioning candidates were designed for anticancer purposes. For example, selumetinib and vemurafenib are in A

Pediatric MATCH Treatment Trial (Phase 2) led by NIH to evaluate their efficacy on treating with solid tumors (e.g., NB), non-Hodgkin lymphoma, or histiocytic disorders with MAPK pathway activation mutations (clinicaltrial.gov IDs: NCT03213691 and NCT03220035). Trametinib, a MEK inhibitor, is in clinical phase 1 for its potential to treat refractory solid tumors, including NB (NCT02780128).

Some literature reports also show some preclinical evidence of repositioning candidates for NB treatments. One example is MK-2206, an allosteric AKT inhibitor is in clinical Phase 2 designed for colorectal cancer, breast cancer, and other solid tumors. Preclinical human *in vitro* studies suggested a combination of an allosteric Akt Inhibitor MK-2206 with etoposide or rapamycin to enhances the antitumor growth effect in high-risk NB patients (Li et al., 2012). Another example is AZD8055, an mTOR inhibitor, which was designed for treating recurrent gliomas in clinical phase 2. It was reported that AZD8055 could induce cell cycle arrest, autophagy, and apoptosis and had strong antitumor activity on NB in both human *in vitro* and mouse *in vivo* models. It may be worth further investigating for clinical application for high-risk NB treatment (Xu et al., 2018).

We also enriched one repositioning candidate named SN-38, which was initially approved for small cell lung cancer. SN-38 is a topoisomerase I inhibitor and the active metabolite of irinotecan. It has a solubility issue that makes it hard for patient administration. However, the efficacy was much better than anti-neuroblastoma drug irinotecan. A novel nano carrier-based strategy for tumor-targeted delivery of a prodrug of SN-38 was developed and verified in mouse xenografts, which solve the poor blood-brain barrier (BBB) concentration of SN-38 for neuroblastoma treatment (Iyer et al., 2015).

More than 20 different gene fusion detection tools based on RNA-seq data have been developed (Kumar et al., 2016; Haas et al., 2019). The performances of fusion detection tools on sensitivity, specificity, and required computational resource (e.g., memory size and computational time) varies among different datasets. It was suggested the critical influence of calling performance also highly relies on the RNA-Seq read length, read number, and the quality of the reads. The optimization of parameters in each pipeline based on the properties of their RNA-Seq datasets along statistical methods may improve the calling performance (Dehghannasiri et al., 2019). Another comparative study also suggested the consensus calling results from different detection tools may decrease the false-positive rates. The standardization of calling results from various detection tools is the key to avoiding the specific tools and establishing a consensus calling result (Liu et al., 2016). However, a ground truth set verified from orthogonal technologies was needed to assess the calling performances from different detection tools objectively. As a proof-of-concept paper, the points described above were out of the scope of the current study. However, we highly recommend further assessing the reliable fusion detection results from multiple detection tools to establish the reproducible and high-confidence fusion calling to enable a real-world application. It is worthwhile to consider some additional work to further enhance and confirm the findings in our studies. First, the current studies were focused on RNA-seq data. As our known, DNA-seq has been well-established for SVs detection. Further investigation on DNA-seq or combined with RNA-seq data could improve the precision of gene fusion detection and establish satisfactory technical reproducibility (Liu et al., 2019). The DNA-seq provides the unbiased characterization and most comprehensive of the potential gene fusions and tumor suppressor genes disrupted by genomic rearrangement. However, it requires in-depth coverage, ample storage, and long computational analysis time. The RNA-seq only sequences the genome regions that are transcribed and spliced into mature mRNA. Therefore, only relatively high expressed fusions can be detected. However, RNA-seq data requires less storage, space, and analysis time. Furthermore, the read length for RNA-seq data can be either short or long with different sequencing platforms (Schröder et al., 2019). Second, as a proof-of-concept study, we explored the discrimination power of gene fusion profiles and their potential for patient stratification and treatment development. However, an integrative approach by incorporating various genetic elements beyond mRNA could yield a better survival curve and more precise treatment development. Third, to reuse the drugs for pediatric tumor treatment is a challenging task. In this study, the enriched repositioning candidates are mainly anticancer candidates, which could serve as a good starting point. The further PK/PB property optimization, dosage adjustment, and safety profile prioritization should be taken into consideration for further investigation (Liu Z et al., 2016; Liu et al., 2017). Fourth, the unbalance number of collected neuroblastoma patient samples across different clinical predefined INSS cancer stages might influence the performance of gene-fusion-based patient stratification. The more consortium efforts with more balanced patients distribution in different cancer stages may further verify and improve the performance. Finally, in the current study, we only focused on three popular fusion detection algorithms to elaborate the points and establish the framework for translating genetic findings into therapy development. Other fusion detection algorithms and patient stratification strategies may also be considered to improve the results. In conclusion, we carried out an exploratory study to investigate how to apply genetic findings such as gene fusions for clinical applications. The framework developed it straightforward and may also serve as a strategy for treatment development for other diseases.

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 authors.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: ZL. Analyzed the data: ZL. The first version of the manuscript: ZL. Contributed reagents/

materials/analysis tools: ZL, RH, MM, and XC. Wrote the manuscript: ZL, RR, and WT. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

XC is grateful to the National Center for Toxicological Research (NCTR) of the U.S. Food and Drug Administration (FDA) for postdoctoral support through the Oak Ridge Institute for Science and Education (ORISE).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.608778/full#supplementary-material.

REFERENCES

Almstedt, E., Elgendy, R., Hekmati, N., Rosén, E., Wärn, C., Olsen, T. K., et al. (2020). Integrative discovery of treatments for high-risk neuroblastoma. *Nat. Commun.* 11 (1), 71. doi:10.1038/s41467-019-13817-8

Beccuti, M., Carrara, M., Cordero, F., Lazzarato, F., Donatelli, S., Nadalin, F., et al. (2014). Chimera: a Bioconductor package for secondary analysis of fusion products. *Bioinformatics* 30 (24), 3556–3557. doi:10.1093/bioinformatics/btu662

Boeva, V., Louis-Brennetot, C., Peltier, A., Durand, S., Pierre-Eugène, C., Raynal, V., et al. (2017). Heterogeneity of neuroblastoma cell identity defined by transcriptional circuitries. *Nat. Genet.* 49, 1408–1413. doi:10.1038/ng.3921

Bueno, R., Stawiski, E. W., Goldstein, L. D., Durinck, S., De Rienzo, A., Modrusan, Z., et al. (2016). Comprehensive genomic analysis of malignant pleural mesothelioma identifies recurrent mutations, gene fusions and splicing alterations. *Nat. Genet.* 48 (4), 407–416. doi:10.1038/ng.3520

Chédotal, A., Kerjan, G., and Moreau-Fauvarque, C. (2005). The brain within the tumor: new roles for axon guidance molecules in cancers. *Cell Death Differ.* 12, 1044. doi:10.1038/sj.cdd.4401707

Chen, T., and Guestrin, C. (2016). "Xgboost: a scalable tree boosting system," in Proceedings of the 22nd acm sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco California, USA. August, 2016, 785–794.

Corsello, S. M., Bittker, J. A., Liu, Z., Gould, J., McCarren, P., Hirschman, J. E., et al. (2017). The Drug Repurposing Hub: a next-generation drug library and information resource. *Nat. Med.* 23 (4), 405–408. doi:10.1038/nm.4306

Coughlin, C. M., Fleming, M. D., Carroll, R. G., Pawel, B. R., Hogarty, M. D., Shan, X., et al. (2006). Immunosurveillance and survivin-specific T-cell immunity in children with high-risk neuroblastoma. *Jco* 24 (36), 5725–5734. doi:10.1200/jco.2005.05.3314

Cully, M. (2015). Advancing precision medicine in silico. *Nat. Rev. Drug Discov.* 14, 311. doi:10.1038/nrd4619

Davidson, N. M., Majewski, I. J., and Oshlack, A. (2015). JAFFA: high sensitivity transcriptome-focused fusion gene detection. *Genome Med.* 7 (1), 43. doi:10.1186/s13073-015-0167-x

Dehghannasiri, R., Freeman, D. E., Jordanski, M., Hsieh, G. L., Damljanovic, A., Lehnert, E., et al. (2019). Improved detection of gene fusions by applying statistical methods reveals oncogenic RNA cancer drivers. *Proc. Natl. Acad. Sci. U.S.A.* 116 (31), 15524–15533. doi:10.1073/pnas.1900391116

Druker, B. J. (2008). Translation of the Philadelphia chromosome into therapy for CML. *Blood* 112 (13), 4808–4817. doi:10.1182/blood-2008-07-077958

Duan, Q., Reid, S. P., Clark, N. R., Wang, Z., Fernandez, N. F., Rouillard, A. D., et al. (2016). L1000CDS2: LINCS L1000 characteristic direction signatures search engine. *Npj Syst. Biol. Appl.* 2, 16015. doi:10.1038/npjsba.2016.15

Fletcher, J. I., Ziegler, D. S., Trahair, T. N., Marshall, G. M., Haber, M., and Norris, M. D. (2018). Too many targets, not enough patients: rethinking neuroblastoma clinical trials. *Nat. Rev. Cancer* 18 (6), 389–400. doi:10.1038/s41568-018-0003-x

Forbes, S. A., Bhamra, G., Bamford, S., Dawson, E., Kok, C., Clements, J., et al. (2008). The Catalogue of somatic mutations in cancer (COSMIC). *Curr. Protoc. Hum. Genet.* Chapter 10, Unit 10.11. doi:10.1002/0471142905.hg1011s57

Gao, Q., Liang, W. W., Foltz, S. M., Mutharasu, G., Jayasinghe, R. G., Cao, S., et al. (2018). Driver fusions and their implications in the development and treatment of human cancers. *Cell Rep.* 23 (1), 227–238.e3. doi:10.1016/j.celrep.2018.03.050

Haas, B. J., Dobin, A., Li, B., Stransky, N., Pochet, N., and Regev, A. (2019). Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol.* 20 (1), 213. doi:10.1186/s13059-019-1842-9

Heng, T. S. P., and Painter, M. W.(2008). The immunological genome ProjectThe immunological genome project: networks of gene expression in immune cells. *Nat. Immunol.* 9, 1091–1094. doi:10.1038/ni1008-1091

Heyer, E. E., Deveson, I. W., Wooi, D., Selinger, C. I., Lyons, R. J., Hayes, V. M., et al. (2019). Diagnosis of fusion genes using targeted RNA sequencing. *Nat. Commun.* 10 (1), 1388. doi:10.1038/s41467-019-09374-9

Hishiki, T., Mise, N., Harada, K., Ihara, F., Takami, M., Saito, T., et al. (2018). Invariant natural killer T infiltration in neuroblastoma with favorable outcome. *Pediatr. Surg. Int.* 34 (2), 195–201. doi:10.1007/s00383-017-4189-x

Hu, X., Wang, Q., Tang, M., Barthel, F., Amin, S., Yoshihara, K., et al. (2018). TumorFusions: an integrative resource for cancer-associated transcript fusions. *Nucleic Acids Res.* 46 (D1), D1144–D1149. doi:10.1093/nar/gkx1018

Huang, M., and Weiss, W. A. (2013). Neuroblastoma and MYCN. *Cold Spring Harbor Perspect. Med.* 3 (10), a014415. doi:10.1101/cshperspect.a014415

Iyer, M. K., Chinnaiyan, A. M., and Maher, C. A. (2011). ChimeraScan: a tool for identifying chimeric transcription in sequencing data. *Bioinformatics* 27 (20), 2903–2904. doi:10.1093/bioinformatics/btr467

Iyer, R., Croucher, J. L., Chorny, M., Mangino, J. L., Alferiev, I. S., Levy, R. J., et al. (2015). Nanoparticle delivery of an SN38 conjugate is more effective than irinotecan in a mouse model of neuroblastoma. *Cancer Lett.* 360 (2), 205–212. doi:10.1016/j.canlet.2015.02.011

Ji, X., Tong, W., Liu, Z., and Shi, T. (2019). Five-feature model for developing the classifier for synergistic vs. Antagonistic drug combinations built by XGBoost. *Front. Genet.* 10, 600. doi:10.3389/fgene.2019.00600

Jia, W., Qiu, K., He, M., Song, P., Zhou, Q., Zhou, F., et al. (2013). SOAPfuse: an algorithm for identifying fusion transcripts from paired-end RNA-Seq data. *Genome Biol.* 14 (2), R12. doi:10.1186/gb-2013-14-2-r12

Kidd, B. A., Wroblewska, A., Boland, M. R., Agudo, J., Merad, M., Tatonetti, N. P., et al. (2016). Mapping the effects of drugs on the immune system. *Nat. Biotechnol.* 34, 47–54. doi:10.1038/nbt.3367

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14 (4), R36. doi:10.1186/gb-2013-14-4-r36

Kim, D., and Salzberg, S. L. (2011). TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. *Genome Biol.* 12 (8), R72. doi:10.1186/gb-2011-12-8-r72

Kumar, S., Vo, A. D., Qin, F., and Li, H. (2016). Comparative assessment of methods for the fusion transcripts detection from RNA-Seq data. *Sci. Rep.* 6, 21597. doi:10.1038/srep21597

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi:10.1038/nmeth.1923

Lee, M., Lee, K., Yu, N., Jang, I., Choi, I., Kim, P., et al. (2017). ChimerDB 3.0: an enhanced database for fusion genes from cancer transcriptome and literature data mining. *Nucleic Acids Res.* 45, D784–D789. doi:10.1093/nar/gkw1083

Li, R., Yu, C., Li, Y., Lam, T.-W., Yiu, S.-M., Kristiansen, K., et al. (2009). SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25 (15), 1966–1967. doi:10.1093/bioinformatics/btp336

Li, Z., Yan, S., Attayan, N., Ramalingam, S., and Thiele, C. J. (2012). Combination of an allosteric Akt inhibitor MK-2206 with etoposide or rapamycin enhances the antitumor growth effect in neuroblastoma. *Clin. Cancer Res.* 18 (13), 3603–3615. doi:10.1158/1078-0432.ccr-11-3321

Liu, S., Tsai, W.-H., Ding, Y., Chen, R., Fang, Z., Huo, Z., et al. (2016a). Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. *Nucleic Acids Res.* 44 (5), e47. doi:10.1093/nar/gkv1234

Liu, Z., Delavan, B., Roberts, R., and Tong, W. (2017). Lessons learned from two decades of anticancer drugs. *Trends Pharmacol. Sci.* 38 (10), 852–872. doi:10.1016/j.tips.2017.06.005

Liu, Z., Fang, H., Slikker, W., and Tong, W. (2016b). Potential reuse of oncology drugs in the treatment of rare diseases. *Trends Pharmacol. Sci.* 37 (10), 843–857. doi:10.1016/j.tips.2016.06.010

Liu, Z., Zhu, L., Roberts, R., and Tong, W. (2019). Toward clinical implementation of next-generation sequencing-based genetic testing in rare diseases: where are we? *Trends Genet.* 35 (11), 852–867. doi:10.1016/j.tig.2019.08.006

Ma, X., Liu, Y., Liu, Y., Alexandrov, L. B., Edmonson, M. N., Gawad, C., et al. (2018). Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature* 555 (7696), 371-376. doi:10.1038/nature25795

Matthay, K. K., Maris, J. M., Schleiermacher, G., Nakagawara, A., Mackall, C. L., Diller, L., et al. (2016). Neuroblastoma. *Nat. Rev. Dis. Primers* 2, 16078. doi:10.1038/nrdp.2016.78

Mertens, F., Johansson, B., Fioretos, T., and Mitelman, F. (2015). The emerging complexity of gene fusions in cancer. *Nat. Rev. Cancer* 15, 371–381. doi:10.1038/nrc3947

Moroz, V., Machin, D., Faldum, A., Hero, B., Iehara, T., Mosseri, V., et al. (2011). Changes over three decades in outcome and the prognostic influence of age-at-diagnosis in young patients with neuroblastoma: a report from the International Neuroblastoma Risk Group Project. *Eur. J. Cancer* 47 (4), 561–571. doi:10.1016/j.ejca.2010.10.022

Mossé, Y. P., Laudenslager, M., Longo, L., Cole, K. A., Wood, A., Attiyeh, E. F., et al. (2008). Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature* 455, 930–935. doi:10.1038/nature07261

Oberthuer, A., Berthold, F., Warnat, P., Hero, B., Kahlert, Y., Spitz, R., et al. (2006). Customized oligonucleotide microarray gene expression-based classification of neuroblastoma patients outperforms current clinical risk stratification. *J. Clin. Oncol.* 24 (31), 5070–5078. doi:10.1200/jco.2006.06.1879

Olsson, M., Beck, S., Kogner, P., Martinsson, T., and Carén, H. (2016). Genome-wide methylation profiling identifies novel methylated genes in neuroblastoma tumors. *Epigenetics* 11 (1), 74–84. doi:10.1080/15592294.2016.1138195

Orentas, R. J., Yang, J. J., Wen, X., Wei, J. S., Mackall, C. L., and Khan, J. (2012). Identification of cell surface proteins as potential immunotherapy targets in 12 pediatric cancers. *Front. Oncol.* 2, 194. doi:10.3389/fonc.2012.00194

Painter, M. W., Davis, S., Hardy, R. R., Mathis, D., and Benoist, C. (2011). Transcriptomes of the B and T lineages compared by multiplatform microarray profiling. *J. Immunol.* 186 (5), 3047–3057. doi:10.4049/jimmunol.1002695

Peifer, M., Hertwig, F., Roels, F., Dreidax, D., Gartlgruber, M., Menon, R., et al. (2015). Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature* 526, 700–714. doi:10.1038/nature14980

Picco, G., Chen, E. D., Alonso, L. G., Behan, F. M., Gonçalves, E., Bignell, G., et al. (2019). Functional linkage of gene fusions to cancer cell fitness assessed by pharmacological and CRISPR-Cas9 screening. *Nat. Commun.* 10 (1), 2198. doi:10.1038/s41467-019-09940-1

Pugh, T. J., Morozova, O., Attiyeh, E. F., Asgharzadeh, S., Wei, J. S., Auclair, D., et al. (2013). The genetic landscape of high-risk neuroblastoma. *Nat. Genet.* 45, 279–284. doi:10.1038/ng.2529

Robinson, M. D., and Smyth, G. K. (2008). Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* 9 (2), 321–332. doi:10.1093/biostatistics/kxm030

Rorie, C. J., Thomas, V. D., Chen, P., Pierce, H. H., O'Bryan, J. P., and Weissman, B. E. (2004). The ews/fli-1 fusion gene switches the differentiation program of neuroblastomas to ewing sarcoma/peripheral primitive neuroectodermal tumors. *Cancer Res.* 64 (4), 1266–1277. doi:10.1158/0008-5472.can-03-3274

Schröder, J., Kumar, A., and Wong, S. Q. (2019). "Overview of fusion detection strategies using next-generation sequencing," in *Tumor profiling: methods and protocols*. Editor S. S. Murray (New York, NY: Humana Press), 125–138. doi:10.1007/978-1-4939-9004-7_9

Semeraro, M., Rusakiewicz, S., Minard-Colin, V., Delahaye, N. F., Enot, D., Vély, F., et al. (2015). Clinical impact of the NKp30/B7-H6 axis in high-risk neuroblastoma patients. *Sci. Transl. Med.* 7 (283), 283ra55. doi:10.1126/scitranslmed.aaa2327

Szanto, C. L., Cornel, A. M., Vijver, S. V., and Nierkens, S. (2020). Monitoring immune responses in neuroblastoma patients during therapy. *Cancers* 12 (2), 519. doi:10.3390/cancers12020519

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi:10.1038/nbt.1621

Van Battum, E. Y., Brignani, S., and Pasterkamp, R. J. (2015). Axon guidance proteins in neurological disorders. *Lancet Neurol.* 14 (5), 532–546. doi:10.1016/S1474-4422(14)70257-1

Wang, K., Diskin, S. J., Zhang, H., Attiyeh, E. F., Winter, C., Hou, C., et al. (2011). Integrative genomics identifies LMO1 as a neuroblastoma oncogene. *Nature* 469 (7329), 216–220. doi:10.1038/nature09609

Wang, Q., Xia, J., Jia, P., Pao, W., and Zhao, Z. (2013). Application of next generation sequencing to human gene fusion detection: computational tools, features and perspectives. *Brief. Bioinform.* 14 (4), 506–519. doi:10.1093/bib/bbs044

Wienke, J., Dierselhuis, M. P., Tytgat, G. A. M., Künkele, A., Nierkens, S., and Molenaar, J. J. (2021). The immune landscape of neuroblastoma: challenges and opportunities for novel therapeutic strategies in pediatric oncology. *Eur. J. Cancer* 144, 123–150. doi:10.1016/j.ejca.2020.11.014

Xu, D.-Q., Toyoda, H., Yuan, X.-J., Qi, L., Chelakkot, V. S., Morimoto, M., et al. (2018a). Anti-tumor effect of AZD8055 against neuroblastoma cells *in vitro* and *in vivo*. *Exp. Cel Res.* 365 (2), 177–184. doi:10.1016/j.yexcr.2018.02.032

Xu, T., Wang, H., Huang, X., Li, W., Huang, Q., Yan, Y., et al. (2018b). Gene fusion in malignant glioma: an emerging target for next-generation personalized treatment. *Translational Oncol.* 11 (3), 609–618. doi:10.1016/j.tranon.2018.02.020

Zeid, R., Lawlor, M. A., Poon, E., Reyes, J. M., Fulciniti, M., Lopez, M. A., et al. (2018). Enhancer invasion shapes MYCN-dependent transcriptional amplification in neuroblastoma. *Nat. Genet.* 50 (4), 515–523. doi:10.1038/s41588-018-0044-9

Zhang, J., White, N. M., Schmidt, H. K., Fulton, R. S., Tomlinson, C., Warren, W. C., et al. (2016). INTEGRATE: gene fusion discovery using whole genome and transcriptome data. *Genome Res.* 26 (1), 108–118. doi:10.1101/gr.186114.114

Zhang, W., Yu, Y., Hertwig, F., Thierry-Mieg, J., Zhang, W., Thierry-Mieg, D., et al. (2015). Comparison of RNA-seq and microarray-based models for clinical endpoint prediction. *Genome Biol.* 16 (1), 133. doi:10.1186/s13059-015-0694-1

Disclaimer: The views presented in this article do not necessarily reflect current or future opinion or policy of the US Food and Drug Administration and National Institute of Health. Any mention of commercial products is for clarification and not intended as an endorsement.

Conflict of Interest: RR is co-founder and co-director of ApconiX, an integrated toxicology and ion channel company that provides expert advice on non-clinical aspects of drug discovery and drug development to academia, industry and not-for-profit organizations.

frontiers in Genetics

ORIGINAL RESEARCH
published: 12 May 2021
doi: 10.3389/fgene.2021.638749

Identification of a Novel Variant in *MT-CO3* Causing MELAS

*Manting Xu[1,2,3], Robert Kopajtich[2,3], Matthias Elstner[4], Zhaoxia Wang[5], Zhimei Liu[1], Junling Wang[1], Holger Prokisch[2,3] and Fang Fang[1]**

[1] Department of Neurology, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China, [2] Institute of Human Genetics, School of Medicine, Technical University of Munich, Munich, Germany, [3] Institute of Neurogenomics, Helmholtz Zentrum München, Munich, Germany, [4] Department of Neurology, School of Medicine, Technical University of Munich, Munich, Germany, [5] Department of Neurology, Peking University First Hospital, Beijing, China

Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS) is a maternally inherited mitochondrial disease. Most cases of MELAS are caused by the m.3243A > G variant in the *MT-TL1* gene encoding tRNALeu$^{(UUR)}$. However, the genetic cause in 10% of patients with MELAS is unknown. We investigated the pathogenicity of the novel mtDNA variant m.9396G > A/*MT-CO3* (p.E64K), which affects an extremely conserved amino acid in the CO3 subunit of mitochondrial respiratory chain (MRC) complex IV (CIV) in a patient with MELAS. Biochemical assays of a muscle biopsy confirmed remarkable CIV deficiency, and pathological examination showed ragged red fibers and generalized COX non-reactive muscle fibers. Transfer of the mutant mtDNA into cybrids impaired CIV assembly, followed by remarkable mitochondrial dysfunction and ROS production. Our findings highlight the pathogenicity of a novel m.9396G > A variant and extend the spectrum of pathogenic mtDNA variants.

Keywords: MELAS, novel mitochondrial DNA variant, *MT-CO3* gene, complex IV of respiratory chain, mitochondrial diseases

OPEN ACCESS

Edited by:
Ruth Roberts, ApconiX, United Kingdom

Reviewed by:
Massimo Zeviani, University of Padua, Italy
Kei Murayama, Chiba Children's Hospital, Japan
Claudia Nesti, University of Pisa, Italy

****Correspondence:***
Fang Fang
fangfang@bch.com.cn

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

Received: *11 December 2020*
Accepted: *09 February 2021*
Published: *12 May 2021*

Citation:
Xu M, Kopajtich R, Elstner M, Wang Z, Liu Z, Wang J, Prokisch H and Fang F (2021) Identification of a Novel Variant in MT-CO3 Causing MELAS. Front. Genet. 12:638749. doi: 10.3389/fgene.2021.638749

INTRODUCTION

Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS) (OMIM 540000) is one of the most common maternally inherited mitochondrial disorders. Its clinical manifestations include stroke-like episodes, dementia, epilepsy, myopathy, recurrent headaches, hearing impairment, diabetes, and short stature (El-Hattab et al., 2015). The diagnostic criteria for MELAS include the presence of these typical manifestations and two of the following three indications of mitochondrial dysfunction: (i) a biochemical defect in mitochondrial energy metabolism, (ii) mitochondrial abnormalities determined by muscle biopsy, and (iii) a pathogenic variant associated with MELAS (Hirano et al., 1992; Yatsuga et al., 2012; Lorenzoni et al., 2015).

Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes is a genetically heterogeneous mitochondrial disorder (El-Hattab et al., 2015). The m.3243A > G variant in the *MT-TL1* gene encoding tRNALeu$^{(UUR)}$ is found in approximately 80% of patients with MELAS. Other variants, such as m.3271T > C and m.3252A > G in *MT-TL1*, are also associated with MELAS. Moreover, numerous mitochondrial DNA (mtDNA) encoded genes have other rare variants that can cause MELAS; these include mtDNA encoded tRNAs (*MT-TS1, MT-TS2, MT-TW, MT-TC, MT-TL2, MT-TK, MT-TH, MT-TQ, MT-TF*, and *MT-TV*) and mitochondrial respiratory

chain (MRC) complex I (*MT-ND1, MT-ND5*, and *MT-ND6*), complex III (*MT-CYB*), and complex IV (CIV) subunits (*MT-CO2* and *MT-CO3*). In addition, variants in nuclear genes, such as *POLG* and *FASTKD2*, have been associated with a MELAS-like phenotype (Cheldi et al., 2013). However, the molecular pathomechanisms of these variants are highly complex and far from well-documented, and in 10% of patients with MELAS the genetic cause is unknown (Cheldi et al., 2013; El-Hattab et al., 2015). Therefore, it is important to identify novel pathogenic variants and elucidate how they directly affect the function of mitochondria.

Here, we report a patient with MELAS carrying a novel m.9396G > A variant in *MT-CO3*. This is the fourth reported MELAS case with a variant of the genes encoding the CIV subunit. We confirmed the pathogenicity of this variant based on clinical characteristics, muscle pathology results, biochemical analyses, and transmitochondrial cytoplasmic hybrid (cybrid) cell studies.

SUBJECTS AND METHODS

Clinical Description

The young female patient was the second child of Chinese non-consanguineous parents, born at full-term after a normal pregnancy (birth weight 3.5 kg, height 50 cm). Her early neurodevelopment was entirely normal. However, she often complained of fatigue and had short stature from about 3 years of age.

At 8 years of age, she developed a hearing impairment after a respiratory infection (never use aminoglycoside antibiotics) and was given a cochlear implant. After another infection at 9 years of age, she experienced a stroke-like episode that initially presented with frequent convulsions and headache, followed by visual impairment, vomiting, and gradual loss of consciousness. Cerebrospinal fluid lactate (8 mmol/L, normal range < 2.8 mmol/L) and serum lactate (10.7 mmol/L, normal range < 2.8 mmol/L) were elevated. All other laboratory results were normal, including blood cell counts, creatine kinase, serum glucose, ammonia, liver and kidney function tests, electrocardiogram, and echocardiography. Computed tomography (CT) showed large hypodense lesions in the right occipital, parietal, and temporal lobes, suggestive of cerebral edema and infarction. The midline of the occipital lobe was shifted slightly to the left, indicating brain herniation. Multiple calcifications in the bilateral basal ganglia were also evident on CT (**Figure 1A**). An electroencephalogram showed abnormally slow background activity in the occipital region and epileptiform discharges in both frontal and temporal regions. Perimetry presented left side hemianopsia (**Figure 1B**). Based on the clinical course, MELAS was suspected, and L-arginine, levetiracetam, and antioxidants were prescribed. After 23 days, she regained consciousness with no headache, convulsions, or visual impairment and was discharged from hospital.

One and a half months later, she experienced a similar stroke-like episode after fatigue, initially with a headache, vomiting, and left side hemianopsia, which progressed to unconsciousness. CT showed enlarged cerebral infarction lesions in the right occipital, parietal and temporal lobes, and right cerebellar hemisphere. More severe cerebral edema and cerebral herniation were also seen (**Figure 1C**). After treatment for several days, she improved, but the hemianopsia and hemiparesis remained.

The patient is now 12 years of age and undergoes continued treatment with levetiracetam and antioxidants. She still complains of intermittent headache and suffers from hemiparesis, dementia, poor emotional control, deafness, and hemianopsia. Repeat CT shows cerebral atrophy and multiple calcifications (**Figure 1D**). Her parents (both over 50 years of age) and 18-year-old brother are asymptomatic.

Histopathological, Histochemical, and Biochemical Analyses of Muscle Biopsy

The right bicep of the 11-year-old patient was biopsied, and fresh specimens were prepared for standard histopathology staining with hematoxylin and eosin (HE), modified Gomori trichrome (MGT), and Oil Red O, as well as histochemical staining with succinate dehydrogenase (SDH), and COX-SDH reactions analyses (Parikh et al., 2015). The activities of the MRC complexes and the mitochondrial marker enzyme citrate synthase (CS) were evaluated in muscle homogenates as described previously (Ogawa et al., 2017).

Molecular Genetic Analyses

Total DNA was extracted from peripheral blood, and mtDNA from skeletal muscle, skin fibroblasts, blood, urine, saliva, and fingernail samples, as described previously (Ji et al., 2014). Mitochondrial genome next-generation sequencing and whole-exome sequencing (WES) were performed to detect causative genes using an Illumina HiSeq X Ten sequencer (Illumina, San Diego, CA, United States). Sanger sequencing was conducted using 2 × Taq Master Mix (Thermo Fisher Scientific, Waltham, MA, United States) and Applied Biosystems thermal cyclers (Applied Biosystems, Foster City, CA, United States). Mutation loads were determined by quantitative pyrosequencing; quantification of the heteroplasmy level of each variant was achieved using PyroMark Q24 software (Grady et al., 2018).

The PolyPhen2 and PROVEAN were used to evaluate the pathogenicity of the mtDNA variants (Adzhubei et al., 2013; Choi and Chan, 2015). PyMOL software was used to visualize the protein structure and model amino acid changes in the three-dimensional structure.

Generation of Transmitochondrial Cybrids

To generate transmitochondrial cybrids carrying the m.9396G > A variant, platelets were isolated from blood samples and fused with mtDNA-less human osteosarcoma 143B cells as described previously (Chomyn, 1996). The cybrid clones were cultured in Dulbecco's modified Eagle's medium (Thermo Fisher Scientific) supplemented with 1% penicillin and 10%

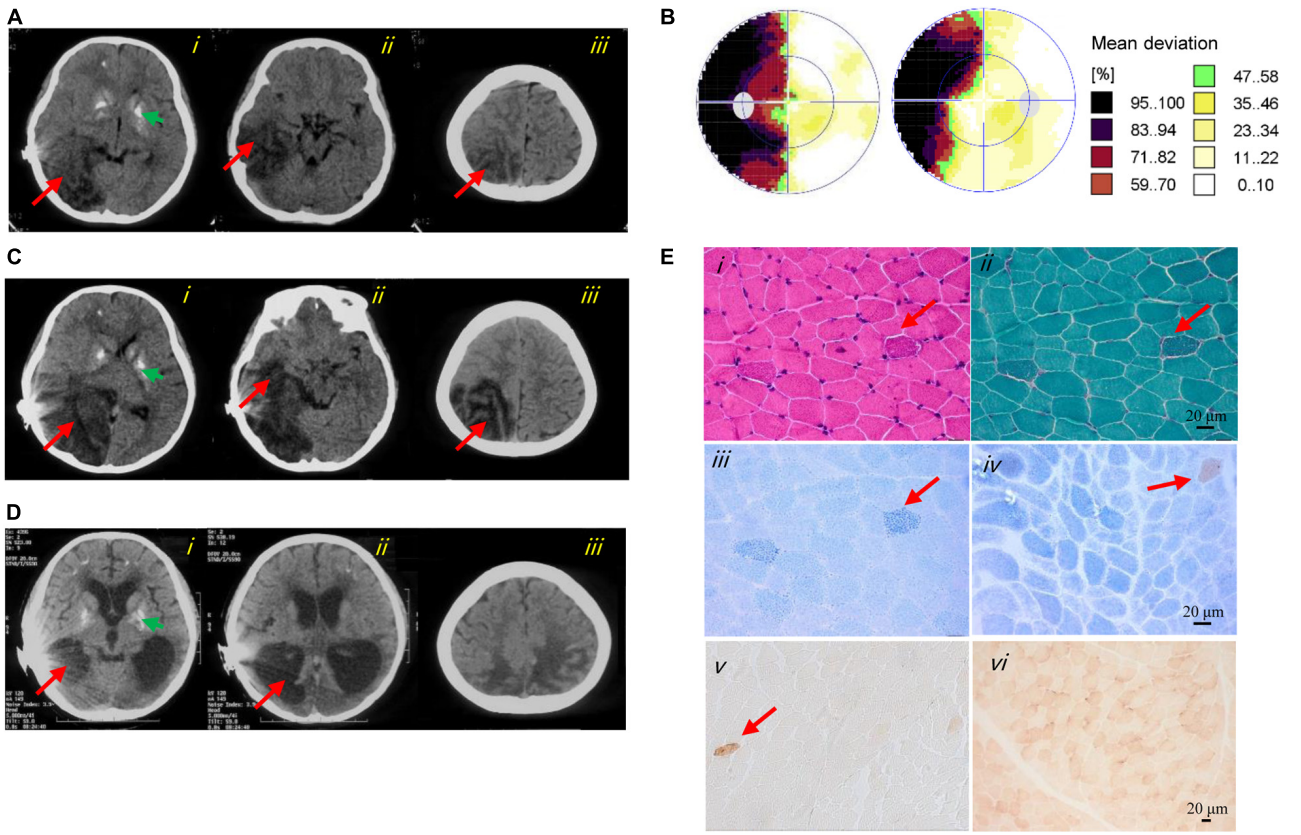

FIGURE 1 | Clinical images of the patient. **(A)** Axial CT images of the first stroke-like episode revealed cerebral edema and brain herniation in the right occipital (*i*), temporal (*ii*), and parietal (*iii*) lobes (red arrows). The green arrow indicated basal ganglia calcification (*i*). **(B)** Perimetry showed left-side hemianopsia. **(C)** Axial CT images of the second stroke-like episode revealed more severe cerebral infarction and cerebral edema in the right occipital (*i*), temporal (*ii*), and parietal (*iii*) lobes, and in the right cerebellar hemisphere (*i*) (red arrows). The green arrow indicated basal ganglia calcification (*i*). **(D)** Axial CT images at 12 years of age showed cerebral atrophy (*i–iii*), an enlarged lateral cerebroventricle (*i–ii*) and multiple calcifications (*i*). **(E)** Image *i* showed several subsarcolemmal basophilic substance (red arrow; HE staining). Image *ii* showed several RRFs (red arrow; MGT staining). Image *iii* showed ragged blue fibers (red arrow; SDH reaction). Images *iv* and *v* showed general COX-deficient fibers (COX-SDH reaction and COX reaction, respectively). The red arrow in image *iv* and *v* indicated a COX-positive fiber. Image *vi* showed a normal muscle sample from an age matched control (COX reaction).

fetal bovine serum (both Thermo Fisher Scientific) at 37°C in a 5% CO_2 atmosphere. After hybridization, we selected cybrid clones with more than 95% mutant mtDNA, 78.6% mutant mtDNA, and those lacking this variant [wild-type (WT)] for further experiments.

RNA Isolation and Quantitative Reverse-Transcription (qRT)-PCR

Total RNA was extracted using the TRIzol reagent (Invitrogen, Carlsbad, CA, United States) according to the manufacturer's procedure, and reverse-transcribed into complementary DNA (cDNA) using a cDNA Reverse Transcription Kit (Applied Biosystems). Quantitative PCR was conducted with SYBR Green Master Mix (Applied Biosystems) on the 7500 Fast Real-Time PCR System (Applied Biosystems). The primer sequences for *MT-CO3* were as follows: sense, 5′-CAC ATC CGT ATT ACT CGC ATC-3′ and antisense, 5′-GAA GTA CTC TGA GGC TTG TAG-3′; the primers for *GAPDH* (internal control) were sense, 5′-CGC TGA GTA CGT CGT GGA GTC-3′ and antisense, 5′-GCT GAT GAT CTT GAG GCT GTT GTC-3′. The data were analyzed by the $2^{-\Delta\Delta Ct}$ method.

Western Blots and Blue Native Polyacrylamide Gel Electrophoresis

Protein extraction from cybrids for sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and Western blotting was carried out as described previously (Ji et al., 2020). Briefly, cells were lysed in radioimmunoprecipitation assay (RIPA) buffer (Sigma-Aldrich, St. Louis, MO, United States). The protein extracts were resolved in a 10% SDS polyacrylamide gel and transferred to a polyvinylidene difluoride membrane. Membranes were blocked with 5% milk for 1 h and incubated with the following primary antibodies at 4°C overnight: anti-MT-CO3 (cat #ab110259; Abcam, Cambridge, United Kingdom) and VDAC (cat #4661; Cell Signaling Technology, Danvers, MA, United States);

the latter served as a loading control. Blue native (BN)-PAGE was performed with mitochondrial proteins isolated from cybrids as described previously (Delmiro et al., 2013), using antibodies against MT-COXIV (ab202554; Abcam), SDHA (ab14715; Abcam), VDAC (cat#4661; Cell Signaling Technology), UQCRC2 (ab14745; Abcam), NDUFS2 (ab110249; Abcam), and ATPB (ab14730; Abcam). The in-gel activity (IGA) of CIV was performed as described previously (Jha et al., 2016). Briefly, after running the native PAGE, incubate the gel in 20 ml of complex IV substrate. Appearance of brown bands is indicative of CIV activity.

Measurement of Oxygen Consumption Rate and MRC Complex Activities of the Cybrids

Cybrid cells were seeded at a density of 4×10^4 cells per well on Seahorse XFe24 polystyrene cell culture plates. The oxygen consumption rate (OCR) of the cybrids was assayed using the XF96 Extracellular Flux Analyzer (Agilent Technologies, Santa Clara, CA, United States), as described previously (Kremer and Prokisch, 2017; Ogawa et al., 2017). The activities of MRC complexes I, II, II + III, and IV and the mitochondrial marker enzyme CS were evaluated in cybrids as described previously (Ogawa et al., 2017).

Assays of ATP and Reactive Oxygen Species

The ATP levels of cybrids were measured using the ATP Bioluminescence Assay Kit HS II (Roche, Indianapolis, IN, United States). Briefly, cybrids were incubated in a solution [156 mM NaCl, 3 mM KCl, 2 mM $MgSO_4$, 1.25 mM KH_2PO_4, 2 mM $CaCl_2$, 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES); pH 7.35] with glucose (total ATP production) or 2-deoxy-D-glucose (2-DG) plus pyruvate (oxidative ATP production). Relative mitochondrial ATP synthesis was defined as the ratio relative to total ATP production of cybrids WT.

Reactive oxygen species (ROS) levels were measured as described elsewhere (Wang et al., 2011). Briefly, approximately 1×10^6 cybrids were harvested and suspended in phosphate-buffered saline (PBS) with 10 μM 2',7'-dichlorodihydrofluorescein diacetate. After incubation (37°C, 5% CO_2) for 20 min, the cybrids were washed and resuspended in PBS. Finally, the fluorescence signals of cybrids were determined using excitation and emission wavelengths of 488 and 529 nm, respectively.

Statistical Analysis

Statistical parameters are reported in the figure legends. *In vitro* experiments were repeated at least twice. Statistical analysis was performed using Prism 8.0 for Windows (GraphPad software, Inc., San Diego, CA, United States). Differences between mutant and WT cybrid cells were evaluated by one-way analysis of variance (ANOVA) followed by the Tukey's honestly significant difference *post hoc* test.

TABLE 1 | MRC enzyme assays of the muscle biopsy.

	CI	CII	CII + III	CIII	CIV
% of CS	78.2 ± 2.9	77.1 ± 3.6	80.1 ± 8.1	64.9 ± 7.1	16.5 ± 6.5
% of CII	98.6 ± 6.1	–	100.8 ± 9.4	82.4 ± 6.0	20.3 ± 8.1

Enzyme activities are expressed as percentages of normalized mean values relative to CS and CII. CI, complex I; CII, complex II; CIII, complex III; CIV, complex IV; CS, citrate synthase.

RESULTS

Muscle Biopsy Analyses Revealed Mitochondrial Dysfunction and a CIV Defect

Muscle histopathology showed slightly increased variation in fiber size and subsarcolemmal basophilic substance on HE staining. Moderately increased fat deposition was noted in some fibers by Oil Red O staining. Ragged red fibers (RRFs) and ragged blue fibers were detected in 2–5% of the total biopsy on MGT staining and SDH reaction, suggestive of abnormal mitochondrial accumulation. Remarkable mitochondrial abnormalities were detected, characterized by COX deficiency affecting >90% of all fibers with the COX-SDH and COX reaction, respectively. In addition, a few arterioles were identified in the COX-deficient muscle section (**Figure 1E**). The activities of CIV in the patient's muscles were reduced to 16.5 ± 6.5% and 20.3 ± 8.1% compared with those of CS and complex II, respectively (**Table 1**).

Identification of a Probably Pathological Variant in *MT-CO3*

Based on pathological and biochemical analyses, the patient was confirmed to have MELAS with CIV deficiency. Although no known pathogenic mtDNA variant was identified by sequencing mitochondrial genes, and there was no potential biallelic rare variant of any gene associated with mitochondrial disease, CIV subunits or assembly factors were detected by WES.

The homoplasmic m.9396G > A (p.E64K) variant of *MT-CO3* was detected in the patient's urinary sediment. Skeletal muscle, skin fibroblasts, saliva, fingernail and blood samples were screened for the presence of the m.9396G > A variant and exhibited heteroplasmy at this position of 94.0, 91.0, 73.5, 77.0, and 38.9%, respectively. The novel m.9396G > A variant was not detected in saliva, urinary sediment, fingernail or blood DNA samples from the patient's mother, aunt, or older brother, implying that the variant had arisen *de novo* during embryogenesis and was not maternally-inherited (**Figure 2A**). No other suspected pathogenic variant was identified in the mtDNA (**Supplementary Table S1**), as all other base substitutions represented mtDNA polymorphisms in the MITOMAP database[1].

Phylogenetic conservation analysis of the p.64 aminoacidic residue in eight species (human, rhesus, mouse, elephant, chicken, X. tropicalis, zebrafish, and lamprey) revealed that this

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

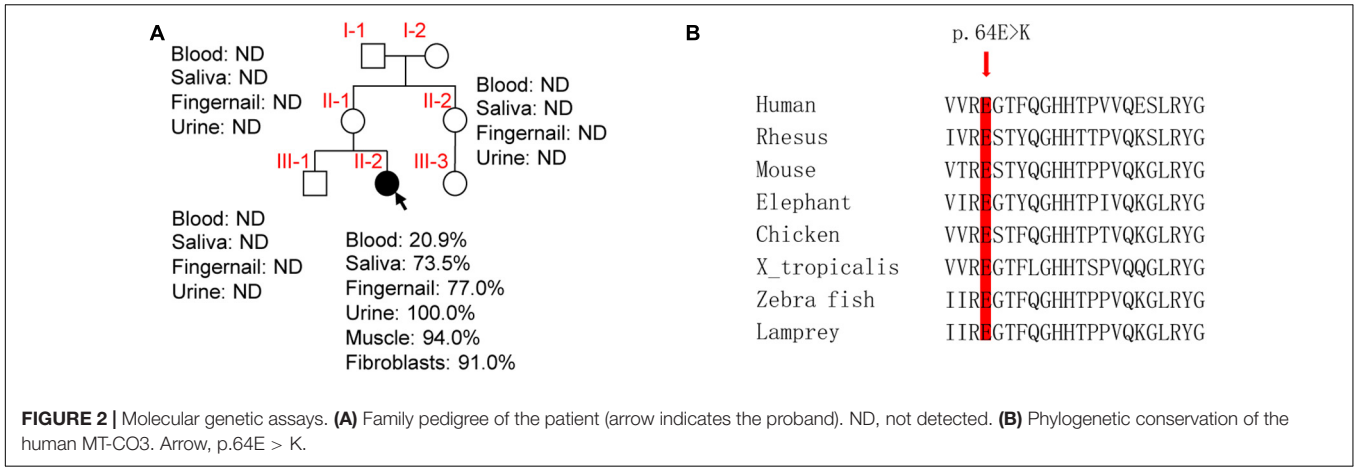

FIGURE 2 | Molecular genetic assays. **(A)** Family pedigree of the patient (arrow indicates the proband). ND, not detected. **(B)** Phylogenetic conservation of the human MT-CO3. Arrow, p.64E > K.

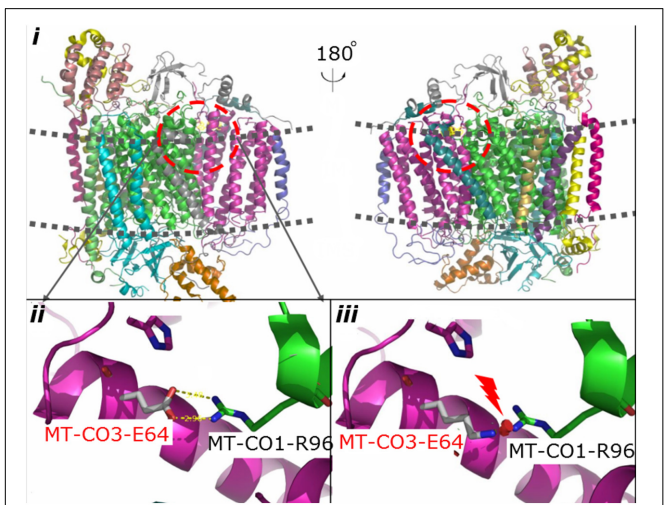

FIGURE 3 | Predicted structural model of MT-CO3. Image *i* is an overview of complex IV. Image *ii* shows the hydrogen bond between the WT glutamate at position 64 of MT-CO3 and an arginine in MT-CO1. Image *iii* shows the broken hydrogen bond caused by the change from glutamate to lysine at position 64 of MT-CO3.

position is highly conserved (**Figure 2B**). The PolyPhen2 score was 1.000, predicted to be "damaging." The PROVEAN score was -3.72, predicted to have a "deleterious" effect on the protein's function. The predicted structure revealed that the MT-CO3 protein was not truncated, but the stable combination of MT-CO3 with MT-CO1 was probably altered due to destruction of the hydrogen bond between the evolutionary conserved glutamate of MT-CO3 at position 64 and an arginine of MT-CO1 at position 96, resulting in incomplete assembly and dysfunction of CIV (**Figure 3**).

Effect of the m.9396G > A Variant on Transcript and Protein Levels

qRT-PCR analysis showed almost equal *MT-CO3* mRNA levels between two cybrid clones with more than 95% mutant mtDNA (M1 and M2) and one without this variant (WT) (**Figure 4A**). Consistently, Western blotting showed almost equal MT-CO3 protein levels in cybrids M1 and M2 compared with WT cybrids (**Figure 4B**).

The steady-state levels of MRC complexes were determined by BN-PAGE. The amount of mature CIV and the IGA of CIV in M1 and M2 clones were significantly reduced compared with the WT cybrids (**Figure 4C**), consistent with the CIV enzymatic defect detected in the patient. However, all other complexes exhibited similar expression levels in M1 and M2 clones compared with the WT cybrids.

The m.9396G > A Variant Results in Mitochondrial Dysfunction

The basal OCR of cybrids M1 and M2 was 79.9 and 73.7% relative to the WT cybrids, respectively. The ATP-linked OCR in cybrids M1 and M2 was significantly reduced, to 56.8 and 57.2%, respectively, of that in WT cybrid cells. The maximal OCR in the two mutant cybrid clones was significantly reduced to 56.5 and 62.2%, respectively, compared with the value in WT cybrids (**Figure 5A**). Biochemical assays were performed to determine the activities of MRC complexes. Cybrids M1 and M2 exhibited a significant reduction in CIV activity compared with the WT cybrids. However, no significant difference was observed in complex I, II, II + III, or III activities between the mutant and WT cybrids (**Table 2**).

Cybrids were cultured in glucose or 2-DG and pyruvate. The ATP levels of cybrids M1 and M2 were similar to those of the WT cybrids in glucose medium. However, in 2-DG and pyruvate medium, the ATP levels were significantly reduced in cybrids M1 and M2. Relative mitochondrial ATP production was 39.1% in the WT cybrids compared to 18.5 and 20.1% in the mutant cybrids (**Figure 5B**). Moreover, the levels of mitochondrial ROS production in cybrids M1 and M2 were significantly increased, by 1.61- and 1.84-fold, relative to the WT cybrids (**Figure 5C**).

A cybrid clone with the mutation load of 78.6% for m.9396G > A variant (M3) was also selected for ATP and ROS production experiments. The ATP levels of cybrids M3 were similar to those of WT cybrids in glucose medium, while significantly reduced in 2-DG and pyruvate medium. Relative

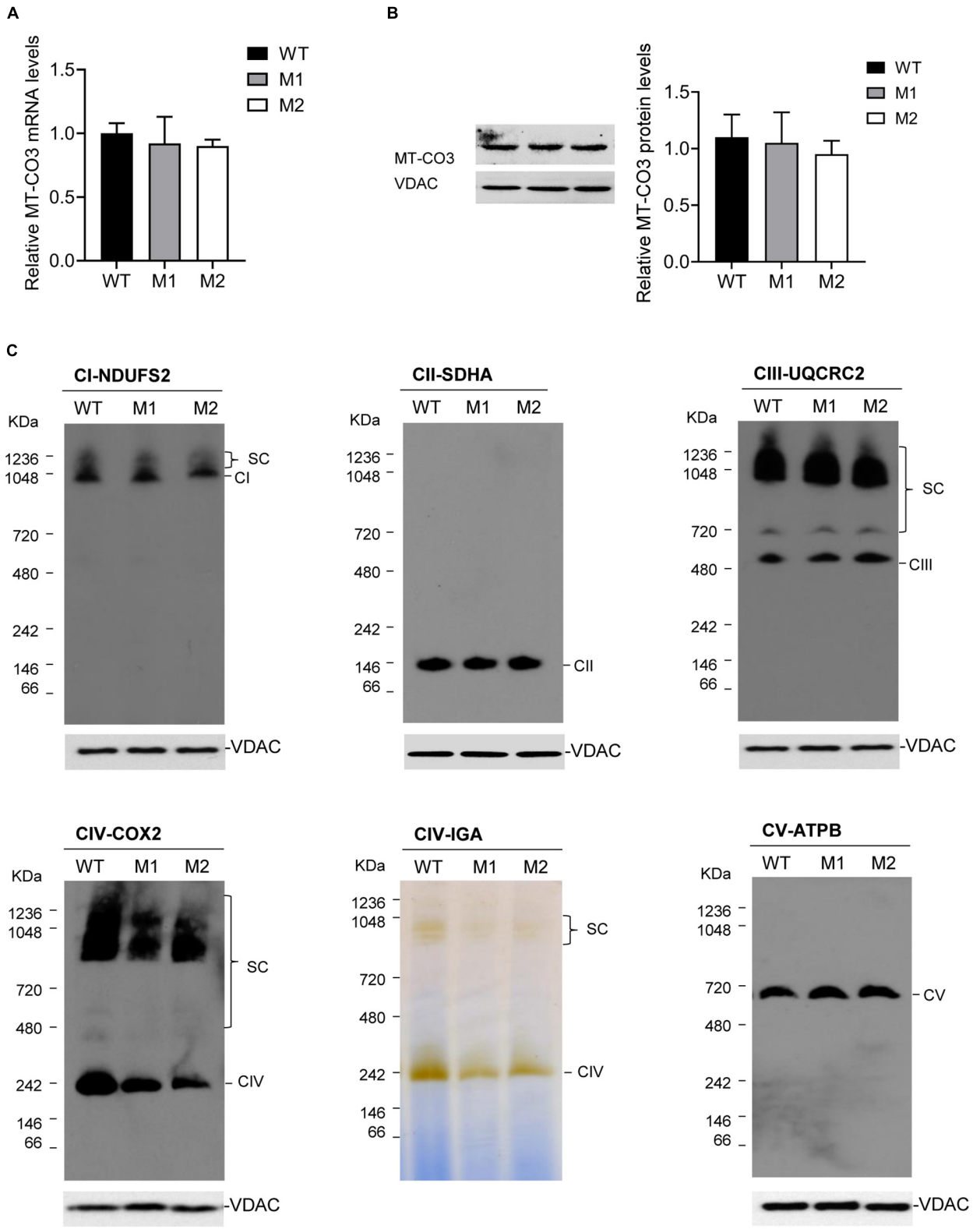

FIGURE 4 | Effect of the m.9396G > A variant on transcript and protein levels. **(A)** Relative *MT-CO3* mRNAs levels in cybrids M1 and M2 relative to that in WT cybrids. **(B)** Relative MT-CO3 protein levels in cybrids WT, M1, and M2 determined by Western blotting. VDAC was used as the internal control. **(C)** For cybrids WT, M1, and M2, IGA analysis of CIV and BN-PAGE/immunoblot analysis of MRC complexes using antibodies against subunits of complexes I (NDUFS2), II (SDHB), III (UQCRC2), IV (COXIV), and V (ATPB). VDAC was used as the internal control. CI, complex I; CII, complex II; CIII, complex III; CV, complex V; SC, supercomplex.

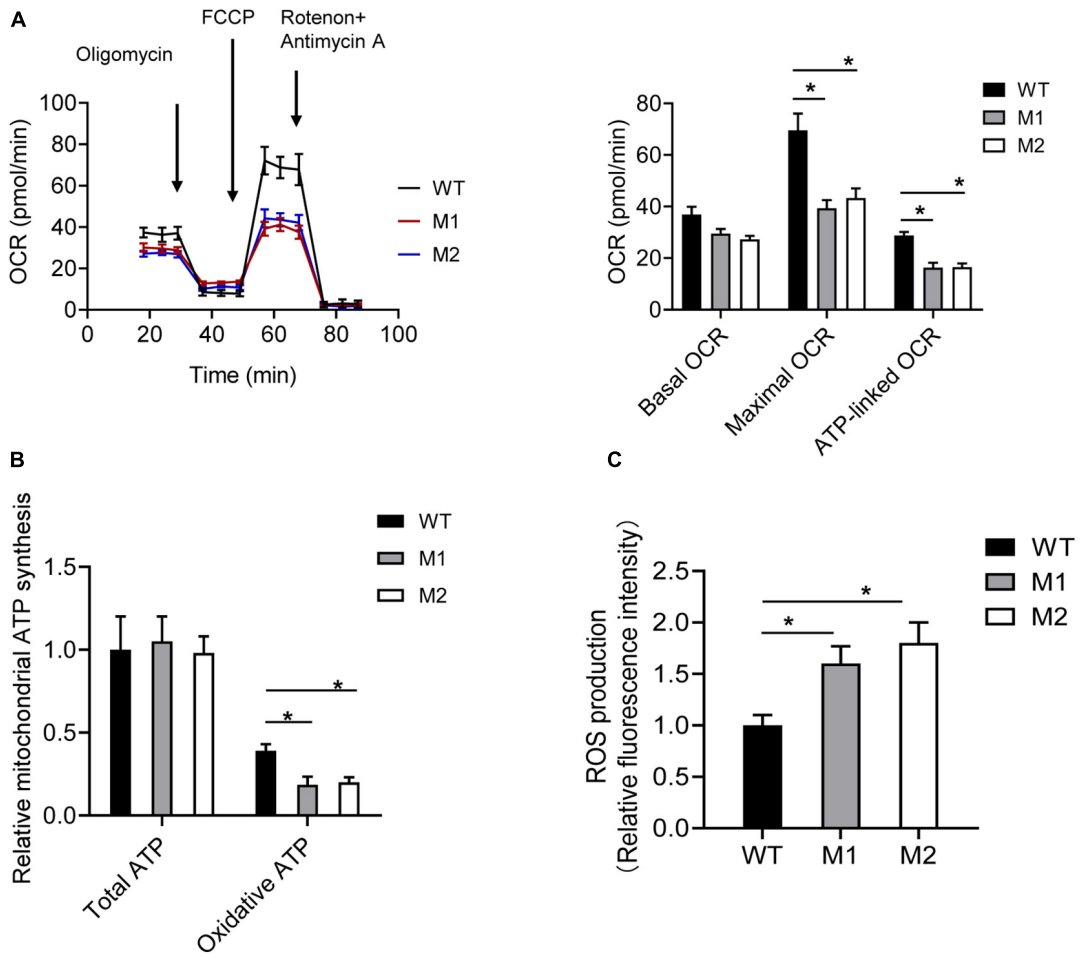

FIGURE 5 | The m.9396G > A variant results in mitochondrial dysfunction. **(A)** Analysis of OCR in WT (black), M1 (blue), and M2 (red) cybrids using the inhibitors oligomycin, FCCP, antimycin A, and rotenone. The basal OCR, maximal OCR, and ATP-linked OCR in cybrids WT, M1, and M2 are shown. Values are means ± SD. *$p < 0.05$ by one-way ANOVA. **(B)** Total and oxidative ATP levels in cybrids WT, M1, and M2. Values are means ± SD. *$p < 0.05$ by one-way ANOVA. **(C)** ROS production in cybrids WT, M1, and M2. Values are means ± SD. *$p < 0.05$ by one-way ANOVA.

TABLE 2 | MRC enzyme assays in WT and mutant cybrids.

	WT cybrids	Mutant cybrids	
		M1	M2
CI/CS	83.3 ± 8.1	76.1 ± 7.4	80.5 ± 10.9
CII/CS	70.9 ± 1.6	65.3 ± 15.2	57.1 ± 12.5
CII + III/CS	83.1 ± 10.7	91.9 ± 8.5	84.3 ± 10.2
CIII/CS	107.2 ± 16.1	114.2 ± 13.6	119.6 ± 10.2
CIV/CS	86.4 ± 14.0	36.3 ± 10.5	41.5 ± 7.3

Enzyme activities are expressed as percentages of normalized mean values relative to CS. CI, complex I; CII, complex II; CIII, complex III; CIV, complex IV; CS, citrate synthase.

mitochondrial ATP production was 48.1% in the WT cybrids compared to 31.4% in cybrid M3 (**Supplementary Figure S1A**). Moreover, the levels of mitochondrial ROS production in cybrids M3 was significantly increased, by 1.45-fold, relative to the WT cybrids (**Supplementary Figure S1B**).

DISCUSSION

Here, we describe a young female patient who met the diagnostic criteria of MELAS (Hirano et al., 1992; Yatsuga et al., 2012). The patient presented with typical manifestations of MELAS, including recurrent stroke-like episodes, seizures, visual impairment, hemiparesis, exercise intolerance, short stature, deafness and intermittent headache. She also had a variety of biochemical defects in mitochondrial energy metabolism, including an increased lactate level in cerebrospinal fluid and blood, and CIV activity deficiency in muscle. Finally, she exhibited mitochondrial pathological abnormalities including RRFs and COX-deficient fibers on muscle biopsy.

Whole-exome sequencing revealed no potential rare biallelic variant of genes for any known diseases, CIV subunits, or assembly factors. By mtDNA sequencing, only the novel m.9396G > A variant was detected. The evidence supporting the pathogenicity of this variant is as follows. This variant is not listed as a single-nucleotide polymorphism in the MITOMAP

database[2], and was not detected in our Chinese (>900 mtDNA sequences) and European (>20,507 mtDNA sequences) in-house database. This variant was nearly homoplasmic in clinically affected skeletal muscle, but only presented 38.9% in blood. The variant arose *de novo* and was not detected in the patient's clinically unaffected mother or maternal relatives. The variant resulted in substitution of a highly evolutionarily conserved glutamate at position 64 with a lysine, which probably affected the assembly of MT-CO3 with other CIV subunits and resulted in CIV dysfunction, as predicted *in silico*. Marked CIV deficiency was indicated by the results of biochemical assays, and generalized COX non-reactive muscle fibers were detected in histochemical assays, which were consistent with the location of the variant in a gene encoding one of the CIV subunits. Variants in *MT-CO3* are reportedly associated with MELAS (Rossmanith et al., 2008). The most important evidence for this is the marked reduction of mature CIV and mitochondrial dysfunction in the mutant cybrid cells, excluding nuclear variants and confirming the pathogenicity of the m.9396 G > A variant.

Although MELAS is a common mitochondrial disorder in which many pathogenic mtDNA variants have been reported, MELAS caused by variants in CIV encoding genes is rare (Manfredi et al., 1995; Choi et al., 2008; Rossmanith et al., 2008). Only three other cases have been reported previously, including one MELAS case with m.7630delT in *MT-CO2* and two with m.9957T > C in *MT-CO3* (**Supplementary Table S2**). The clinical and neuroradiological findings of these cases were characteristic. The pathological results of a muscle biopsy, however, were more distinctive than normal for MELAS, as first described by Rossmanith (Manfredi et al., 1995; Choi et al., 2008; Rossmanith et al., 2008). The muscle biopsy showed pronounced COX-negative fibers, which are less common in typical MELAS phenotypes (B. O. Manfredi et al., 1995; Choi et al., 2008; Rossmanith et al., 2008). Furthermore, only a few RRFs were seen, whereas the proportion of RRFs is usually high in 3243A > G MELAS (Manfredi et al., 1995; Choi et al., 2008; Rossmanith et al., 2008). In addition, intramuscular arteries, which are typically COX-positive in 3243A > G MELAS, are COX-negative in MELAS patients with a genetic CIV deficiency (Manfredi et al., 1995; Choi et al., 2008; Rossmanith et al., 2008). Our patient was clinically compatible with typical MELAS and pathologically consistent with the condition discussed by Rossmanith et al. (2008), supporting the pathological muscle peculiarities of rare MELAS caused by deleterious variants in the essential CIV subunit gene.

Complex IV in humans is composed of 14 subunits, including 3 core subunits (MT-CO1, MT-CO2, and MT-CO3) encoded by mtDNA, and 11 accessory subunits encoded by nuclear DNA. MT-CO3, with seven transmembrane helices, stabilizes the other two core subunits and acts as an initial proton acceptor in one of the transfer pathways. Compatibly, the m.9396G > A variant, located in one of the highly conserved transmembrane helices of MT-CO3, impairs the synthesis of CIV as studied in patient's mutant cybrids, leading to mitochondrial dysfunction. The ATP-linked and maximal OCR were significantly reduced in the mutant cybrids compared with the WT cybrids. The maximal OCR may allow cells to protect mitochondria from oxidative insults. Impairment of OCR promotes electron leakage, elevating the oxidative stress level. Consistently, ROS production in cybrids carrying this variant was significantly increased compared with that in WT cybrids. Mutant cybrids were unable to proliferate normally on 2-DG medium, indicating a deleterious effect of the variant on viability when cells were forced to obtain ATP from the OXPHOS system. The mitochondrial dysfunction mentioned above may have contributed to the pathogenesis of MELAS in this patient.

[2]http://www.mitomap.org/

CONCLUSION

We report a patient with a rare form of MELAS associated with the novel m.9396G > A variant in *MT-CO3* and demonstrated the molecular pathomechanisms of this variant, extending the spectrum of pathogenic variants of this gene.

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: Genome Sequence Archive in National Genomics Data Center, accession no: GVM000119.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by 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

MX contributed to the drafting of the original manuscript. FF contributed to the study concept and design, and critical revision of the manuscript for important intellectual content. All authors contributed to the acquisition and interpretation of data, and critical revision of the manuscript.

FUNDING

This work was supported by the Capital Health Research and Development Fund (2018-2-2096), the National Natural Science Foundation of China (81541115), the German BMBF and Horizon 2020 through the ERA PerMed project PerMiM (01KU2016A) and the European Joint Programme on Rare Diseases project GENOMIT (01GM1920A).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.638749/full#supplementary-material

Supplementary Figure 1 | The cybrid clone with a mutation load of 78.6% for m.9396G > A variant (M3) results in mitochondrial dysfunction. **(A)** Total and oxidative ATP levels in cybrids WT and M3. Values are means ± SD. *$p < 0.05$ by one-way ANOVA. **(B)** ROS production in cybrids WT and M3. Values are means ± SD. *$p < 0.05$ by one-way ANOVA.

Supplementary Table 1 | List of all the variants in mtDNA sequencing for the patient.

Supplementary Table 2 | A summary of reported patients with MELAS with pathogenic variants in the mitochondrial genes encoding CIV subunits.

REFERENCES

Adzhubei, I., Jordan, D. M., and Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* 76, 7.20.1–7.20.41. doi: 10.1002/0471142905.hg0720s76

Cheldi, A., Ronchi, D., Bordoni, A., Bordo, B., Lanfranconi, S., Bellotti, M. G., et al. (2013). POLG1 mutations and stroke like episodes: a distinct clinical entity rather than an atypical MELAS syndrome. *BMC Neurol.* 13:8. doi: 10.1186/1471-2377-13-8

Choi, B. O., Hwang, J. H., Kim, J., Cho, E. M., Cho, S. Y., Hwang, S. J., et al. (2008). A MELAS syndrome family harboring two mutations in mitochondrial genome. *Exp. Mol. Med.* 40, 354–360. doi: 10.3858/emm.2008.40.3.354

Choi, Y., and Chan, A. P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* 31, 2745–2747. doi: 10.1093/bioinformatics/btv195

Chomyn, A. (1996). Platelet-mediated transformation of human mitochondrial DNA-less cells. *Methods Enzymol.* 264, 334–339. doi: 10.1016/s0076-6879(96)64031-2

Delmiro, A., Rivera, H., García-Silva, M. T., García-Consuegra, I., Martín-Hernández, E., Quijada-Fraile, P., et al. (2013). Whole-exome sequencing identifies a variant of the mitochondrial MT-ND1 gene associated with epileptic encephalopathy: west syndrome evolving to Lennox-Gastaut syndrome. *Hum. Mut.* 34, 1623–1627. doi: 10.1002/humu.22445

El-Hattab, A. W., Adesina, A. M., Jones, J., and Scaglia, F. (2015). MELAS syndrome: clinical manifestations, pathogenesis, and treatment options. *Mol. Genet. Metab.* 116, 4–12. doi: 10.1016/j.ymgme.2015.06.004

Grady, J. P., Pickett, S. J., Ng, Y. S., Alston, C. L., Blakely, E. L., Hardy, S. A., et al. (2018). mtDNA heteroplasmy level and copy number indicate disease burden in m.3243A>G mitochondrial disease. *EMBO Mol. Med.* 10:e8262. doi: 10.15252/emmm.201708262

Hirano, M., Ricci, E., Koenigsberger, M. R., Defendini, R., Pavlakis, S. G., DeVivo, D. C., et al. (1992). Melas: an original case and clinical criteria for diagnosis. *Neuromuscul. Disord.* 2, 125–135. doi: 10.1016/0960-8966(92)90045-8

Jha, P., Wang, X., and Auwerx, J. (2016). Analysis of mitochondrial respiratory chain supercomplexes using blue native polyacrylamide gel electrophoresis (BN-PAGE). *Curr. Protoc. Mouse Biol.* 6, 1–14. doi: 10.1002/9780470942390.mo150182

Ji, K., Wang, W., Lin, Y., Xu, X., Liu, F., Wang, D., et al. (2020). Mitochondrial encephalopathy due to a novel pathogenic mitochondrial tRNA(Gln) m.4349C>T Variant. *Ann. Clin. Transl. Neurol.* 7, 980–991. doi: 10.1002/acn3.51069

Ji, K., Zheng, J., Sun, B., Liu, F., Shan, J., Li, D., et al. (2014). Novel mitochondrial C15620A variant may modulate the phenotype of mitochondrial G11778A mutation in a Chinese family with Leigh syndrome. *Neuromol. Med.* 16, 119–126. doi: 10.1007/s12017-013-8264-8

Kremer, L. S., and Prokisch, H. (2017). Identification of disease-causing mutations by functional complementation of patient-derived fibroblast cell lines. *Methods Mol. Biol.* 1567, 391–406. doi: 10.1007/978-1-4939-6824-4_24

Lorenzoni, P. J., Werneck, L. C., Kay, C. S., Silvado, C. E., and Scola, R. H. (2015). When should MELAS (Mitochondrial myopathy, Encephalopathy, Lactic Acidosis, and Stroke-like episodes) be the diagnosis? *Arq. Neuropsiquiatr.* 73, 959–967. doi: 10.1590/0004-282x20150154

Manfredi, G., Schon, E. A., Moraes, C. T., Bonilla, E., Berry, G. T., Sladky, J. T., et al. (1995). A new mutation associated with MELAS is located in a mitochondrial DNA polypeptide-coding gene. *Neuromuscul. Disord.* 5, 391–398. doi: 10.1016/0960-8966(94)00079-o

Ogawa, E., Shimura, M., Fushimi, T., Tajika, M., Ichimoto, K., Matsunaga, A., et al. (2017). Clinical validity of biochemical and molecular analysis in diagnosing Leigh syndrome: a study of 106 Japanese patients. *J. Inherit. Metab. Dis.* 40, 685–693. doi: 10.1007/s10545-017-0042-6

Parikh, S., Goldstein, A., Koenig, M. K., Scaglia, F., Enns, G. M., Saneto, R., et al. (2015). Diagnosis and management of mitochondrial disease: a consensus statement from the Mitochondrial Medicine Society. *Genet. Med.* 17, 689–701. doi: 10.1038/gim.2014.177

Rossmanith, W., Freilinger, M., Roka, J., Raffelsberger, T., Moser-Thier, K., Prayer, D., et al. (2008). Isolated cytochrome c oxidase deficiency as a cause of MELAS. *J. Med. Genet.* 45, 117–121. doi: 10.1136/jmg.2007.052076

Wang, S., Li, R., Fettermann, A., Li, Z., Qian, Y., Liu, Y., et al. (2011). Maternally inherited essential hypertension is associated with the novel 4263A>G mutation in the mitochondrial tRNAIle gene in a large Han Chinese family. *Circ. Res.* 108, 862–870. doi: 10.1161/circresaha.110.231811

Yatsuga, S., Povalko, N., Nishioka, J., Katayama, K., Kakimoto, N., Matsuishi, T., et al. (2012). MELAS: a nationwide prospective cohort study of 96 patients in Japan. *Biochim. Biophys. Acta* 1820, 619–624. doi: 10.1016/j.bbagen.2011.03.015

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 reviewer KM declared a past co-authorship with the authors HP and FF to the handling editor.

frontiers in Genetics

CASE REPORT
published: 13 May 2021
doi: 10.3389/fgene.2021.605778

Case Report: A Variant Non-ketotic Hyperglycinemia With *GLRX5* Mutations: Manifestation of Deficiency of Activities of the Respiratory Chain Enzymes

*Wei-xing Feng†, Xiu-wei Zhuo†, Zhi-mei Liu, Jiu-wei Li, Wei-hua Zhang, Yun Wu, Tong-li Han and Fang Fang**

Neurology Department, National Center for Children's Health China, Beijing Children Hospital Affiliated to Capital Medical University, Beijing, China

OPEN ACCESS

Edited by:
Ruth Roberts,
ApconiX, United Kingdom

Reviewed by:
Gerarda Cappuccio,
University of Naples Federico II, Italy
Moncef Feki,
Tunis El Manar University, Tunisia

***Correspondence:**
Fang Fang
13910150389@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 Genetics

Received: *24 October 2020*
Accepted: *08 April 2021*
Published: *13 May 2021*

Citation:
Feng W-x, Zhuo X-w, Liu Z-m, Li J-w, Zhang W-h, Wu Y, Han T-l and Fang F (2021) Case Report: A Variant Non-ketotic Hyperglycinemia With GLRX5 Mutations: Manifestation of Deficiency of Activities of the Respiratory Chain Enzymes. Front. Genet. 12:605778. doi: 10.3389/fgene.2021.605778

Objective: Variant non-ketotic hyperglycinaemia (NKH) is a rare disorder characterized by variable clinical, biochemical, and imaging features. The variant form of NKH is rare and characterized by variable clinical, biochemical and imaging features.

Subjects: Herein, we report a girl with variant NKH with two mutations in glutaredoxin 5 (*GLRX5*), which has been described in only three patients.

Results: The clinical and biochemical phenotypes of the patient are also described. She suffered from developmental regression associated with spasticity, developmental delay, anemia and optic atrophy. The mitochondrial leukoencephalopathy was used to designate these disorders. An increased T2 signal from the medulla oblongata to the C6 spinal region was also observed on spinal cord MRI. Tandem mass analysis of a dried blood sample revealed elevated levels of glycine. The patient has two compound heterozygous mutations (c.151_153 del AAG and c.196C>T) in the *GLRX5* gene. The c.196C>T mutation led to a stop codon (p.Q66Ter). Activities of mitochondrial respiratory chain (MRC) complexes II+III in the patient's fibroblasts were abnormal.

Conclusions: We present the case of a girl with variant NKH who manifested spasticity and bilateral cavitating leukoencephalopathy. The patient had a deficiency of a respiratory chain enzyme, and this is the first report. Genetic testing is important for physicians to evaluate suspected variant NKH patients and to provide proper genetic counseling.

Keywords: variant non-ketotic hyperglycinaemia, glutaredoxin 5, stop codon, compound heterozygous mutations, pediatric patient

INTRODUCTION

Non-ketotic hyperglycinaemia (NKH) is a disorder that affects the mitochondrial glycine cleavage enzyme system. NKH is characterized by elevated glycine in serum, cerebrospinal fluid and other body compartments. It can be classified into classic (or typical) and variant (or atypical) forms (Dinopoulos et al., 2005). Classic NKH occurs in the neonatal period; severe infantile encephalopathy, hypotonia, seizures, and progressive apnoea are the major symptoms.

Considerably elevated glycine levels in the plasma and cerebrospinal fluid (CSF) of these patients have been observed (Poothrikovil et al., 2019). Approximately 70∼75% of classic NKH patients carry mutations in the *GLDC* gene, whereas 20% have mutations in *AMT*, and 1% have mutations in *GCSH* (Toone et al., 2001). The variant form of NKH is rare and characterized by variable clinical, biochemical and imaging features. These cases differ in the age of onset and presentation. Deficiencies in the biosynthesis of lipoyl-H lipoate synthase (*LIAS*), glutaredoxin 5(*GLRX5*), and BolA type 3(*BOLA3*) can cause variant NKH (Kikuchi et al., 2008). A novel homozygous missense mutation in the *GLDC* gene of an atypical NKH patient has also been reported (Baker et al., 2014).

Only three variant NKH patients have been reported to carry mutations in *GLRX5*, as described by Chiong et al. (2007), Wei et al. (2011), and Baker et al. (2014). (*GLRX5*; OMIM 616859). The three cases were two Lebanese girls (onset ages, 2.5 years and 7 years) and one Chinese boy in Taiwan (onset age, 2.5 years). In the two Lebanese girls, *GLRX5* exhibited an exon 1,c.151_153delAAG; p.K51del, homozygous deletion. In the Chinese boy, the mutation c.151_153delAAG was found on the maternal allele, and the 8-bp insertion c.82_83insGCGTGCGG; p.G28Gfs*25 was found on the paternal allele, resulting in a premature stop codon at amino acid 52, located 25 amino acid residues downstream. Patient samples were genotyped using Affymetrix SNP 6.0 chips. Herein, we report a Chinese girl with variant NKH with mutations in *GLRX5*. She suffered development retrogression associated with spasticity and developmental delay. Unfortunately, she died of respiratory failure after the infection. MRI showed signal abnormalities involving the bilateral cavitating leukoencephalopathy and long tract-like lesions of the spinal cord. Mutations in *GLRX5* were detected, namely, c.151_153 del AAG [p.K51del] (of maternal origin) and c.196C>T [p.Q66Ter] (of paternal origin). The c.196C>T [p.Q66Ter] mutation was characterized as a stop codon and protein truncation. This mutation is pathogenic.

PATIENT AND RESULT

The girl was Han Chinese in Shandong Province, and she was born normally. Her developmental history was also standard until 1 year and 3 months of age. She was able to sit steadily, crawl, stand with support, and say "mum" and "dad." Then, the patient began to exhibit clumsy and uncoordinated movements. She was subsequently admitted to the local hospital. Multifocal lesions were detected on MRI by increased signal intensity on the T2 sequence of the bilateral corona radiata, centrum ovale, and corpus callosum. The cystic lesions were observed in the abnormal white matter (**Figure 1A**). Besides, increased signal intensity in the frontal, parietal and occipital white matter was noted on T2-weighted and diffusion-weighted imaging. The bilateral cavitating leukoencephalopathy was used to designate these disorders. Tandem mass analysis of a dried blood sample revealed elevated levels of glycine (812 μM vs. normal 20–760 μM). No treatment was administered initially at the local hospital to which she was admitted. However, her condition deteriorated after 2 months, with difficulty in movements and language ability. She was unable to raise her head or sit steadily. She presented with stiffness and spasticity in all extremities. She only screamed and no longer said "mum" and "dad." At this point, she was admitted to our hospital. She suffered from slight encephalopathy, was irritable, and had been febrile for 8 days. Neurological examination showed spastic paralysis of the four extremities, reduced muscle strength (3 to 4 degrees), clasp-knife-like muscular tension, hyperactivity of the tendon reflexes, bilateral clonus, ankle clonus, and the Babinski sign. Several blood tests showed below-normal hemoglobin levels (**Appendix 1**), which suggested hypochromic microcytic anemia. She suffered from intermittent mild anemia after 6 months of age, but her parents did not attend to it. The cytology test of cerebrospinal fluid (CSF) and CSF total protein were normal. Routine blood biochemical examination revealed elevated blood lactic acid (3.47 mmol/L, normal 2 mmol/L) and increased CK (457 U/L) levels. The urinary organic acid analysis and dried blood spot analysis using GC/MS and tandem mass analysis showed normal results.

Re-examination by MRI revealed enlarged lesions in the frontal, parietal and occipital regions of the deep white matter and the corpus callosum, and the cystic lesions also were observed in the abnormal white matter. Encephalatrophy was observed in the cerebrum and cerebellum. Spinal MRI indicated an increased T2 signal from the medulla oblongata to the C6 spinal region (**Figure 1A**). The MRA, MRV and SWI were normal. Slow waves were detected by electroencephalography (EEG). Leopard skin-like fundus oculi and bilateral optic atrophy were observed in the ophthalmological examinations.

As the girl had been febrile for 8 days. We provided the treatment of acyclovir and ceftriaxone, as we can't exclude central nervous system infections in the early stage of hospitalization. And patient was treated with baclofen, nitrazepam and trihexyphenidyl were administered for spasticity. She was diagnosed with mitochondrial disease. Vitamins B1, B2, C, and E, coenzyme Q10, biotin and iron dextran were prescribed. On day 21, she was released from the hospital with a slightly improved condition. Her muscular tension was slightly decreased. The clinical symptoms, MRI imaging and next-generation sequencing (NGS) result provided important evidence for the diagnosis of the variant non-ketotic hyperglycinaemia. Lipoic acid and 5 ALAs were administered for treatment. We then followed up with the patient's family until now. The patient's condition showed improvement in terms of her intelligence, and she once again could say "mum" and "dad" when she was followed up for 1 and 1 and a half year. However, she died of infection when 4 years old in the local hospital.

We collaborated with SINOPATH Medical Inspection Inc. to perform whole-exome sequencing utilizing NGS with the proposita sample on Hiseq X (Illumina, CA). Then, Sanger sequencing was performed to confirm the findings of NGS. Sequencing data were aligned to the human genome reference. Next Gene V2.3.4 software was used to evaluate the coverage and mean read depth of the target regions and to identify variants. The mean read depth was 151.24 × 97, and 95% of the target regions covered at least 20x. Two point mutations

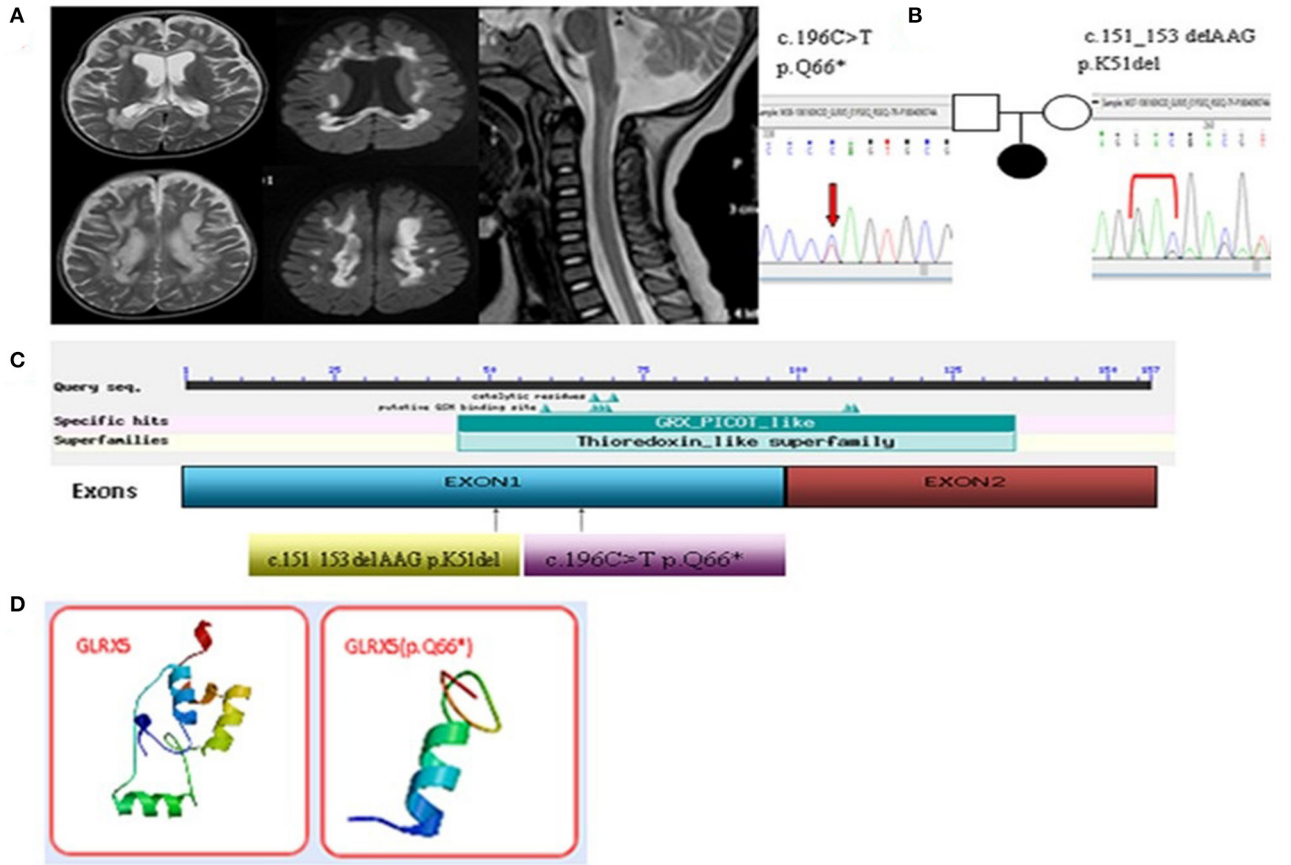

FIGURE 1 | (A) Enlarged lesions in the frontal, parietal and occipital of the white matter and the corpus callosum in T2 and the diffusion weighted images were observed. Spinal MRI images showed central lesions of the upper spinal cord. **(B)** Two compound heterozygous missense mutations were identified in the GLRX5 gene. **(C)** The mutation sites of the GLRX5. **(D)** The difference of the noted stop codon c.196C>T[p.Q66*] of the GLRX5 protein in the three-dimensional modeling prediction.

of *GLRX5*, c.151_153 del AAG [p.K51del] (of maternal origin) and c.196C>T [p.Q66Ter] (of paternal origin), were found in the DNA of the proband (**Figures 1B,C** and **Appendix 2**). The parents were healthy. This c.196C>T mutation led to a new stop codon that was found to be deleterious via *in silico* analysis by SIFT, PolyPhen and Mutation_Taster software programs. The position of the c.196C>T residue was conserved among different species (http://genome.ucsc.edu/index.html). The p.K51del and p.Q66Ter is located in the thioredoxin superfamily domain. **Figure 1D** indicates the difference in the p.Q66Ter of the GLRX5 protein in the three-dimensional modeling prediction (https://swissmodel.expasy.org/). This suggests a pathogenic role for the non-sense variant resulting in a premature stop codon and protein truncation.

Skin fibroblasts from the patient and control fibroblasts (fHDF, Toyobo) were cultured in Dulbecco's modified Eagle medium (DMEM; Gibco®, Life Technologies, Thermo Fisher Scientific) in the presence of 1% penicillin/streptomycin and 10% fetal bovine serum (both Gibco®, Life Technologies, Thermo Fisher Scientific) at 37°C and 5% CO_2. Activities of mitochondrial respiratory chain (MRC) complexes I, II, II+III, III, IV, and the mitochondrial marker enzyme citrate synthase (CS) were assayed in isolated mitochondria obtained from skin fibroblasts, as described previously (Kirby et al., 1999; Frazier and Thorburn, 2012; Ogawa et al., 2017). Enzyme activities of each complex were presented as the percentage of normal control mean relative to appropriate reference enzyme activities, such as citrate synthase. Oxygen comsumption rate (OCR) was measured in the patient's cultured fibroblasts with XF96 Extracellular Flux Analyzer (Seahorse Bioscience, Billerica, MA, USA). Samples were prepared as reported (Ogawa et al., 2017, Kremer and Prokisch, 2017).

Activities of MRC complexes II+III in the patient's fibroblasts was abnormal (CS ratio 37%), and complex II and complex III were mild declines, CS ratio 47.3% and 41.3, respectively. MRR (maximum respiration rate) was significantly lower than the control in glucose-containing and galactose-containing medium, respectively (37 and 33% of the control, respectively). Basal OCR also decreased compared with the control (**Appendix 3**). MRR was expressed as a percentage relative to the average of controls, in which a reduction to <71.6% is considered to represent a significant decline (Ogawa et al., 2017).

DISCUSSION

Non-ketotic hyperglycinaemia (NKH) is an autosomal recessive disorder in the glycine cleavage system (GCS), which is a multienzyme complex located in the inner mitochondrial membrane of the brain, liver and kidney. The variant NKH patients have less severe presentations, variable ages of onset, and milder neurological symptoms than classic NKH (Hiraga et al., 1981; Kremer and Prokisch, 2017). Non-ketotic hyperglycinemia (NKH) is a well-recognized metabolic cause of life-threatening illness in the neonate. The variant NKH most patients with normal development for 6 months. Then they lost all milestones gradually, and often signifies a poor prognosis (Baker et al., 2014). In this case, the girl died of respiratory failure after the infection which may be associated with deficiency of activities of the respiratory chain enzymes. It was not reported before.

Diagnosing variant NKH is more difficult than diagnosing the classic form because the clinical manifestations are non-specific. Variant NKH may be diagnosed according to the developmental retrogression associated with spasticity and the MRI. Our MRI findings matched the pathology of variant NKH. The white matter involvement was recognized as a manifestation of mitochondrial disorders and the cavitating leukoencephalopathy were used to designate these disorders. They are mainly defined by the MRI characteristics such as cystic lesions in the abnormal white matter (Bindu et al., 2018). In our case, the brain MRI revealed the coincident and typically high T2 signal and diffusion restriction in the frontal, parietal and occipital white matter, as well as the corpus callosum and spine. These findings matched the pathology of variant NKH. The cranial MRI of the Lebanese girl also revealed abnormalities in the frontal and parietal white matter, which then progressed to the deep white matter, periaqueductal gray matter, and around the central canal in the cervical spine (Chiong et al., 2007). The variant non-ketotic hyperglycinaemia patients can have defective glycine transport systems in both the brain and the spinal cord. Glycine transporters may play an important role in brain and spinal cord function. Lipoate is an essential co-factor for the glycine cleavage system, and lipoate biosynthesis requires lipoate synthase enzyme (LIAS), which need [Fe-S]. It has therefore been speculated that the *GLRX5* mutation may impair [Fe-S] transfer to the LIAS protein. The subcortical white matter and the genu of the corpus callosum were the regions with the highest lesion burden (Liu et al., 1993; Zafra et al., 1995).

Our patient showed reduced activities of MRC complexs II+III, and mild decline of complex II and complex III. Iron-sulfur clusters are incorporated in many enzymes including respiratory chain complexes I, II and III and the aconitases (Baker et al., 2014). In theory, deficiency in Iron-sulfur clusters can have reduced activities of these respiratory chain enzymes, and some previously reported BOLA3 mutation patients had reduced activities of some respiratory chain enzymes (Cameron et al., 2011; Haack et al., 2013). None of the reported *GLRX5* mutation patients had a deficiency of a respiratory chain enzyme, and this is the first report. And the girl died of respiratory failure which may be associated with it.

Mutations in *GLRX5* have been associated with the variant NKH in humans because of disorders in the biosynthesis of lipoyl-H. *GLRX5*, c.151_153 del AAG [p.K51del] (of maternal origin) and c.196C>T [p.Q66Ter] (of paternal origin) were detected in the patient. The stop codon, c.196C>T [p.Q66Ter], which was not reported in the Human Gene Mutation Database, and it was found to be deleterious *in silico* analysis. This may lead to a poor prognosis. The deleted lysine 51 lies in the highly conserved glutaredoxin domain (Baker et al., 2014). The mutation c.151_153 delAAG [p.K51del] in *GLRX5* was also discovered in two Lebanese girls and the Chinese boy. It was reported that cells expressing the K51del mutation had decreased PDH and alpha-KGDH complex activities (Liu et al., 2016). The gene *GLRX5* encodes a mitochondrial protein, which is evolutionarily conserved. It is involved in the biogenesis of iron-sulfur clusters, which are required for normal iron homeostasis. In this case, the mitochondrial respiratory chain enzyme defect was also confirmed.

This study reported a case of variant NKH with the *GLRX5* mutation. Mutations in *GLRX5* have been reported in sideroblastic anemia as well as variant non-ketotic hyperglycinaemia in humans (Liu et al., 2014; Ogawa et al., 2017). In this case, intermittent mild anemia was observed at 6 months of age, and hypochromic microcytic anemia was confirmed by blood tests (**Appendix 1**). These symptoms were associated with the *GLRX5* mutation (Lill et al., 2012). GLRX5 is a mitochondrial protein that plays an essential role in mitochondrial iron-sulfur cluster transfer. Interestingly, Gang Liu reported a Chinese congenital sideroblastic anemia patient with two compound heterozygous missense mutations in *GLRX5*(c.301A>C;K101Q and c.443T>C;L148S). GLRX5 plays a major role in mitochondrial [Fe-S] transfer to apoproteins.The heme biosynthesis pathway is more vulnerable to GLRX5 dysfunction.And that patient did not exhibit central nervous system symptoms. Thus, different mutations, even though they occur in the same *GLRX5* gene, may manifest different phenotypes (Liu et al., 2014).

As a limitation, we did not get the data on glycine determination in CSF. And this case may be described little information as one patient (No.31) in 37 patients diagnosed with cavitating leukoencephalopathies (Zhang et al., 2019). That report was an excellent study. However, it focused on 37 patients. It's unrealistic to describe each patient in detail. And we provide explicit clinical information and follow-up of the patient. Importantly, we found the activities of MRC complexes II+III in the patient's fibroblasts was abnormal and complex II and complex III were mild declines.

In conclusion, an association between *GLRX5* gene mutations and variant NKH has been identified and described. The c.196C>T mutation was identified as a stop codon which may associate a poor prognosis. This is the first report of the deficiency of activities of the respiratory chain enzymes of the patient with *GLRX5* mutations. It is of great importance to characterize any suspected cases at the genetic level to enhance the awareness of variant NKH.

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 the Ethics Committee of Beijing Children's Hospital. 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

FF and W-xF: conceived and designed the manuscript. X-wZ, J-wL, and W-hZ: clinical data acquisition. T-lH, YW, and Z-mL: analyzed the clinical and genetic data. W-xF, FF, and YW: wrote the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the National Natural Science Foundation of China (No. 81541115) and the Beijing Municipal administration of Hospitals incubating Program (PX2017065).

ACKNOWLEDGMENTS

We would like to thank the patients and the many clinicians and the clinical laboratory scientists who contributed to this research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.605778/full#supplementary-material

REFERENCES

Baker, P. R. 2nd., Friederich, M. W., Swanson, M. A., Shaikh, T., Bhattacharya, K., Scharer, J., et al. (2014). Variant non ketotic hyperglycinemia is caused by mutations in LIAS, BOLA3 and the novel gene GLRX5. *Brain* 137, 366–379. doi: 10.1093/brain/awt328

Bindu, P. S., Sonam, K., Chiplunkar, S., Govindaraj, P., Nagappa, M., Vekhande, C., et al. (2018). Mitochondrial leukoencephalopathies: a border zone between acquired and inherited white matter disorders in children? *Mult. Scler. Relat. Disord.* 20, 84–92. doi: 10.1016/j.msard.2018.01.003

Cameron, J. M., Janer, A., Levandovskiy, V., Mackay, N., Rouault, T. A., Tong, I., et al. (2011). Mutations in iron-sulfur cluster scaffold genes NFU1 and BOLA3 cause a fatal deficiency of multiple respiratory chain and 2-oxoacid dehydrogenase enzymes. *Am. J. Hum. Genet.* 89, 486–495. doi: 10.1016/j.ajhg.2011.08.011

Chiong, M. A., Procopis, P., Carpenter, K., and Wilcken, B. (2007). Late-onset nonketotic hyperglycinemia with leukodystrophy and an unusual clinical course. *Pediatr. Neurol.* 37, 283–286. doi: 10.1016/j.pediatrneurol.2007.05.016

Dinopoulos, A., Matsubara, Y., and Kure, S. (2005). Atypical variants of nonketotic hyperglycinemia. *Mol. Genet. Metab.* 86, 61–69. doi: 10.1016/j.ymgme.2005.07.016

Frazier, A. E., and Thorburn, D. R. (2012). Biochemical analyses of the electron transport chain complexes by spectrophotometry. *Methods Mol. Biol.* 837, 49–62. doi: 10.1007/978-1-61779-504-6_4

Haack, T. B., Rolinski, B., Haberberger, B., Zimmermann, F., Schum, J., Strecker, V., et al. (2013). Homozygous missense mutation in BOLA3 causes multiple mitochondrial dysfunctions syndrome in two siblings. *J. Inherit. Metab. Dis.* 36, 55–62. doi: 10.1007/s10545-012-9489-7

Hiraga, K., Kochi, H., Hayasaka, K., Kikuchi, G., and Nyhan, W. L. (1981). Defective glycine cleavage system in nonketotic hyperglycinemia. Occurrence of a less active glycine decarboxylase and an abnormal aminomethyl carrier protein. *J. Clin. Invest.* 68, 525–534. doi: 10.1172/JCI110284

Kikuchi, G., Motokawa, Y., Yoshida, T., and Hiraga, K. (2008). Glycine cleavage system: reaction mechanism, physiological significance, and hyperglycinemia. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* 84, 246–263. doi: 10.2183/pjab.84.246

Kirby, D. M., Crawford, M., Cleary, M. A., Dahl, H. H., Dennett, X., and Thorburn, D. R. (1999). Respiratory chain complex I deficiency: an underdiagnosed energy generation disorder. *Neurology* 52, 1255–1264. doi: 10.1212/WNL.52.6.1255

Kremer, L. S., and Prokisch, H. (2017). Identification of disease-causing mutations by functional complementation of patient-derived fibroblast cell lines. *Methods Mol. Biol.* 1567, 391–406. doi: 10.1007/978-1-4939-6824-4_24

Lill, R., Hoffmann, B., Molik, S., Pierik, A. J., Rietzschel, N., Stehling, O., et al. (2012). The role of mitochondria in cellular iron-sulfur protein biogenesis and iron metabolism. *Biochim. Biophys. Acta* 1823, 1491–1508. doi: 10.1016/j.bbamcr.2012.05.009

Liu, G., Guo, S., Anderson, G. J., Camaschella, C., Han, B., and Nie, G. (2014). Heterozygous missense mutations in the GLRX5 gene cause sideroblastic anemia in a Chinese patient. *Blood* 124, 2750–2751. doi: 10.1182/blood-2014-08-598508

Liu, G., Wang, Y., Anderson, G., J., Camaschella, C., Chang, Y., and Nie, G. (2016). Functional analysis of GLRX5 mutants reveals distinct functionalities of GLRX5 protein. *J. Cell Biochem.* 117, 207–217. doi: 10.1002/jcb.25267

Liu, Q. R., Lopez-Corcuera, B., Mandiyan, S., Nelson, H., and Nelson, N. (1993). Cloning and expression of a spinal cord- and brain-specific glycine transporter with novel structural features. *J. Biol. Chem.* 268, 22802–22808. doi: 10.1016/S0021-9258(18)41598-0

Ogawa, E., Shimura, M., Fushimi, T., Tajika, M., Ichimoto, K., Matsunaga, A., et al. (2017). Clinical validity of biochemical and molecular analysis in diagnosing Leigh syndrome: a study of 106 Japanese patients. *J. Inherit. Metab. Dis.* 40, 685–693. doi: 10.1007/s10545-017-0042-6

Poothrikovil, R. P., Al Thihli, K., Al Futaisi, A., and Al Murshidi, F. (2019). Nonketotic hyperglycinemia: two case reports and review. *Neurodiagn. J.* 59, 142–151. doi: 10.1080/21646821.2019.1645549

Toone, J. R., Applegarth, D. A., Coulter-Mackie, M. B., and James, E. R. (2001). Recurrent mutations in P- and T-proteins of the glycine cleavage complex and a novel T-protein mutation (N145I): a strategy for the molecular investigation of patients with nonketotic hyperglycinemia (NKH). *Mol. Genet. Metab.* 72, 322–325. doi: 10.1006/mgme.2001.3158

Wei, S. H., Weng, W. C., Lee, N. C., Hwu, W. L., and Lee, W. T. (2011). Unusual spinal cord lesions in late-onset non-ketotic hyperglycinemia. *J. Child Neurol.* 26, 900–903. doi: 10.1177/0883073810393965

Zafra, F., Gomeza, J., Olivares, L., Aragon, C., and Gimenez, C. (1995). Regional distribution and developmental variation of the glycine transporters GLYT1 and GLYT2 in the rat CNS. *Eur. J. Neurosci.* 7, 1342–1352. doi: 10.1111/j.1460-9568.1995.tb01125.x

Zhang, J., Liu, M., Zhang, Z., Zhou, L., Kong, W., Jiang, Y., et al. (2019). Genotypic spectrum and natural history of cavitating leukoencephalopathies in childhood. *Pediatr. Neurol.* 94, 38–47. doi: 10.1016/j.pediatrneurol.2019.01.002

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

frontiers in Pediatrics

ORIGINAL RESEARCH
published: 13 May 2021
doi: 10.3389/fped.2021.635703

Detection of Disease-Causing SNVs/Indels and CNVs in Single Test Based on Whole Exome Sequencing: A Retrospective Case Study in Epileptic Encephalopathies

Dan Sun[1†], Yan Liu[2†], Wei Cai[3†], Jiehui Ma[1], Kun Ni[1], Ming Chen[4], Cheng Wang[5], Yongchu Liu[6], Yuanyuan Zhu[6], Zhisheng Liu[1] and Feng Zhu[5,7*]*

[1] Department of Neurology, Wuhan Children's Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [2] Department of Pediatrics, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [3] Department of Hematology, Wuhan Children's Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [4] Xiangyang Central Hospital, Affiliated Hospital of Hubei University of Arts and Science, Xiangyang, China, [5] Department of Cardiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [6] Aegicare Technology Co., Ltd. Shenzhen, China, [7] Clinic Center of Human Gene Research, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

OPEN ACCESS

Edited by:
Tieliu Shi,
East China Normal University, China

Reviewed by:
Jeffrey Dennis Calhoun,
Northwestern University, United States
Cristina Skrypnyk,
Arabian Gulf University, Bahrain

***Correspondence:**
Zhisheng Liu
liuzsc@126.com
Feng Zhu
zhufeng@hust.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 Pediatrics

Received: *30 November 2020*
Accepted: *24 March 2021*
Published: *13 May 2021*

Citation:
Sun D, Liu Y, Cai W, Ma J, Ni K, Chen M, Wang C, Liu Y, Zhu Y, Liu Z and Zhu F (2021) Detection of Disease-Causing SNVs/Indels and CNVs in Single Test Based on Whole Exome Sequencing: A Retrospective Case Study in Epileptic Encephalopathies. Front. Pediatr. 9:635703. doi: 10.3389/fped.2021.635703

Background: Epileptic encephalopathies (EEs) are a pediatric entity with highly phenotypic and genetic heterogeneity. Both single nucleotide variants (SNVs)/Indels and copy number variations (CNVs) could be the causes. Whole exome sequencing (WES) is widely applied to detect SNVs/Indels, but the bioinformatics approach for detecting CNVs is still limited and weak. In the current study, the possibility of profiling both disease-causing SNVs/Indels and CNVs in a single test based on WES in EEs was evaluated.

Methods: The infants diagnosed with EEs were enrolled from a single pediatric epilepsy center between January 2018 and February 2020. Demographic and clinical data were collected. In WES data, the pathogenic SNVs were identified through an in-house pipeline, and pathogenic CNVs were identified by CNVkit. The diagnostic rate was evaluated, and the molecular findings were characterized.

Results: A total of 73 infants were included; 36 (49.32%) of them were males. The median age was 7 months. Thirty-two (43.84%) infants had been diagnosed with epilepsy syndrome. The most common type of syndrome was West syndrome (22/73, 30.1%), followed by Dravet syndrome (20/77, 27.4%). Fifty-four (73.97%) had intellectual development delay. The genetic cause of EEs, pathogenic or likely pathogenic variants, were successfully discovered in 46.6% (34/73) of the infants, and 29 (39.7%) infants carried SNVs/Indels, while 5 (6.8%) carried CNVs. The majority of the disease-causing variants were inherited in *de novo* pattern (25, 71.4%). In addition to showing that the variants in the ion channel encoding genes accounted for the main etiology, we discovered and confirmed two new disease-causing genes, *CACNA1E* and *WDR26*. Five discovered CNVs were deletions of 2q24.3, 1p36, 15q11-q13, 16p11.2, and 17p13.3, and all were confirmed by array comparative genomic hybridization.

Conclusion: The application of both SNVs/Indels and CNVs detection in a single test based on WES yielded a high diagnosis rate in EEs. WES may serve as a first-tier test with cost-effective benefit in EEs.

Keywords: epileptic encephalopathies, whole exome sequencing, genetics, copy number variation, variant

INTRODUCTION

Epileptic encephalopathies (EEs) are characterized by frequent severe seizures, severe electroencephalography (EEG) abnormalities, and intellectual/developmental disabilities (1–3). EEs commonly occur during infancy and early childhood, with poor clinical outcome. They are highly heterogeneous in clinical features, including West syndrome (WS), Dravet syndrome (DS), Ohtahara syndrome (OS), migrating partial epilepsy of infancy (MPEI), and other related epilepsy syndromes (4–6). A range of etiologic spectra in epilepsy has been recognized, including structural, genetic, infectious, metabolic, and immune etiologies (7). Among them, genetic factors play an important role in the pathogenesis of EEs. Many genes have been linked to EEs, and a significant proportion of them is ion channel genes which are particularly relevant to epilepsy (1, 8). Identifying the etiology of EEs is a primary clinical objective in the management of the disorder. Up to now, over 30 genes with precision medicine implications have been described in existing published literature (9, 10).

Sequencing-based studies have identified risk genes for rare and severe epilepsies and revealed a role of rare deleterious variation in common epilepsies (11). Whole exome sequencing (WES) is used routinely to detect sequence variants among the infants with EEs [single nucleotide variants (SNVs) and short insertions and deletions, SNPs/Indels] in clinical genetic laboratories (12, 13). However, the sequencing-based methods for identifying SNPs/Indels are not optimized for detecting the potential copy number variation (CNV), which is also involved in the pathogenesis of epilepsy, particularly early infantile epileptic encephalopathy (EIEE) (14, 15). This deficiency requires additional samples and application of genetic tests such as SNP arrays or array comparative genomic hybridization (aCGH) to discover the CNVs. Currently, many algorithms have been developed to obtain copy number information directly from the WES data (14, 16). This exome-based CNV analytic strategy has been proven to improve the diagnostic yields with cost-effective benefit (17). In this study, we evaluated the clinical application of profiling both SNPs/Indels and CNVs in a cohort of 73 infants with EEs who underwent WES as a diagnostic test.

METHODS

Study Populations

A retrospective case study was performed in a single pediatric epilepsy center at Wuhan Children's Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China. The infants diagnosed with EEs were recruited between January 2018 and December 2019. All infants underwent a neurological examination, electromyography (EMG), and magnetic resonance imaging (MRI) for brain. And intellectual disabilities were assessed according to the criteria of the Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5) (18). All these clinical data were collected and independently reviewed by two neurologists (DS and YL).

WES

Genomic DNA samples were extracted from peripheral blood using QIAamp® Blood Mini Kit (Qiagen, Hilden, Germany). The quality of genomic DNA was evaluated by agarose gel electrophoresis analysis, and the quantity was measured by NanoDrop2000 and Qubit3.0. DNA was sheared with M220 Focused-ultrasonicator (Covaris, Woburn, MA, USA). DNA target regions were captured by hybridizing the genomic DNA sample library with the xGen® Exome Research Panel v1.0 (IDT, USA). The captured and amplified DNA samples were sequenced using Illumina NovaSeq6000 (Illumina, San Diego, CA, USA) with 150 base-paired end reads.

Detection of SNVs/Indels From WES Data

To identify disease-associated SNVs/Indels (short insertions and deletions smaller than 50 bp), sequencing data were analyzed according to an in-house pipeline. Both public software and commercial packages were implemented during bioinformatics analysis. Specifically, raw FASTQ was processed with FASTP (https://github.com/OpenGene/fastp) to cut adapters and filter out bad reads of low quality. The clean reads were then aligned against human reference (GRCh37) with BWA (19). SNPs/Indels were discovered by HaplotypeCaller tool of GATK after the necessary post-processes on primary alignment, including removal of duplicated reads, realignment, and base recalibrating (20). VEP(Ensembl Variant Effect Predictor) was employed to identify the effect of all discovered variants followed by variant annotation with AnnoVar (21). Notably, each variant was compared against several public databases, gnomAD, 1000 genomes project, NHLBI Exome Sequencing Project 6500 (ESP6500), and Exome Aggregation Consortium (ExAC) to achieve allele frequency in the general population. To identify the known and reported pathogenic variants, each variant was also compared against ClinVar (https://www.ncbi.nlm.nih.gov/clinvar) and Human Gene Mutation Database (HGMD) (www.hgmd.cf.ac.uk). In terms of possible influence on the protein function, variants were evaluated by several popular prediction tools, including InterVar, SIFT, PolyPhen, ClinPred, and GERP++ (22–26). Based on the variant annotations, series of filtering strategies were applied to identify candidate SNVs/Indels associated with EEs, which had been previously described (27).

Detection of CNVs From WES Data

CNVs were detected by CNVkit (28). To create a stable and reliable CNV reference, in-house samples over the same sequencing protocol were selected for reference training in an iterative manner. Specifically, a set of 80 samples was used to create an initial CNV reference. CNV calling process was then run over this reference for each sample in the training set. Samples with any CNV event larger than 1 Mbp were excluded from the next iteration. In the next iteration, new samples were added in to make up the training set to be 80 samples. The iteration process was ended once no <50 samples were qualified. During the reference creation process, circular binary segmentation algorithm was chosen for CNV event segmentation, and the threshold parameter for copy number calling was set to be "−1.6, −0.8, 0.5, 1." For better visualization investigation of CNV events, a tool was designed to plot copy number aligned with B-allele frequencies along the chromosome coordinates. The copy number was shown as $\log_2$ ratio obtained from bin level and segmented level CNV calling results. The B-allele frequencies were computed by samtools mpileup tool. AnnotSV and its annotation databases were locally installed to annotate detected CNV events for each tested sample for the following clinical interpretation (29).

Confirmation and Interpretation of the Candidate Variants

To confirm the candidate SNVs/Indels, PCR amplification of the genomic DNA fragments of infants and their parents was performed, then the samples were sequenced by Sanger sequencing. The candidate CNVs identified were further verified by aCGH such as chromosomal microarray analysis (CMA) or methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA). CMA was performed on Affymetrix GeneChip System 3000Dx v.2 (Thermo Fisher Scientific, MA, USA) by using CytoScan™ HD Array Kit (Thermo Fisher Scientific, MA, USA). Array data were analyzed with Affymetrix Chromosome Analysis Suite Software (Thermo Fisher Scientific, MA, USA). MS-MLPA was carried out using MS-MLPA Kit ME028 (MRC Holland, Amsterdam, Netherlands), and data analyses were performed using Coffalyser Software (MRC Holland, Amsterdam, Netherlands). The American College of Medical Genetics and Genomics/Association for Molecular Pathology (ACMG/AMP) criteria were applied in the interpretation of the pathogenicity of all identified variants (30).

Statistical Analysis

Quantitative variables were expressed as median and interquartile range (IQR). Categorical variables were presented as numbers and percentages. Comparison of categorical variables was analyzed by Fisher's exact-test or chi-squared-test where appropriate. All statistical analyses were performed using SPSS software (version 23.0; SPSS Inc., Chicago, IL, USA).

RESULT

Demographic and Clinical Features of the Infants With EEs

Seventy-three infants with EEs admitted between January 2018 and February 2020 were enrolled in the single pediatric epilepsy center for the study. The median age was 7 months (range 0.03–12 months); 36 (49.32%) of them were males. Clinical features of the 73 infants are listed in **Table 1**. Thirty-two (43.84%) infants had been diagnosed with epilepsy syndrome. The most common type of syndrome was WS (22/73, 30.1%), followed by DS (20/77, 27.4%), OS (1/73, 1.4%), and MPEI (1/73, 1.4%). The remaining 29 patients (39.7%) were diagnosed as unclassified EIEE due to non-specific clinical manifestations. Fifty-four (73.97%) had intellectual development delay. Sixty-three (86.30%) infants had an abnormal record of EEG, and 25 (34.2%) had brain MRI abnormality. Focal seizure (31/73, 42.5%) was the most common onset feature, followed by tonic-clonic seizures (23/73, 31.5%) and tonic seizures (15/73, 20.5%).

Identification of Pathogenic SNVs/Indels

For each sample, an average of 43.9 M (±5.9 M) pairs of 150 bp raw reads of WES were obtained. The exonic target region capture efficiency was 84.5% (±2.5%) on average. The mean sequencing depth on exonic target regions with flanking 10 bp on either side was 151-folds (±20-folds), providing a fairly high coverage of targeted exonic regions. Such high average coverage also ensured the detection of CNVs according to the demand of CVNkit. On average, 80,753 variants were called in each sample, consisting of 72,522 SNVs and 8,231 Indels. After excluding low quality variants with depths lower than 20-folds, about 68,741 SNVs and 7,382 Indels were left per sample.

When interpreted by ACMG recommended standards and analyzed for clinical concordance, 30 SNVs/Indels in 29 infants (39.7%) were graded pathogenic or likely pathogenic variants associated with EEs (**Table 2**). Among these variants, 15 (50.0%) were missenses, and 15 (50.0%) were protein-truncating variants including 4 (13.3%) non-senses, 10 (33.3%) frameshift Indels, and 1 (3.3%) splicing site variant. The majority of the variants were inherited in *de novo* pattern (20, 66.7%), defined as variants present in the infants but not in the parents. Two variants (6.7%) were inherited in X-linked pattern, and one variant (3.3%) in the *PCDH19* gene was inherited in male-sparing pattern. Moreover, 16 (53.3%) variants were previously unreported in variant databases.

These pathogenic or likely pathogenic SNVs/Indels were located in 14 different genes. The genes, encoding sodium and potassium channels of neuron and maintaining neuron excitability, were the prevalent disease genes and account for 22 (30.1%) infants with EEs. Variants of the *SCN1A* gene, which encodes sodium voltage-gated channel alpha subunit 1, accounted for the largest proportion (12/73, 16.4%) of positively detected infants. And variants of the other ion channels genes, *SCN2A*, *SCN8A*, *KCNQ2*, and *KCNT1*, accounted for 2.7%

TABLE 1 | Demographic and baseline clinical characteristics of the patients with EEs.

Demographics	Total ($n = 73$)		Patients with P or LP variants ($n = 36$)		Patients without P or LP variants ($n = 37$)		*P*-value
	n	%	*n*	%	*n*	%	
Male sex	36	49.3	19	52.8	17	45.9	0.642
Intellectual/developmental disabilities	54	74.0	27	75.0	27	73.0	1.000
Multiple congenital anomalies	10	13.7	7	19.4	3	8.1	0.190
Abnormality in EEG	63	86.3					
Abnormality in brain MRI	25	34.2	15	41.6	10	27.0	0.223
Epilepsy diagnosis							
DS	20	27.4	11	30.6	9	24.3	0.607
WS	22	30.1	8	22.2	14	37.8	0.203
OS	1	1.4	1	2.7	0	0	0.493
MPEI	1	1.4	1	2.7	0	0	0.493
EIEE	29	39.7	15	41.6	14	37.8	0.813
Epilepsy features							
Focal seizures	31	42.5	20	55.6	11	29.7	0.034
Tonic seizures	15	20.5	5	13.8	10	27	0.247
Tonic-clonic seizures	23	31.5	10	27.8	13	35.1	0.616
Myoclonic seizures	4	5.5	1	2.78	3	8.1	0.615
Absence seizures	2	2.7	0	0	2	5.4	0.493
Infantile spasms	6	8.2	1	2.78	5	13.5	0.199

DS, Dravet syndrome; WS, West syndrome; OS, ohtahara syndrome; MPEI, malignant migrating partial seizures; EIEE, early infantile epileptic encephalopathy; P, pathogenic; LP, likely pathogenic.

(2/73), respectively. One pathogenic variant in *de novo* status was found in the *CACNA1E* gene, which encodes calcium voltage-gated channel subunit alpha1 E. The disease-causing variants in *CACNA1E* has recently been identified as a cause of developmental and epileptic encephalopathies (31). One variant each was found in both *GABRA1* and *GABRA2* genes, encoding Gamma-aminobutyric acid (GABA) type A receptor subunits. GABA type A receptors are pentameric chloride channels, which are the principal receptors that mediate the inhibitory synaptic transmission in the mammalian brain. Mutations in the genes encoding the GABA receptor subunits have been associated with a spectrum of neurological disorders including EEs. From the current study, four variants were located in the epilepsy-related genes, the *DPYD* gene, *NF1* gene, and *WDR26* gene. Two compound heterozygous variants were found in the *DPYD* gene, which encodes the rate-limiting enzyme for fluoropyrimidine catabolism and eliminates over 80% of administered 5-FU (32). The pathogenic homozygous or compound heterozygous variants within the *DPYD* gene are associated with dihydropyrimidine dehydrogenase (DPD) deficiency, and DPD-deficient infants may develop intellectual disability, motor retardation, and seizures. One non-sense variant was detected within the *NF1* gene in the infant with WS. The *NF1* gene is a tumor suppressor gene; therefore, loss of function due to a mutation leads to increase in cell proliferation and results in the development of neurofibromatosis type 1. Apart from tumors, patients with neurofibromatosis type 1 have an increased risk for WS (33). One novel *de novo* fame-shift variant was identified in the *WDR26* gene, and *WDR26* haploinsufficiency has been linked to a syndrome characterized by intellectual disability, seizures, abnormal gait, and distinctive facial features.

There were 21 candidate variants in 15 infants (20.5%) that were evaluated to be variants of uncertain significance (VUS) by ACMG/AMP criteria (**Supplementary Table 1**). These variants were kept through filtering steps because of pathogenicity prediction inconsistency in multiple perdition tools and limited published literature information.

Identification of Pathogenic CNVs

Through CNV analysis, six CNVs were linked to the clinical phenotype among six cases (8.2%) (**Table 2**). All of these variants were in *de novo* status. Five discovered CNVs were deletions of 2q24.3, 1p36, 15q11–q13, 16p11.2, and 17p13.3. Patient 73, who was a 10-month-old female, was admitted for infantile spasms, and brain MRI disclosed the lissencephaly. Both the CNV analysis and CMA found a *de novo* heterozygous deletion of 17p13.3 which was indicated by a value $\log_2$ ratio $= -1$ and missing value of 0.5 in BAF (**Figure 1**). In the deletion region, the *YWHAE* and *CRK* genes on the telomeric end of chromosome 17p have been considered as epilepsy-causing genes in previous clinical reports, and the deletion was interpreted as pathogenic (34, 35). The clinical findings and genetic test supported the diagnosis of Miller–Dieker syndrome. All these five deletions

TABLE 2 | Profiles of identified pathogenic or likely pathogenic sequence variants in the patients with EEs.

ID	Age (months)	Gender	Clinical phenotype	Coordinates (GRCh37/hg19)	Ref	Gene	Nucleotide substitution	Amino acid substitution	Parental origin	Interpretation	Novel/reported
EE2	3	M	DS	chr2:166900484	NM_001165963	SCN1A	c.1738C>T	p.R580X	*De novo*	P	Reported
EE5	8	M	EIEE	chr1:224619250-224619256	NM_001115113	WDR26	c.550_556delCCTTTAG	p.P184Cfs*9	*De novo*	P	Novel
EE8	2	F	EIEE	chr1:97915746	NM_000110	DPYD	c.1774C>T	p.R592W	Mother	LP	Reported
				chr1:97547896	NM_000110	DPYD	c.2897C>T	p.S966F	Father	LP	Novel
EE7	4	M	DS	chr2:166896016	NM_001165963	SCN1A	c.2506G>T	p.D836Y	*De novo*	LP	Novel
EE9	6	F	DS	chr2:166866296	NM_001165963	SCN1A	c.3934dupA	p.I1312fs	Father	LP	Novel
EE11	7	M	DS	chr2:166870339	NM_001165963	SCN1A	c.3620T>C	p.L1207P	*De novo*	LP	Reported
EE12	6	F	DS	chr2:166896032	NM_001165963	SCN1A	c.2490delA	p.G831Afs*10	NA	LP	Novel
EE14	8	M	WS	chr9:101216454	NM_005458	GABBR2	c.1045G>A	p.V349M	NA	LP	Reported
EE18	8	M	WS	chr1:181745289	NM_001205294	CACNA1E	c.5135G>A	p.R1712Q	*De novo*	LP	Novel
EE20	5	F	DS	chr2:166848251-166848254	NM_001165963	SCN1A	c.5531_5534delTTTG	p.P1844fs	*De novo*	P	Novel
EE21	12	F	EIEE	chr2:166237178	NM_001040143	SCN2A	c.4385delT	p.F1462fs	*De novo*	P	Novel
EE23	6	F	EIEE	chr2:166848621	NM_001165963	SCN1A	c.5164A>G	p.T1722A	*De novo*	LP	Reported
EE32	9	M	EIEE	chr2:166898844	NM_001165963	SCN1A	c.2050C>T	p.R684X	*De novo*	P	Reported
EE33	12	F	EIEE	chrX:99657797	NM_001184880	PCDH19	c.2341delA	p.I781fs	*De novo*	P	Reported
EE39	2	F	MPEI	chr9:138660694	NM_020822	KCNT1	c.1421G>A	p.R474H	Mother	LP	Reported
EE40	1	F	OS	chr20:62070961	NM_004518	KCNQ2	c.917C>T	p.A306V	NA	LP	Reported
EE42	12	M	DS	chr2:166892659-166892660	NM_001165963	SCN1A	c.3327_3328delTG	p.P1109fs	*De novo*	P	Novel
EE43	11	M	EIEE	chr2:166908355	NM_001165963	SCN1A	c.837dupA	p.W280fs	NA	LP	Novel
EE46	1	M	EIEE	chr2:166245951	NM_001040143	SCN2A	c.5635A>G	p.M1879V	*De novo*	LP	Novel
EE47	3	F	EIEE	chr2:166898801	NM_001165964	SCN1A	c.2092+1G>T	NA	*De novo*	P	Novel
EE48	4	F	EIEE	chr17:29588751	NM_000267	NF1	c.4537C>T	p.R1513X	Father	P	Reported
EE49	1	M	EIEE	chr20:62071057	NM_004518	KCNQ2	c.821C>T	p.T274M	*De novo*	P	Reported
EE51	6	M	EIEE	chr2:166019317	NM_001081676	SCN3A	c.716C>A	p.A239D	*De novo*	LP	Novel
EE60	7	F	EIEE	chr12:52082568	NM_001177984	SCN8A	c.641G>A	p.G214D	*De novo*	P	Reported
EE61	2	M	EIEE	chrX:18646629-18646630	NM_001323289	CDKL5	c.2635_2636delCT	p.L879fs	Mother	P	Novel
EE63	3	F	WS	chr5:161317979	NM_000806	GABRA1	c.779C>T	p.P260L	*De novo*	P	Reported
EE65	6	F	EIEE	chr12:52159789	NM_001177984	SCN8A	c.2879T>A	p.V960D	*De novo*	P	Reported
EE66	6	M	DS	chr2:166908376	NM_001202435	SCN1A	c.817C>G	p.L273V	*De novo*	LP	Novel
EE70	3	M	WS	chr9:138660693	NM_020822	KCNT1	c.1420C>T	p.R474C	*De novo*	LP	Reported

DS, Dravet syndrome; WS, West syndrome; OS, ohtahara syndrome; MPEI, malignant migrating partial seizures; EIEE, early infantile epileptic encephalopathy; P, pathogenic; LP, likely pathogenic.

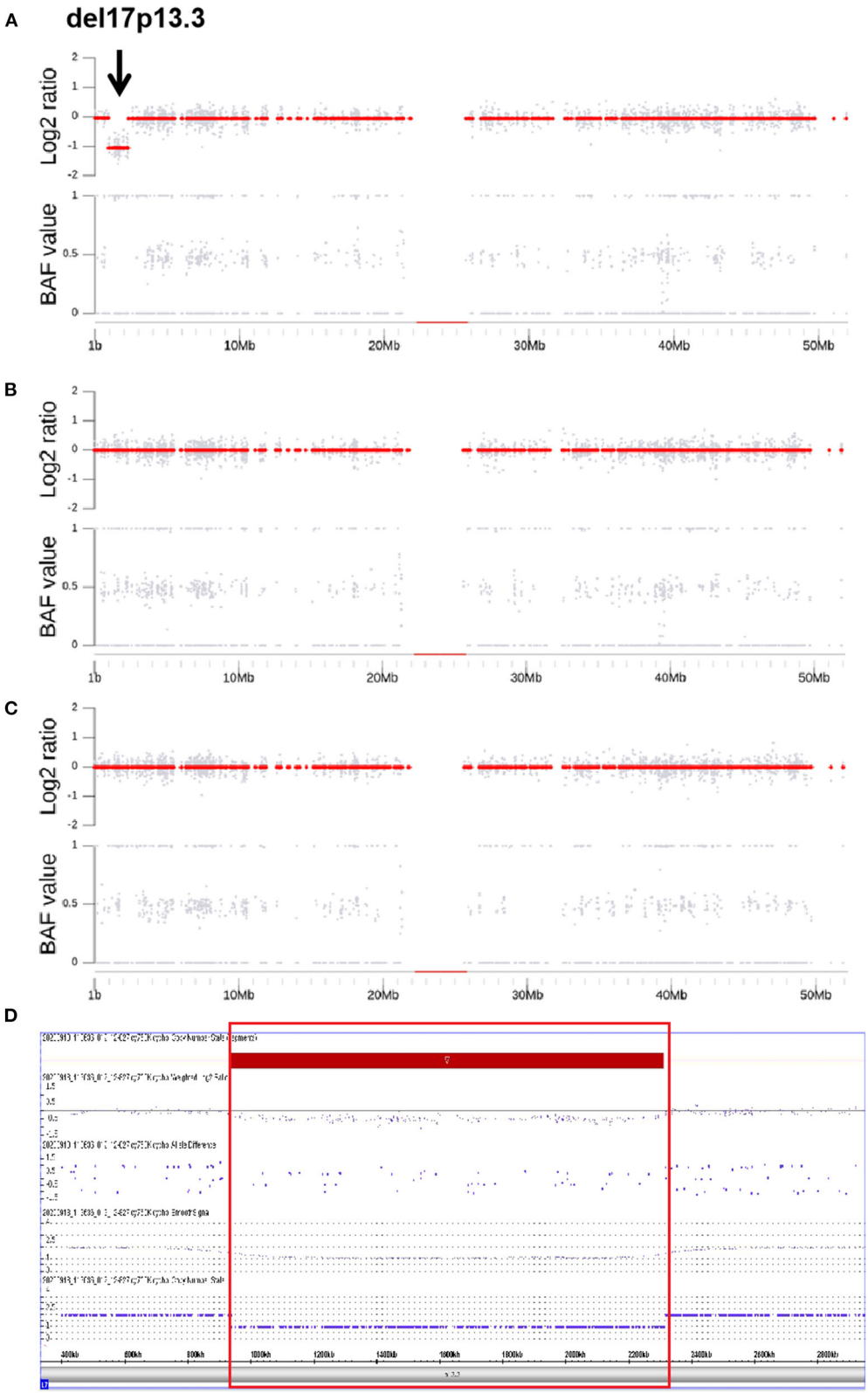

FIGURE 1 | B allele Frequency, Log_2 Ratio plots and CMA of a *de novo* heterozygous deletion of 17p13.3 in the patient with Miller-Dieker syndrome. In **(A–C)**, the upper and the lower subplots are the copy number estimation

(Continued)

FIGURE 1 | and the B allele Frequency (BAF, allelic copy ratio) in the relevant regions. The copy number estimation is provided as $\log_2$ ratio (the upper subplots) which is calculated by $\log_2$ (N_e/N_d), where N_e is the estimated copy number and N_d is the normal copy number of the diploid (i.e., 2 for autosome). The gray dots show the Log_2 Ratios of the bins and the red dots are the ones of the segments. **(A–C)** are the copy number Log_2 Ratio and BAF plots for the proband, father and mother, respectively. The proband **(A)**, the values of Log_2 Ratio in detected region are around −1 and being recognized as a segment (red dots of −1). Meanwhile, the BAF are all closing to 0 or 1, indicating the homozygosity of this segment. A deletion of 17p13.3 is indicated by BAF band along with the lack of change in Log_2 Ratio (black arrow, **A**). Trio analysis confirms that the deletion has occurred *de novo* in the proband. In **(D)**, the deletion was verified by MS-MLPA and the loss in 17p13.3 is indicated by red rectangle.

were classified as pathogenic, which included genes that are well-known to cause epilepsy (**Table 2**).

One large duplication of 4q11–q13.1 was detected in a 12-month-old female with generalized tonic-clonic seizures and growth retardation. A previous study had detected duplication at 4q11–q13.1 in patients with idiopathic short stature, and *TMEM165* and *POLR2B* in this region were suspected as candidate genes for growth retardation (36, 37). The duplication also included the *SRD5A3* gene, the cause of Kahriz syndrome, and the null mutation of *SRD5A3* having been reported to be associated with the epileptic phenotype, WS (38). In the duplicated regions, we also considered the REST gene as another candidate disease-causing gene because of its potential to control fundamental transcription patterns that drive circuit excitability, seizures, and epilepsy (39). However, due to the large size of duplication, it was difficult to determine the disease-causing gene(s), and the duplication was interpreted as VUS.

Altogether, the genetic causes (pathogenic or likely pathogenic variants) of EEs were successfully discovered for 34 (46.6%) infants. Twenty-nine (39.7%) infants carried the SNVs/Indels, and five (6.8%) carried CNVs related to the EEs. And the majority of the disease-causing variants were inherited in *de novo* pattern (25, 71.4%).

Genotype to Phenotype Correlation

In 20 clinical cases of DS, half of them (10/20, 50%) were found to arise from pathogenic or likely pathogenic *SCN1A* variants, and none of the disease-causing variants were identified in the other genes. In contrast, in 22 infants with WS, 8 cases (7/22, 31.8%) were caused by pathogenic or likely pathogenic variants from a variety of genes, including *GABRA1*, *GABRA2*, *KCNT1*, *SCN3A*, *CACNA1E*, *PHGDH*, and *PRRT2* (in del 16p11.2). One infant was diagnosed with OS, and a likely pathogenic variant was found in *KCNQ2*. Additionally, another one infant was diagnosed with MPEI, and a likely pathogenic variant was found in *KCNT1*. As shown in **Table 3**, three pathogenic CNVs del1p36, del15q11–q13, and del17p13.3 corresponded to 1p36 deletion syndrome, Angelman syndrome, and Miller–Dieker syndrome, respectively.

DISCUSSION

In the study, 73 infants with EEs for both SNPs/Indels and CNVs were evaluated by analyzing the WES data generated from a commercial human exome panel, and 34 mutations were found. SNVs/Indels accounted for 39.7% cases, and the detection of CNVs by using CNVkit increased a diagnostic rate of 6.8%. This is the first report of detecting both SNPs/Indels and CNVs in a single test for genetic diagnosis of EEs.

Through the current study, pathogenic and likely pathogenic SNPs/Indels were identified in 14 different disease-causing genes, and pathogenic CNVs were associated with at least 9 potential disease-causing genes. Altogether, 23 genes were involved in the pathogenesis of EEs in the current cohort. Similar to previous studies, ion channel-related genes, especially sodium ion channels and potassium ion channels, accounted for the large proportion of pathogenic and likely pathogenic variants. Indeed, in half of these genes, variants have been recorded to produce truncating proteins (40). One novel pathogenic variant (c.5135G>A/p.R1712Q) in *de novo* status was identified in a calcium channel-related gene, *CACNA1E*, which encodes a functionally critical subunit of a high voltage-activated, rapidly inactivating R-type calcium channel and initiates rapid synaptic transmission in the brain. The infant harboring the variant had intellectual disability and hypsarrhythmia on EEG, consistent with a diagnosis of WS. Helbig et al. had identified *de novo CACNA1E* variants in 30 individuals with EEs, characterized by refractory infantile-onset seizures, severe hypotonia, and profound developmental impairment (31). Most of the pathogenic variant clusters within the cytoplasmic ends of all the four S6 segments (Domains I, II, III, and IV) form the presumed CaV2.3 channel activation gate. The p.R1712Q variants also located in the S6 segment (Domain IV) are supposed to play a gain-of-function effect comprising facilitated voltage-dependent activation and slowed inactivation (31). One novel fame-shift deletion (c.550_556delCCTTTAG/p.P184Cfs*9) in the *WDR26* gene was detected in an infant with absence seizures and febrile seizures, ventriculomegaly, intellectual/developmental disabilities, congenital heart disease, facial dysmorphism, and yellow hair. As WDR26 is located in chromosomal region 1q42, this showed that the patient shared some clinical features with 1q41q42 microdeletion syndrome, such as seizures, intellectual/developmental disabilities, and facial dysmorphism (41). Therefore, the truncation was supposed to lead to a haploinsufficiency of *WDR26* to cause EEs like pathogenic mechanism of 1q41q42 microdeletion (41).

The majority of positive cases with EEs resulted to *de novo* pathogenic or likely pathogenic variants. Twenty SNPs/Indels and five CNVs in *de novo* status accounted for 71.4% positive findings and explained about one third of the infants with EEs overall. However, the current study was unable to exclude the possible parental mosaicism, which remains undetected by current methods. Parental mosaicism had been shown to present significantly higher recurrence risk in future pregnancies than the apparent *de novo* variants. It was identified in ∼10% of families with children with an apparent *de novo SCN1A* variant

TABLE 3 | Profiles of identified rare CNVs in the patients with EEs.

ID	Age at onset (m)	Gender	Cytoband	Clinical diagnosis	Coordinates (GRCh37/hg19)	CNV	Size	Relevant genes	Parental origin	Interpretation	Novel/reported
EE22	8	M	1p36	1p36 deletion syndrome	chr1:736,369-5660,269	Del	4.92 Mb	*GABRD, GNB1, SKI*	*De novo*	P	Reported
EE27	12	F	4q11-q13.1	NA	chr4:52,685,118-65,914,334	Dup	13.23Mb	*SRD5A3, REST*	*De novo*	VUS	Novel
EE38	6	F	2q24.3	Dravet syndrome	chr2:166858942-166850970	Del	7.972kb	*SCN1A*	*De novo*	P	Novel
EE58	10	F	15q11-q13	Angelman syndrome	chr15:22471433-29444050	Del	6.972Mb	*UBE3A,GABRB3*	*De novo*	P	Reported
EE72	3	M	16p11.2	16p11.2 deletion syndrome	chr16:29,675,049-30,199,897	Del	524.848Kb	*PRRT2*	*De novo*	P	Reported
EE73	10	F	17p13.3	Miller-Dieker syndrome	chr17:934,739-2,304,030	Del	1.369Mb	*YWHAE, CRK*	*De novo*	P	Reported

P, pathogenic variant; VUS, a variant of uncertain significance.

in EEs (42). For further genetic counseling with apparent *de novo* variants in EEs, there is a need to consider higher recurrence risk especially for families who may wish to have another baby, who should therefore be subjected to more tests to confirm the status of *de novo*.

In our cohort, WS showed great heterogeneity in genetic etiology, with seven genes accounting for seven cases. And both SNP/Indels or CNVs were identified as causatives in WS. Currently, the routine treatment includes adrenocorticotropic hormone, vigabatrin, and corticosteroids. However, with the understanding of disease-causing genes and related pathogenesis, new therapies that target specific pathways of pathogenesis to correct the underlying molecular dysfunction will be developed in the near future. For example, there are a number of trials including rapamycin/everolimus, targeting the mTOR pathway, for the treatment of tuberous sclerosis complex and related WS (43); retigabine for the treatment of *KCNQ2*-related WS (44); and topiramate for the treatment of *CACNA1E*-related WS (31).

The detection of CNVs by using CNVkit increased the diagnostic rate by 6.8% in our cohort. The traditional approach for detecting genome-wide CNVs includes karyotyping, fluorescence *in situ* hybridization (FISH), SNP array, and aCGH. Among them, aCGH has been proven as an indispensable approach to screen CNVs associated with epilepsy, intellectual disability, developmental delay, or congenital anomalies in children (45). In the past decade, WES, which refers to sequencing of all protein-coding exons, has been widely used in genetic testing in Mendelian disorders as well as other complex diseases (46). It is powerful to detect SNVs/Indels in enriched protein-coding regions. Recently, many bioinformatics tools have been developed to do CNVs calling from WES data. These included XHMM (47), using hidden Markov model, and CoNIFER (48), using singular value decomposition, whereas others such as cn.MOPS use Bayesian inference (49). CNVkit is a toolkit that uses both the targeted reads and the non-specifically captured off-target reads to achieve highly accurate and reliable copy ratio estimates across the genome (28). All these CNV calling tools based on WES data could be considered as a complementary way with only computational effort to improve the diagnostic yields in EEs. In our study, aCGH served as a confirmatory tool rather than a screening tool and could bring cost-effective benefit for clinical diagnosis.

There are various limitations of the current study that need to be addressed. First, phenotypic details for the parents and familial disease history were lacking in some cases, which is a common issue with retrospective studies. Second, the current study approach was to identify known reported disease-causing genes and their rare variants and cloud not identify the variants in new disease-causing genes and common variants as genetic risks for EEs. Third, there were 22 variants (21 SNPs/Indels and 1 CNV) in 16 infants that were interpreted as VUS in EEs. Most of them were novel findings when they were compared with disease variants database or published literature. As more evidence emerges through function analysis, some VUS will be reclassified as the pathogenic variants. All these limitations may influence genetic diagnostic rate under the current study.

In summary, the application of both SNPs/Indels and CNVs detection in a single test based on WES yielded diagnosis in 46.6% of the infants with EEs, which demonstrates the utility of this pipeline as a diagnostic test for pediatric patients with a disease presenting a highly phenotypic and genetic heterogeneity. Importantly, a number of novel variants involved in EEs are being reported here. Although there are technical challenges with CNVs calling as well as challenges for pathogenic classification of identified variants, WES may serve as a first-tier test with cost-effective benefit in EEs.

DATA AVAILABILITY STATEMENT

The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive in National Genomics Data Center, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (https://bigd.big.ac.cn/). After publication of study findings, the data will be available for others to request. The research team will provide accession number of the database once the data are approved to be shared with others. The proposal with description of study objectives will be needed for evaluation of the reasonability to request for the data. The corresponding authors and the Ethical Committee of Wuhan Children's Hospital will make a decision based on these materials.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Ethical Committee of Wuhan Children's Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

FZ and DS conceived, supervised this study, and wrote the manuscript. DS, YaL, WC, and ZL evaluated and cared for the patients. JM, KN, and MC collected clinical data. CW and YZ performed the WES test. FZ and YoL performed the bioinformatics analysis of WES data. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (No. 81570348) and the National Key Research and Development Program of China (No. 2016YFC1306202).

ACKNOWLEDGMENTS

We are thankful to the patients and their families for their participation in the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fped.2021.635703/full#supplementary-material

REFERENCES

1. McTague A, Howell KB, Cross JH, Kurian MA, Scheffer IE. The genetic landscape of the epileptic encephalopathies of infancy and childhood. *Lancet Neurol*. (2016) 15:304–16. doi: 10.1016/S1474-4422(15)00250-1
2. Hamdan FF, Myers CT, Cossette P, Lemay P, Spiegelman D, Laporte AD, et al. High rate of recurrent *de novo* mutations in developmental and epileptic encephalopathies. *Am J Hum Genet*. (2017) 101:664–85. doi: 10.1016/j.ajhg.2017.09.008
3. Helbig I, Tayoun AA. Understanding genotypes and phenotypes in epileptic encephalopathies. *Mol Syndromol*. (2016) 7:172–81. doi: 10.1159/000448530
4. Auvin S, Cilio MR, Vezzani A. Current understanding and neurobiology of epileptic encephalopathies. *Neurobiol Dis*. (2016) 92:72–89. doi: 10.1016/j.nbd.2016.03.007
5. Trivisano M, Specchio N. "Early-onset epileptic encephalopathies," *Clinical Electroencephalography* ed M. Oriano (New York, NY: Springer), (2019) 405–411. doi: 10.1007/978-3-030-04573-9
6. Shbarou R, Mikati MA. The expanding clinical spectrum of genetic pediatric epileptic encephalopathies. *Semin Pediatr Neurol*. (2016) 23:134–42. doi: 10.1016/j.spen.2016.06.002
7. Scheffer IE, Berkovic S, Capovilla G, Connolly MB, French J, Guilhoto L, et al. ILAE classification of the epilepsies: position paper of the ILAE commission for classification and terminology. *Epilepsia*. (2017) 58:512–21. doi: 10.1111/epi.13709
8. Kaplan DI, Isom LL, Petrou S. Role of sodium channels in epilepsy. *Cold Spring Harb Perspect Med*. (2016) 6:a022814. doi: 10.1101/cshperspect.a022814
9. Truty R, Patil N, Sankar R, Sullivan J, Millichap J, Carvill G, et al. Possible precision medicine implications from genetic testing using combined detection of sequence and intragenic copy number variants in a large cohort with childhood epilepsy. *Epilepsia Open*. (2019) 4:397–408. doi: 10.1002/epi4.12348
10. Lin Lee V, Kar Meng Choo B, Chung YS, Kundap UP, Kumari Y, Shaikh MF. Treatment, therapy and management of metabolic epilepsy: a systematic review. *Int J Mol Sci*. (2018) 19:871. doi: 10.3390/ijms19030871
11. Milh M, Riccardi F, Denis J. Genetics of neonatal onset epilepsies: an overview. *Rev Neurol (Paris)*. (2020) 176:2–9. doi: 10.1016/j.neurol.2019.01.396
12. Ostrander BEP, Butterfield RJ, Pedersen BS, Farrell AJ, Layer RM, Ward A, et al. Whole-genome analysis for effective clinical diagnosis and gene discovery in early infantile epileptic encephalopathy. *NPJ Genom Med*. (2018) 3:22. doi: 10.1038/s41525-018-0061-8
13. Palmer EE, Schofield D, Shrestha R, Kandula T, Macintosh R, Lawson JA, et al. Integrating exome sequencing into a diagnostic pathway for epileptic encephalopathy: evidence of clinical utility and cost effectiveness. *Mol Genet Genomic Med*. (2018) 6:186–99. doi: 10.1002/mgg3.355
14. Tsang MH, Leung GK, Ho AC, Yeung KS, Mak CC, Pei SL, et al. Exome sequencing identifies molecular diagnosis in children with drug-resistant epilepsy. *Epilepsia Open*. (2019) 4:63–72. doi: 10.1002/epi4.12282
15. Epi25 Collaborative. Electronic address: s.berkovic@unimelb.edu.au. Ultra-rare genetic variation in the epilepsies: a whole-exome

sequencing study of 17,606 individuals. *Am J Hum Genet*. (2019) 105:267–82. doi: 10.1016/j.ajhg.2019.05.020
16. Striano P, Coppola A, Paravidino R, Malacarne M, Gimelli S, Robbiano A, et al. Clinical significance of rare copy number variations in epilepsy: a case-control survey using microarray-based comparative genomic hybridization. *Arch Neurol*. (2012) 69:322–30. doi: 10.1001/archneurol.2011.1999
17. Epilepsy Phenome/Genome Project Epi KC. Copy number variant analysis from exome data in 349 patients with epileptic encephalopathy. *Ann Neurol*. (2015) 78:323–8. doi: 10.1002/ana.24457
18. Association AP. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub (2013).
19. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* (Philadelphia: Oxford, England). (2009) 25:1754–60. doi: 10.1093/bioinformatics/btp324
20. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. (2010) 20:1297–303. doi: 10.1101/gr.107524.110
21. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The ensembl variant effect predictor. *Genome Biol*. (2016) 17:122. doi: 10.1186/s13059-016-0974-4
22. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*. (2003) 31:3812–4. doi: 10.1093/nar/gkg509
23. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. (2010) 7:248–9. doi: 10.1038/nmeth0410-248
24. Li Q, Wang K. InterVar: clinical interpretation of genetic variants by the 2015 ACMG-AMP guidelines. *Am J Hum Genet*. (2017) 100:267–80. doi: 10.1016/j.ajhg.2017.01.004
25. Alirezaie N, Kernohan KD, Hartley T, Majewski J, Hocking TD. ClinPred: prediction tool to identify disease-relevant nonsynonymous single-nucleotide variants. *Am J Hum Genet*. (2018) 103:474–83. doi: 10.1016/j.ajhg.2018.08.005
26. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol*. (2010) 6:e1001025. doi: 10.1371/journal.pcbi.1001025
27. Zhou P, He N, Zhang JW, Lin ZJ, Wang J, Yan LM, et al. Novel mutations and phenotypes of epilepsy-associated genes in epileptic encephalopathies. *Genes Brain Behav*. (2018) 17:e12456. doi: 10.1111/gbb.12456
28. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput Biol*. (2016) 12:e1004873. doi: 10.1371/journal.pcbi.1004873
29. Geoffroy V, Herenger Y, Kress A, Stoetzel C, Piton A, Dollfus H, et al. AnnotSV: an integrated tool for structural variations annotation. *Bioinformatics (Oxford, England)*. (2018) 34:3572–4. doi: 10.1093/bioinformatics/bty304
30. 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
31. Helbig KL, Lauerer RJ, Bahr JC, Souza IA, Myers CT, Uysal B, et al. *De novo* pathogenic variants in CACNA1E cause developmental and epileptic encephalopathy with contractures, macrocephaly, and dyskinesias. *Am J Hum Genet*. (2018) 103:666–78. doi: 10.1016/j.ajhg.2018.09.006
32. Wigle TJ, Tsvetkova EV, Welch SA, Kim RB. DPYD and fluorouracil-based chemotherapy: mini review and case report. *Pharmaceutics*. (2019) 11:199. doi: 10.3390/pharmaceutics11050199
33. Ruggieri M, Iannetti P, Clementi M, Polizzi A, Incorpora G, Spalice A, et al. Neurofibromatosis type 1 and infantile spasms. *Childs Nerv Syst*. (2009) 25:211–6. doi: 10.1007/s00381-008-0706-5
34. Noor A, Bogatan S, Watkins N, Meschino WS, Stavropoulos DJ. Disruption of YWHAE gene at 17p13.3 causes learning disabilities and brain abnormalities. *Clin Genet*. (2018) 93:365–7. doi: 10.1111/cge.13056
35. Blazejewski SM, Bennison SA, Smith TH, Toyo-Oka K. Neurodevelopmental genetic diseases associated with microdeletions and microduplications of chromosome 17p13.3. *Front Genet*. (2018) 9:80. doi: 10.3389/fgene.2018.00080
36. Schulte Althoff S, Gruneberg M, Reunert J, Park JH, Rust S, Muhlhausen C, et al. TMEM165 deficiency: postnatal changes in glycosylation. *JIMD Rep*. (2016) 26:21–9. doi: 10.1007/8904_2015_455
37. Hu G, Fan Y, Wang L, Yao RE, Huang X, Shen Y, et al. Copy number variations in 119 Chinese children with idiopathic short stature identified by the custom genome-wide microarray. *Mol Cytogenet*. (2016) 9:16. doi: 10.1186/s13039-016-0225-0
38. Tuysuz B, Pehlivan D, Ozkok A, Jhangiani S, Yalcinkaya C, Zeybek CA, et al. Phenotypic expansion of congenital disorder of glycosylation due to SRD5A3 null mutation. *JIMD Rep*. (2016) 26:7–12. doi: 10.1007/8904_2015_478
39. Hwang JY, Zukin RS. REST, a master transcriptional regulator in neurodegenerative disease. *Curr Opin Neurobiol*. (2018) 48:193–200. doi: 10.1016/j.conb.2017.12.008
40. Oyrer J, Maljevic S, Scheffer IE, Berkovic SF, Petrou S, Reid CA. Ion channels in genetic epilepsy: from genes and mechanisms to disease-targeted therapies. *Pharmacol Rev*. (2018) 70:142–73. doi: 10.1124/pr.117.014456
41. Skraban CM, Wells CF, Markose P, Cho MT, Nesbitt AI, Au PYB, et al. WDR26 haploinsufficiency causes a recognizable syndrome of intellectual disability, seizures, abnormal gait, and distinctive facial features. *Am J Hum Genet*. (2017) 101:139–48. doi: 10.1016/j.ajhg.2017.06.002
42. Epi KC. *De novo* mutations in SLC1A2 and CACNA1A are important causes of epileptic encephalopathies. *Am J Hum Genet*. (2016) 99:287–98. doi: 10.1016/j.ajhg.2016.06.003
43. Franz DN, Belousova E, Sparagana S, Bebin EM, Frost MD, Kuperman R, et al. Long-term use of everolimus in patients with tuberous sclerosis complex: final results from the EXIST-1 study. *PLoS ONE*. (2016) 11:e0158476. doi: 10.1371/journal.pone.0158476
44. Kalappa BI, Soh H, Duignan KM, Furuya T, Edwards S, Tzingounis AV, et al. Potent KCNQ2/3-specific channel activator suppresses *in vivo* epileptic activity and prevents the development of tinnitus. *J Neurosci*. (2015) 35:8829–42. doi: 10.1523/JNEUROSCI.5176-14.2015
45. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet*. (2010) 86:749–64. doi: 10.1016/j.ajhg.2010.04.006
46. Han JY, Lee IG. Genetic tests by next-generation sequencing in children with developmental delay and/or intellectual disability. *Clin Exp Pediatr*. (2020) 63:195–202. doi: 10.3345/kjp.2019.00808
47. Fromer M, Purcell SM. Using XHMM software to detect copy number variation in whole-exome sequencing data. *Curr Protoc Hum Genet*. (2014) 81:7.23.1–21. doi: 10.1002/0471142905.hg0723s81
48. Krumm N, Sudmant PH, Ko A, O'Roak BJ, Malig M, Coe BP, et al. Copy number variation detection and genotyping from exome sequence data. *Genome Res*. (2012) 22:1525–32. doi: 10.1101/gr.138115.112
49. Klambauer G, Schwarzbauer K, Mayr A, Clevert DA, Mitterecker A, Bodenhofer U, et al. cn.MOPS: mixture of poissons for discovering copy number variations in next-generation sequencing data with a low false discovery rate. *Nucleic Acids Res*. (2012) 40:e69. doi: 10.1093/nar/gks003

Conflict of Interest: YoL and YZ are employees of company Aegicare Technology Co., Ltd. (Shenzhen, China).

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

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 07 June 2021
doi: 10.3389/fphar.2021.658072

Comparative Analysis for the Performance of Long-Read-Based Structural Variation Detection Pipelines in Tandem Repeat Regions

*Mingkun Guo[1], Shihai Li[1], Yifan Zhou[1], Menglong Li[1] and Zhining Wen[1,2]**

[1]College of Chemistry, Sichuan University, Chengdu, China, [2]Medical Big Data Center, Sichuan University, Chengdu, China

There has been growing recognition of the vital links between structural variations (SVs) and diverse diseases. Research suggests that, with much longer DNA fragments and abundant contextual information, long-read technologies have advantages in SV detection even in complex repetitive regions. So far, several pipelines for calling SVs from long-read sequencing data have been proposed and used in human genome research. However, the performance of these pipelines is still lack of deep exploration and adequate comparison. In this study, we comprehensively evaluated the performance of three commonly used long-read SV detection pipelines, namely PBSV, Sniffles and PBHoney, especially the performance on detecting the SVs in tandem repeat regions (TRRs). Evaluated by using a robust benchmark for germline SV detection as the gold standard, we thoroughly estimated the precision, recall and F1 score of insertions and deletions detected by the pipelines. Our results revealed that all these pipelines clearly exhibited better performance outside TRRs than that in TRRs. The F1 scores of Sniffles in and outside TRRs were 0.60 and 0.76, respectively. The performance of PBSV was similar to that of Sniffles, and was generally higher than that of PBHoney. In conclusion, our findings can be benefit for choosing the appropriate pipelines in real practice and are good complementary to the application of long-read sequencing technologies in the research of rare diseases.

Keywords: structural variation, tandem repeats, detection pipelines evaluation, rare diseases, long-read sequencing

OPEN ACCESS

Edited by:
Zhichao Liu,
National Center for Toxicological Research (FDA), United States

Reviewed by:
Zhimin Liu,
Janssen Pharmaceuticals, Inc., United States
Bohu Pan,
National Center for Toxicological Research (FDA), United States

****Correspondence:***
Zhining Wen
w_zhining@163.com

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *25 January 2021*
Accepted: *14 May 2021*
Published: *07 June 2021*

Citation:
Guo M, Li S, Zhou Y, Li M and Wen Z (2021) Comparative Analysis for the Performance of Long-Read-Based Structural Variation Detection Pipelines in Tandem Repeat Regions. Front. Pharmacol. 12:658072. doi: 10.3389/fphar.2021.658072

INTRODUCTION

Previous studies typically defined structural variations as genomic changes at least 50 base pairs (bp) in size. SVs are closely related to diverse human diseases Weischenfeldt et al. (2013); Lupski, (2015), such as autism Pinto et al. (2010); Sanders et al. (2012); Chen et al. (2017) and schizophrenia (Sebat et al., 2007; Stefansson et al., 2008; Walsh et al., 2008; Kirov et al., 2012). Compared with single-nucleotide variations (SNVs), SVs contain more nucleotides and are considered to be higher correlated with evolution, genetic diversity and disease-causing mutations (Stankiewicz and Lupski, 2010; Weischenfeldt et al., 2013; Abel et al., 2020).

Since the size of SV can exceed 1,000 bp, SV detection will be limited by the size of DNA fragments in sequencing. Furthermore, if SVs occur in repetitive regions with high mutation rate, it will be more difficult for detection (Hills et al., 2007; Hastings et al., 2009; Hodgkinson et al., 2012).

In view of the above problems, short-read data may have some difficulties while long-read data can be a good solution (Pollard et al., 2018; Liu et al., 2019). In recent years, long-read technologies

have been developed rapidly Amarasinghe et al. (2020) and used in the discovery of SVs with complex forms (Aneichyk et al., 2018; Song et al., 2018; Ishiura et al., 2019; Zeng et al., 2019; Logsdon et al., 2020). The size of DNA fragment sequenced by long-read technologies is usually larger than 1,000 bp, which can cover the range of large SV and contain much context information (Chaisson et al., 2015). It ensures the advantages of long-read technologies in SV detection, especially in the complex repetitive regions of the genome. Characterized by high incidence rate of SVs and high complexity, repetitive regions are an important and challenging problem in SV detection (Sudmant et al., 2015; Zook et al., 2020). However, the performance of SV detection pipelines based on long-read data applied in repetitive regions still need to be analyzed.

Therefore, in this study, we selected three commonly used long-read-based pipelines Kosugi et al. (2019); Logsdon et al. (2020), namely PBSV Wenger et al. (2019), Sniffles Sedlazeck et al. (2018) and PBHoney English et al. (2014), and comprehensively evaluated their performance on SV detection. Using the benchmark established by the Genome in a Bottle (GIAB) Consortium Zook et al. (2020) as the gold standard, we evaluated the precision, recall and F1 score of these pipelines. The comparison included the comparison between insertions and deletions, the comparison among four size ranges of SVs and the comparison between SVs in TRRs and SVs outside TRRs. The F1 scores of Sniffles were 0.60 in TRRs and 0.76 outside TRRs. Similarly, The F1 scores of PBSV were 0.59 and 0.74 in and outside TRRs, respectively. The performances of the two pipelines were generally higher than that of PBHoney. For the three pipelines, the performances in TRRs were lower than those outside TRRs, which indicated that SV detection in TRRs was more difficult than that outside TRRs. Concerning the type of SVs, it was found that large insertions (> 1,000 bp) were the most difficult to detect while large deletions were easy to precisely detect, especially in TRRs. In addition, we also analyzed the potential performance of three pipelines on detecting *de novo* SVs. The results suggested that long-read technologies and the SV detection pipelines still need further development for the precise detection of *de novo* SVs.

MATERIALS AND METHODS

Datasets

The long-read sequencing data of an Ashkenazim Jewish trio Zook et al. (2016) were used in our study. Subreads datasets of the son (HG002), the father (HG003) and the mother (HG004) were downloaded from GIAB (https://ftp-trace.ncbi.nlm.nih.gov/giab/ftp/data/AshkenazimTrio/). The average coverages of the trio are approximately 69X, 32X and 30X, and their N50 subread lengths are 11,087, 10,728, and 10,629 bp.

Benchmark

The benchmark is established by GIAB for HG002 on GRCh37, which was downloaded from GIAB FTP site (https://ftp-trace.ncbi.nlm.nih.gov/ReferenceSamples/giab/data/AshkenazimTrio/analysis/NIST_SVs_Integration_v0.6/). The benchmark dataset contains close to 100% of true insertions and deletions in the specific regions. According to the guidance of the benchmark, we used the SVs with the FILTER field "PASS" in the Tier 1 vcf, including 12,745 isolated, sequence-resolved insertion (7,281) and deletion (5,464) calls. The benchmark regions include 34,830 large regions, of which 15% are within 1,000–10,000 bp and 82% are over 10,000 bp. Through the manual inspection in the benchmark work, it was found that approximately 5% of true insertions in the benchmark regions might be missing. Therefore, when comparing callsets (especially from long-read data) with the benchmark, it is possible to misjudge some true insertions. When making the comparison, we first selected the SVs in the benchmark regions, and then compared these SVs with the benchmark SVs.

Structural Variation Detection Pipelines

In this study, we used three long-read-based pipelines named PBSV (version 2.2.2; https://github.com/PacificBiosciences/pbsv), Sniffles (version 1.0.11; https://github.com/fritzsedlazeck/Sniffles) and PBHoney (in PBSuite-15.8.24; http://sourceforge.net/projects/pb-jelly/). For PBSV and Sniffles, subreads were aligned to reference genomes GRCh37 by PBMM2 and NGMLR, respectively. After the help of SAMtools, SVs were called by PBSV and Sniffles. PBHoney includes two parts of results, namely Tails (based on interrupted mapping) and Spots (based on intra-read discordance). There were too few results in the Tails part to compare with other pipelines, thus the result of the Tails part was separately shown in **Supplementary Figure S1**. Because of the complexity of parameter optimization in Spots and the time-consume of recommended aligner BLASR, following a previous work Kosugi et al. (2019), we used NGMLR to align the subreads and detected SVs with custom-made parameters for insertions and deletions. SVs with < 0.2 of the value, which was calculated by dividing the szCount tag with the coverage tag, were filtered out.

In these callsets, we only summarized the variations ≥ 50 bp. SVs with the type "BND" (breakpoint end) were excluded. In this study, we only focused on the SVs on the autosomes and sex chromosomes.

The Metrics for Comparison

During comparison, we mainly considered the type consistency, the distance between breakpoints and the proportion of the reciprocal overlap. For compared insertions, if the distance of breakpoints was within 200 bp, they were considered the same. For compared deletions, the called SV needed to exhibit ≥ 50% reciprocal overlap with the reference SV. When comparing the callset with regions (i.e. the benchmark regions and tandem repeat regions), it was only required that breakpoints overlapped with these regions. When comparing the overlap among pipelines, the callset with more SVs was chosen as the comparison benchmark. The code used for comparison are available at GitHub (https://github.com/cic-gmk/DNSV).

When comparing the callsets with the benchmark, the precision, recall and F1 score were calculated *via* the following equations:

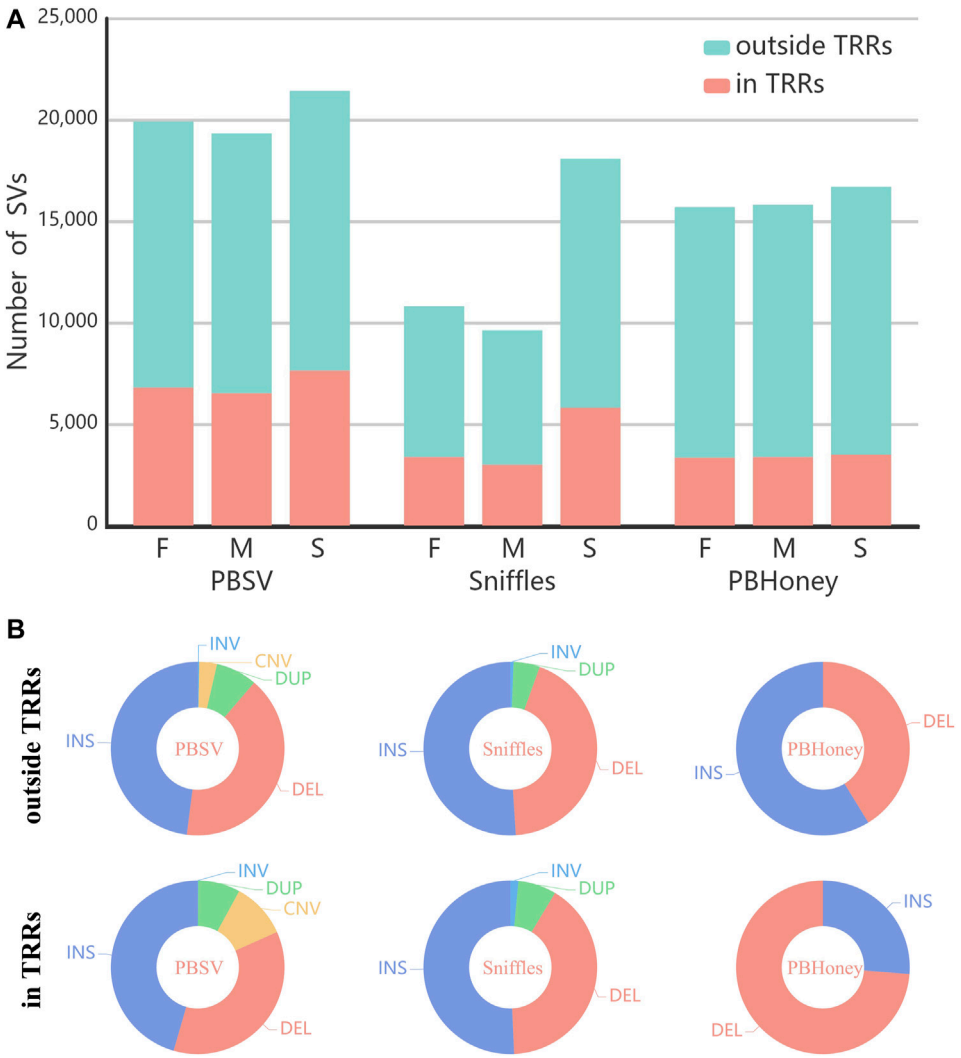

FIGURE 1 | (A) The number of SVs in the callsets of the parents and the son (F, Father; M, Mother; S, Son) detected by PBSV, Sniffles and PBHoney. **(B)** The type distribution of the whole callsets of the trio detected by three pipelines in/outside TRRs.

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

where TP, FP and FN are the numbers of true positives, false positives and false negatives. TP + FP is equal to the number of the called SVs. TP + FN is equal to the number of the benchmark SVs.

Tandem Repeat Regions

The repeats used in our study were annotated in the annotation file of hg19, which can be obtained at the download site of UCSC Genome Browser (Fernandes et al., 2020; Navarro Gonzalez et al., 2020) (http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/rmsk.txt.gz). "Simple repeats" and "Satellites" were selected as the TRRs from the file. "Simple repeats" are short pattern tandem repeats and "Satellites" are medium to long pattern tandem repeats. SVs were divided into two parts according to whether they were in TRRs or not.

RESULTS

The Landscape of Structural Variation Callsets

The numbers of SVs detected by the three pipelines are shown in **Figure 1A**. Among the three pipelines, PBSV detected the largest number of SVs and Sniffles detected the least number of SVs. For all the pipelines, the numbers of detected SVs of the son were more than those of the parents mainly due to the higher coverage of the sequencing data of the son. For Sniffles, largest difference in

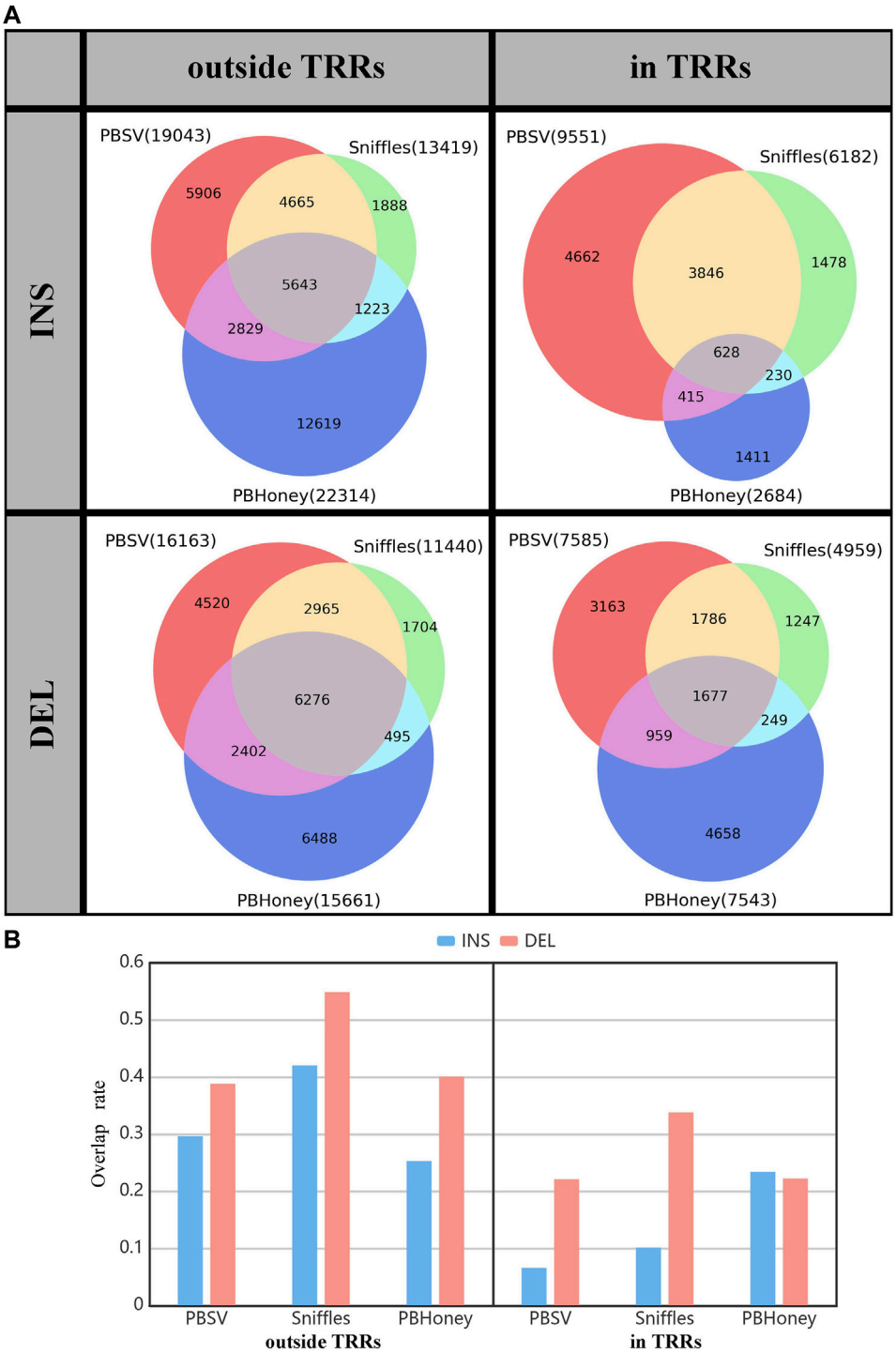

FIGURE 2 | (A) The overlap among callsets detected by three pipelines divided by the type and the location relation with TRRs of SVs. **(B)** The overlap rate of three pipelines. The gray part in **Figure 2A** shows the overlap SVs.

the numbers of detected SVs existed between the son and the parents. Because of the high mutation rate of TRRs Hills et al. (2007); Hastings et al. (2009), although the abundance of TRRs accounts for only about 10% of the human genome Benson, (1999), a number of SVs were still detected in TRRs in the callsets of the trio by three pipelines (PBSV 35%, Sniffles 32%, PBHoney 21%). **Figure 1B** shows the type distribution of all SVs detected by each of the three pipelines. Although insertions are more difficult

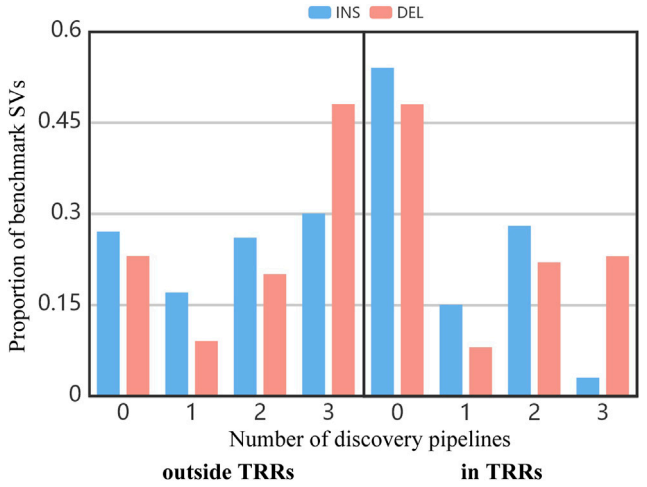

FIGURE 3 | The proportion of benchmark SVs concurrently detected by different number of pipelines.

to detect Zook et al. (2020), the proportions of insertions detected by three pipelines were higher than those of deletions, except for PBHoney in TRRs. The size distribution of SVs detected by each of the three pipelines is provided in **Supplementary Figure S2**. It was found that the number of detected SVs decreased fast as the size of SVs increased. Insertions were generally in the majority when the size < 1,000 bp, but the proportion of deletions increased with the increase of size. In addition, we also investigated the distribution of the percentage of SVs across chromosomes for pipelines (**Supplementary Figure S3**).

We summed the overlap among the callsets of each person detected by three pipelines for comparison (**Figure 2**). For the SVs outside TRRs, the overlap proportion of SVs detected by Sniffles was the highest, and close to 42% (5,643/13,419) of insertions and 55% (6,267/11,440) of deletions can be detected by the other two pipelines. For SVs in TRRs, when comparing Sniffles with PBSV, about 72% ((3,846 + 628)/6,182) of insertions and 70% ((1,786 + 1,677)/4,959) of deletions identified by Sniffles can be detected by PBSV. However, only 628 insertions detected by PBHoney in TRRs were involved in the callsets of PBSV and Sniffles due to the insufficient ability of PBHoney for detecting insertion in TRRs. It can be seen from **Figure 2B** that, except for PBHoney in TRRs, the overlap rates of insertions were lower than those of deletions. For the three pipelines, the overlap rates in TRRs were lower than those outside TRRs, suggesting that the difference among the callsets from different pipelines in TRRs was large.

Evaluation on the Performance of Pipelines

The benchmark used in our study defines the comparing regions, in which the benchmark contains close to 100% of true insertions and deletions. Therefore, we compared the callsets detected by three pipelines with the benchmark callset in the comparing regions.

Figure 3 shows the proportion of benchmark SVs concurrently detected by different number of pipelines. In the whole benchmark callset, close to 25% of benchmark SVs outside TRRs and 51% of benchmark SVs in TRRs cannot be discovered by any pipeline (the bar of "0"). It suggested that SVs in TRRs were more difficult to detect. Similarly, the proportion of benchmark SVs concurrently detected by three pipelines in TRRs was obviously lower than that outside TRRs (the bar of "3"), which agreed with the overlap results among three pipelines (**Figure 2**).

Using the benchmark as the golden standard, the precision, recall and F1 score of three pipelines are shown in **Figure 4**. The precisions achieved by PBSV and Sniffles were higher than 80% both in and outside TRRs, indicating that the SVs detected by these two pipelines were relatively precise. The precision of PBHoney was the lowest, suggesting that more false positives existed in the callset of PBHoney. For all the pipelines, the recalls were under 80 and 50% outside and in TRRs, respectively. It suggested that there were still a number of SVs omitted by the three pipelines. The recall of insertions detected by PBHoney in TRRs was especially low (8%), suggesting that its detection ability of insertions in TRRs was suboptimal. For all the three pipelines, the F1 scores in TRRs were obviously lower than those outside TRRs, indicating the detection of SVs in TRRs was more challenging. In addition, because the son's SVs are inherited from the parents, we also made comparison between the callsets of the parents and the benchmark (**Supplementary Figure S4**). The nominal precisions of the callsets of the parents were clearly lower than those of the son mainly because the benchmark was constructed only based on the sequencing data of the son. It suggested that the benchmark construction in future need to consider the diversity of the population.

Figure 5 shows the impact of the size of SVs on the detection ability of three pipeline. The F1 scores of PBSV and Sniffles were relatively stable with the increase of the size of SVs. However, the size of SVs induced a clear impact on PBHoney. Especially when the size of SVs was more than 1,000 bp, PBHoney can hardly detect true insertions. For all the three pipelines, the F1 scores of large insertion detection (>1,000 bp) were obviously lower than those of large deletion detection, suggesting that the detection for large insertions were more challenging.

Potential Performance of Three Pipelines on Detecting *de novo* Structural Variations

The mutations that only occurred in the child rather than the parents are generally called *de novo* mutations (Conrad et al., 2011; Veltman and Brunner, 2012). We calculated the rate of "*de novo*" SVs by dividing the number of the SVs only detected from the son by the number of all SVs detected from the son. As shown in **Figure 6**, the "*de novo*" rate of Sniffles was 32%, which was higher than that of PBSV (18%) and PBHoney (14%). These rates were much higher than the actual *de novo* rate (Veltman and Brunner, 2012). It indicated that a large number of false positive *de novo* SVs existed in the callsets, which may also be attributed to Mendelian inheritance errors Pilipenko et al. (2014); Kothiyal et al. (2019), the false positive SVs of the son and the false negative SVs of the parents. Our results suggested that long-read

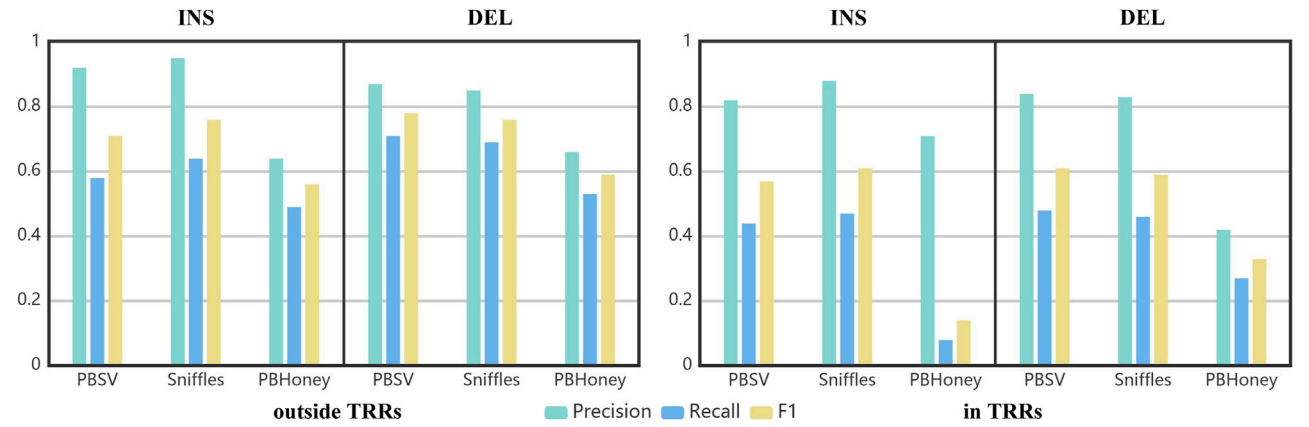

FIGURE 4 | The precision, recall and F1 of three pipelines in detecting insertions and deletions in/outside TRRs.

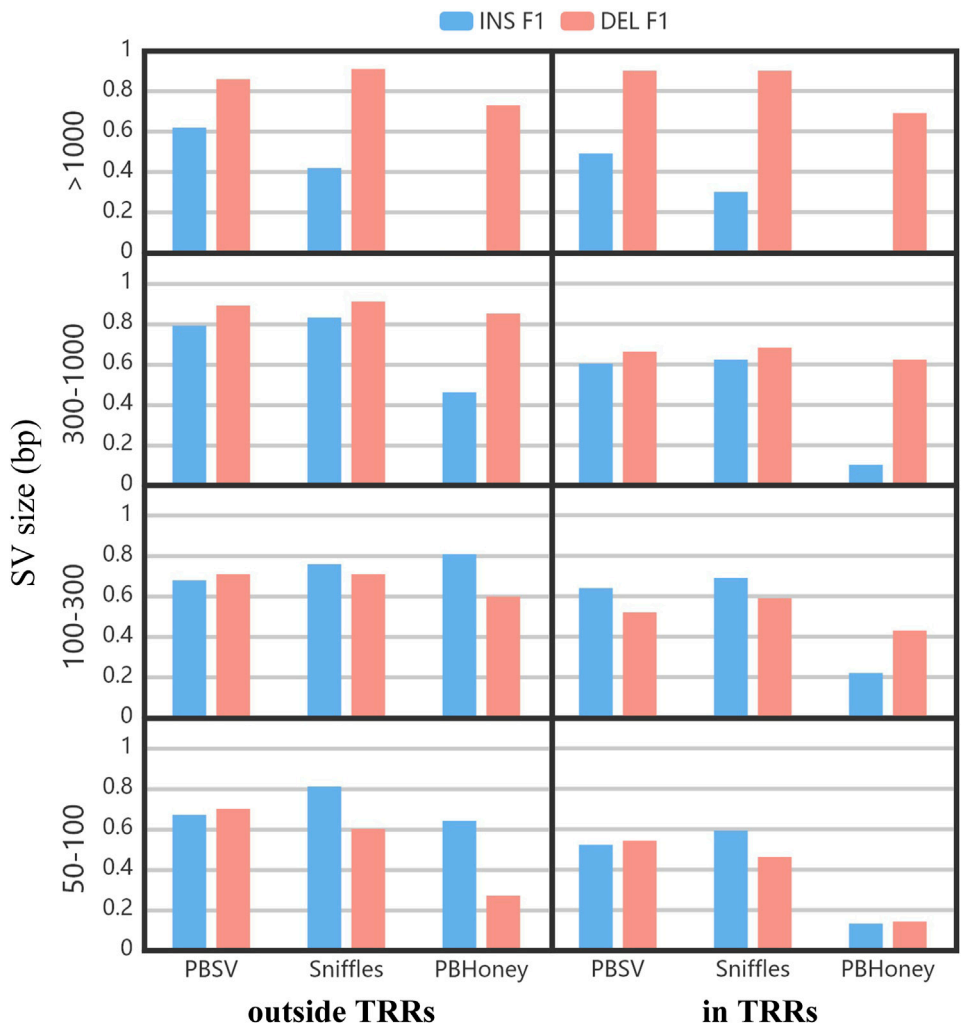

FIGURE 5 | The F1 of insertions and deletions divided by the size and the location relation with TRRs of SVs. The size ranges included 50–100 bp, 100–300 bp, 300–1,000 bp and > 1,000 bp.

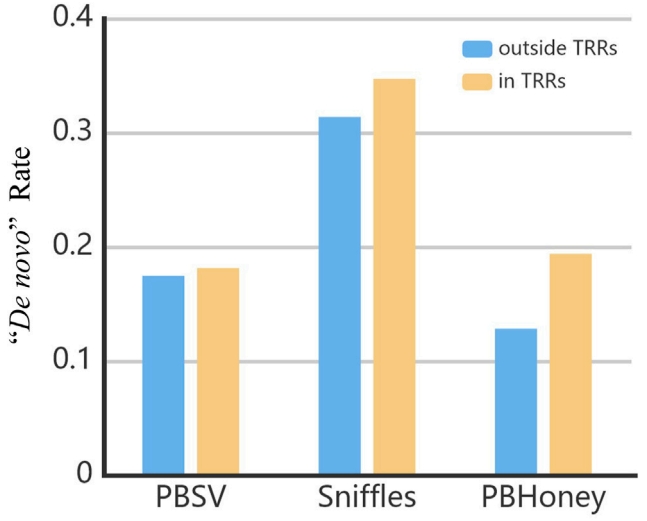

FIGURE 6 | The "*de novo*" rate in the callsets of three pipelines in and outside TRRs.

technologies and the detection pipelines still need further improvement for detecting *de novo* SVs when applying to the exploration of mechanisms of rare diseases. It is a considerable challenge to reduce the false positives and false negatives in SV detection in future study.

DISCUSSION

The relatively large size of SVs and the complex repetitive context make SV detection challenging. Because long-read sequencing data contain abundant context information, it can perform well in SV detection. Therefore, we comprehensively analyzed the performance of three commonly used long-read SV detection pipelines. Our results showed that the overlap proportion among the callsets of the pipelines in TRRs was generally lower than that outside TRRs. Comparing callsets with the benchmark, the precisions, recalls and F1 scores of these pipelines in TRRs were obviously lower than those outside TRRs. These results suggested that the detection of SVs in TRRs was more difficult than that outside TRRs.

As shown in **Figure 4**, the F1 scores of PBSV and Sniffles were similar, and higher than that of PBHoney. With the default recommended parameter, preferable results can be obtained by PBSV and Sniffles. As shown in **Figure 5**, the F1 scores of PBSV and Sniffles did not change a lot with the increase of the size of SVs except for the detection of insertions larger than 1,000 bp. But the detection of both insertions and deletions with PBHoney was clearly influenced by the size of SVs. In fact, as shown in **Supplementary Figure S2**, it was difficult for PBHoney to detect SVs above 4,000 bp. For PBHoney, it was necessary to make proper settings and filter process for SVs with different types. In addition, there were 3% of the son's callset of PBSV and 45% of the son's callset of Sniffles marked with the label "IMPRECISE", which indicated the probably insufficient precision of SVs. These SVs were mainly composed of insertions (95% for PBSV and 71% for Sniffles). Interestingly, the precision of SVs tagged "IMPRECISE" in PBSV was really low (15%), but for Sniffles, the precision was still high (81%), which meant there was no need to filter these SVs in Sniffles.

In our study, we selected SVs (≥ 50 bp) from the benchmark for comparison. If more variations was selected, such as using the cutoff of variations ≥ 30 bp, it would lead to higher precision and lower recall (Kosugi et al., 2019). Our results showed that, even with high precision, no pipeline can achieve very high recall in SV detection. Therefore, it is necessary to integrate different pipelines for generating a comprehensive callset. Integrating Sniffles and PBHoney, NextSV Fang et al. (2018) had been developed to detect SVs from low-coverage long-read sequencing data and achieved better performance than a single pipeline. In addition, a pipeline with multiple algorithms can be developed to optimize and simplify the procession of SV detection.

Previous studies have found that long-read sequencing can identify pathogenic SVs of rare genetic diseases which cannot be identified by short-read sequencing, such as the pathogenic SVs of Carney complex Merker et al. (2018) and progressive myoclonic epilepsy (Mizuguchi et al., 2019). In the study of Carney complex, the pathogenic SV was identified by pipeline detection followed by manual screening and analysis. Since the SV was not detected from the parents, the pathogenic SV was also proved to be a *de novo* SV. However, it was difficult to identify pathogenic SVs from "*de novo*" SVs using long-read sequencing. It is known that the number of *de novo* mutations in heredity is very small (Conrad et al., 2011; Veltman and Brunner, 2012). But compared with the number of *de novo* single-nucleotide variations detected from short-read data Liang et al. (2019), the number of "*de novo*" SVs detected from long-read data was too large. And the precision was too high when comparing the "*de novo*" SVs with the benchmark (**Supplementary Figure S5**). Therefore, these SVs cannot be simply regarded as true *de novo* SVs. In order to analyze true *de novo* SVs, more true SVs need to be detected from parents. As shown in **Figure 1A**, in each pipeline, the numbers of SVs of the parents were obviously smaller than that of the son due to the lower coverages of the parents. It suggested that the sequencing coverages of the trio need to be ensured. Also, for high precision and recall of the detection of *de novo* SVs, the precision and recall of SV detection pipelines still need to be improved.

CONCLUSION

In this study, we thoroughly compared three commonly used SV detection pipelines and found that the precisions of PBSV and Sniffles were generally similar, and higher than PBHoney. The recalls of the three pipelines were still suboptimal. The performances of PBSV and Sniffles were relatively stable with the increase of the size of SVs, while the performance of PBHoney varied largely. The performances of the three pipelines in TRRs were obviously lower than those outside TRRs, indicating that SV detection in TRRs was more difficult. Comparing insertions with deletions, the detection of large insertions was obviously more

difficult than that of large deletions. Our findings can be helpful for conducting the SV detection in the mechanism exploration of rare diseases.

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.

AUTHOR CONTRIBUTIONS

ZW designed the experiments. MG did the entire experiment. MG, SL, and YZ analyzed the data. ZW and MG wrote the main manuscript text and prepared all the figures. All authors discussed the results and revised the manuscript. All authors have read and approved the final manuscript.

FUNDING

This project was supported by the grant from the National Natural Science Foundation of China (No. 21575094).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.658072/full#supplementary-material

REFERENCES

Abel, H. J., Larson, D. E., Larson, D. E., Regier, A. A., Chiang, C., Das, I., et al. (2020). Mapping and Characterization of Structural Variation in 17,795 Human Genomes. *Nature* 583, 83–89. doi:10.1038/s41586-020-2371-0

Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., and Gouil, Q. (2020). Opportunities and Challenges in Long-Read Sequencing Data Analysis. *Genome Biol.* 21 (1), 30. doi:10.1186/s13059-020-1935-5

Aneichyk, T., Hendriks, W. T., Yadav, R., Shin, D., Gao, D., Vaine, C. A., et al. (2018). Dissecting the Causal Mechanism of X-Linked Dystonia-Parkinsonism by Integrating Genome and Transcriptome Assembly. *Cell* 172, 897–909. doi:10.1016/j.cell.2018.02.011

Benson, G. (1999). Tandem Repeats Finder: a Program to Analyze DNA Sequences. *Nucleic Acids Res.* 27, 573–580. doi:10.1093/nar/27.2.573

Chaisson, M. J. P., Huddleston, J., Dennis, M. Y., Sudmant, P. H., Malig, M., Hormozdiari, F., et al. (2015). Resolving the Complexity of the Human Genome Using Single-Molecule Sequencing. *Nature* 517, 608–611. doi:10.1038/nature13907

Chen, C. H., Chen, H. I., Chien, W. H., Li, L. H., Wu, Y. Y., Chiu, Y. N., et al. (2017). High Resolution Analysis of Rare Copy Number Variants in Patients with Autism Spectrum Disorder from Taiwan. *Scientific Rep.* 7 (1), 11919. doi:10.1038/s41598-017-12081-4

Conrad, D. F., Keebler, J. E., Depristo, M. A., Lindsay, S. J., Zhang, Y., Casals, F., et al. (2011). Variation in Genome-wide Mutation Rates within and between Human Families. *Nat. Genet.* 43, 712–714. doi:10.1038/ng.862

English, A. C., Salerno, W. J., and Reid, J. G. (2014). PBHoney: Identifying Genomic Variants via Long-Read Discordance and Interrupted Mapping. *Bmc Bioinf.* 15, 180. doi:10.1186/1471-2105-15-180

Fang, L., Hu, J., Wang, D., and Wang, K. (2018). NextSV: a Meta-Caller for Structural Variants from Low-Coverage Long-Read Sequencing Data. *Bmc Bioinf.* 19 (1), 180. doi:10.1186/s12859-018-2207-1

Fernandes, J. D., Zamudio-Hurtado, A., Clawson, H., Kent, W. J., Haussler, D., Salama, S. R., et al. (2020). The UCSC Repeat Browser Allows Discovery and Visualization of Evolutionary Conflict across Repeat Families. *Mobile DNA* 11, 13. doi:10.1186/s13100-020-00208-w

Hastings, P. J., Lupski, J. R., Rosenberg, S. M., and Ira, G. (2009). Mechanisms of Change in Gene Copy Number. *Nat. Rev. Genet.* 10, 551–564. doi:10.1038/nrg2593

Hills, M., Jeyapalan, J. N., Foxon, J. L., and Royle, N. J. (2007). Mutation Mechanisms that Underlie Turnover of a Human Telomere-Adjacent Segmental Duplication Containing an Unstable Minisatellite. *Genomics* 89, 480–489. doi:10.1016/j.ygeno.2006.12.011

Hodgkinson, A., Chen, Y., and Eyre-Walker, A. (2012). The Large-Scale Distribution of Somatic Mutations in Cancer Genomes. *Hum. Mutat.* 33, 136–143. doi:10.1002/humu.21616

Ishiura, H., Shibata, S., Yoshimura, J., Suzuki, Y., Qu, W., Doi, K., et al. (2019). Noncoding CGG Repeat Expansions in Neuronal Intranuclear Inclusion Disease, Oculopharyngodistal Myopathy and an Overlapping Disease. *Nat. Genet.* 51, 1222, 1232. doi:10.1038/s41588-019-0458-z

Kirov, G., Pocklington, A. J., Holmans, P., Ivanov, D., Ikeda, M., Ruderfer, D., et al. (2012). De Novo CNV Analysis Implicates Specific Abnormalities of Postsynaptic Signalling Complexes in the Pathogenesis of Schizophrenia. *Mol. Psychiatry* 17, 142–153. doi:10.1038/mp.2011.154

Kosugi, S., Momozawa, Y., Liu, X. X., Terao, C., Kubo, M., and Kamatani, Y. (2019). Comprehensive Evaluation of Structural Variation Detection Algorithms for Whole Genome Sequencing. *Genome Biol.* 20 (1), 117. doi:10.1186/s13059-019-1720-5

Kothiyal, P., Wong, W. S. W., Bodian, D. L., and Niederhuber, J. E. (2019). Mendelian Inconsistent Signatures from 1314 Ancestrally Diverse Family Trios Distinguish Biological Variation from Sequencing Error. *J. Comput. Biol.* 26, 405–419. doi:10.1089/cmb.2018.0253

Liang, Y., He, L., Zhao, Y. R., Hao, Y. Y., Zhou, Y. F., Li, M. L., et al. (2019). Comparative Analysis for the Performance of Variant Calling Pipelines on Detecting the De Novo Mutations in Humans. *Front. Pharmacol.* 10, 358. doi:10.3389/fphar.2019.00358

Liu, Z., Zhu, L., Roberts, R., and Tong, W. (2019). Toward Clinical Implementation of Next-Generation Sequencing-Based Genetic Testing in Rare Diseases: Where Are We? *Trends Genet.* 35, 852–867. doi:10.1016/j.tig.2019.08.006

Logsdon, G. A., Vollger, M. R., and Eichler, E. E. (2020). Long-read Human Genome Sequencing and its Applications. *Nat. Rev. Genet.* 21, 597–614. doi:10.1038/s41576-020-0236-x

Lupski, J. R. (2015). Structural Variation Mutagenesis of the Human Genome: Impact on Disease and Evolution. *Environ. Mol. Mutagen.* 56, 419–436. doi:10.1002/em.21943

Merker, J. D., Wenger, A. M., Sneddon, T., Grove, M., Zappala, Z., Fresard, L., et al. (2018). Long-read Genome Sequencing Identifies Causal Structural Variation in a Mendelian Disease. *Genet. Med.* 20, 159–163. doi:10.1038/gim.2017.86

Mizuguchi, T., Suzuki, T., Abe, C., Umemura, A., Tokunaga, K., Kawai, Y., et al. (2019). A 12-kb Structural Variation in Progressive Myoclonic Epilepsy Was Newly Identified by Long-Read Whole-Genome Sequencing. *J. Hum. Genet.* 64, 359–368. doi:10.1038/s10038-019-0569-5

Navarro Gonzalez, J., Zweig, A. S., Speir, M. L., Schmelter, D., Rosenbloom, K. R., and Raney, B. J. (2020). The UCSC Genome Browser Database: 2021 Update. *Nucleic Acids Res.* 49 (D1), D1046–D1057. doi:10.1093/nar/gkaa1070

Pilipenko, V. V., He, H., Kurowski, B. G., Alexander, E. S., Zhang, X., Ding, L., et al. (2014). Using Mendelian Inheritance Errors as Quality Control Criteria in Whole Genome Sequencing Data Set. *BMC Proc.* 8, S21. doi:10.1186/1753-6561-8-s1-s21

Pinto, D., Pagnamenta, A. T., Klei, L., Anney, R., Merico, D., Regan, R., et al. (2010). Functional Impact of Global Rare Copy Number Variation in Autism Spectrum Disorders. *Nature* 466, 368–372. doi:10.1038/nature09146

Pollard, M. O., Gurdasani, D., Mentzer, A. J., Porter, T., and Sandhu, M. S. (2018). Long Reads: Their Purpose and Place. *Hum. Mol. Genet.* 27, R234–R241. doi:10.1093/hmg/ddy177

Sanders, S. J., Murtha, M. T., Gupta, A. R., Murdoch, J. D., Raubeson, M. J., Willsey, A. J., et al. (2012). De Novo mutations Revealed by Whole-Exome Sequencing Are Strongly Associated with Autism. *Nature* 485, 237–241. doi:10.1038/nature10945

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., et al. (2007). Strong Association of De Novo Copy Number Mutations with Autism. *Science* 316, 445–449. doi:10.1126/science.1138659

Sedlazeck, F. J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., Von Haeseler, A., et al. (2018). Accurate Detection of Complex Structural Variations Using Single-Molecule Sequencing. *Nat. Methods* 15, 461–468. doi:10.1038/s41592-018-0001-7

Song, J. H. T., Lowe, C. B., and Kingsley, D. M. (2018). Characterization of a Human-specific Tandem Repeat Associated with Bipolar Disorder and Schizophrenia. *Am. J. Hum. Genet.* 103, 421–430. doi:10.1016/j.ajhg.2018.07.011

Stankiewicz, P., and Lupski, J. R. (2010). Structural Variation in the Human Genome and its Role in Disease. *Annu. Rev. Med.* 61, 437–455. doi:10.1146/annurev-med-100708-204735

Stefansson, H., Rujescu, D., Cichon, S., Pietiläinen, O. P., Ingason, A., Steinberg, S., et al. (2008). Large Recurrent Microdeletions Associated with Schizophrenia. *Nature* 455, 232–236. doi:10.1038/nature07229

Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., et al. (2015). An Integrated Map of Structural Variation in 2,504 Human Genomes. *Nature* 526**,** 75, 81. doi:10.1038/nature15394

Veltman, J. A., and Brunner, H. G. (2012). De Novo mutations in Human Genetic Disease. *Nat. Rev. Genet.* 13, 565–575. doi:10.1038/nrg3241

Walsh, T., Mcclellan, J. M., Mccarthy, S. E., Addington, A. M., Pierce, S. B., Cooper, G. M., et al. (2008). Rare Structural Variants Disrupt Multiple Genes in Neurodevelopmental Pathways in Schizophrenia. *Science* 320, 539–543. doi:10.1126/science.1155174

Weischenfeldt, J., Symmons, O., Spitz, F., and Korbel, J. O. (2013). Phenotypic Impact of Genomic Structural Variation: Insights from and for Human Disease. *Nat. Rev. Genet.* 14, 125–138. doi:10.1038/nrg3373

Wenger, A. M., Peluso, P., Rowell, W. J., Chang, P.-C., Hall, R. J., Concepcion, G. T., et al. (2019). Accurate Circular Consensus Long-Read Sequencing Improves Variant Detection and Assembly of a Human Genome. *Nat. Biotechnol.* 37, 1155–1162. doi:10.1038/s41587-019-0217-9

Zeng, S., Zhang, M.-Y., Wang, X.-J., Hu, Z.-M., Li, J.-C., Li, N., et al. (2019). Long-read Sequencing Identified Intronic Repeat Expansions inSAMD12from Chinese Pedigrees Affected with Familial Cortical Myoclonic Tremor with Epilepsy. *J. Med. Genet.* 56, 265–270. doi:10.1136/jmedgenet-2018-105484

Zook, J. M., Catoe, D., Mcdaniel, J., Vang, L., Spies, N., Sidow, A., et al. (2016). Extensive Sequencing of Seven Human Genomes to Characterize Benchmark Reference Materials. *Scientific Data* 3, 160025. doi:10.1038/sdata.2016.25

Zook, J. M., Hansen, N. F., Olson, N. D., Chapman, L., Mullikin, J. C., Xiao, C., et al. (2020). A Robust Benchmark for Detection of Germline Large Deletions and Insertions. *Nat. Biotechnol.* 38 (11), 1347–1355. doi:10.1038/s41587-020-0538-8

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

frontiers in Genetics

CASE REPORT
published: 30 June 2021
doi: 10.3389/fgene.2021.649556

Case Report: Causative *De novo* Variants of *KCNT2* for Developmental and Epileptic Encephalopathy

Pan Gong[1], *Xianru Jiao*[1], *Dan Yu*[2] *and Zhixian Yang*[1]*

[1] Department of Pediatrics, Peking University First Hospital, Beijing, China, [2] Department of Pediatrics, West China Second University Hospital, Sichuan University, Chengdu, China

Objective: *KCNT2* gene mutations had been described to cause developmental and epileptic encephalopathies (DEEs). In this study, we presented the detailed clinical features and genetic analysis of two unrelated patients carrying two *de novo* variants in *KCNT2* and reviewed eight different cases available in publications.

Methods: Likely pathogenic variants were identified by whole exome sequencing; clinical data of the patients were retrospectively collected and analyzed.

Results: Our two unrelated patients were diagnosed with Ohtahara syndrome followed by infantile spasms (IS) and possibly the epilepsy of infancy with migrating focal seizures (EIMFS), respectively. They both manifested dysmorphic features with hirsute arms, thick hair, prominent eyebrows, long and thick eyelashes, a broad nasal tip, and short and smooth philtrum. In the eight patients reported previously, two was diagnosed with IS carrying a 'change-of-function' mutation and a gain-of-function mutation, respectively, two with EIMFS-like carrying a gain-of-function mutation and a loss-of-function mutation, respectively, one with EIMFS carrying a loss-of-function mutation, three with DEE without functional analysis. Among them, two patients with gain-of-function mutations both exhibited dysmorphic features and presented epilepsy phenotype, which was similar to our patients.

Conclusion: Overall, the most common phenotypes associated with *KCNT2* mutation were IS and EIMFS. Epilepsy phenotype associated with gain- and loss-of-function mutations could overlap. Additional *KCNT2* cases will help to make genotype-phenotype correlations clearer.

Keywords: epilepsy, *KCNT2*, genetic, seizures, encephalopathy

OPEN ACCESS

Edited by:
Zhichao Liu,
National Center for Toxicological Research (FDA), United States

Reviewed by:
Catarina M. Quinzii,
Columbia University, United States
Nicole J. Lake,
Yale University, United States

***Correspondence:**
Zhixian Yang
zhixian.yang@163.com

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

Received: *27 January 2021*
Accepted: *12 April 2021*
Published: *30 June 2021*

Citation:
Gong P, Jiao X, Yu D and Yang Z (2021) Case Report: Causative De novo Variants of KCNT2 for Developmental and Epileptic Encephalopathy. Front. Genet. 12:649556. doi: 10.3389/fgene.2021.649556

INTRODUCTION

Developmental and epileptic encephalopathies (DEEs) comprise a heterogeneous group of sever neurological disorders with onset in infancy and childhood, which is characterized by refractory seizures, frequent epileptic activity, and developmental regression or further slowing (Scheffer et al., 2017). Increasing number of genes identified as the cause of DEEs and channelopathies represent an important and broad class (Kumar et al., 2016). A systematic review of neurological disorders and potassium channelopathies revealed pathogenic variants in 19 potassium channel genes, including

KCNMA1, KCNN3, KCNT1, KCNT2, KCNB1, KCNJ6, KCNJ10, KCNJ11, KCNA2, KCNA4, KCND3, KCNH1, KCNQ2, KCNAB1, KCNQ3, KCNQ5, KCNC1, KCNC3, and *KCTD3* (Kessi et al., 2020). There is a large phenotypic and genetic heterogeneity and the majority of genetic defects are still unknown. Recently, pathogenic variants in *KCNT2* gene that encodes the $K_{Na}1.2$ subunit (Slick or Slo2.1) have been identified in eight cases (Gururaj et al., 2017; Ambrosino et al., 2018; Alagoz et al., 2020; Inuzuka et al., 2020; Mao et al., 2020). The *KCNT2*-associated DEEs comprises West syndrome, Lennox-Gastaut syndrome (LGS) as well as epilepsy of infancy with migrating focal seizures (EIMFS). *In vitro* functional analysis suggested that both gain and loss of function variants in *KCNT2* may lead to DEEs (McLachlan et al., 2019). Here, we reported two *de novo KCNT2* variants in two unrelated patients diagnosed with DEE charactering by profound developmental delay and intractable infantile-onset seizure disorders.

CASE PRESENTATION

In total, two unrelated boys were enrolled in our study. Clinical features of affected individuals with *KCNT2* variants were summarized in **Table 1**.

Patient #1 was a 5 month-old boy. He was born at 38 weeks of gestation with a birth weight of 2,950 g. He was the first child of healthy non-consanguineous parents. There was no familial history of any neuropsychiatric disease including epilepsy or febrile seizures. He was hypotonic and had a profound delayed development. From 8 days of age, the boy presented with daily clusters of epileptic spasms. The EEG showed burst suppression. At the current age of 5 months, the boy was unable to hold his head with poor visual contact. His body measures were 66 cm and 6.7 kg. The seizures were still uncontrolled on the combination of phenobarbitone and topiramate at the last follow-up. During the course, the semiology evolved from Ohtahara syndrome to West syndrome with severe developmental delay and frequent epileptic spasms. EEG displayed with hypsarrhythmia. Brain magnetic resonance imaging (MRI) at 18 days was normal. The patients manifested dysmorphic features with hirsute arms, thick hair, prominent eyebrows, and long and thick eyelashes and had a broad nasal tip, short and smooth philtrum (**Figure 1**).

Patient #2 was a 9 year-old boy. He was born at full term of gestation with a birth weight of 3,300 g. He was the first child of healthy non-consanguineous parents. There was no familial history of any neuropsychiatric disease including epilepsy or febrile seizures. He had a profound delayed developmental milestone with raising his head at the age of 1 years and starting to walk at the age of 5 years. According to his parents and medical records, the boy presented with daily seizures of awaking from sleep with the head swinging from side to side and the limbs jerking at 45 days of age. Seizures were mainly focal and migrating. EEG showed multifocal epileptic discharges with left temporal predominance. At the current age of 9 years, the boy had severe intellectual disability, was able to walk with assistance, but had no language and no verbal response. His body measures were 118 cm and 20 kg. He still had frequent seizures under the combination of valproate, topiramate, and clonazepam. The current EEG displayed generalized discharges. Brain MRI at 4 years was normal.

The patients manifested dysmorphic features with hirsute arms, thick hair, prominent eyebrows, and long and thick eyelashes and had a broad nasal tip, short and smooth philtrum (**Figure 2**).

GENETIC ANALYSIS

Trio-based whole-exome sequencing revealed *de novo* missense variants of [Chr1:196,395,112A>T, c.991T>A, p.(Tyr331Asn)] and [Chr1:196,465,339G>C, c.592C>G, p.(Gln198Glu)] in *KCNT2* gene (NM_198503.5, GRCh37/hg19), respectively, in two patients. Both variants were confirmed by conventional Sanger sequencing. The variants were novel and absent in control populational databases, including 1000 genomes, ExAC, and gnomAD. The pathogenicity of two novel variants was further analyzed using PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/). Y331N and Q198E substitutions were shown to be possibly damaging (score 0.619 and 0.598, respectively). MutationTaster program (http://www.mutationtaster.org/) showed that both variants were predicted to cause change in protein function and were disease causing. Y331N was localized in the C-terminal part of $K_{Na}1.2$ and Q198E was localized in S4-S5 of the protein (**Figure 3**). Moreover, the phenotype of the patients was consistent with that of *KCNT2*-associated disease. Therefore, both variants were classified as likely pathogenic based on the American College of Medical Genetics (ACMG) guidelines, both mutations were classified as pathogenic with the PS2, PM2, and PP3 criteria (Richards et al., 2015).

DISCUSSION

As a recently discovered gene associated with epilepsy disorders, a total of eight patients bearing five *KCNT2* variants had been reported (Gururaj et al., 2017; Ambrosino et al., 2018; Alagoz et al., 2020; Inuzuka et al., 2020; Mao et al., 2020). The *KCNT2*-associated DEEs comprises West syndrome as well as epilepsy of infancy with migrating focal seizures (EIMFS). In 2017, Gururaj et al. firstly identified a *KCNT2* mutation in an early onset epileptic encephalopathy (EOEE) patient with intractable epileptic spasms by exome sequencing and confirmed the mutation with "change-of-function" effect (Gururaj et al., 2017). At the last follow-up age of 4 years, the patient presented with prolonged tonic seizures, myoclonic jerks and atypical absences with a disorganized background, decrements, multifocal epileptogenic activity or hypsarrhythmia on EEG (Gururaj et al., 2017). We assumed that the semiology of this patient was likely to evolved from infantile spasms (IS) to Lennox-Gastaut syndrome (LGS). In 2018, Ambrosino et al. described two female patients with IS followed by LGS or with DEE with migrating focal seizures. *In vitro* analysis further suggested that the variants found in patients had gain-of-function effects (Ambrosino et al., 2018). Seizure discharges in the patient with DEE with migrating focal seizures had been observed to migrate from the left to

TABLE 1 | Patients with *KCNT2* variants in this publication and previous literatures.

	This publication		Gururaj et al. (2017)	Ambrosino et al. (2018)	
	Patient 1	**Patient 2**	**Patient 3**	**Patient 4**	**Patient 5**
Gender/Age	Male/5 m	Male/9 y 10 m	Male/4 y	Female/9 y	Female/14 y
Variant	c.991T>A, p.(Tyr331Asn)	c.592C>G, p.(Gln198Glu)	c.720T>A, p.(Phe240Leu)	c.569G>A, p.(Arg190His)	c.569G>C, p.(Arg190Pro)
Transcript	NM_198503.5	NM_198503.5	NM_001287819.1	NM_001287820.2	NM_001287820.2
Functional analysis	NA	NA	Change-of-function	Gain-of-function	Gain-of-function
Diagnosis	Ohtahara syndrome followed by IS	DEE with migrating focal seizures (EIMFS-like)	IS	IS followed by LGS	DEE with migrating focal seizures (EIMFS-like)
Age at epilepsy onset	8 d	45 d	3 m	8 m	1st day of life
Seizure type	ES	Focal and migrating seizures	Focal seizures, myoclonus, ES, tonic seizures, atypical absence	ES, nocturnal tonic and bilateral tonic-clonic seizures, non-convulsive SE	Generalized seizures, absences
Antiepileptic Treatment	PB, TPM	VPA, TPM, LTG, NZP	TPM, NZP, LEV, LTG, VBG, ESX, PLP, LCM, ketogenic diet	Sultiame, VPA, VBG, PLP, Rufinamide, methy-prednisolone, ketogenic diet	PB
Outcome	Uncontrolled	Uncontrolled	Daily seizures	Uncontrol	Controlled by PB, isolated GTCS at 6 m with PB withdrawn, frequent absences without medication
EEG	Burst suppression evolving into hypsarrhythmia	Multifocal epileptic discharges evolving into generalized epileptic discharges	Persistently abnormal with a disorganized background, decrements, multifocal epileptic activity or hypsarrhythmia	Hypsarrhythmia evolving to intermittent sharp-slow waves and general slowing, current EEG showing bilateral spike waves with central maximum	Rhythmical activity which was initially observed over the left parieto-auricular region and right temporal area evolving into clear focal sharp wave discharges over the left temporo-parietal area
Neurological features					
Before seizure onset	–	Profound developmental delay at birth	Profound developmental delay at birth	Mildly motor developmental delay at birth	–
After seizure onset	–	Regression in development	Regression in development	Developmental arrest	Delayed milestones of development
Current development	Profound developmental delay, poor visual contact, head deviation to one side	Sever intellectual disability, no language, no verbal responses, walk with significant assistance	Truncal instability, limited visual attention, no verbal responses, walk with significant assistance	Sever intellectual disability, no language	Sever language delay and learning disability, aggressive and uncooperative
Imaging	Normal	Normal	A generalized reduction in white matter and thinning of the corpus callosum	Stable supratentorial mild volume loss and slightly delayed myelination	Normal
Additional features					
Dysmorphic features	Hirsute arms, thick hair, prominent eyebrows, and long and thick eyelashes, broad nasal tip, short and smooth philtrum	Hirsute arms, thick hair, prominent eyebrows, and long and thick eyelashes, broad nasal tip, short and smooth philtrum	No	Hirsute arms, thick hair, prominent eyebrows, and long and thick eyelashes, broad nasal tip, short and smooth philtrum with prominent upper lip, mild tooth displacement with diastema	Hirsute arms, thick hair, prominent eyebrows, and long and thick eyelashes, broad nasal tip, short and smooth philtrum without prominent upper lip, mild tooth displacement with diastema, and slightly long, spatulate fingers, with slightly deep-set nails
Other medical issues	No	No	No	No	No

(Continued)

TABLE 1 | Continued

	Mao et al. (2020)		Alagoz et al. (2020)		Inuzuka et al. (2020)
	Patient 6	**Patient 7**	**Patient 8**	**Patient 9**	**Patient 10**
Gender/Age	Female/3m	Female/29y	Male/6y	Male/5y	Male/17y
Variant	c.1690A>T, p.(Lys564*)	c.143-144del, p.(Leu48Glnfs*43)	c.544A>T, p.(Asn182Ile)	c.2638C>A, p.(Leu880Met)	c.725C>A, p.(Thr242Asn)
Transcript	NM_198503.2	NM_198503.2	NM_198503.2	NM_198503.4	NA
Functional analysis	Loss-of-function	Loss-of-function	NA	NA	NA
Diagnosis	EIMFS	DEE with migrating focal seizures (EIMFS-like)	DEE	DEE	DEE with clinical features of frontal lobe epilepsy
Age at epilepsy onset	2 m	4 m	Unknown	Unknown	5 m
Seizure type	Focal seizures	Focal and migrating seizures	Unknown	Unknown	Tonic motor seizures, hyperkinetic focal motor seizures
Antiepileptic Treatment	VPA, LTG, LEV	Unknown	Unknown	Unknown	CBZ, LEV, OXC
Outcome	Uncontrol	Uncontrol	Unknown	Unknown	Hyperkinetic focal motor seizures twice a month
EEG	Symmetric slow background pattern, multifocal spikes and seizures arising from different regions independently and migrating from one hemisphere to the other at time	Unknown	Sharp and slow waves in the right frontotemporal region	Normal at the age of 1 year old	Disorganized background, multifocal epileptiform discharges (predominantly frontocentral)
Neurological features					
Before seizure onset	Unknown	Unknown	Unknown	Unknown	Profound developmental delay at birth
After seizure onset	Unknown	Unknown	Regression in development	Regression in development	Regression in development
Current development	Sever neurologic impairment with the poor visual following	Mild intellectual disability	Delayed neural development	Delayed neural development	Sever intellectual disability with autistic features, mild spasticity, ataxic gait, partially dependent for daily life activities
Imaging	Normal	Unknown	Diffusely thin corpus callosum, dilated lateral ventricles and partial colpocephaly		Normal
Additional features					
Dysmorphic features	No	No	No	No	No
Other medical issues	No	No	Hypotonia	Hypotonia	No

IS, infantile spasms; DEE, development and epileptic encephalopathy; EIMFS, the epilepsy of infancy with migrating focal seizures; LGS, Lennox-Gastaut syndrome; ES, epileptic spasms; EEG, electroencephalogram; PB, phenobarbitone; TPM, topiramate; VPA, valproic acid; LTG, lamotrigine; NZP, nitrazepam; LEV, levetiracetam; VGB, vigabatrin; ESX, ethosuximide; PLP, pyridoxal phosphate; LCM, lacosamide; GTCS, generalized tonic clonic seizure; CBZ, carbamazepine; OXC, oxcarbazepine.

the right hemisphere, which likely corresponded to EIMFS (Ambrosino et al., 2018). Most recently, Mao et al. identified two novel *de novo* non-sense and frameshift mutations of the *KCNT2* gene in two patients diagnosed as EIMFS and early onset epileptic encephalopathy (EOEE) with migrating focal seizures, respectively (Mao et al., 2020). Due to the fact that the latter case was an aged case, diagnosis was not be firmly ascertained and the author described the patient as EIMFS-like EOEE (Mao et al., 2020). They investigated the functional consequence of the two mutations, which showed that both mutations reduced whole-cell potassium current (Mao et al., 2020). Subsequently, there were two reports describing three patients with EOEE caused by mutations in *KCNT2* without functional analysis (Alagoz et al., 2020; Inuzuka et al., 2020). Here, our patients showed DEE in common with previous studies. Patient #1 was diagnosed as Ohtahara syndrome subsequently followed by IS with EEG showing burst suppression evolving to hypsarrhythmia. Patient #2 was diagnosed as DEE with migrating focal seizures manifesting profound developmental delay at birth and intractable seizure with developmental regression. Due to incomplete medical records, there was no definitive diagnosis but EIMFS-like EOEE of Patient #2. In general, both gain and loss of function mutations in the *KCNT2* gene could lead to DEE, and the most common phenotypes were IS and EIMFS. The epilepsy phenotype associated with gain- and loss-of-function mutations could overlap.

EIMFS is a rare EOEE characterized by polymorphous focal seizures and cognitive, sensory and motor impairment, with arrest of psychomotor development in the first 6 months of life (Coppola et al., 1995, 2006). EIMFS have a genetic origin in which the *de novo* gain-of-function *KCNT1* variants are the most common cause (Barcia et al., 2012; McTague et al., 2013; Rizzo et al., 2016; Villa and Combi, 2016). *KCNT1* and *KCNT2* respectively encode the $K_{Na}1.1$ (Slack) and $K_{Na}1.2$ (Slick) subunits. Both genes belongs to the SLO2 family of the sodium-dependent voltage-gated potassium channel K_{Na} (Bhattacharjee et al., 2003). Functional analysis of mutant channels associated with EIMFS mostly revealed gain-of-function effects (Barcia et al., 2012; Rizzo et al., 2016; Villa and Combi, 2016). Similarly, Ambrosino et al. identified a gain-of-function mutation in the *KCNT2* gene in a patient with EIMFS-like DEE. It suggested that pathogenic variants in *KCNT1* and *KCNT2* therefore might contribute to a similar and overlapping spectrum of DEEs (Ambrosino et al., 2018; Kessi et al., 2020). Meantime, Mao et al. also reported EIMFS in patients carrying loss-of-function mutations in the *KCNT2* gene (Mao et al., 2020), which indicated that EIMFS might be caused not only by an increase but also by a decrease in the function of K_{Na}.

Moreover, it was noticed that the two patients confirmed to carry gain-of-function mutations both exhibited not only severe developmental delay, but also dysmorphic features with hirsute arms, thick hair, prominent eyebrows, and long and thick eyelashes, broad nasal tip, short and smooth philtrum with prominent upper lip, mild tooth displacement with diastema (Ambrosino et al., 2018). Similarly, both of our patients had prominent dysmorphic features with hirsute arms, thick hair, prominent eyebrows, and long and thick eyelashes. We presumed that the dysmorphic features might be the distinguishing characterizations in patients with gain-of-function mutations in *KCNT2*. However, we could not confirm the mutations in our patients in the absence of functional analysis. Given the small number of patients reported so far, it is difficult to make genotype-phenotype correlations. Additional *KCNT2* cases will help clarify the potential association between gain-of-function and dysmorphic features.

Seizures in all patients were intractable and there was a poor prognosis in development. The oldest patient at the age of 29 years still had seizures and showed mild intellectual disability (Mao et al., 2020). Quinidine as an underlying precision medicine for epilepsy syndromes due to gain of function mutations in *KCNT1* aroused particular clinical interest (Bearden et al., 2014). Therefore, Ambrosino et al. confirmed by functional analysis that the patients carried quinidine-responsive gain-of-function mutations in

FIGURE 1 | Facial Gestalt of the Patient #1 with *de novo KCNT2* variant **(A)** at birth and **(B)** at the age of 3 months.

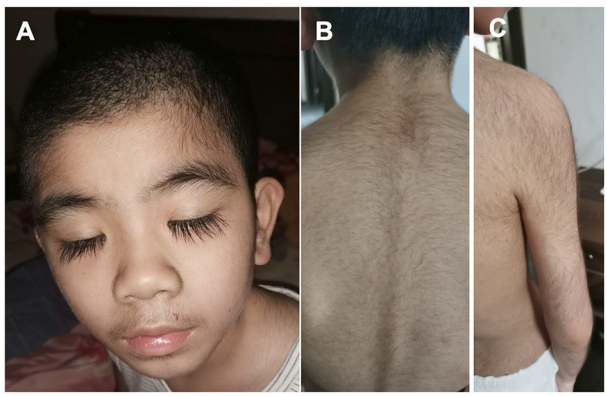

FIGURE 2 | Dysmorphic features of the Patient #2 with *de novo KCNT2* variant at the age of 10 years. **(A)** Facial Gestalt of thick hair, prominent eyebrows, and long and thick eyelashes. **(B)** Thick hair on the back. **(C)** Thick hair on the right arm.

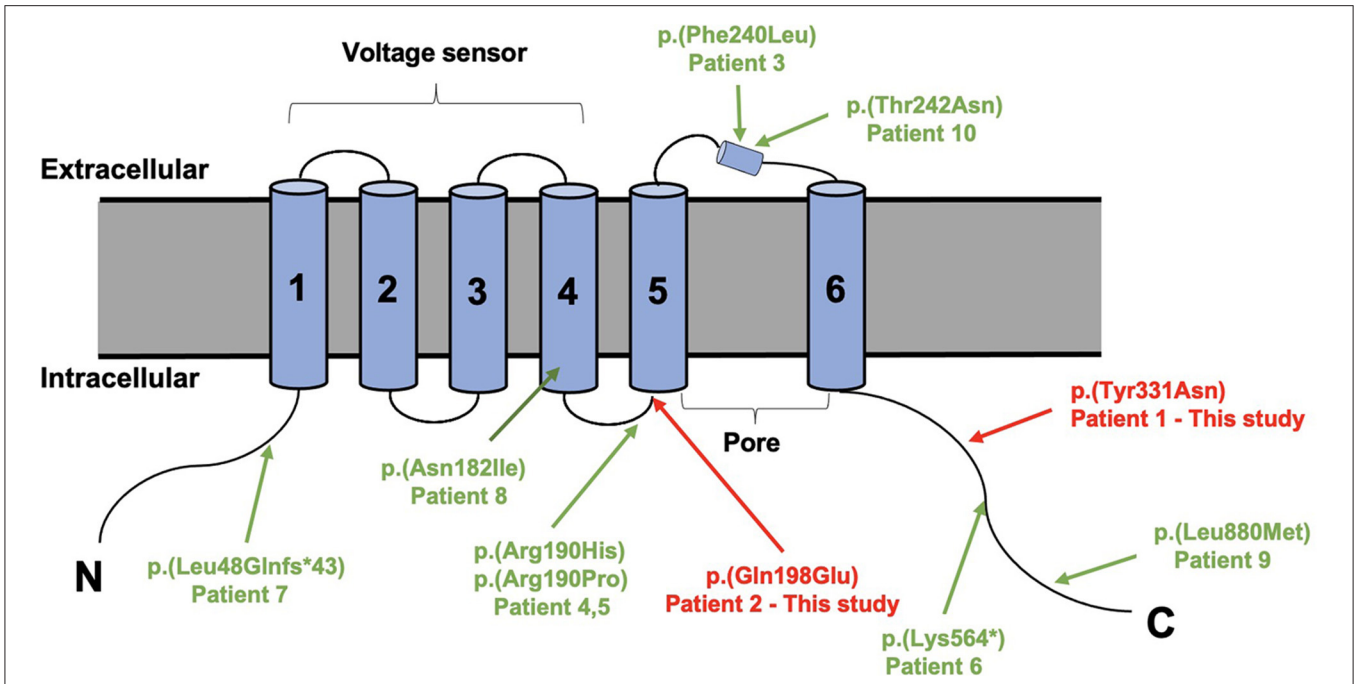

FIGURE 3 | Structure of the sodium-activated potassium ($K_{Na}1.2$) channel (9606 [NCBI]) with all variants published variants in *KCNT2*.

the *KCNT2* gene and treated one patient with quinidine add-on therapy and achieved marked clinical improvements, including EEG, vigilance, alertness as well as developmental progression (Ambrosino et al., 2018). It suggested that quinidine could be a potential personalized medicine approach for *KCNT2*-related DEE.

CONCLUSION

In this study, we presented the detailed clinical features and genetic analysis of two unrelated patients with *KCNT2*-related DEE and provided a comprehensive outline of available publications regarding *KCNT2* mutations. Our data further expanded the spectrum of *KCNT2* mutation. So far, with our case, a total of ten cases carrying *KCNT2* mutations have been reported. IS and EIMFS were the most common phenotypes caused by pathogenic mutations in *KCNT2*. Both gain and loss of function mutations could lead to EIMFS. Additional *KCNT2* cases will help to make genotype-phenotype correlations clearer. Quinidine could be a potential personalized medicine approach for *KCNT2*-related DEE. Clinical follow-up of additional patients will further define the clinical spectrum of *KCNT2*-related DEE and the long-time efficacy on seizures and development of quinidine as an underlying precision medicine for epilepsy syndromes due to gain-of-function mutations.

DATA AVAILABILITY STATEMENT

The sequencing data are available on https://www.ncbi.nlm.nih.gov/bioproject/PRJNA689060.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of Peking University First Hospital. 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

ZY conceptualized and designed the study, coordinated the study overall, and revised the manuscript. PG co-designed the study, drafted the initial manuscript, and revised the manuscript. XJ and DY helped to collect and summarize data and revised the manuscript. All authors approved the final revision of the article.

FUNDING

This work was supported by National Nature Science Foundation of China (81771393), Beijing Municipal Science & Technology Commission (Z171100001017125), Beijing Natural Science Foundation (7202210), Capital's Funds for Health Improvement and Research (2020-2-4077).

ACKNOWLEDGMENTS

We extend our appreciation to the patients and their families who participated in this study. We would like to thank Dr. Xiaodong Wang from Cipher Gene, Ltd. for assisting in manuscript editing.

REFERENCES

Alagoz, M., Kherad, N., Bozkurt, S., and Yuksel, A. (2020). New mutations in KCNT2 gene causing early infantile epileptic encephalopathy type 57: case study and literature review. *Acta Biochim. Pol.* 67, 431–434. doi: 10.18388/abp.2020_5364

Ambrosino, P., Soldovieri, M. V., Bast, T., Turnpenny, P. D., Uhrig, S., Biskup, S., et al. (2018). *De novo* gain-of-function variants in KCNT2 as a novel cause of developmental and epileptic encephalopathy. *Ann. Neurol.* 83, 1198–1204. doi: 10.1002/ana.25248

Barcia, G., Fleming, M. R., Deligniere, A., Gazula, V. R., Brown, M. R., Langouet, M., et al. (2012). *De novo* gain-of-function KCNT1 channel mutations cause malignant migrating partial seizures of infancy. *Nat. Genet.* 44, 1255–1259. doi: 10.1038/ng.2441

Bearden, D., Strong, A., Ehnot, J., DiGiovine, M., Dlugos, D., and Goldberg, E. M. (2014). Targeted treatment of migrating partial seizures of infancy with quinidine. *Ann. Neurol.* 76, 457–461. doi: 10.1002/ana.24229

Bhattacharjee, A., Joiner, W. J., Wu, M., Yang, Y., Sigworth, F. J., and Kaczmarek, L. K. (2003). Slick (Slo2.1), a rapidly-gating sodium-activated potassium channel inhibited by ATP. *J. Neurosci.* 23, 11681–11691. doi: 10.1523/JNEUROSCI.23-37-11681.2003

Coppola, G., Plouin, P., Chiron, C., Robain, O., and Dulac, O. (1995). Migrating partial seizures in infancy: a malignant disorder with developmental arrest. *Epilepsia* 36, 1017–1024. doi: 10.1111/j.1528-1157.1995.tb00961.x

Coppola, G., Veggiotti, P., Del Giudice, E. M., Bellini, G., Longaretti, F., Taglialatela, M., et al. (2006). Mutational scanning of potassium, sodium and chloride ion channels in malignant migrating partial seizures in infancy. *Brain Dev.* 28, 76–79. doi: 10.1016/j.braindev.2005.05.002

Gururaj, S., Palmer, E. E., Sheehan, G. D., Kandula, T., Macintosh, R., Ying, K., et al. (2017). A *De novo* mutation in the sodium-activated potassium channel KCNT2 Alters ion selectivity and causes epileptic encephalopathy. *Cell Rep.* 21, 926–933. doi: 10.1016/j.celrep.2017.09.088

Inuzuka, L. M., Macedo-Souza, L. I., Della-Ripa, B., Monteiro, F. P., Ramos, L., Kitajima, J. P., et al. (2020). Additional observation of a *de novo* pathogenic variant in KCNT2 leading to epileptic encephalopathy with clinical features of frontal lobe epilepsy. *Brain Dev.* 42, 691–695. doi: 10.1016/j.braindev.2020.05.003

Kessi, M., Chen, B., Peng, J., Tang, Y., Olatoutou, E., He, F., et al. (2020). Intellectual disability and potassium channelopathies: a systematic review. *Front. Genet.* 11:614. doi: 10.3389/fgene.2020.00614

Kumar, P., Kumar, D., Jha, S. K., Jha, N. K., and Ambasta, R. K. (2016). Ion channels in neurological disorders. *Adv. Prot. Chem. Struct. Biol.* 103, 97–136. doi: 10.1016/bs.apcsb.2015.10.006

Mao, X., Bruneau, N., Gao, Q., Becq, H., Jia, Z., Xi, H., et al. (2020). The epilepsy of infancy with migrating focal seizures: identification of *de novo* mutations of the KCNT2 gene that exert inhibitory effects on the corresponding heteromeric KNa1.1/KNa1.2 potassium channel. *Front. Cell Neurosci.* 14:1. doi: 10.3389/fncel.2020.00001

McLachlan, F., Sires, A. M., and Abbott, C. M. (2019). The role of translation elongation factor eEF1 subunits in neurodevelopmental disorders. *Hum. Mutat.* 40, 131–141. doi: 10.1002/humu.23677

McTague, A., Appleton, R., Avula, S., Cross, J. H., King, M. D., Jacques, T. S., et al. (2013). Migrating partial seizures of infancy: expansion of the electroclinical, radiological and pathological disease spectrum. Brain. 136(Pt. 5), 1578–1591. doi: 10.1093/brain/awt073

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

Rizzo, F., Ambrosino, P., Guacci, A., Chetta, M., Marchese, G., Rocco, T., et al. (2016). Characterization of two *de novo* KCNT1 mutations in children with malignant migrating partial seizures in infancy. *Mol. Cell Neurosci.* 72, 54–63. doi: 10.1016/j.mcn.2016.01.004

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

Villa, C., and Combi, R. (2016). Potassium channels and human epileptic phenotypes: an updated overview. *Front. Cell Neurosci.* 10:81. doi: 10.3389/fncel.2016.00081

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

frontiers in Pediatrics

CASE REPORT
published: 08 July 2021
doi: 10.3389/fped.2021.626657

DNM1L-Related Mitochondrial Fission Defects Presenting as Encephalopathy: A Case Report and Literature Review

Xingmiao Liu†, Zhongbin Zhang†, Dong Li†, Meifang Lei, Qing Li, Xiaojun Liu and Peiyuan Zhang**

Department of Pediatric Neurology, Tianjin Children's Hospital/Tianjin University Children's Hospital, Tianjin, China

OPEN ACCESS

Edited by:
Shraddha Thakkar, United States Food and Drug Administration, United States

Reviewed by:
Muhammad Umair, King Abdullah International Medical Research Center (KAIMRC), Saudi Arabia
Andrea Legati, Fondazione IRCCS Istituto Neurologio Carlo Besta, Italy

***Correspondence:**
Xiaojun Liu
xiaojun_liu88@sina.com
Peiyuan Zhang
13132004076@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 Pediatrics

Received: *06 November 2020*
Accepted: *11 June 2021*
Published: *08 July 2021*

Citation:
Liu X, Zhang Z, Li D, Lei M, Li Q, Liu X and Zhang P (2021) DNM1L-Related Mitochondrial Fission Defects Presenting as Encephalopathy: A Case Report and Literature Review. Front. Pediatr. 9:626657. doi: 10.3389/fped.2021.626657

Background: Mitochondrial dynamics, including mitochondrial fission and fusion, transport and distribution, biogenesis and degradation, are critical to neuronal function. The dynamin-1 like *(DNM1L)* gene encodes dynamin-related protein 1 (DRP1/DLP1), which is an evolutionarily conserved member of the dynamin family and is responsible for mitochondrial division. *DNM1L* variants can lead to mitochondrial fission dysfunction and neurological disorders.

Methods: We report a case of *DNM1L*-related mitochondrial disease admitted to Tianjin Children's Hospital. We searched for similar reported cases in the PubMed database using the terms "*DNM1L*" and "mitochondrial," reviewed recent literature to summarize the clinical and genetic characteristics, and analyzed genotype-phenotype correlations.

Results: The patient presented with psychomotor retardation, motor disturbance (muscle weakness with paroxysmal hypermyotonia), and a *de novo* variant (c.116G>A, g.22229G>A, p.S39N) in the GTPase domain of *DNM1L* (reference sequence NM_012062), which has not previously been reported in the literature. This case was combined with an additional 35 cases identified in 20 relevant references in order to analyze a total of 36 patients. The male-to-female ratio was 1:1.06, and the median age of onset was 6 months (range, neonatal period to 9 years). The cardinal symptoms included psychomotor retardation in 77.8% (28/36), limb paralysis in 66.7% (18/27), dystonia in 82.8% (24/29), and epilepsy in 59.4% (19/32). The clinical manifestations of variants in the GTPase domain of DRP1 were milder than those identified in the middle domain.

Conclusion: This case report describes a new variant of the *DNM1L* gene, and summarizes previously reported cases. Furthermore, the clinical phenotype and the genotype of *DNM1L* gene-associated mitochondrial disease was analyzed to improve the understanding of this disease.

Keywords: *DNM1L*, DRP1, mitochondrial disease, psychomotor retardation, dyskinesia, hypertonia, follow-up study

INTRODUCTION

Mitochondrial diseases are a group of multi-system disorders with heterogeneous clinical manifestations and genetic characteristics. They are defined by mitochondrial structural abnormalities and dysfunction caused by variants in mitochondrial DNA or nuclear DNA (1). In addition to structural defects in the mitochondrial respiratory chain complex and mitochondria, mutations affecting a group of genes enconding for mitochondrial dynamics-related proteins, which act on mitochondrial fusion, fission, transport, distribution, biogenesis, degradation, and for other related proteins, can also lead to mitochondrial dysfunction (2). Mitochondrial fission and fusion play vital roles in maintaining functional mitochondria when cells encounter environmental and metabolic stresses. Mitochondrial fission can create new mitochondria, but it also contributes to quality control by enabling the removal of damaged mitochondria and facilitating apoptosis to maintain metabolism during cellular stress. Disruptions in these processes affect normal physiological processes and have been implicated in mitochondrial diseases (3).

The most important mediator of mitochondrial and peroxisome division is DRP1, which is encoded by the *DNM1L* gene (4, 5). DRP1 is recruited from the cytosol to the outer mitochondrial membrane, where it oligomerizes, hydrolyzes guanosine triphosphate (GTP), and forms spirals around the mitochondria to constrict both outer and inner membranes to complete mitochondrial fission (3, 6).

The neurological disorder associated with *DNM1L* variants is referred to as encephalopathy due to defective mitochondrial and peroxisomal fission (EMPF1), which has been described as a fatal encephalopathy (7). EMPF1 is characterized by the early onset of a range of symptoms, including psychomotor delay, limb paralysis, dystonia (hypertonia or hypotonia), epilepsy (most refractory seizures, status epilepticus, consistent with epileptic encephalopathy), ataxia, nystagmus, optic atrophy, nystagmus, dysarthria, microcephalus, pain insensitivity, sensory and motor axonal neuropathy, respiratory distress, and even death in childhood.

There have been relatively few reports on *DNM1L*-related mitochondrial diseases; moreover, most of the reported cases are in medical records, which lack details of the disease characteristics. Here, we report a case of *DNM1L*-related mitochondrial disease admitted to Tianjin Children's Hospital and review recent literature to summarize the clinical and genetic characteristics, as well as the phenotype-genotype correlation to deepen our understanding of this genetic disease.

CLINICAL REPORT

Case Presentation

Clinical Manifestation

The patient, a 3-year-old boy, was born at term through an uneventful delivery with normal birth parameters of consanguineous Chinese parents. He was the first child in a second pregnancy, after the first resulted in an induced abortion due to fetal death. The grandparents of the patients were cousins. The pedigree of the family was as follows (**Figure 1A**). The age at the first visit was 5 months. The first symptoms observed were motor and cognitive retardation. He could not attain head control, turn over, or actively grasp objects. He had normal eyesight and hearing. The second symptom, paroxysmal hypertonia, began as early as 1 month after birth. Initially, the symptoms were clenched fists with both hands and straightened limbs, which gradually progressed to flexion of both upper limbs, fisted hands, and straightened lower limbs. During the course, consciousness was clear and the symptoms lasted for a few seconds to a few minutes before spontaneous relief. Physical examination revealed microcephaly, white skin, yellow hair, slightly long face, high palatine arch, right hand penetration, small penis, and small hands and feet. His muscle force was grade three, and he had paroxysmal hypertonia, negative limb deep reflection, and bilateral symmetric flexor plantar reflex.

Auxiliary Examination

Magnetic resonance imaging (MRI) showed ventricular broadening and left lateral fissure widening in 2017. 1 year later, MRI revealed symmetric hyperintensity in the bilateral basal ganglia on T2-weighted imaging, T2-FLAIR, and DWI sequences (**Figure 1B**). Video electroencephalogram showed a normal background and recorded paroxysmal limb rigidity with no epileptiform discharges. Liver and renal function, myocardial enzyme, and ceruloplasmin levels were normal, but metabolism screening showed mild ketonuria and hyperlactacidemia (2.55–3.36). G-banding karyotype analysis: 46, XY. Gesell Developmental Scale: Adaptability: DQ 26.7, DA 2.5, severe delay; Gross motor: DQ 23.7, DA 2.2, extremely severe delay; Fine motor: DQ 24.7, DA 2.3, extremely severe delay; Communication: DQ 49.5, DA 4.6, moderately delayed; personal-social skills: DQ 37.6, DA 3.5, severe delay.

Genetic Analysis

DNA was extracted from peripheral blood using the QIAamp DNA Blood Mini Kit (QIAGEN). A total of 3 μg of patient DNA was sheared to produce 200–300 base pair (bp) fragments. Whole-exome enrichment was performed using an Agilent SureSelect Target Enrichment system (Agilent Technologies). Coding exons and flanking intronic regions were enriched with the Agilent SureSelect Human All Exon V6 reagent (Agilent Technologies), according to the manufacturer's protocol. Captured libraries were loaded onto a HiSeq 2500 platform (Illumina). Base calling and assessments of sequence read quality were performed using Illumina Sequence Control Software (SCS; Illumina), which provided real-time analyses. Reads with average quality scores <25 were removed, and bases with quality scores <20 were trimmed. The mean read depth of each sample was 100×. The reads were aligned to the human reference genome (UCSC GRCh37/hg19) using the Burrows-Wheeler Aligner (BWA v0.7.15) BWA-MEM algorithm. Reads with low mapping quality scores were removed. Presumed PCR duplicates were identified using Picard's MarkDuplicates and removed. Local alignment optimization and base quality recalibration were performed using the Genome Analysis Toolkit (GATK). Single-nucleotide variants (SNVs) and InDels were

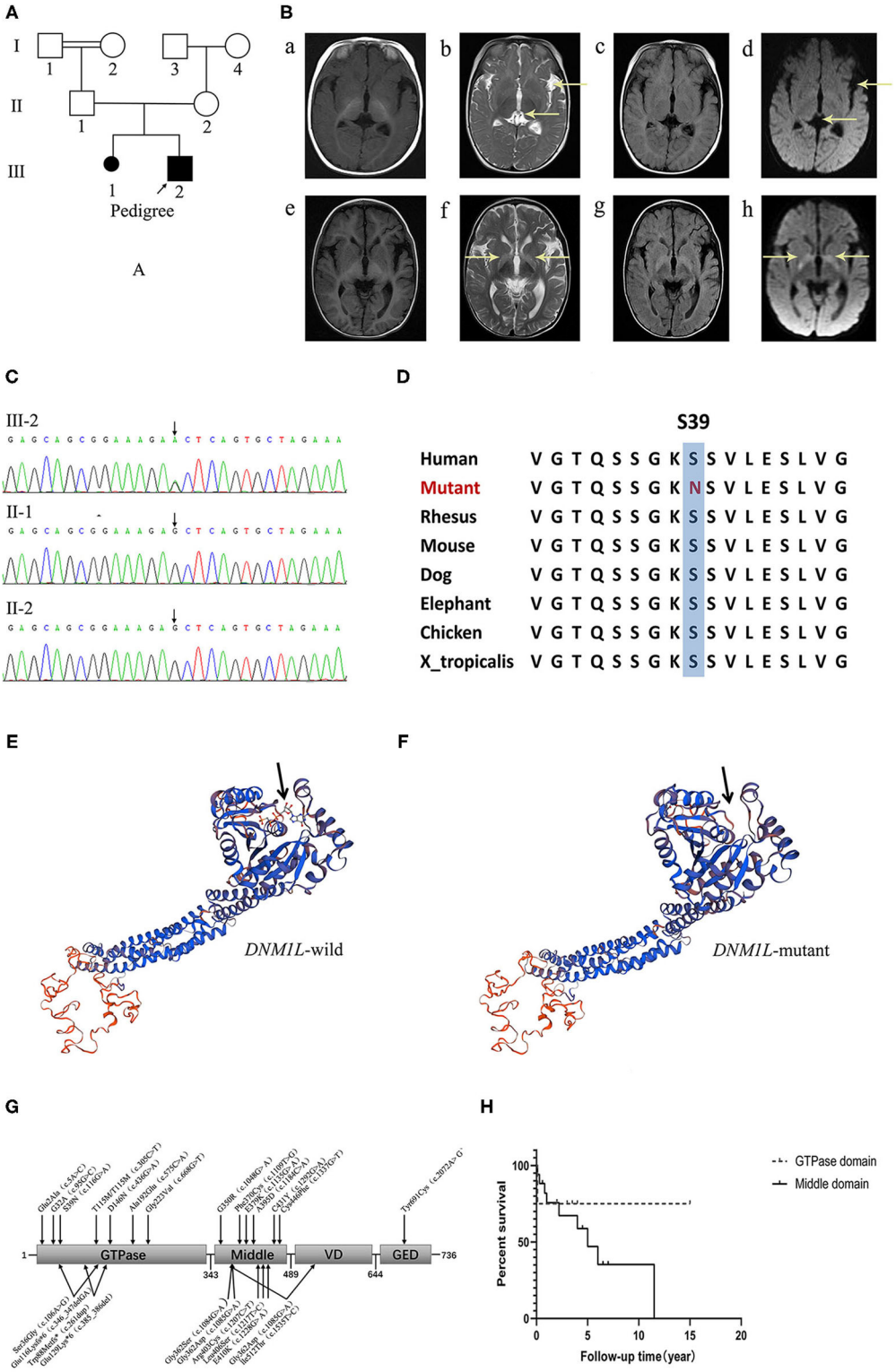

FIGURE 1 | (A) The pedigree of the family. **(B)** Brain MR images at 5-, and 18-months of age. (a–d) show that the ventricles of the brain and the left lateral fissure were widened on T1WI, T2WI, FLAIR and DWI sequences at 5-months of age. (e–h) disclose symmetrical hyperintensity in the bilateral basal ganglia on T2WI, FLAIR, and DWI sequences at 18-months of age. **(C)** Electropherograms of the regions containing the *DNM1L* variant identified in genomic DNA from the patient. **(D)** The conservation of the identified amino acid across different species. **(E,F)** The protein 3D modeling for the *DNM1L*-wild and identified missense variant. **(G)** The schematic chart and the locations of identified variants in DRP1. **(H)** The survival curve of patients of the GTPase domain group and the middle domain group of DRP1.

identified with Mutect2, respectively, and saved in a variant call format (8). Variants were functionally annotated and filtered with ANNOVAR (http://annovar.openbioinformatics.org/en/latest), which provided built-in public databases (OMIM, InterVar, ClinVar, HGMD, Cosmic70, dbSNP, 1000G, ESP, ExAC, and gnomAD), and the HGMD Professional database (9). We identified a *de novo* variant (c.116 G>A, p.S39N) in exon 2 of *DNM1L*, which has not been reported previously. Furthermore, gene variation was validated using Sanger sequencing (**Figure 1C**). The conservation of the identified amino acid across different species and the protein 3D modeling for the *DNM1L*-wild and identified missense variant were shown in **Figures 1D–F**. The pathogenicity score was checked using the following online tools (**Table 1**).

Treatment and Follow-Up

Benzhexol (0.5 mg BID) and Medopar (31.25 mg QID) were used to reduce muscle tension, and a cocktail therapy (VitB1 10 mg TID, VitC 0.5 g BID, VitE 100 mg QD, CoQ 10 mg TID and L-carnitine 100 mg BID) was administered to treat the primary illness, but none provided significant clinical benefit. At the last follow-up, the patient was 3 years and 4 months of age. He was 85 cm tall and weighed 8 kg, with a head circumference of 46 cm. He was admitted to our hospital with symptoms including spastic tetraparesis, no postural control, paroxysmal hypermyotonia and tremor, and severe cognitive impairment but no seizures. He could only follow a voice and light, occasionally respond to communication, and make vowel sounds, but he could not cry or laugh out loud.

Literature Review

We searched for reported cases in the PubMed database using the terms "*DNM1L*" and "mitochondrial." Thirty-five cases from 20 relevant references (5–7, 10–26) were retrieved; these included the first case reported by Waterham in 2007 and other cases reported from January 2015 to date. A total of 36 cases were analyzed together with the case reported by us. Fifteen males, 16 females, and two patients without reference to sex and three pedigrees (both male and female) were included. The median age of onset was 6 months (range, neonatal period: 9 years).

Clinical Data Summary

Clinical manifestations included psychomotor retardation in 77.8% (28/36) of patients, limb paralysis in 66.7% (18/27), dystonia in 82.8% (24/29), and epilepsy in 59.4% (19/32). Other symptoms included ataxia in six cases, nystagmus in three cases, optic atrophy in six cases, dysarthria in four cases, microcephaly in four cases, pain insensitivity in two cases, and sensory and motor axonal neuropathy in two cases.

On auxiliary examination, 83.3% (25/30) of the cases had abnormalities in brain MR images, including abnormal hyperintensity in the cortex (10 cases), basal ganglia/thalamus (10 cases), subcortical white matter (three cases), brain atrophy (12 cases), lactate peak (four cases), and thinning/absence of the corpus callosum (three cases). Seventy-five percent (15/20) of the cases showed abnormalities in EEG, such as epileptic discharge or a weak background. Hyperlactacidemia was identified in 60.7% (17/28) of the cases, and decreased respiratory chain enzyme activity was detected in 28.6% (4/14) of cases. Muscle biopsy was performed in 22 cases, ragged red fibers were observed in four cases, and long tangled or swollen mitochondria were observed in 12 cases.

Prognosis

At the last follow-up, the 23 surviving patients had a median age of 7 years (2–27 years), whilst 12 had died with a median age at death of 1.5 years (from the neonatal period to 13 years). The median course of follow-up was 3.5 years (range, 7–24 years), and the median course of disease from illness to death was 1 year (range, 8–11.5 years). At the last follow-up, barring three families with optic atrophy as an isolated symptom, the rest had varying degrees of dyskinesia (four patients could walk with assistance, two could stand with assistance, one could sit alone, 11 could not control body position, and the rest were not mentioned), cognitive impairment, dystonia, or epileptic seizure.

Genetic Analysis

De novo variants were identified in 72.2% (26/36) of the cases, 13.9% (5/36) had heterozygous dominant variants, 11.1% (4/36) had compound heterozygous variants, and 2.8% (1/36) had homozygous variants. A total of 24 *DNM1L* variants were identified, including 21 missense variants, two deletion variants, and one duplication variant (**Table 2**).

Correlation Between Domains and Phenotypes

Clinical Phenotype

The clinical manifestations of the GTPase domain group were milder than those of the middle domain group (**Figure 1G**, **Table 3**). The prevalence of psychomotor retardation and epilepsy in the GTPase domain group was lower than that in the middle-domain group (Psychomotor retardation: GTPase domain group, 50.0% [6/12] vs. middle domain group, 91.3% [21/23]; epilepsy: GTPase domain group, 8.3% [1/12] vs. middle domain group, 89.5% [17/19]); these differences were statistically significant ($P < 0.05$). In the five patients with heterozygous variants, three pedigrees with variants in the GTPase domain presented with optic atrophy as the isolated symptom, which

TABLE 1 | The pathogenicity prediction using different online tools.

Predicting tools	Polyphen2	VariantTaster	SIFT	ACMG
pathogenicity	probably damage	Disease causing	Deleterious	Pathogenic PS2+PM1+PM2+PP1+PP2+PP3

TABLE 2 | Related patient's information and variants in *DNM1L*.

Patients	Country/District	Gender	Variant		Variant type	Reference
Pedigrees 1–2	France	Pedigree 1: four male, four female Pedigree 2: six male, three female	c.5A>C	p.Glu2Ala	Heterozygous variant	Gerber et al. (18)
Patient 3	USA/Canada	Female	c.95G>C	p.G32A	*De novo*	Whitley et al. (22)
Patient 4–5 (A pair of brothers)	Italy	two male	c.106A>G; c.346_347delGA	p.Ser36Gly; p.Glu116Lysfs*6	Compound heterozygous variants	Nasca et al. (5)
Patient 6	China	Male	c.116G>A	p.S39N	*De novo*	—
Patient 7–8 (A pair of siblings)	Canada	Female Male	c.261dup; c.385_386del	p.Trp88Metfs*; p.Glu129Lys*6	Compound heterozygous variants	Yoon et al. (15)
Patient 9	Canada	Female	c.305C>T	p.T115M/T115M	Homozygous variants	Hogarth et al. (19)
Patient 10	Italy	Female	c.436G>A	p.D146N	*De novo*	Longo et al. (26)
Pedigree 11	France	three male, one female	c.575C>A	p.Ala192Glu	Heterozygous variant	Gerber et al. (18)
Patient 12	Italy	Female	c.668G>T	p.Gly223Val	*De novo*	Verrigni et al. (6)
Patient 13	USA	Male	c.1048G> A	p.G350R	Heterozygous variant (Maternal mosaic)	Chao et al. (11)
Patient 14–15	Italy Arab	Female Male	c.1084G>A	p.Gly362Ser	*De novo*	Verrigni et al. (6) Sheffer et al. (13)
Patient 16	Canada	Female	c.1085G>A	p.Gly362Asp	*De novo*	Vanstone et al. (14)
Patient 17	Italy	Male	c.1085G>A c.1535T>C	p.Gly362Asp p.lle512Thr	Heterozygous variant; the two variants were on the same maternal allele, c.1085G>A was present in the mother's blood DNA at a low level	Verrigni et al. (6)
Patient 18	Italy	Female	c.1109T>G	p.Phe370Cys	*De novo*	Verrigni et al. (6)
Patient 19	USA	Female	c.1135G> A	p.E379K	*De novo*	Chao et al. (11)
Patient 20–21	UK USA/Canada	Female Female	c.1184C>A	p.A395D	*De novo*	Waterham et al. (10) Whitley et al. (22)
Patient 22–31	USA Italy USA /UK USA Turkey USA USA/Canada	two male one male two not mentioned one male one male one female one male, one female	c.1207C>T	p.Arg403Cys	*De novo*	Fahrner et al. (12) Fahrner et al. (6) Ladds et al. (20) Ryan et al. (21) Schmid et al. (24) Nolan et al. (23) Whitley et al. (22)
Patient 32	Japan	Male	c.1217T>C	p.Leu406Ser	*De novo*	Zaha et al. (16)
Patient 33	USA	Female	c.1228G> A	p.E410K	*De novo*	Vandeleur et al. (25)
Patient 34	USA/Canada	Female	c.1292G>A	p.C431Y	*De novo*	Whitley et al. (22)
Patient 35	Spain	Female	c.1337G>T	p.Cys446Phe	*De novo*	Zaha et al. (16)
Patient 36	USA	Female	c.2072A> G	p.Tyr691Cys	*De novo*	Assia Batzir et al. (7)

was the mildest phenotype, and the family pedigrees were shown together with the segregation of the *DNM1L* variants in the affected individuals. The variants of the two other patients were located in the middle domain; one patient had a c.1048G>A, p.G350R variant, and variant analysis by Sanger sequencing suggested a low level (6–8%) of maternal mosaicism. The patient presented with developmental delay, hypotonia, status epilepticus, and elevated serum lactate levels, and he died of status epilepticus at the age of 5. Another patient had two heterozygous *DNM1L* variants (c.1085G>A, p.Gly362Asp; c.1535T>C, p.lle512Thr) on the same allele (maternal allele). His mother with the variant (c.1535T>C, p.lle512Thr) was healthy, but the other variant (c.1085G>A, p.Gly362Asp) was not present in either parent. However, further analysis of the *DNM1L* gene using a next-generation sequencing approach revealed that the variant (c.1085G>A, p.Gly362Asp) was present in the mother's

TABLE 3 | The clinical characteristics in patients with *DNM1L* variants in different domains.

Clinical data %(n/n)or n		GTPase domain (*n* = 12)	Middle domain (*n* = 23)	*P*-value
Age of onset		0.5 (0.1–5)	0.5 (0.1–9)	/
Sex	Male	44.4% (4/9)	52.4% (11/21)	1.000
	Female	55.6% (5/9)	47.6% (10/21)	
The last follow-up	Survive	83.3% (10/12)	56.5% (13/23)	0.149
	death	16.7% (2/12)	43.5% (10/23)	
Psychomotor retardation		50.0% (6/12)	91.3% (21/23)	0.011
Limb paralysis		66.7% (8/12)	64.3% (9/14)	1.000
Dependent ambulation		2	2	
Stand		–	1	
Sitting		–	1	
No postural control		6	5	
Dystonia		72.7% (8/11)	88.2% (15/17)	0.353
Epilepsy		8.3% (1/12)	89.5% (17/19)	0.000
Status epilepticus		1	12	
Ataxia		5	1	/
Nystagmus		1	1	/
Optic atrophy		4	1	/
Dysarthria		3	1	/
Microcephalus		1	3	/
Pain insensitivity		–	2	/
Sensory and motor axonal neuropathy		1	1	/
Abnormality in MR image		62.5% (5/8)	90.9% (20/22)	0.102
Hyperintense in cortex		1	9	/
Hyperintense in white matter		1	2	/
Hyperintense in basal ganlia/ hypothalamus		4	6	/
Cerebral atrophy		1	11	/
thinning/absence of corpus callosum		0	3	/
lactate peak		1	3	/
Abnormal EEG		20.0% (1/5)	92.9% (13/14)	0.006
Epileptiform discharge		1	12	
Weak background		–	7	
Hyperlactacidemia		57.1% (4/7)	65.0% (13/20)	0.661
decreased respiratory chain enzyme activity		0.0% (0/3)	36.4% (4/11)	0.505
Ragged red fibers		1	3	/
Abnormal morphology of mitochondria		5	6	/

Data are shown as prevalence %(n/n) or n(number). "–" indicates that it was not mentioned in the literature. Statistical analysis included the chi-square test or Fisher's exact probability, as appropriate. "/" indicates that no statistical analysis was performed.

blood DNA at a very low level. Furthermore, the patient's half-brother had died at the age of 3 years. Therefore, these findings supported maternal germline mosaicism for a dominant variant (c.1085G>A, p.Gly362Asp). The patient manifested developmental delay prior to the onset of illness. At 5 years and 5 months of age, he presented with partial motor status epilepticus after a viral illness. 4 months later, he presented with spastic quadriplegia and hyperkinesis but poor voluntary movements, no postural control, and severe cognitive impairment. However, there was no statistically significant difference in the survival rate in the different domain groups (log-rank test, $P = 0.39$, **Figure 1H**).

Auxiliary examination revealed that 83.3% (25/30) of the cases had abnormalities in brain MR images; 62.5% (5/8) of the GTPase domain group vs. 90.9% (20/22) of the middle domain group ($P = 0.102$), with no statistically significant difference. Seventy-five percent (15/20) of the cases showed abnormality in EEG; 20.0% (1/5) of the GTPase domain group vs. 92.9% (13/14) of the middle domain group ($P = 0.006$), with statistically significant differences.

DISCUSSION

Mitochondrial diseases are a group of disorders with heterogeneous clinical and genetic characteristics caused by mitochondrial structural or function disturbances due to nuclear or mitochondrial DNA variations. There have been several studies on mitochondrial respiratory chain complex defects and mitochondrial diseases, but there are few reports on mitochondrial diseases caused by abnormal mitochondrial dynamics. Mitochondria are highly dynamic organelles that undergo continuous fission, fusion, and migration. The genes that encode these mitochondrial dynamic-related proteins are mitochondrial nuclear genes, and variants in these genes can lead to mitochondrial diseases: (1) mitochondrial membrane fusion abnormality: gene variants in *MFN2* encoding a GTPase joining adjacent mitochondria leads to Charcot–Marie–Tooth disease type 2A and hereditary motor and sensory neuropathy VI, and optic atrophy or Leigh-like infantile encephalopathies are caused by variants in *OPA1*, which encodes a GTPase that mediates the fusion of the inner mitochondrial membrane (5, 6). (2) Mitochondrial fission abnormality: gene variants in *DNM1L* involving mitochondrial fission lead to encephalopathy; gene variants in the mitochondrial fission factor (*MFF*), the receptor gene of *DNM1L*, cause Leigh-like encephalopathy (27). (3) Mitochondrial migrating abnormality: gene variants in *Miro* are associated with Parkinson's disease, Alzheimer's disease, and amyotrophic lateral sclerosis (28, 29).

DNM1L, encoding DRP1, is located in chromosomal region 12p11.21. DRP1 is a highly conserved GTPase that mediates mitochondrial fission and contains an *N*-terminal GTPase domain, a middle domain, a non-conserved variable domain, and a C-terminal GTPase effector domain (GED) (19). DRP-1, a cytosolic protein that can be recruited to the mitochondria in response to cellular signals, is the most important mediator of mitochondrial and peroxisomal division. Different sites in different domains can bind to the receptor on the mitochondrial outer membrane. The GTPase domain is responsible for binding to receptors on the outer membrane of the mitochondria. The middle domain mainly mediates the oligomerization of DRP1, with the fission process executed by DRP1 forming a concentric ring-like structure around the scission site, followed by GTPase-dependent constriction (15, 19). GED stimulates GTPase activity and mediates the formation and stability of the DRP1 homodimer complex (7). *DNM1L* variants impair the oligomerization of DRP1, thereby affecting the contraction of GTPase and leading disrupting mitochondrial fission.

DNM1L-related mitochondrial diseases are rarely reported, and limited gene variants of *DNM1L* are included in the Human Gene Mutation Database. We identified a *DNM1L* variant that had not previously been reported, and enriched the genetic information from the Chinese population to improve the understanding of rare diseases. The clinical phenotype of this patient was consistent with that previously reported.

The clinical symptoms of *DNM1L*-related mitochondrial disease vary. The median age of onset was 6 months (from the neonatal period to 9 years), and the main symptoms included psychomotor retardation, limb paralysis, dystonia (hypermyotonia or hypomyotonia), epilepsy, and other symptoms, including ataxia, nystagmus, optic atrophy, dysarthria, microcephaly, pain insensitivity, and sensory and motor axonal neuropathy.

Most *DNM1L* variants were located in the middle domain ($n = 23$) and GTPase domain ($n = 12$) of DRP1. The severity of clinical symptoms was related to the domains in which the *DNM1L* variants were located. The clinical symptoms of the GTPase domain group were milder than those in the middle domain group. In the GTPase domain group, three families presented with optic atrophy as an isolated symptom, which was the mildest phenotype, and the family pedigrees were shown with the segregation of *DNM1L* variants in the affected individuals. The prevalence of psychomotor retardation and epilepsy in the GTPase domain group was lower than that in the middle domain group, and the proportion of status epilepticus and refractory epilepsy was higher in the middle domain group. However, there was no significant difference in occurrence of limb paralysis or dystonia between the two groups. The variant c.116G>A (p.S39N) in our patient was located in the GTPase domain of DRP1, and he presented with psychomotor retardation, limb paralysis, paroxysmal hypermyotonia, and limb tremors, but no epileptic seizures were observed during the 3-year-follow-up. This highlights the importance of the location of variants in determining clinical phenotypes.

As DRP1 is a cytoplasmic protein, it is recruited to the outer mitochondrial membrane by specific receptors. These include mitochondrial fission 1, MFF, and mitochondrial dynamics proteins of 49 and 51 kDa (MID49, MID51), which can all independently recruit DRP1 to mediate mitochondrial division (12, 30, 31). After binding with its receptors, DRP1 polymerizes to a concentric ring-like structure, encircles the mitochondria, and causes mechanochemical contraction of mitochondria. Taking MID49 as example, each chain of DRP1 binds MID49 through four different surfaces in different domains, and each MID49 in turn binds four DRP1, generating a huge interaction network (31). Variants in the GTPase domain can destroy the binding of the GTPase domain, while there are multiple interaction sites in oligomers that may hinder their cohesion; thus, the clinical phenotype is relatively milder (19) (Kalia et al., 2018). In addition to *de novo* variants, reported homozygous variants and compound heterozygous variants were all located in the GTPase domain, while recessive genes included null alleles and hypomorphic alleles, which could explain the different clinical phenotypes (5, 21). However, variations in the middle domain can impair the oligomerization of DRP1 and cause mitochondrial fission defects; therefore, the clinical phenotype is typically worse.

The prognosis of *DNM1L*-related mitochondrial diseases is poor. At the last follow-up, only 23 patients had survived, with 12 deaths with at a median age of 1.5 years. Although there was no statistically significant difference in the survival rate in different domains, patients with variants in the GTPase domain had a relatively better prognosis and could survive longer, and three families with variants in the GTPase domain showed isolated manifestations of optic atrophy. The remaining patients had

varying degrees of dyskinesia, cognitive impairment, dystonia, or epileptic seizures.

By reviewing the relevant literature, the most common symptoms of *DNM1L*-related mitochondrial diseases were summarized; based on these finding, *DNM1L*-related mitochondrial diseases should be considered in the early stages of severe encephalopathy, including neonatal fatal diseases, psychomotor retardation, motor disorders, cognitive disorders, and epilepsy. The variant sites of *DNM1L* were summarized, and the understanding of the *DNM1L* gene profile was expanded.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Medical Ethics Committee of Tianjin Children's Hospital, Tianjin Children's Hospital. 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

ZZ and DL retrieved and summarized the related literature. XinL wrote the review article. ML created the figures. QL provided statistical analysis. XiaL provided the medical records. PZ revised the review article. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We thank our patient's family for allowing us to share his case study. This study was approved by the ethics committee of our hospital. We would like to thank Editage (www.editage.cn) for English language editing.

REFERENCES

1. Gorman GS, Chinnery PF, Dimauro S, Hirano M, Koga Y, Mcfarland R, et al. Mitochondrial diseases. *Nat Rev Dis Primers.* (2016) 2:16080. doi: 10.1038/nrdp.2016.80
2. Meyer JN, Leuthner TC, Luz ALJT. Mitochondrial fusion, fission, mitochondrial toxicity. *Toxicology*. (2017) 391:42–53. doi: 10.1016/j.tox.2017.07.019
3. Youle RJ, Van Der Bliek AM. Mitochondrial fission, fusion, and stress. *Science*. (2012) 337:1062–5. doi: 10.1126/science.1219855
4. Smirnova E, Griparic L, Shurland DL, Van Der Bliek AM. Dynamin-related protein Drp1 is required for mitochondrial division in mammalian cells. *Mol Biol Cell.* (2001) 12:2245–56. doi: 10.1091/mbc.12.8.2245
5. Nasca A, Legati A, Baruffini E, Nolli C, Moroni I, Ardissone A, et al. Biallelic variants in DNM1L are associated with a slowly progressive infantile encephalopathy. *Hum Mutat.* (2016) 37:898–903. doi: 10.1002/humu.23033
6. Verrigni D, Di Nottia M, Ardissone A, Baruffini E, Nasca A, Legati A, et al. Clinical-genetic features and peculiar muscle histopathology in infantile DNM1L-related mitochondrial epileptic encephalopathy. *Hum Mutat.* (2019) 40:601–18. doi: 10.1002/humu.23729
7. Assia Batzir N, Bhagwat PK, Eble TN, Liu P, Eng CM, Elsea SH, et al. De novo missense variant in the GTPase effector domain (GED) of DNM1L leads to static encephalopathy and seizures. *Cold Spring Harb Mol Case Stud.* (2019) 5:a003673. doi: 10.1101/mcs.a003673
8. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point variants in impure and heterogeneous cancer samples. *Nat Biotechnol*. (2013) 31:213–9. doi: 10.1038/nbt.2514
9. Kai W, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. (2010) 38:e164. doi: 10.1093/nar/gkq603
10. Waterham HR, Koster J, Van Roermund CW, Mooyer PA, Wanders RJ, Leonard JV. A lethal defect of mitochondrial and peroxisomal fission. *N Engl J Med.* (2007) 356:1736–41. doi: 10.1056/NEJMoa064436
11. Chao YH, Robak LA, Xia F, Koenig MK, Adesina A, Bacino CA, et al. Missense variants in the middle domain of DNM1L in cases of infantile encephalopathy alter peroxisomes and mitochondria when assayed in Drosophila. *Hum Mol Genet.* (2016) 25:1846–56. doi: 10.1093/hmg/ddw059
12. Fahrner JA, Liu R, Perry MS, Klein J, Chan DC. A novel de novo dominant negative variant in DNM1L impairs mitochondrial fission presents as childhood epileptic encephalopathy. *Am J Med Genet A*. (2016) 170:2002–11. doi: 10.1002/ajmg.a.37721
13. Sheffer R, Douiev L, Edvardson S, Shaag A, Tamimi K, Soiferman D, et al. Postnatal microcephaly pain insensitivity due to a de novo heterozygous DNM1L variant causing impaired mitochondrial fission function. *Am J Med Genet A*. (2016) 170:1603–7. doi: 10.1002/ajmg.a.37624
14. Vanstone JR, Smith AM, Mcbride S, Naas T, Holcik M, Antoun G, et al. DNM1L-related mitochondrial fission defect presenting as refractory epilepsy. *Eur J Hum Genet.* (2016) 24:1084–8. doi: 10.1038/ejhg.2015.243
15. Yoon G, Malam Z, Paton T, Marshall CR, Hyatt E, Ivakine Z, et al. Lethal disorder of mitochondrial fission caused by variants in DNM1L. *J Pediatr.* (2016) 171:313–6.e311-312. doi: 10.1016/j.jpeds.2015.12.060
16. Zaha K, Matsumoto H, Itoh M, Saitsu H, Kato K, Kato M, et al. DNM1L-related encephalopathy in infancy with Leigh syndrome-like phenotype and suppression-burst. *Clin Genet.* (2016) 90:472–4. doi: 10.1111/cge.12805
17. Díez H, Cortès-Saladelafont E, Ormazábal A, Marmiese AF, Armstrong J, Matalonga L, et al. Severe infantile parkinsonism because of a de novo variant on DLP1 mitochondrial-peroxisomal protein. *Mov Disord.* (2017) 32:1108–10. doi: 10.1002/mds.27021
18. Gerber S, Charif M, Chevrollier A, Chaumette T, Angebault C, Kane MS, et al. Variants in DNM1L, as in OPA1, result in dominant optic atrophy despite opposite effects on mitochondrial fusion and fission. *Brain.* (2017) 140:2586–96. doi: 10.1093/brain/awx219
19. Hogarth KA, Costford SR, Yoon G, Sondheimer N, Maynes JT. DNM1L Variant alters baseline mitochondrial function and response to stress in a patient with severe neurological dysfunction. *Biochem Genet.* (2018) 56:56–77. doi: 10.1007/s10528-017-9829-2
20. Ladds E, Whitney A, Dombi E, Hofer M, Anand G, Harrison V, et al. De novo DNM1L variant associated with mitochondrial epilepsy syndrome with fever sensitivity. *Neurol Genet.* (2018) 4:e258. doi: 10.1212/NXG.0000000000000258
21. Ryan CS, Fine AL, Cohen AL, Schiltz BM, Renaud DL, Wirrell EC, et al. De Novo DNM1L variant in a teenager with progressive paroxysmal dystonia and lethal super-refractory myoclonic status epilepticus. *J Child Neurol.* (2018) 33:651–8. doi: 10.1177/0883073818778203
22. Whitley BN, Lam C, Cui H, Haude K, Bai R, Escobar L, et al. Aberrant Drp1-mediated mitochondrial division presents in humans with variable outcomes. *Hum Mol Genet.* (2018) 27:3710–9. doi: 10.1093/hmg/ddy287
23. Nolan DA, Chen B, Michon AM, Salatka E, Arndt D. A Rasmussen encephalitis, autoimmune encephalitis, and mitochondrial disease mimicker: expanding the DNM1L-associated intractable epilepsy and encephalopathy phenotype. *Epileptic Disord.* (2019) 21:112–6. doi: 10.1684/epd.2019.1036
24. Schmid SJ, Wagner M, Goetz C, Makowski C, Freisinger P, Berweck S, et al. A De Novo dominant negative variant in DNM1L

causes sudden onset status epilepticus with subsequent epileptic encephalopathy. *Neuropediatrics.* (2019) 50:197–201. doi: 10.1055/s-0039-1685217

25. Vandeleur D, Chen CV, Huang EJ, Connolly AJ, Sanchez H, Moon-Grady AJ. Novel and lethal case of cardiac involvement in DNM1L mitochondrial encephalopathy. *Am J Med Genet A*. (2019) 179:2486–9. doi: 10.1002/ajmg.a.61371
26. Longo F, Benedetti S, Zambon AA, Sora MGN, Di Resta C, De Ritis D, et al. Impaired turnover of hyperfused mitochondria in severe axonal neuropathy due to a novel DRP1 variant. *Hum Mol Genet.* (2020) 29:177–88. doi: 10.1093/hmg/ddz211
27. Shamseldin HE, Alshammari M, Al-Sheddi T, Salih MA, Alkhalidi H, Kentab A, et al. Genomic analysis of mitochondrial diseases in a consanguineous population reveals novel candidate disease genes. *J Med Genet.* (2012) 49:234–41. doi: 10.1136/jmedgenet-2012-100836
28. Zhang F, Wang W, Siedlak SL, Liu Y, Liu J, Jiang K, et al. Miro1 deficiency in amyotrophic lateral sclerosis. *Front Aging Neurosci.* (2015) 7:100. doi: 10.3389/fnagi.2015.00100
29. Kay L, Pienaar IS, Cooray R, Black G, Soundararajan M. Understanding Miro GTPases: implications in the treatment of neurodegenerative disorders. *Mol Neurobiol.* (2018) 55:7352–65. doi: 10.1007/s12035-018-0927-x
30. Serasinghe MN, Chipuk JE. Mitochondrial fission in human diseases. *Handb Exp Pharmacol.* (2017) 240:159–88. doi: 10.1007/164_2016_38
31. Kalia R, Wang RY, Yusuf A, Thomas PV, Agard DA, Shaw JM, et al. Structural basis of mitochondrial receptor binding and constriction by DRP1. *Nature.* (2018) 558:401–5. doi: 10.1038/s41586-018-0211-2

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

frontiers in Genetics

CASE REPORT
published: 09 July 2021
doi: 10.3389/fgene.2021.666225

Case Report: Diagnosis of Human Alveolar Echinococcosis via Next-Generation Sequencing Analysis

Ke Li[1], *Yubao Ma*[2], *Rui Ban*[3] *and Qiang Shi*[2]*

[1] Department of Neurology, The Second Medical Center, Chinese PLA General Hospital, Beijing, China, [2] Department of Neurology, The First Medical Center, Chinese PLA General Hospital, Beijing, China, [3] Department of Neurology, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China

OPEN ACCESS

Edited by:
Ruth Roberts,
ApconiX, United Kingdom

Reviewed by:
Tatiana Kuester,
University of Bern, Switzerland
Urmas Saarma,
University of Tartu, Estonia
Klaudija Viskovic,
University Hospital for Infectious Diseases "Dr Fran Mihaljevic", Croatia
Bita Geramizadeh,
Shiraz University of Medical Sciences, Iran

***Correspondence:**
Qiang Shi
shiq301@163.com

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

Received: 09 February 2021
Accepted: 16 June 2021
Published: 09 July 2021

Citation:
Li K, Ma Y, Ban R and Shi Q (2021) Case Report: Diagnosis of Human Alveolar Echinococcosis via Next-Generation Sequencing Analysis. Front. Genet. 12:666225. doi: 10.3389/fgene.2021.666225

Introduction: Alveolar echinococcosis (AE) is a rare parasitic disease caused by the infection of *Echinococcus multilocularis*. AE may mimic malignancy both in clinical presentation and radiological imaging, which is often misdiagnosed as metastatic tumor. Recently, next-generation sequencing (NGS) technologies are increasingly being used to address a diverse range of biological questions. Here, we describe a rare case of alveolar echinococcosis diagnosed by pan-pathogen screening, using next-generation sequencing. To the best of our knowledge, this is the first reported case of AE which was definitely diagnosed relying NGS of cerebrospinal fluid (CSF).

Case Presentation: A 33-year-old man presented with repeat seizure and progressive headache for six months. Head magnetic resonance imaging (MRI) showed multiple masses with edema. Lung and abdominal computer tomography (CT) revealed multiple masses in bilateral lung, liver and the right adrenal gland. Bacterial, tuberculosis and fungal infection were excluded by CSF examination. Repeated target biopsy on the masses in the lung and liver showed as fibrous connective tissue without positive findings. NGS of CSF was performed and detected nucleic acid sequences of *E. multilocularis*. Consequently, the patient has accepted 1-year albendazole therapy. His case was followed up through imaging procedures.

Conclusion: The next-generation sequencing of CSF is a reliable and sensitive diagnostic method for the detection of pathogenic microorganisms, and may allow the accurate diagnosis of alveolar echinococcosis. In view of this case, we recommend NGS as a potential tool for diagnosis of cerebral AE, especially if repeated biopsies are negative.

Keywords: alveolar echinococcosis, next-generation sequencing, adrenal gland, *Echinococcus multilocularis*, diagnose

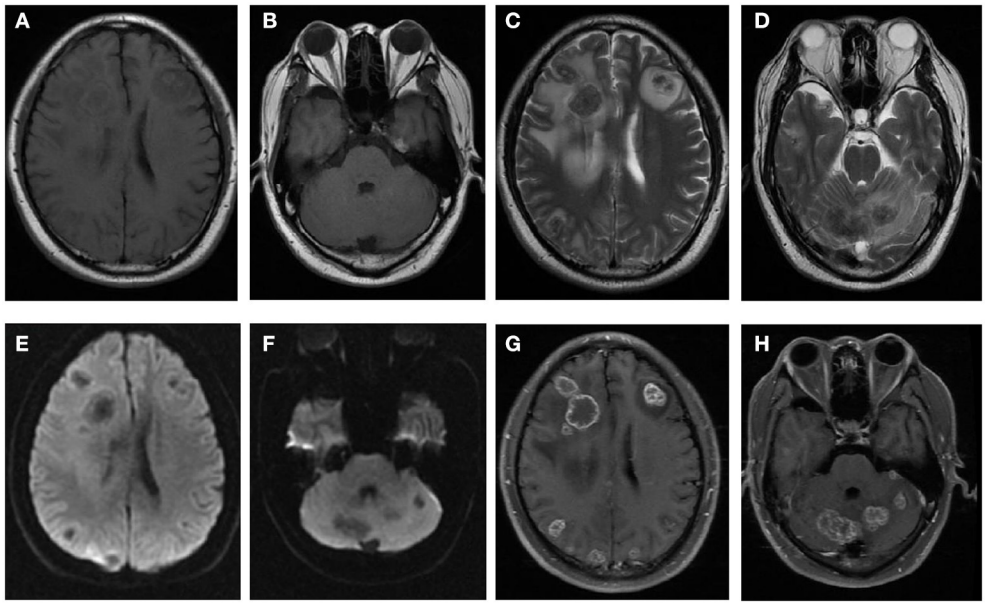

FIGURE 1 | Brain MRI images of the patient. Multiple lesions revealed with isointensity on T1WI **(A,B)**, hypointensity on T2WI **(C,D)** surrounded by edema, hypointensity on DWI **(E,F)** and irregular ring enhancement after injection of Gd-DTPA **(G,H)**.

INTRODUCTION

Brain alveolar echinococcosis is a fatal parasitic disease caused by *Echinococcus multilocularis* (Deplazes et al., 2017), which is often misdiagnosed as metastatic tumor or intracranial tuberculosis. Diagnosis of AE relies on clinical presentation, imaging examinations, serological test and biopsy if available. However, metastatic tumors and AE are difficult to differentiate through imaging examinations. In addition, serological test sometimes results in biologically false-positive results. Herein, we present a rare case of cerebral AE which was finally confirmed by NGS. This is the first reported case which is conclusively diagnosed as cerebral AE by NGS of CSF.

CASE PRESENTATION

A 33-year-old man presented repeat seizures and progressive headache for six months. A complete blood count (CBC) showed: hemoglobin 123g/L, white blood cell count 7.73×10^9/L, and percentage of eosinophil 0.009. Liver function showed: alanine transaminase (ALT) 92.8 U/L, aspartate aminotransferase (AST) 44.9U/L, alkaline phosphatase (ALP) 135.1U/L, γ-glutamyl transferase (γ-GT) 313.4 U/L, total bilirubin (TBil) 22.7 mmol/L, direct bilirubin (DBil) 19.4 mmol/L. Head MRI (**Figure 1**) showed multiple masses with edema. Lung and abdominal CT (**Figure 2**) presented with several lesions in bilateral lungs, liver and right adrenal gland. Bacterial, tuberculosis and fungal infection were excluded by CSF examination. Serological evaluation of multiple parasite antigen by ELISA were applied. Cysticercosis Immunoglobulin G (IgG) and liver hydatid IgG antibody were both positive. Repeated target biopsy on the masses in the lung and liver showed fibrous connective tissue without positive findings. In addition, we performed positron emission tomography with fluorodeoxyglucose integrated with CT (^{18}F-FDG PET/CT) that showed uptake in all masses with a maximum standardized uptake value of 7.2. Accordingly, it was still hard to draw a definitive pathogenic diagnosis. Therefore, the next-generation sequencing of CSF was performed. The CSF was collected according to standard procedures, and DNA was extracted directly from the sample with TIANamp Micro DNA Kit. The extracted DNA was sonicated to a size of 200–300 bp (Bioruptor Pico protocols). The DNA libraries were constructed and sequencing using the BGISEQ-100 platform. After removing human sequences, the remaining sequencing data were aligned to the microbial databases and detected 161 nucleic acid sequences of *E. multilocularis* (**Table 1**, *E. multilocularis* sequences which detected were provided as **Supplementary Material**). On this basis, the patient was diagnosed as having AE. Consequently, the patient was recommended a 1-year albendazole therapy. During 1 year of follow-up, symptoms and neurological signs were not aggravated, with decreased seizure frequency. Follow-up CT after 1 year of albendazole treatment revealed slightly decreased multiple lesions and partly relieved surrounding edema in brain. The lesion in the liver (red arrow) was evidently diminished and calcification was slightly increased. The thickened right adrenal gland (yellow arrow) has obviously decreased in size (**Figure 3**).

DISCUSSION AND CONCLUSIONS

We reported a special human alveolar echinococcosis case involving the brain, lung, liver, adrenal gland, which was confirmed by NGS. AE is endemic in certain parts of the

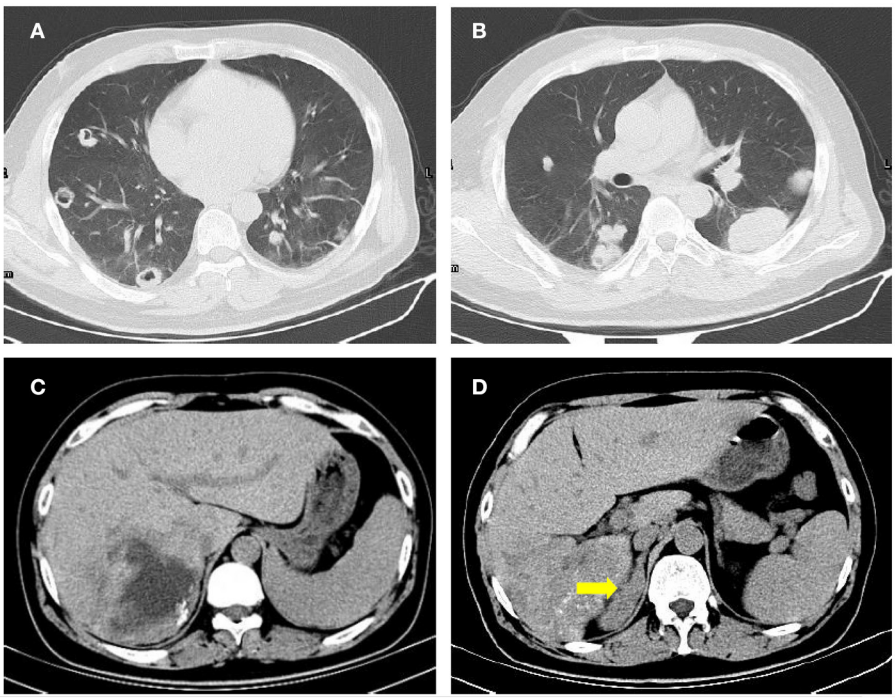

FIGURE 2 | Lung and abdominal CT images of the patient. Multiple Lesions were found in bilateral lung **(A,B)** and the liver **(C)**. The right adrenal gland (yellow arrow) was obviously thickened **(D)**.

TABLE 1 | List of parasites detected.

Genus		Species	
Name	**Numbers of sequences detected**	**Name**	**Numbers of sequences detected**
Echinococcus	318	*Echinococcus multilocularis*	161

world, especially in Europe, Northern America, and Central Asia (Deplazes et al., 2017). A recent meta-analysis indicated that the pooled prevalence of AE in China was 0.96% (Wang et al., 2020). In China, the AE endemic area is restricted to the northwest region (Qian et al., 2017), for example, Qinghai province and Gansu province. In most cases, AE is initially located in the liver and spreads into other organs by infiltration or metastasis formation. Extrahepatic lesions are usually located in the lung and brain (Tappe et al., 2008). Adrenal AE, however, is rare, with only 9 cases reported in the literature so far (Huang and Zheng, 2013; Spahn et al., 2016; Seidel et al., 2017). Our case suggests that AE may involve all parts of the body. We should suspect the diagnosis of AE in a patient with exposure history of parasite and a mass in the adrenal gland.

The diagnosis of AE is complicated by diverse clinical features and mimicking of differential diagnoses. For this case, the patient's biopsy results were negative. Therefore, a biopsy is likely to have lower sensitivity for the diagnosis of AE due to limited tissue specimens. Imaging findings of multiple lesions are helpful in differential diagnosis; however, it can be misinterpreted as other metastatic tumors even by experienced radiologists due to their limited awareness of this condition. Another difficulty in the differential diagnosis of AE and cysticercosis is that serologic test revealed that both of cysticercosis IgG and liver hydatid IgG were all positive. Is it a co-infection or immunological cross-reaction? We prefer immunological cross-reaction since there are common antigens between *Cysticercus cellulosae* and *E. multilocularis*. While mixed infection of 2 pathogens is extremely rare. Furthermore, NGS detected nucleic acid sequences of *E. multilocularis* without sequences of *C. cellulose,* which conclusively exclude cysticercosis.

Finally, AE diagnosis has been confirmed by NGS for this patient. NGS technologies are increasingly being used to address a diverse range of biological and epidemiological questions since NGS provides pathogen identification without prior target knowledge. In addition, other genetic diagnostic methods, such as PCR, is not applicable for screening of rare and unknown pathogens. We believe that this is the first reported case of AE which was definitely diagnosed relying next-generation sequencing of CSF. Clinicians should promptly recognize that NGS of CSF might be provided as a potential test for detecting cerebral AE.

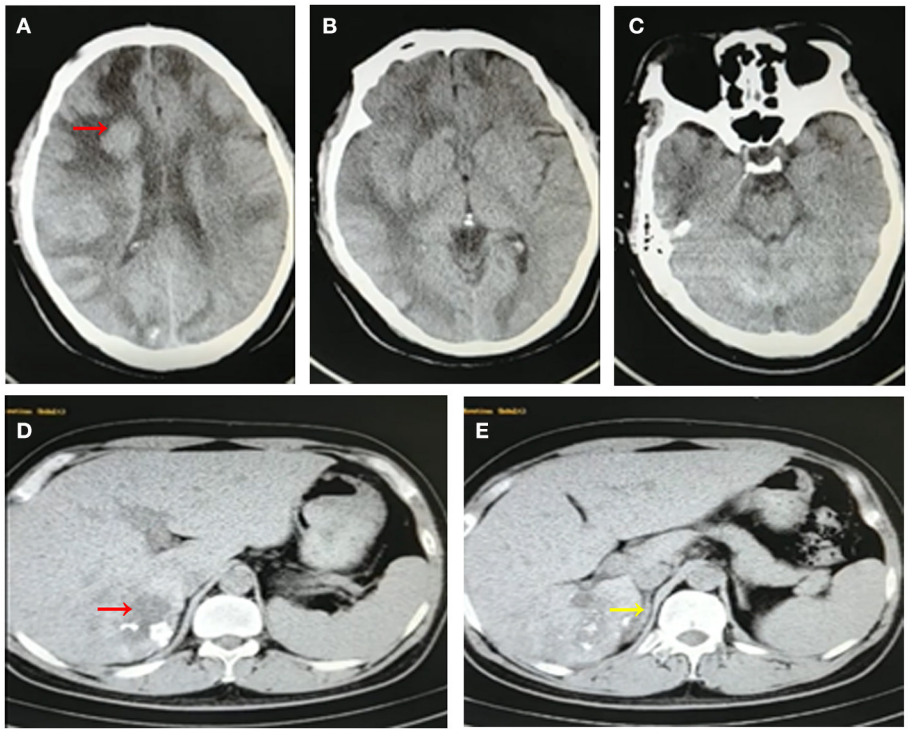

FIGURE 3 | Follow-up CT images of the patient after 1-year albendazole therapy. Brain CT images revealed multiple lesions has slightly shrink and surrounding edema was slightly relieved **(A–C)**. Abdominal CT images showed that the lesion in liver was obviously diminished accompanied with calcification **(D,E)**.

The limitation of this study includes small sample size. Therefore, if there are similar cases in the future, we should expand the sample size and validate the accuracy of NGS method. In conclusion, the NGS of CSF is a reliable and sensitive diagnostic method to detect pathogenic microorganisms, which may allow accurate diagnoses of cerebral AE.

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 This work was supported by National Natural Science Foundation of China (No. 81601086, 81771358). 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

All co-authors have seen and agree with the contents of the manuscript, that the ICMJE requirements for authorship have been met, and that each author believes that the manuscript represents honest work. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by National Natural Science Foundation of China (Grant No.81601086, 81771358).

ACKNOWLEDGMENTS

We would like to thank the patient and his family who are the focus of this report.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.666225/full#supplementary-material

REFERENCES

Deplazes, P., Rinaldi, L., Rojas, C., Torgerson, P. R., and Jenkins, E. J. (2017). Global distribution of alveolar and cystic echinococcosis. *Adv. Parasitol.* 95, 315–493. doi: 10.1016/bs.apar.2016.11.001

Huang, M., and Zheng, H. (2013). Primary alveolar echinococcosis (*Echinococcus multilocularis*) of the adrenal gland: report of two cases. *Int. J. Infect. Dis.* 17, e653–e655. doi: 10.1016/j.ijid.2013.03.011

Qian, M. B., Abela-Ridder, B., Wu, W. P., and Zhou, X. N. (2017). Combating echinococcosis in China: strengthening the research and development. *Infect. Dis. Poverty* 6:161. doi: 10.1186/s40249-017-0374-3

Seidel, A. K., Pless, M., Michel, C., Soll, C., Hochuli, C., and Gubler, J. (2017). A rare differential diagnosis of an adrenal mass: a case report. *Case Rep. Oncol.* 10, 981–986. doi: 10.1159/000481501

Spahn, S., Helmchen, B., and Zingg, U. (2016). Alveolar echinococcosis of the right adrenal gland: a case report and review of the literature. *J. Med. Case Rep.* 10:325. doi: 10.1186/s13256-016-1115-0

Tappe, D., Weise, D., Ziegler, U., Müller, A., Müllges, W., and Stich, A. (2008). Brain and lung metastasis of alveolar echinococcosis in a refugee from a hyperendemic area. *J. Med. Microbiol.* 57(Pt 11), 1420–1423. doi: 10.1099/jmm.0.2008/002816-0

Wang, X., Dai, G., Li, M., Jia, W., Guo, Z., and Lu, J. (2020). Prevalence of human alveolar echinococcosis in China: a systematic review and meta-analysis. *BMC Public Health* 20:1105. doi: 10.1186/s12889-020-08989-8

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

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 16 July 2021
doi: 10.3389/fphar.2021.644352

Mutational Characteristics of Causative Genes in Chinese Hereditary Spherocytosis Patients: a Report on Fourteen Cases and a Review of the Literature

Dong Wang[1†], Li Song[2†], Li Shen[3], Kaihui Zhang[1], Yuqiang Lv[1], Min Gao[1], Jian Ma[1], Ya Wan[1], Zhongtao Gai[1] and Yi Liu[1]**

[1]Pediatric Research Institute, Qilu Children's Hospital of Shandong University, Jinan, China, [2]Pediatric Hematology and Oncology, Qilu Children's Hospital of Shandong University, Jinan, China, [3]Clinical Laboratory, The Fourth Hospital of Jinan, Jinan, China

OPEN ACCESS

Edited by:
Tieliu Shi,
East China Normal University, China

Reviewed by:
Marianthi Georgitsi,
Aristotle University of Thessaloniki, Greece
Yaqiong Jin,
Capital Medical University, China

***Correspondence:**
Zhongtao Gai
gaizhongtao@cinnamon
Yi Liu
y_liu99@sina.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *21 December 2020*
Accepted: *24 May 2021*
Published: *16 July 2021*

Citation:
Wang D, Song L, Shen L, Zhang K, Lv Y, Gao M, Ma J, Wan Y, Gai Z and Liu Y (2021) Mutational Characteristics of Causative Genes in Chinese Hereditary Spherocytosis Patients: a Report on Fourteen Cases and a Review of the Literature. Front. Pharmacol. 12:644352. doi: 10.3389/fphar.2021.644352

Background: Hereditary spherocytosis (HS), characterized by the presence of spherocytic red cells in peripheral blood, hemolysis, splenomegaly, jaundice, and gallstones, is a common form of inherited hemolytic anemia (HA). To date, five causative genes associated with HS have been identified, including *ANK1, SPTB, SPTA1, SLC4A1*, and *EPB42*.

Methods: Clinically suspected patients with HS or undiagnosed HA from 14 Chinese families were enrolled in this study. We presented the patients' clinical features and identified the causative gene variants in these patients using whole exome sequencing (WES), with 10 novel and four reported mutations in the *ANK1* and *SPTB* genes (seven mutations in *ANK1* and seven in *SPTB*), individually. Then, we reviewed all available literature on Chinese HS patients from 2000 to 2020 in PubMed and Chinese Journals with genetic results and clinical information, to delineate gene mutation spectrum and potential correlation with phenotypes.

Results: A total of 158 variants (including 144 in previous reports and 14 in this study) indicated that *ANK1* (46%) and *SPTB* (42%) were the most frequently mutated genes in Chinese HS patients, followed by *SLC4A1* (11%) *and SPTA1* (1%), while no mutations in *EPB42* was reported. Most of the mutations in *ANK1* and *SPTB* were nonsense (26/73 in *ANK1* and 32/66 in *SPTB*) and frameshift (20/73 in *ANK1* and 15/66 in *SPTB*), while missense mutations (14/18) accounted for the majority in *SLC4A1*. The higher mutation frequency of *ANK1* was found in its exon 8, 9, 26, and 28. The majority of mutations in *SPTB* were located in its exon 13, 15, and 18–30, whereas mutations in *SLC4A1* were scattered throughout the entire region of the gene.

Conclusion: Our study expanded the mutation spectrum of *ANK1* and *SPTB*. Furthermore, we clarified the mutational characteristics of causative genes by reviewing all available literature on Chinese patients with HS.

Keywords: hereditary spherocytosis, mutation, ANK1, SPTB, whole exome sequencing

INTRODUCTION

Hereditary spherocytosis (HS) refers to a group of heterogeneous disorders, and is a common form of congenital hemolytic anemia (HA). It is characterized by the presence of sphere-shaped red blood cells (spherocytes) on the peripheral blood smear, anemia, jaundice, and splenomegaly, with wide heterogeneity in severity ranging from virtually asymptomatic conditions to severe forms that require transfusions in early childhood, which can make an exact diagnosis of HS cases quite difficult, particularly for asymptomatic or atypical cases if only depending on clinical manifestations, family history and hematologic laboratory tests (Bolton-Maggs et al., 2012). It has been reported worldwide with a prevalence that varies from 1:2,000 in Caucasians to 1:100,000 in Chinese cases (Perrotta et al., 2008; Wang et al., 2018). The study of the gene mutation spectrum has been subject to some general concerns.

Early diagnosis of HS is necessary to prevent adverse outcomes, and if diagnosed too late, HS cases are prone to risk of long-term complications, such as cholelithiasis, hemolytic episodes, and aplastic crises; furthermore, non-diagnosed HS may lead to a severe neurological complication called kernicterus (Perrotta et al., 2008; Will et al., 2017; Ciepiela, 2018; Brosius, 2019). Splenectomy is curative and considered the standard surgical treatment for those with moderate to severe HS conditions but has a life-long risk of potentially lethal infections (Manciu et al., 2017). Since traditional laboratory tests often fail to diagnose HS, particularly asymptomatic or atypical cases, molecular genetic testing, especially next-generation sequencing is becoming a powerful tool for clinical diagnosis of HS in neonates or infants with its capability of accurately identifying genetic variants (Long and Shen, 2019; Shen and Shi, 2019; Yang et al., 2019).

The molecular basis of HS is due to defects in red cell membrane proteins which result in decreased membrane surface area and resultant deformability of the erythrocytes, thereby accelerating the degradation and hemolysis in the spleen. The severity of the anemia is directly associated with the extent of membrane surface area loss and consequent increase of cell sphericity (Narla and Mohandas, 2017; Iolascon et al., 2019). The deficiency of membrane in the erythrocytes is caused by their corresponding gene mutations, including *ANK1*, *SPTA1*, *SPTB*, *SLC4A1,* and *EPB42*, which encodes ankyrin 1, spectrin *a*-chain, spectrin *ß*-chain, anion exchanger 1 (band 3), and protein 4.2, respectively (Perrotta et al., 2008; Da Costa et al., 2013; Narla and Mohandas, 2017; Iolascon et al., 2019). Autosomal dominant inheritance is the main inherited manner of HS accounting for 75%, while autosomal recessive (AR) and non-dominant inheritance have also been described in about 25% of cases (Delaunay, 1995; Iolascon and Avvisati, 2008; Garcon, 2009; Zamora and Schaefer, 2019). Thus, the definitive diagnosis of HS is determined by molecular diagnostic study. It has been reported that the mutation in *ANK1* (~50%) and spectrin gene (*SPTB*: ~20% and *SPTA1*: ~5%) is the major cause of HS, followed by a mutation in *SLC4A1* (~15%) and *EPB42* (~10%) (An and Mohandas, 2008; Zamora and Schaefer, 2019). Until recently, next-generation sequencing was performed on patients with HS from different countries and regions, such as America, Europe, Brazil, and Korea (Agarwal et al., 2016; Mansour-Hendili et al., 2020; Svidnicki et al., 2020; Shin et al., 2018). There have been some case reports in Chinese HS patients (Wang et al., 2018; Meng et al., 2019; Qin et al., 2020). However, the gene mutation spectrum has not been delineated well to date.

In the present study, we investigated causative gene mutations in a cohort of 14 patients with clinically suspected HS or undiagnosed HA from Shandong Province, a northern area of China, using next generation sequencing. The 14 patients were from 14 unrelated families, including one neonate, six infants, and seven children over one-year-old. We identified 10 novel mutations in ANK1 and *SPTB* genes. In addition, we discussed the clinical features of these patients in the light of the previous reports regarding mutations of the causative genes in Chinese HS patients to provide information for genetic counseling, prenatal screening, and future research of individual therapy.

MATERIALS AND METHODS

Ethics Statement

This work was approved by the Medical Ethics Committee of Qilu Children's Hospital of Shandong University. Clinical and laboratory examinations were performed on the probands and their parents after informed consent was obtained. The information of all patients was anonymized before submission. All the procedures performed in this work were in accordance with the Helsinki Declaration.

Patients

Fourteen hospitalized children (8 boys and six girls) from 14 unrelated Chinese families with clinically suspected HS or undiagnosed HA were recruited for this study in Qilu Children's Hospital, Shandong University (QCHSU) from February 2016 to July 2020. All probands from the Han Chinese population in Shandong Province, China were examined and diagnosed by experienced hematology specialists from the Hematology or Hematology Oncology Department of QCHSU (**Table 1**).

Genetic Analysis

Blood samples collected in EDTA vacutainer were obtained from the probands and their parents for DNA extraction using QIAamp DNA Blood Midi Kit (Qiagen, Shanghai, China) according to the manufacturer's protocol. Whole exome sequencing (WES) with the Human Exome Probes P039-Exome (MyGenostics, Beijing, China) on the Illumina NovaSeq5000 platform (Illumina, United States) was applied for the mutation screening of the probands. The obtained mean exome coverage was more than 95% (>10X coverage; mean depth of over 100X). Paired-end alignment

TABLE 1 | The clinical and laboratory features of the Chinese patients with HS.

Patients		P 1	P 2	P 3	P 4	P 5	P 6	P 7	P 8	P9	P10	P11	P12	P13	P14
Gender		M	F	M	M	M	M	M	F	F	M	M	F	F	F
Age		1 y4 m	1 m7 d	1 m1 d	11 m4 d	1 y3 m	8 d	11 y	5 m20 d	1 y1 m	1 m18 d	7 y	1 m6 d	1 y5 m	6 y
Age at the onset		2 m	3 d	1 d	1 d	NA	1 d	6 y	2 d	1 d	2 d	7 y	2 d	1 y4 m	15 d
Clinical symptoms	Jaundice	–	+	+	+	+	+	+	+	+	+	–	+	+	+
	Anemia	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Splenomegaly	+	–	–	–	NA	–	+	+	+	+	–	–	–	+
Lab tests	RBC ($\times 10^{12}$/L)	2.12	1.95	1.92	1.85	3.18	3.19	2.28	1.98	2.61	1.79	1.95	1.66	3.78	2.48
	Hemoglobin (g/L)	50	60	59	56	92	109	79	57	64	54	54	54	109	76
	MCV (fL)	72.4	86.2	84.2	83.8	81.8	96.6	99.8	82.8	77.8	91.1	77.4	94.1	86.8	98.0
	MCH (pg)	23.4	30.8	29.2	28.9	28.9	34.2	34.7	29.0	24.5	30.2	27.8	32.3	28.8	30.6
	MCHC (g/L)	323	357	347	346	354	354	347	350	315	338	359	343	332	313
	Reticulocytes (%)	4.10	3.56	7.63	5.78	12.92	8.44	14.23	8.47	12.20	6.30	14.40	8.88	11.70	17.02
	Total bilirubin (μmol/L)	22.0	484.7	202.0	272.3	54.4	510.8	87.7	395	38	179	34.1	146.4	148.9	55.7
	LDH (U/L)	349	384	234	471	354	572	401	757	324	541	424	233	239	399
	Spherocytes on the peripheral blood smears	–	+	+	+	+	+	+	+	+	+	+	+	+	+
	Direct antiglobulin test	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Increased osmotic fragility	–	–	–	–	+	+	+	–	+	NA	NA	+	+	+
Family history		HA	HS	HS	–	–	–	–	HS	–	–	HA	HS	–	–
Clinical diagnosis		HA	HS	HS	HS	HS	HS	HS	HS	HS	HS	HS	HS	HS	HS

M, male; F, female; HA, hemolytic anemia; HS, Hereditary spherocytosis; LDH, lactate dehydrogenase; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; RBC, red blood cell. NA: not available; y, year(s); m, month(s); d, day(s); "+": Present; "–": Absent.

was performed using Burrows-Wheeler Aligner software (BWA Version: 0.7.10) to version GRCh37/hg19 of the human genome. SAM files then were sorted and converted to BAM. BAM files were filtered and duplicates were marked with Picard. Local realignment was performed using GATK's RealignerTargetcreator and IndelRealigner, base quality score recalibration using GATK's BaseRecalibrator, and variants were called jointly in all samples using the GATK's HaplotypeCaller in the "GENOTYPE_GIVEN_ALLELES" mode. Then, GATK's VariantFiltration was used for filtering SNPs and Indels. ANNOVAR was used for annotation (http://wannovar.wglab.org/) after variant detection. Variant frequencies were determined in thousands of genomes (http://www.1000genomes.org), ExAC (http://exac.broadinstitute.org/), Exome Variant Server (EVS,http://evs.gs.washington.edu/EVS) and in-house database to remove common variants (suballelic frequency >5%). Then, the variants resulting in frameshift, missense variants, premature stop-gain, or initiation codon loss, and typical splicing site changes were prioritized for study. SIFT, PolyPhen-2, MutationTaster, and REVEL were used to evaluate the novel variants' pathogenicity. Moreover, genetic variations included in HGMD (http://www.hgmd.cf.ac.uk) and ClinVar (http://www.ncbi.nlm.nih.gov/clinvar) will also be further analyzed. The variants identified in this study were classified according to the 2015 American College of Medical Genetics and Genomics (ACMG) guidelines (Richards et al., 2015).

Validation of Gene Mutations

Sanger sequencing was then utilized to validate the definitely and likely pathogenic variants identified by whole exome sequencing in the patients with designed specific primers. Sanger validation primer sets were designed using Primer Premier v5.0 software. PCR amplification was performed using AmpliTaq Gold® 360 DNA polymerase (Applied Biosystems). PCR products were further purified and sequenced using an ABI Prism 3700 automated sequencer (Applied Biosystems, Foster City, CA).

Literature Review and Statistical Analysis

To delineate gene mutation spectrum and potential correlation with phenotypes, we reviewed all available published reports of Chinese HS patients from 2000 to 2020 in PubMed and Chinese Journals and classified all the cases into different groups based on 1) mutated genes: *AN*K1 group, *SPTB* group, and *SLC4*A1 group; 2) types of mutations: missense group, nonsense group, frameshift group, and splicing group. The genotype-phenotype correlation was analyzed by comparing the clinical parameters of the patients, such as hemoglobin, total bilirubin, and reticulocytes among different groups. The association analysis was determined using the Kruskal–Wallis test. The significance of association was estimated by calculating the odds ratio and relative risk with a 95% confidence interval (CI). All tests were two tailed, and p-value < 0.05 was considered statistically significant. All statistical analyses were performed using SPSS v19.0 software.

TABLE 2 | Mutations analyzed by WES and validated by Sanger sequencing in the patients from eight Chinese families.

Patients	Clinical diagnosis	Gene	Exon	DNA change	Effect	Mutation type	dbSNP/1000G/EVS/ExAC	Status	Inheritance	Pathogenicity prediction score (Mutation Taster)	Pathogenic evaluation according to ACMG
P1	HA	*ANK1*	31	c.3754C > T	p.R1252*	Nonsense	0/0/0/0	Reported (Nakanishi et al., 2001)	Paternal	1	Pathogenic
P2	HS	*SPTB*	25	c.5266C > T	p.R1756*	Nonsense	0/0/0/0	Reported (Maciag et al., 2009)	Paternal	1	Pathogenic
P3	HS	*SPTB*	13	c.1912C > T	p.R638*	Nonsense	0/0/0/0	Reported Qin et al. (2020)	Maternal	1	Pathogenic
P4	HS	*SPTB*	15	c.3168dupG	p.L1057Afs*16	Frameshift	0/0/0/0	Novel	Paternal	1	Pathogenic
P5	HS	*SPTB*	24	c.4978C > T	p.Q1660*	Nonsense	0/0/0/0	Novel	*De novo*	1	Pathogenic
P6	HS	*ANK1*	4	c.319C > T	p.Q107*	Nonsense	0/0/0/0	Novel	*De novo*	1	Pathogenic
P7	HS	*SPTB*	27	c.5933_5934delAG	p.E1978G*18	Frameshift	0/0/0/0	Novel	*De novo*	1	Pathogenic
P8	HS	*ANK1*	7	c.709C > T	p.Q237*	Nonsense	0/0/0/0	Novel	Paternal	1	Pathogenic
P9	HS	*ANK1*	26	c.2950C > T	p.Q984*	Nonsense	0/0/0/0	Novel	*De novo*	1	Pathogenic
P10	HA	*ANK1*	31	c.3813_3823del	p.Q1272Lfs*100	Frameshift	0/0/0/0	Novel	*De novo*	1	Pathogenic
P11	HS	*SPTB*	18	c.3984G > A	p.W1328*	Nonsense	0/0/0/0	Novel	Maternal	1	Pathogenic
P12	HS	*SPTB*	15	c.3448dupT	p.W1150L*32	Frameshift	0/0/0/0	Novel	Paternal	1	Pathogenic
P13	HS	*ANK1*	9	c.856C > T	p.R286*	Nonsense	0/0/0/0	Reported Wang et al. (2018)	*De novo*	1	Pathogenic
P14	HS	*ANK1*	31	c.3847delA	p.R1283Gfs*3	Frameshift	0/0/0/0	Novel	*De novo*	1	Pathogenic

The variants are described using NM_020476.2 for ANK1 *and NM_001024858.2 for* SPTB *transcript reference sequences.* De novo: *build up from nothing.*

RESULTS

Clinical Manifestations and Laboratory Tests

All patients age from 8 days to 11 years old manifesting anemia and jaundice (except for P1 and P11) with varying severity. Of them, eight patients (P2, P3, P4, P6, P8, P9, P12, and P14) presented anemia or jaundice in the neonatal period, and six patients (P1, P7, P8, P9, P10, and P14) displayed splenomegaly on abdominal ultrasound. Blood biochemical index showed variable low level hemoglobin (Hb), reticulocytosis, and hyperbilirubinemia (unconjugated). Anisocytosis or spherical red blood cells were observed on the peripheral blood smear of all the patients (except for P1). Autoimmune hemolytic anemia was excluded by a direct antiglobulin test. Seven of the patients presented an increase in erythrocyte osmotic fragility. Moreover, six patients had a family history of HS or HA. Clinical data from them are shown in **Table 1**.

Gene Mutation Test

In total, 14 different heterozygous candidate mutations were identified including 10 novel and four reported, with seven in *ANK1* and seven in *SPTB*, nine nonsense, and five frameshift mutations (**Table 2**). All 10 novel mutations can be considered null mutations and result in a premature terminator codon in the coding sequence of *ANK1* or *SPTB* that may cause a defective protein product lacking key domains of ankyrin one or *ß*-spectrin. The validation of parental genetic tests showed that seven mutations (5 in *ANK1* and two in *SPTB*) were *de novo* and the others were inherited mutations. All the 14 mutations were finally determined to be pathogenic after analyzing according to the latest ACMG guidelines (**Table 2**), all 14 probands were genetically confirmed as HS patients and interventions took place to prevent complications.

The Literature Review and Mutational Characteristics of Chinese HS Patients

We summarized data for 158 cases from 144 Chinese patients with HS (genetically confirmed cases) previously reported in the literature and 14 HS cases identified in this study (**Supplementary Table 1**), and the mutational characteristics of causative genes in all these patients were summarized. In total, 158 mutations consisted of 144 previous reports and 14 mutations in this study with 157 heterozygous in four genes of *ANK1, SPTB, SLC4A1,* and *SPTA1*, while one exception of homozygous in *SPTB*, including 73 (46%) in *ANK1*, 66 (42%) in *SPTB*, 18 (11%) in *SLC4A1*, 1 (1%) in *SPTA1,* but no *EPB42* (0%), which indicates that *ANK1* and *SPTB* mutations are two major causes of HS in the Chinese population. The types of gene mutations varied in the different genes, for example, the mutations in *ANK1* including 26 (36%) nonsense, 20 (27%) frameshift, 13 (18%) missense, 11 (15%) splicing and 3 (4%) nucleotide substitutions in the start codon. The mutations in *SPTB* included 32 (48%) nonsense, 15 (23%) frameshift, 9 (14%) missense, 8 (12%) splicing, 1 (1.5%) nucleotide substitution in

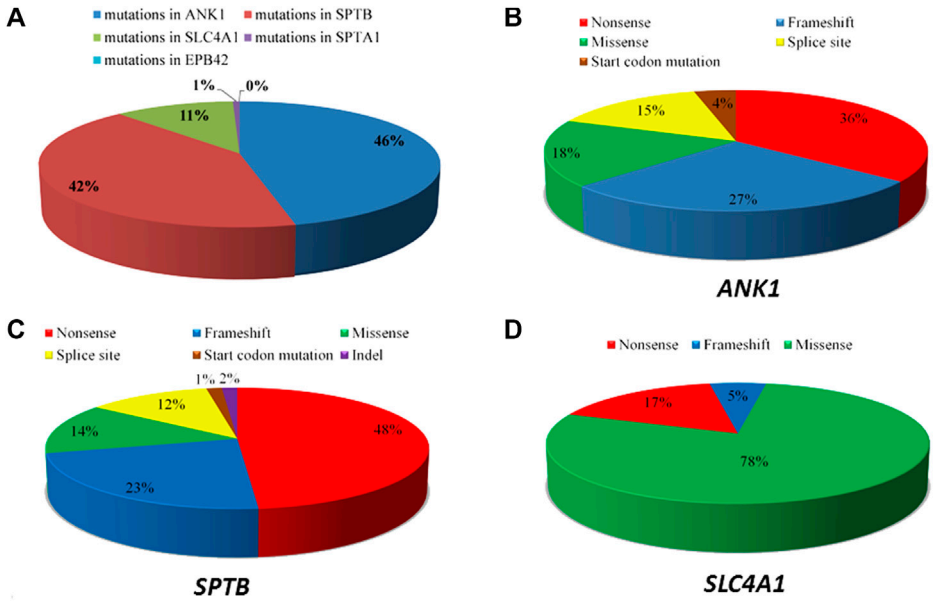

FIGURE 1 | Summary of the causative genes spectrum of Chinese HS patients. **(A)** The distribution of five causative genes (*ANK1, SPTB, SPTA1, SLC4A1*, and *EPB42*) mutations. **(B)** Types of *ANK1* mutations. **(C)** Types of *SPTB* mutations. **(D)** Types of *SLC4A1* mutations.

start condon, and 1 (1.5%) indel mutations. The mutations in *SLC4A1* included 3 (17%) nonsense, 1 (5%) frameshift and 14 (78%) missense mutations. Therefore, the majority of mutation types in both *ANK1* and *SPTB* were nonsense (26/73 in *ANK1* and 32/66 in *SPTB*) followed by frameshift (20/73 in *ANK1* and 15/66 in *SPTB*), but missense mutations (14/18) accounted for the majority of the *SLC4A1* mutations. **Figure 2** shows that exon 8, 9, 26, and 28 of the *ANK1* gene have a higher mutation frequency, and the majority of *SPTB* mutations were located in exon 13, 15, and 18–30, whereas *SLC4A1* mutations were scattered throughout the entire gene. The analysis of all the mutations is shown in **Figure 1** and **Figure 2**.

Genotype–Phenotype Correlation in HS Patients

To investigate the genotype–phenotype association of all HS patients, we first divided the patients into different groups based on the mutated genes and types of mutations, separately, then analyzed their data among the groups by Kruskal–Wallis test which showed no significant differences regarding Hb level, reticulocytes, and total bilirubin among different groups (**Table 3**, **Table 4**).

DISCUSSION

The molecular basis of HS is the mutations in the genes coding red cell membrane components (Delaunay, 2007). Therefore, the detection of these mutations is the principal method to accurately diagnose HS. In the present study, 14 different heterozygous candidate mutations (10 novel and four reported) were identified in all 14 Chinese HS families, and seven families harbored *ANK1* mutations (7/14, 50%), and seven harbored *SPTB* mutations (7/14, 50%), whereas no mutation in the *SPTA1*, *SLC4A1*, and *EPB42* genes were found. According to ACMG standards (Richards et al., 2015), all the 10 novel mutations were classified as pathogenic, and have not been reported in any human gene mutation databases. The detailed evaluation is presented in **Supplementary Table 2**. Non-diagnosed HS may lead to complications, such as kernicterus, hemolytic anemia, and the development of gall stones. Thus, early diagnosis of HS is crucial to reduce the risk of complications later in life (Steward et al., 2014; Christensen et al., 2015). The 14 HS cases diagnosed in this study have an average age of 2.2 years (including 7 cases less than 1 year old), meaning that an earlier diagnosis is more conducive to reducing the risk of complications in these patients.

The analysis of all HS-associated mutations (including 144 from previous reports and 14 from this study) in Chinese patients showed that *ANK1* or *SPTB* was the major causative gene of HS cases, which is consistent with previous reports (Delaunay, 2007; Park et al., 2016; Wang et al., 2018; Qin et al., 2020), but only 5–10% of HS cases in Japan or Brazil were caused by *ANK1* mutations, implying that the geographical distribution of *ANK1* gene mutations is different. In addition, Chinese HS patients who carried *SLC4A1* mutations only account for 14%, which is close to the mutation incidence of Caucasian (15–20%) or Japanese patients (20%). Hughes et al. obtained a mice model by using N-ethyl-N-nitrosourea mutagenesis to generate random point mutations in exon 27 of Ank1 (E924X) in the mouse genome. They found that heterozygous mice have low RBC (MCV), reticulocytosis, and increased osmotic fragility, which are the clinical features of HS, while the homozygous mutant mice displayed severe hemolytic anemia, profound extramedullary

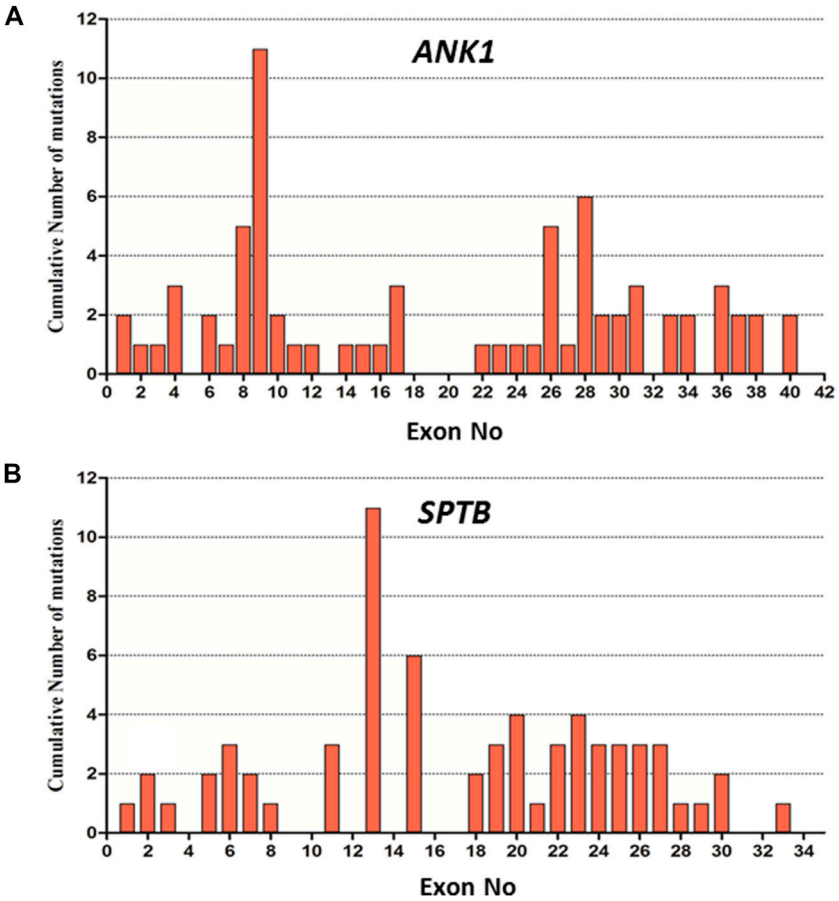

FIGURE 2 | Distribution of mutations in the exon of *ANK1* and *SPTB* gene. **(A)** Cumulative Number of mutations in each exon of the *ANK1* gene. **(B)** Cumulative Number of mutations in each exon of *SPTB* gene.

TABLE 3 | Comparison of clinical features of HS patients with *SPTB*, *ANK1* and *SLC4A1* mutations.

Clinical data	*ANK1* (*n* = 73)	*SPTB* (*n* = 66)	*SLC4A1* (*n* = 18)	*p*-value
Hb (g/L), median (range)	72.4 (26.0–148.0) *n* = 34	76.3 (10.8–125) *n* = 33	96.0 (56.0–120.0) *n* = 4	0.131
Ret (%), median (range)	10.86 (3.09–23.57) *n* = 25	9.82 (1.10–17.44) *n* = 27	12.95 (5.60–20.30) *n* = 2	0.884
T-Bil (μmol/L), median (range)	125.9 (22.0–521.2) *n* = 26	112.0 (25.1–484.7) *n* = 26	125.4 (45.2–257.0) *n* = 3	0.768

p-*values of <0.05 were considered statistically significant. Hb, hemoglobin; Ret, reticulocyte; T-Bil, total bilirubin.*

TABLE 4 | Comparison of clinical features of HS patients with different types of mutation.

Clinical data	Nonsense (*n* = 61)	Frameshift (*n* = 36)	Missense (*n* = 36)	Splicing (*n* = 19)	*p*-value
Hb (g/L), median (range)	74.2 (10.8–125) *n* = 28	76.2 (40.0–148.0) *n* = 19	80.7 (55.0–120.0) *n* = 13	73.5 (56.0–104.0) *n* = 10	0.853
Ret (%), median (range)	10.29 (3.56–17.88) *n* = 20	9.74 (1.93–23.57) *n* = 16	12.06 (1.10–20.30) *n* = 9	12.95 (5.60–20.30) *n* = 7	0.656
T-Bil (μmol/L), median (range)	144.4 (22.0–510.8) *n* = 22	122.5 (39.6–521.2) *n* = 16	72.1 (25.1–177.2) *n* = 8	93.5 (52.3–118.1) *n* = 8	0.385

p-*values of <0.05 were considered statistically significant. Hb, hemoglobin; Ret, reticulocyte; T-Bil, total bilirubin.*

hematopoiesis, stress erythropoiesis, and many pups died within the first 2 weeks of birth.

Ankyrin-1 protein, encoded by *the ANK1* gene, is important for interacting with transmembrane proteins and the membrane skeleton of cells through *ß*-spectrin, band 3, and band 4.2 proteins. High affinity interaction is critical for the deformability and stability of the erythrocytes. Its deficiency leads to a decrease of spectrin assembly on the membrane, which causes anemia. *ANK1,* located on chromosome

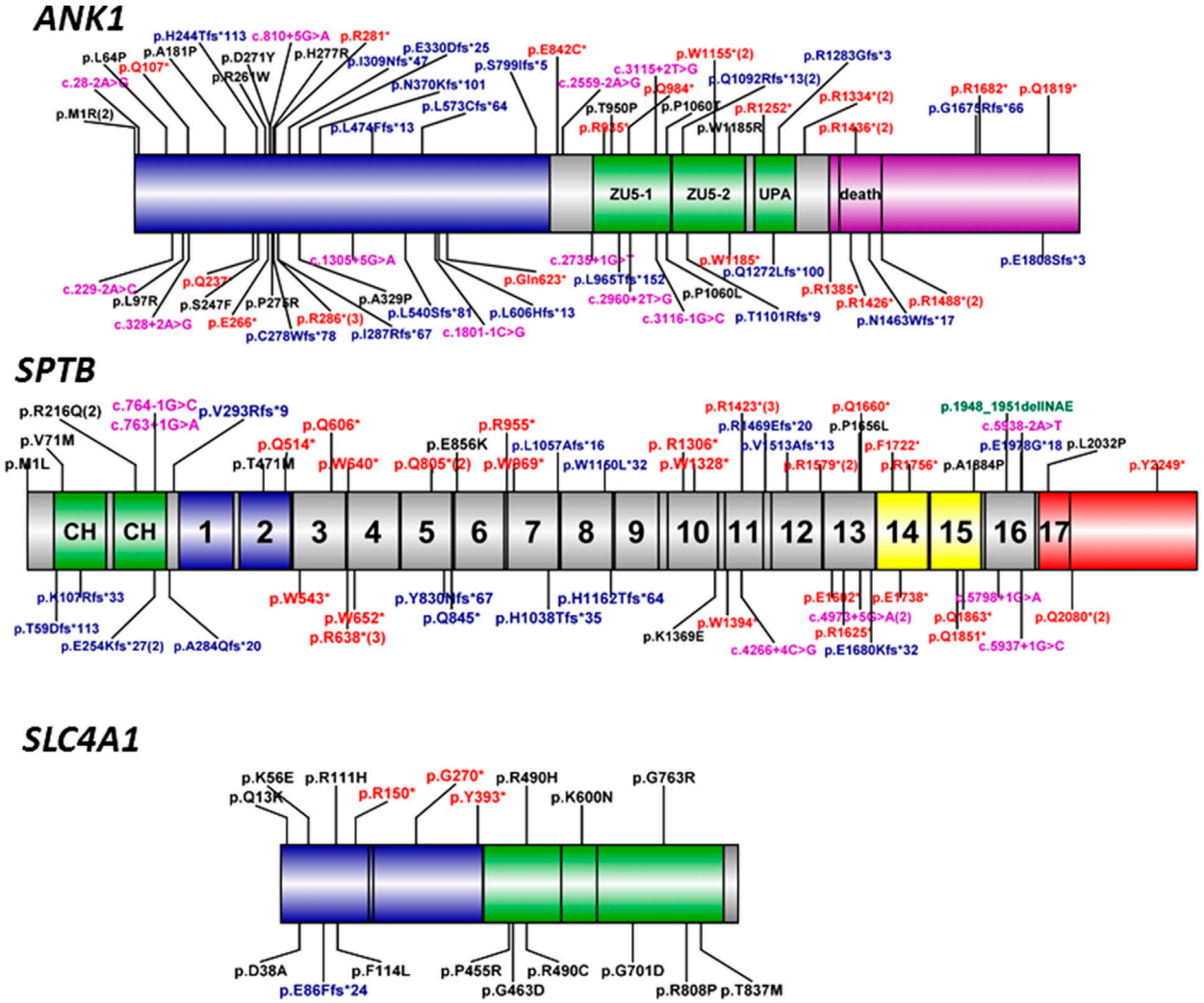

FIGURE 3 | Schematic diagram of ankyrin, *ß*-spectrin, and band three protein domains with *ANK1, SPTB,* and *SLC4A1* mutations. Human erythroid ankyrin encoded by *ANK1* consists of an N-terminal membrane protein binding domain containing ankyrin repeats (blue box), a central spectrin-binding domain that involves two ZU5 and the UPA domains (green box), and a C-terminal regulatory domain that modulates the affinities of the other domains and contains a death domain (violet box). Human erythroid-spectrin protein encoded by *SPTB* mainly consists of two N-terminal actin-binding domains (green box) and seventeen spectrin repeats, of which repeats one and 2 (blue box) mediate α/β dimer formation, repeats 14 and 15 (yellow box) are the ankyrin-binding domain, and the last repeat is a tetramerization domain (red box). Human band three protein encoded by *SLC4A1* mainly consists of an N-terminal cytoplasmic domain (blue box) responsible for *ANK1* binding and the C-terminal domain (green box) spanning the lipid bilayer involved in anion transport. The different mutation types were labeled with different colors, black-missense, red-nonsense, blue-frameshift, violet-splicing.

8p11.2, contains 42 exons and encodes the ankyrin-1 protein (ANK1) with 1,881 amino acids (NM_020476.2), consisting of three main domains: an N-terminal membrane protein domain containing binding site for the band three membrane protein, a central domain binding the spectrin that involves two ZU5 and one UPA domain for interacting with the actin-spectrin cytoskeleton, and a regulatory domain of C-terminal, responsible for modulating the affinities of the other domains (**Figure 3**). In this study, five novel mutations were identified in *ANK1*, including two (p.Q107*, p. Q237*) on the N-terminal domain, three mutations on the central domain including one (p.Q984*) on ZU5-1, and two (p.Q1272Lfs*100, p. R1283Gfs*3) on UPA, which were all deleterious mutations, have been predicted to damage the binding with the band three membrane protein and the actin-sepctrin cytoskeleton resulting in the occurrence of HS.

To our knowledge, more than 190 *ANK1* mutations associated with HS have been described to date, including nonsense, missense, and splicing mutations, small or gross deletions, insertions, regulatory mutations, and complex rearrangements. In this study, we have identified 5 novel and two reported *ANK1* mutations including five nonsense and two frameshift mutations (**Table 2**). We then summarized and analyzed the mutational characteristics of *ANK1* mutations in Chinese HS patients and found that exon 8, 9, 26, and 28 of *ANK1* gene are high frequency mutant exons (**Figure 2**) and types of *ANK1* mutations in Chinese HS patients include nonsense mutations (36%), frameshift mutations (27%, containing small or

gross deletions, insertions), missense mutations (18%), splicing abnormalities (15%) and start codon mutations (4%), but no regulatory mutations and complex rearrangements. All the *ANK1* mutations summarized in this study were heterozygous mutations, although autosomal recessive (AR) and non-dominant inheritance have been described in the HS patients with *ANK1* mutations by a few reports (Eber et al., 1996; Iolascon et al., 2019; Zamora and Schaefer, 2019). In addition, for the first time, we analyzed the distribution of *ANK1* mutations in the ankyrin-1 domain and found that the *ANK1* mutations in Chinese HS patients are mainly distributed in the N-terminal membrane protein binding domain, two ZU5 and the UPA domains, as well as the death domain. However, different types of *ANK1* mutations in different domains were presented as randomly distributed (**Figure 3**).

β-spectrin, one of the principal sepctrins of the assembly, is anchored to membrane proteins *via* ankyrin-1 by which spectrin plays a crucial role in the formation and stability of the erythrocyte membrane. The *SPTB* gene (MIM: 182,870), located on chromosome 14q23-q24.2, has a length of more than 100 kb and encodes *ß*-spectrin with 2,328 amino acids (NM_001024858), forming the cytoskeletal superstructure of the erythrocyte plasma membrane, which is comprised of two N-terminal calponin homology (CH) domains and seventeen spectrin repeats domains containing two domains mediating α/β dimer formation, two ankyrin-binding domains and a tetramerization domain (**Figure 3**) (Ipsaro et al., 2009). 185 *SPTB* mutations have been reported in the HGMD database. Unlike *ANK1* mutations, which were identified only in HS patients, *SPTB* mutations could result in other diseases including hereditary elliptocytosis and hereditary pyropoikilocytosis. Here, we have identified 5 novel and two reported *SPTB* mutations including four nonsense and four frameshift mutations only in the HS patients (**Table 2**). The five novel pathogenic mutations included two nonsense (p.W1328*, p. Q1660*) and three fromeshit (p.L1057Afs*16, p. W1150L*32, p. E1978G*18) located on the spectrin repeats domains of 7, 8, 10, 13, and 16, which resulted in truncation or premature termination of the protein causing loss of normal function thereby the occurrence of HS.

Some reports have pointed out that the *SPTB* mutations in HS patients are mainly distributed outside the tetramerization domain of the C-terminus in *ß*-spectrin (Park et al., 2016; Wang et al., 2018). Until now, 66 *SPTB* mutations (including the seven mutations identified in this study, see **Supplementary Table 1**) have been reported in 66 unrelated Chinese HS patients. Almost all these mutations were distributed outside the tetramerization domain except four mutations (c.6095T > C, c.6747C > G and two c.6238C > T, see **Supplementary Table 1** and **Figure 3**), which might be associated with the HE phenotype, and only two mutations located on the spectrin repeats one and two domain. In addition, Most of these *SPTB* mutations are nonsense (32/66) and frameshift (20/73) mutations, which can be considered as null mutations and cause haploinsufficiency of *SPTB* and *ß*-spectrin deficiency in HS patients, and, interestingly, most of the nonsense mutations are mainly distributed in repeats 3–15 of erythroid-spectrin protein (**Figure 3**).

Band three or anion exchanger 1 (AE1) encoded by *SLC4A1* contains three distinct domains: an N-terminal band three cytoplasmic domain (residue 1–403) that is responsible for *ANK1* binding, a C-terminal domain (residues 404–882) that consists of 12–14 segments spanning the lipid bilayer involved in anion transport, and a short C-terminal cytoplasmic tail at the extreme (residues 883–911) with the function of binding to carbonic anhydrase II (Reithmeier et al., 2016). In this study, no *SLC4A1* mutation was found in our 14 cases. Until now, 176 *SLC4A1* mutations have been reported in the professional version of the HGMD database. Mariani et al. reported that band three deficiencies were the most common protein abnormalities (54%) in European HS patients (Mariani et al., 2008); however, only 18 mutations (11%) dispersed throughout the whole gene have been reported in Chinese HS patients (**Supplementary Table 1** and **Figure 3**), and most of these *SLC4A1* mutations were missense mutations, this result is consistent with some previous reports (Yawata et al., 2000; Barneaud-Rocca et al., 2011; Reithmeier et al., 2016).

In addition, only one compound heterozygous mutation in the *SPTA1* gene and no *EPB42* mutations were reported in Chinese HS patients, indicating that pathogenic mutations of the two genes are relatively rare in the Chinese population.

This study found no significant differences in Hb level, reticulocytes, and total bilirubin among different groups of *ANK1*, *SPTB,* and *SLC4A1* genes. These results are consistent with a previous report (Qin et al., 2020). In addition, we also compared the clinical features among different types of mutations, but no significant differences were found among them either.

CONCLUSION

This study has reported on 14 Chinese patients with suspected clinical features of HS and identified 14 pathogenic gene mutations (10 novel and four reported in *ANK1* and *SPTB*) by whole-exome sequencing. We clarified the mutational characteristics in Chinese patients with HS: *ANK1* gene mutation is the most major cause of HS followed by the *SPTB* gene. Findings also indicated that most mutations in *ANK1* or *SPTB* are heterozygous nonsense or frameshift mutations, which are also known as null mutations.

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: GenBank (https://www.ncbi.nlm.nih.gov/Genbank) with accession number from MW812426 to MW812439.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Medical Ethics Committee of Qilu Children's Hospital of Shandong 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

ZG and YL conceived and designed this study. Experiments were conducted by DW, KZ, YL, MG, JM, and YW. Data were analyzed by DW, LS, and YL. Clinical diagnosis of the patients was undertaken by LS, LS, and ZG. The paper was written by DW and YL.

ACKNOWLEDGMENTS

We thank the patients and their families for their participation and contribution to the work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.644352/full#supplementary-material

REFERENCES

An, X., and Mohandas, N. (2008). Disorders of Red Cell Membrane. *Br. J. Haematol.* 141 (3), 367–375. doi:10.1111/j.1365-2141.2008.07091.x

Barneaud-Rocca, D., Pellissier, B., Borgese, F., and Guizouarn, H. (2011). Band 3 Missense Mutations and Stomatocytosis: Insight into the Molecular Mechanism Responsible for Monovalent Cation Leak. *Int. J. Cell Biol.* 2011, 136802. doi:10.1155/2011/136802

Bolton-Maggs, P. H., Langer, J. C., Iolascon, A., Tittensor, P., and King, M. J. (2012). Guidelines for the Diagnosis and Management of Hereditary Spherocytosis--2011 Update. *Br. J. Haematol.* 156 (1), 37–49. doi:10.1111/j.1365-2141.2011.08921.x

Brosius, J. (2019). Exaptation at the Molecular Genetic Level. *Sci. China Life Sci.* 62 (4), 437–452. doi:10.1007/s11427-018-9447-8

Christensen, R. D., Yaish, H. M., and Gallagher, P. G. (2015). A Pediatrician's Practical Guide to Diagnosing and Treating Hereditary Spherocytosis in Neonates. *Pediatrics* 135 (6), 1107–1114. doi:10.1542/peds.2014-3516

Ciepiela, O. (2018). Old and New Insights into the Diagnosis of Hereditary Spherocytosis. *Ann. Transl. Med.* 6 (17), 339. doi:10.21037/atm.2018.07.35

Da Costa, L., Galimand, J., Fenneteau, O., and Mohandas, N. (2013). Hereditary Spherocytosis, Elliptocytosis, and Other Red Cell Membrane Disorders. *Blood Rev.* 27 (4), 167–178. doi:10.1016/j.blre.2013.04.003

Delaunay, J. (2007). The Molecular Basis of Hereditary Red Cell Membrane Disorders. *Blood Rev.* 21 (1), 1–20. doi:10.1016/j.blre.2006.03.005

Delaunay, J. (1995). Genetic Disorders of the Red Cell Membrane. *Crit Rev Oncol. Hematol.* 19 (2), 79–110. doi:10.1016/1040-8428(94)00139-k

Eber, S. W., Gonzalez, J. M., Lux, M. L., Scarpa, A. L., Tse, W. T., Dornwell, M., et al. (1996). Ankyrin-1 Mutations Are a Major Cause of Dominant and Recessive Hereditary Spherocytosis. *Nat. Genet.* 13 (2), 214–218. doi:10.1038/ng0696-214

Garçon, L. (2009). [Biological Diagnosis of Red Cell Membrane Disorders]. *Arch. Pediatr.* 16 (6), 553–555. doi:10.1016/S0929-693X(09)74064-3

Iolascon, A., and Avvisati, R. A. (2008). Genotype/phenotype Correlation in Hereditary Spherocytosis. *Haematologica* 93 (9), 1283–1288. doi:10.3324/haematol.13344

Iolascon, A., Andolfo, I., and Russo, R. (2019). Advances in Understanding the Pathogenesis of Red Cell Membrane Disorders. *Br. J. Haematol.* 187 (1), 13–24. doi:10.1111/bjh.16126

Ipsaro, J. J., Huang, L., and Mondragón, A. (2009). Structures of the Spectrin-Ankyrin Interaction Binding Domains. *Blood* 113 (22), 5385–5393. doi:10.1182/blood-2008-10-184358

Long, M., and Shen, B. (2019). Evolution of Genes and Genomes: an Emerging Paradigm in Life Science. *Sci. China Life Sci.* 62 (4), 435–436. doi:10.1007/s11427-018-9448-x

Manciu, S., Matei, E., and Trandafir, B. (2017). Hereditary Spherocytosis - Diagnosis, Surgical Treatment and Outcomes. A Literature Review. *Chirurgia (Bucur)* 12, 110–116. doi:10.21614/chirurgia.112.2.110

Mariani, M., Barcellini, W., Vercellati, C., Marcello, A. P., Fermo, E., Pedotti, P., et al. (2008). Clinical and Hematologic Features of 300 Patients Affected by Hereditary Spherocytosis Grouped According to the Type of the Membrane Protein Defect. *Haematologica* 93 (9), 1310–1317. doi:10.3324/haematol.12546

Narla, J., and Mohandas, N. (2017). Red Cell Membrane Disorders. *Int. J. Lab. Hematol.* 39 (Suppl. 1), 47–52. doi:10.1111/ijlh.12657

Park, J., Jeong, D. C., Yoo, J., Jang, W., Chae, H., Kim, J., et al. (2016). Mutational Characteristics of ANK1 and SPTB Genes in Hereditary Spherocytosis. *Clin. Genet.* 90 (1), 69–78. doi:10.1111/cge.12749

Perrotta, S., Gallagher, P. G., and Mohandas, N. (2008). Hereditary Spherocytosis. *Lancet* 372 (9647), 1411–1426. doi:10.1016/S0140-6736(08)61588-3

Qin, L., Nie, Y., Zhang, H., Chen, L., Zhang, D., Lin, Y., et al. (2020). Identification of New Mutations in Patients with Hereditary Spherocytosis by Next-Generation Sequencing. *J. Hum. Genet.* 65 (4), 427–434. doi:10.1038/s10038-020-0724-z

Reithmeier, R. A., Casey, J. R., Kalli, A. C., Sansom, M. S., Alguel, Y., and Iwata, S. (2016). Band 3, the Human Red Cell Chloride/bicarbonate Anion Exchanger (AE1, SLC4A1), in a Structural Context. *Biochim. Biophys. Acta* 1858 (7 Pt A), 1507–1532. doi:10.1016/j.bbamem.2016.03.030

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

Shen, Y., and Shi, T. (2019). Innovation for Better Health of Children. *Sci. China Life Sci.* 62 (12), 1561–1562. doi:10.1007/s11427-019-1589-2

Steward, S. C., Chauvenet, A. R., and O'Suoji, C. (2014). Hereditary Spherocytosis: Consequences of Delayed Diagnosis. *SAGE Open Med.* 2, 2050312114547093. doi:10.1177/2050312114547093

Wang, R., Yang, S., Xu, M., Huang, J., Liu, H., Gu, W., et al. (2018). Exome Sequencing Confirms Molecular Diagnoses in 38 Chinese Families with Hereditary Spherocytosis. *Sci. China Life Sci.* 61 (8), 947–953. doi:10.1007/s11427-017-9232-6

Will, A., Henderson, C. A., Jnah, A. J., and Newberry, D. (2017). Hereditary Spherocytosis in the Neonatal Period: A Case Report. *Neonatal. Netw.* 36 (5), 280–288. doi:10.1891/0730-0832.36.5.280

Yang, N., Wu, S., and Yan, J. (2019). Structural Variation in Complex Genome: Detection, Integration and Function. *Sci. China Life Sci.* 62 (8), 1098–1100. doi:10.1007/s11427-019-9664-4

Yawata, Y., Kanzaki, A., Yawata, A., Doerfler, W., Ozcan, R., and Eber, S. W. (2000). Characteristic Features of the Genotype and Phenotype of Hereditary Spherocytosis in the Japanese Population. *Int. J. Hematol.* 71 (2), 118–135.

Zamora, E. A., and Schaefer, C. A. (2019). *Hereditary Spherocytosis*. Treasure Island (FL): StatPearls

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 23 July 2021
doi: 10.3389/fgene.2021.652718

A Novel Germline Compound Heterozygous Mutation of *BRCA2* Gene Associated With Familial Peripheral Neuroblastic Tumors in Two Siblings

Yeran Yang[1,2,3†], Jiwei Chen[4†], Hong Qin[5†], Yaqiong Jin[1,2,3], Li Zhang[4], Shen Yang[5], Huanmin Wang[5], Libing Fu[6], Enyu Hong[1,2], Yongbo Yu[1,2,3], Jie Lu[1,2,3], Yan Chang[1,2,3], Xin Ni[1,2,3], Min Xu[7], Tieliu Shi[4]* and Yongli Guo[1,2,3]**

[1] Beijing Key Laboratory for Pediatric Diseases of Otolaryngology, Head and Neck Surgery, MOE Key Laboratory of Major Diseases in Children, National Center for Children's Health, Beijing Children's Hospital, Beijing Pediatric Research Institute, Capital Medical University, Beijing, China, [2] Biobank for Clinical Data and Samples in Pediatric, National Center for Children's Health, Beijing Children's Hospital, Beijing Pediatric Research Institute, Capital Medical University, Beijing, China, [3] Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beihang University, Capital Medical University, Beijing, China, [4] Center for Bioinformatics and Computational Biology, School of Life Sciences, Institute of Biomedical Sciences, East China Normal University, Shanghai, China, [5] Department of Surgical Oncology, National Center for Children's Health, Beijing Children's Hospital, Capital Medical University, Beijing, China, [6] Department of Pathology, National Center for Children's Health, Beijing Children's Hospital, Capital Medical University, Beijing, China, [7] Department of Surgery, Shanghai Children's Medical Center, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

Objectives: To investigate the genetic variants that are responsible for peripheral neuroblastic tumors (PNTs) oncogenesis in one family case.

Materials and Methods: One family was recruited, including the healthy parents, sister affected by neuroblastoma (NB), and brother who suffered from ganglioneuroma (GN). Whole-genome sequencing (WGS) of germline DNA from all the family members and RNA-seq of tumor RNA from the siblings were performed. Mutants were validated by Sanger sequencing and co-IP was performed to assess the impact of the mutant on chemosensitivity in the SH-SY5Y cell line.

Results: A novel compound heterozygous mutation of *BRCA2* was locked as the cause of carcinogenesis. One allele was BRCA2-S871X (stop-gain) from the siblings' mother, the other was BRCA2-N372H (missense) from their father. This novel compound heterozygous mutations of the *BRCA2* gene associated with PNTs by disordering DNA damage and response (DDR) signal pathway. Moreover, chemosensitivity was reduced in the NB cell line due to the BRCA2-N372H mutant.

Conclusion: In summary, these results revealed a novel germline compound heterozygous mutation of the *BRCA2* gene associated with familial PNTs.

Keywords: peripheral neuroblastic tumors, neuroblastoma, ganglioneuroma, whole-genome sequencing, RNA-Seq, *BRCA2*, cancer predisposition gene

OPEN ACCESS

Edited by:
Nadeem Abbas,
Riphah International University, Pakistan

Reviewed by:
Yifan Zhang,
University of Arkansas at Little Rock, United States
Ting Li,
University of Arkansas at Little Rock, United States

****Correspondence:***
Min Xu
jackxm1236@126.com
Tieliu Shi
tieliushi@yahoo.com;
tlshi@bio.ecnu.edu.cn
Yongli Guo
guoyongli@bch.com.cn;
jackxm1236@126.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 January 2021*
Accepted: *31 May 2021*
Published: *23 July 2021*

Citation:
Yang Y, Chen J, Qin H, Jin Y, Zhang L, Yang S, Wang H, Fu L, Hong E, Yu Y, Lu J, Chang Y, Ni X, Xu M, Shi T and Guo Y (2021) A Novel Germline Compound Heterozygous Mutation of BRCA2 Gene Associated With Familial Peripheral Neuroblastic Tumors in Two Siblings. Front. Genet. 12:652718. doi: 10.3389/fgene.2021.652718

INTRODUCTION

Peripheral neuroblastic tumors (PNTs), which account for 7–10% of all tumors in children, arise from primitive sympathogonia (Luksch et al., 2016). The PNTs encompass the histologic variants neuroblastoma (NB), ganglioneuroblastoma (GNB), and ganglioneuroma (GN). As tumors of the sympathetic nervous system, they arise wherever sympathetic tissue exists with a predilection for the adrenal gland and retro peritoneum. The biggest difference among these three tumors is the differentiation of tumor cells and prognosis. NB is the most undifferentiated and malignant tumor of the three. As a rare pediatric cancer, NB affects 10.2 per million children under 15 years of age and accounts for 15% of cancer deaths in children (Cheung and Dyer, 2013). The most benign tumor is GN, which is composed of well-differentiated gangliocytes and mature stroma (Lonergan et al., 2002). A familial history of NB is very rare, which is reported in about 1% of patients (Janoueix-Lerosey et al., 2010). Missense and frameshift mutations of *PHOX2B* and activating mutations of *ALK* have been identified in NB families firstly (Trochet et al., 2004; Mosse et al., 2008). Furthermore, an increasing number of susceptibility genes have been observed in NB families, such as *NBPF1*, *NBPF23*, *BARD*, and *BRCA2* (Ritenour et al., 2018).

BRCA2 was first revealed to be involved in the homologous recombination (HR) repair pathway depending on interacting with the recombination protein RAD51. For DNA double-strand breaks (DSBs), HR is a high-fidelity mechanism of repair. As one of fifteen Fanconi anemia (FA) genes, BRCA2 also participates in the HR step in inter-strand crosslinks (ICLs) repair depending on RAD51. According to a previous study, about half of the cases of early-onset breast cancer are due to different inherited mutations in the *BRCA2* gene (Valencia et al., 2017). In pediatrics, two BRCA2 germline mutations associated with NB have been reported to date. One is BRCA2 p. W2830_E20splice detected from 56 children with NB (Zhang et al., 2015); the other is BRCA2 p. Y2215fs* identified from one NB family (Cai et al., 2017).

In this study, a new compound heterozygous mutation of the *BRCA2* gene was discovered in two siblings with PNTs by applying WGS technology. One of the biallelic mutations was *BRCA2* rs397507634 chr13: 32911104C > A (S871Ter), which was previously found from the screening of the BRCA1/2 genes (Alter, 2014). However, there are no reports on the relationship between this stop-gain mutant and diseases including PNTs. The other was a missense mutant *BRCA2* rs144848 chr13:32906729A > C (N372H). Although it was identified as highly related to breast cancer (Healey et al., 2000), ovarian cancer (Wenham et al., 2003), and other many kinds of cancers (Rudd et al., 2006; Ramus et al., 2008), there is no report on whether this missense mutant is associated with PNTs. The healthy parents harbored one different allele mutant of *BRCA2* separately. Unfortunately, the two children both inherited the two mutants from their parents and the novel compound heterozygous mutant of *BRCA2* led to PNTs. The evidence in this study may explain how both of the children suffer from the disease despite the absence of PNTs in their parents.

MATERIALS AND METHODS

Patients and Samples

This study included two young patients who were an elder sister (age at diagnosis was 30 months, II-1), and younger brother (age at diagnosis was 36 months, II-2), and their healthy parents (age at 30 s, I-1 and I-2), as shown in **Figure 1A**. All patients' parents signed informed consent and this study obtained approval from the Institution's Research Ethics Board of Hospital. DNeasy Blood and Tissue Kit (QIAGEN) was used for extracting genomic DNA from blood samples of the younger brother (II-2) and the parents (I-1 and I-2), and the tumor tissue of the siblings (II-1 and II-2). The elder sister's tissue DNA was substituted for her genomic DNA from blood to analyze germline mutants accompanying other family members because her blood was unavailable because she had passed away when this study was conducted. DNA concentration was measured by Qubit® DNA Assay Kit in Qubit® 2.0 Fluorometer (Life Technologies, Carlsbad, CA, United States). RNA of the two patients' tumor tissue was extracted by TRzol (Geng et al., 2019).

Whole Genome Sequencing

A total of 0.5 μg of DNA per sample was used as input material for the DNA library preparations. Genomic libraries were prepared using the Illumina Truseq Nano DNA HT Sample Prep Kit following the manufacturer's instructions. Libraries were analyzed for size distribution by Agilent 2100 Bioanalyzer. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using Hiseq X PE Cluster Kit V2.5 (Illumina, San Diego, CA, United States) according to the manufacturer's instructions. Then, the DNA libraries were sequenced on the Illumina Hiseq platform and 150 bp paired-end reads were generated.

Read Mapping and Variant Calling

Reads after quality control were aligned to the UCSC human reference genome (GRCh37/hg19 assembly) using BWA 0.7.12-r1039 mem mode. Samtools-0.1.18 was used for sorting and removing PCR duplicates, and building an index for the bam files. Variants were called using the VarScan (version 2.3.9) pipeline (McKenna et al., 2010).

Variant Annotation and Prioritization

The resulting variants were annotated and prioritized by ANNOVAR (Wang et al., 2010). A threshold of minor allele frequency (MAF < 0.01) from the 1000 Genomes Project Asian (Devuyst, 2015) and ExAC (Lek et al., 2016) non-TCGA cohorts was used to screen rare variants. The pathogenicity of rare missense variants was evaluated by SIFT (Li and Durbin, 2009), PolyPhen2 (Adzhubei et al., 2013), MutationTaster (Schwarz et al., 2014), and M-CAP (Jagadeesh et al., 2016; Li et al., 2019). Other rare variants, including frameshift, prematurely truncating, and initial codon variants were also retained for further analysis.

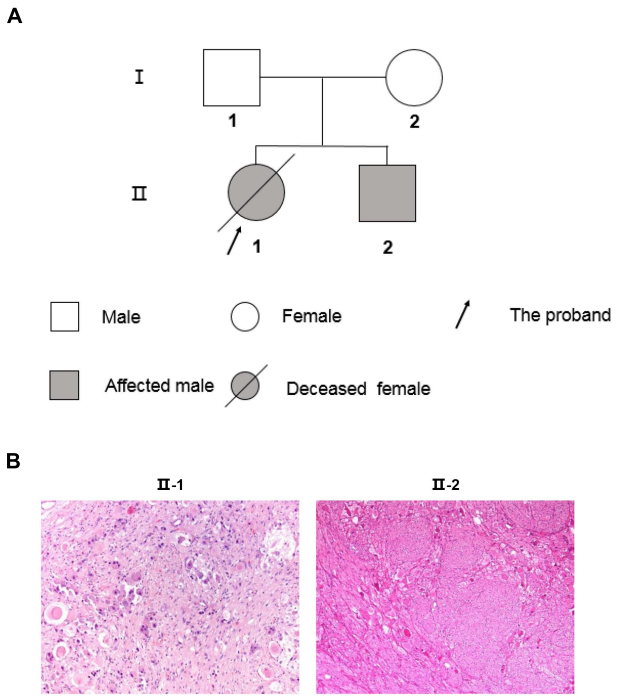

FIGURE 1 | The two-generation pedigree of the family and pathological diagnosis of the PNTs siblings. **(A)** Solid symbols (squares = males, circles = females) indicate clinically affected individuals, and open symbols indicate unaffected individuals. **(B)** Histological study confirmed the type of PNTs. II-1 shows the immature neuroblastic tissue from NB after chemotherapy (H&E 100 ×). II-2 shows mature ganglion cells with abundant cytoplasm in a GN (H&E 50×).

Amplification and Sanger Sequencing of Mutation Sites

The two mutations in *BRCA2* were amplified by PCR with a Veriti 96-well Thermal Cycler (Applied Biosystems, Thermo Fisher Scientific). The primer sequences used were as follows: BRCA2-N372H (rs144848) F: 5′-CTGAAG TGGAACCAAATGATACTGA-3′, R: 5′-AGACGGTACAACTT CCTTGGAGAT-3′ (Wenham et al., 2003); and BRCA2-S871X (rs397507634) F: 5′- AGTGGAATACAGTGATACTGAC -3′, R: 5′- TCGTTTACACAAGTCAAGTCTG -3′. Mutations were confirmed by Sanger sequencing (Bao et al., 2019).

Differential Expression Analysis and KEGG Pathway Enrichment Analysis

RNA-seq data were aligned to the UCSC human reference genome (GRCh37/hg19 assembly) using Hisat2 2.0.1 (Kim et al., 2015). Expression of genes (raw count) was quantified by StingTie 1.3.5 (Pertea et al., 2015). Differentially expressed genes between different conditions were selected using R package DESeq2 based on the criteria of parameter: fold change > 2 and FDR < 0.05 (Love et al., 2014). KEGG pathway enrichment analysis of differentially expressed genes was carried out using the R package clusterProfiler (Yu et al., 2012). Only those pathways with adjusted P-value < 0.05 were considered statistically significant.

Plasmids and Reagents

The p3xFlag-GV141-P/CAF plasmid was purchased from Shanghai GeneChem Co., Ltd. The fragment of BRCA2-290-453aa-wildtype was cloned from pDEST26 -BRCA2 gifted from Dr. Dongyi Xu (Peking University). Using BM seamless cloning kit (Biomed, China), the truncation was cloned into pXJ40-HA. The construct with a mutation in BRCA2-290-453aa-N372H was generated as described previously (Yang et al., 2015).

Anti-Flag M2 agarose affinity gel and mouse monoclonal antibody against Flag was purchased from Sigma (St. Louis, MO, United States). Antibody against HA and antibody against 53BP1 were from Abcam.

Cell Culture and Reagents

SH-SY5Y cells were obtained from the American Type Culture Collection (Rockville, MD, United States). The cell line was grown in DMEM medium supplemented with 10% fetal bovine serum at 37°C in the presence of 5% CO2. For transient transfection experiments, cells were transfected with indicated constructs, using Lipofectamin 3000 (Invitrogen) following the manufacturer's protocols.

Coimmunoprecipitation and Western Blotting

SH-SY5Y cells transfected with p3xFlag-GV141-P/CAF and pXJ40-HA- BRCA2-290-453aa-WT or pXJ40-HA-BRCA2-290-453aa-N372H were harvested and immunoprecipitation with anti-Flag M2 agarose was performed using whole cell lysates as described previously (Wang et al., 2010). For the drug treatment group, 1 μM final concentration of Adriamycin (ADR) (Sigma) was added into the cell culture medium for 6 h before harvesting the whole cell lysis. Samples were separated by SDS-PAGE and detected by immunoblotting with indicated antibodies.

RESULTS

Clinical Features

A 30-month-old Chinese girl (the proband, elder sister, and II-1; **Figure 1**) was diagnosed with NB, which primary site was retroperitoneum, in Shanghai Children's Medical Center (SCMC) Affiliated to Shanghai Jiao Tong University. Serum neuron-specific enolase (NSE), serum lactate dehydrogenase (LDH), urinary vanillylmandelic acid (VMA), and urinary homovanillic acid (HVA) were all significantly elevated. The patient was diagnosed with high risk NB at stage M, with metastasis to bone and bone marrow, according to the international Neuroblastoma Risk Group (INRG) staging and risk system. Complying with SCMC-NB-2009 treatment protocol (Cai et al., 2017), the proband was treated with chemotherapy before removing the primary tumor. After surgery, the proband was determined to be NB by pathological examination (**Figure 1B**) according to international Neuroblastoma Pathology Classification (INPC). *MYCN* status, 1p36, and 11q23 were all normal without

TABLE 1 | Clinical characteristics of two patients with peripheral neuroblastic tumors.

	Patient 1 (Older, II 1)	**Patient 2 (Younger, II 2)**
Gender	Female	Male
Age at diagnosis (months)	30	36
Primary site	Retroperitoneal	Adrenal
Serum NSE	Elevated	Normal
Serum LDH	Elevated	Normal
Urinary VMA and HVA	Elevated	Normal
INRG stage	M (bone marrow, bone)	L1
INPC	Neuroblastoma	Ganglioneuroma
MYCN status	Not amplified	Not amplified
1p36	Normal	Normal
11q23	Normal	Normal
INRG risk	High	Very low
Treatment schedules	Chemotherapy, surgery, myeloablative therapy and autologous stem cell transplant, radiation, and isotretinoin	Surgery and observation
Recurrent time (months)	26	–
Rescue therapies	Additional chemotherapy, and allogeneic hematopoietic stem cell transplantation	–
Prognosis	Died of disease recurrence and progression	Alive without disease
Follow-up time (months)	53	So far

NSE, neuron-specific enolase;
LDH, lactate dehydrogenase;
VMA, vanillylmandelic acid;
HVA, homovanillic acid;
INRG, International Neuroblastoma Risk Group;
INPC, International Neuroblastoma Pathology Classification.

amplification or deletion. Although the myeloablative therapy, autologous stem cell transplant, radiation, and isotretinoin were following applied, the girl was recurrent after 26 months from the first diagnosis with NB. Then the proband accepted additional chemotherapy and allogeneic hematopoietic stem cell transplantation. Unfortunately, the patient died 53 months after the final onset of illness (**Table 1**).

The second case was the proband's younger brother (63 months younger than his sister, II-2, and **Figure 1**). He was 36 months old at the time of diagnosis with GN, which is the same origin as NB but one benign type of PNTs in Beijing Children Hospital Affiliated with Capital Medical University. Unlike his sister, the tumor was in the adrenal gland. Serum NSE, serum LDH, urinary VMA, and urinary HVA were all normal. Besides, *MYCN* status, 1p36, and 11q23 were also normal without amplification or deletion, which was the same as his sister. The patient was diagnosed with very low risk GN at stage L1, according to the INRG staging and risk system. On account of this being a benign tumor, the treatment schedule only included surgery and observation, and no recurrence as yet (**Table 1**).

Information on the three generations of this family was collected. The grandfather of the proband died from liver cancer, and the grandmother is alive but suffering from gastric carcinoma and thyroid cancer. The maternal grandfather of the proband had Alzheimer's, and the maternal grandmother lives with benign thyroid nodules. Other family members are healthy, including the father's sister and her two children. The parents stated normal pregnancies and deliveries without any adverse event for the siblings.

Whole-Genome Sequencing Analysis and Pathogenic Variants Validating

The fact that the siblings in the same family suffered from PNTs strongly suggested that their carcinogenesis was caused by germline variants. Therefore, to identify the causative variants contributing to the carcinogenesis, we performed high-coverage and high-quality WGS of all family members (father, mother, sister, and brother). Specifically, more than 91.4% of the whole genome had 10-fold coverage and showed high consistency among the experiments. In addition, more than 89% of sequenced bases reached Q_{30} based on the analysis of the Phred-scaled quality score (**Supplementary Tables 1, 2**).

The variant calling analysis identified about 4.4 million variants in each family member (**Supplementary Table 2**). Of these, a total of shared 47088 rare variants of the siblings and their mother were selected based on MAF < 0.01. Furthermore, exonic non-synonymous variants were analyzed and 185 rare variants inherited from the mother were selected. Finally, among rare variants from the mother, as cancer driver genes, a missense mutant *ALK* (chr2: 30143039G > T, V163L, rs55697431), and a stop-gain mutant *BRCA2* (chr13: 32911104C > A, S871Ter, rs397507634) were focused through screening using the IntOGen database (**Figure 2A**). For ALK-V163L, multiple databases predicted that it was a benign mutation although *ALK* is one of the most famous susceptibility genes for NB (**Supplementary Table 3**).

Meanwhile, the same strategy was performed to select shared rare variants of the siblings and their father (**Figure 2A**). However, there was no same gene between the mother and father by this strategy (data not shown). Knudson's "two-hit" theory, one susceptibility gene carried different mutants from mother and father separately. Further analysis revealed that they shared total variants gained from father (no MAF value limitation), and a missense mutant *BRCA2* (chr13:32906729A > C, N372H, rs144848) was selected. Although BRCA2-N372H is a common variant (C = 0.279642, GnomAD_exome) inherited from the father, it induced a novel compound heterozygous mutation of *BRCA2*, and combining with a stop-gain mutant BRCA2-S871Ter inherited from the mother. In other words, one allele was BRCA2-S871Ter (stop-gain) from the sibling's mother, and the other was BRCA2-N372H (missense) from the father (**Figure 2B** and **Supplementary Table 3**). These two mutants of *BRCA2* were confirmed in the four persons by Sanger sequencing (**Figure 3A**). In a word, the novel compound heterozygous mutation *BRCA2* gene was locked as the cause of carcinogenesis.

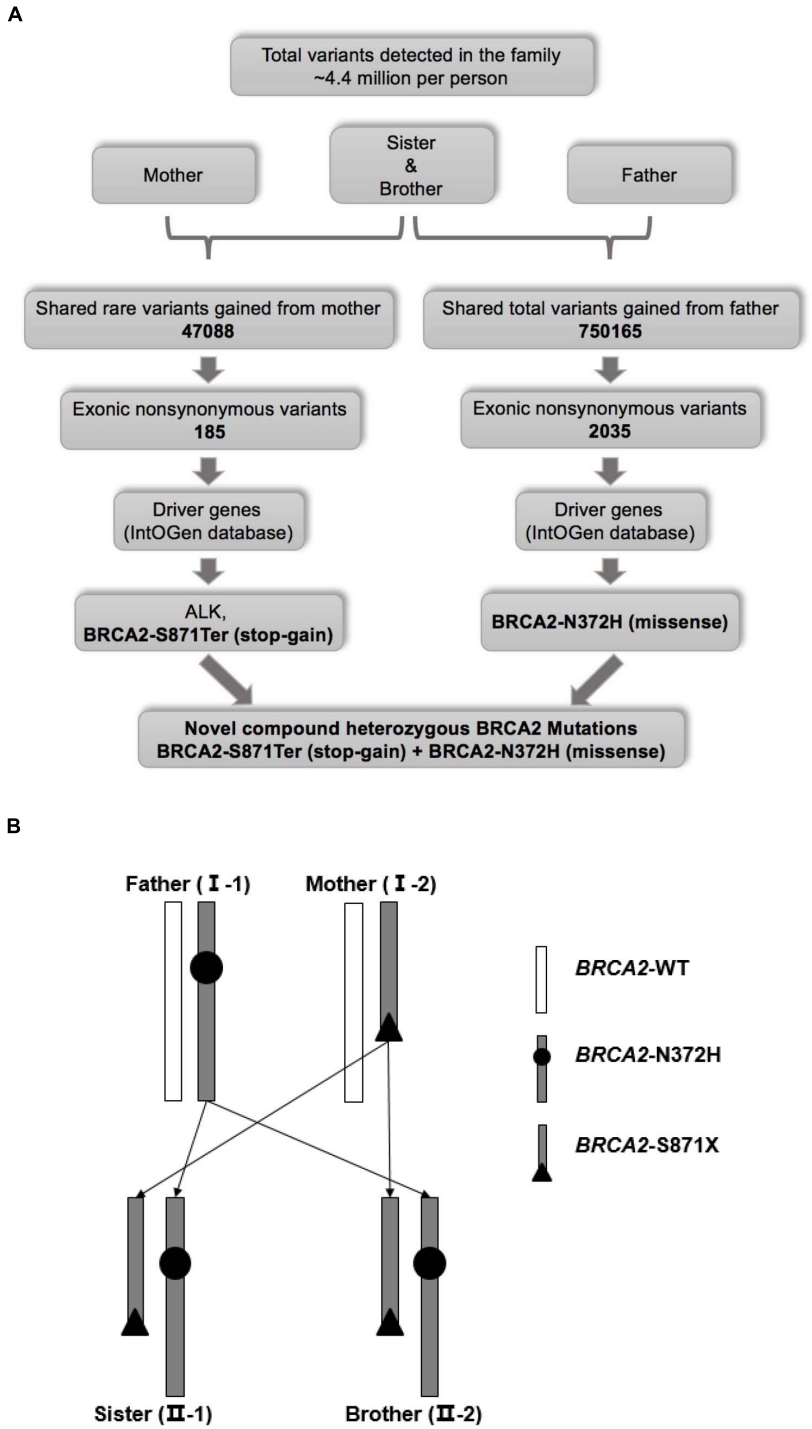

FIGURE 2 | (A) Workflow for the identification of pathogenic mutations. **(B)** Schematic diagram of genetic pattern.

BRCA2-S871Ter Mutant Mechanistic Association With PNTs

As a stop-gain mutation, rs397507634 (S871Ter) in *BRCA2* was found first by quantitative polymerase chain reaction and high-resolution melting curve analysis (Coulet et al., 2010). This variation resulted in premature truncation at the 871st amino acid of the BRCA2 protein, which only includes the PALB2 and P/CAF binding domains (**Figure 3B**). Based on ClinVar annotation, rs397507634 in *BRCA2* is related to hereditary breast and ovarian cancer syndrome (RCV000590670.1), breast-ovarian

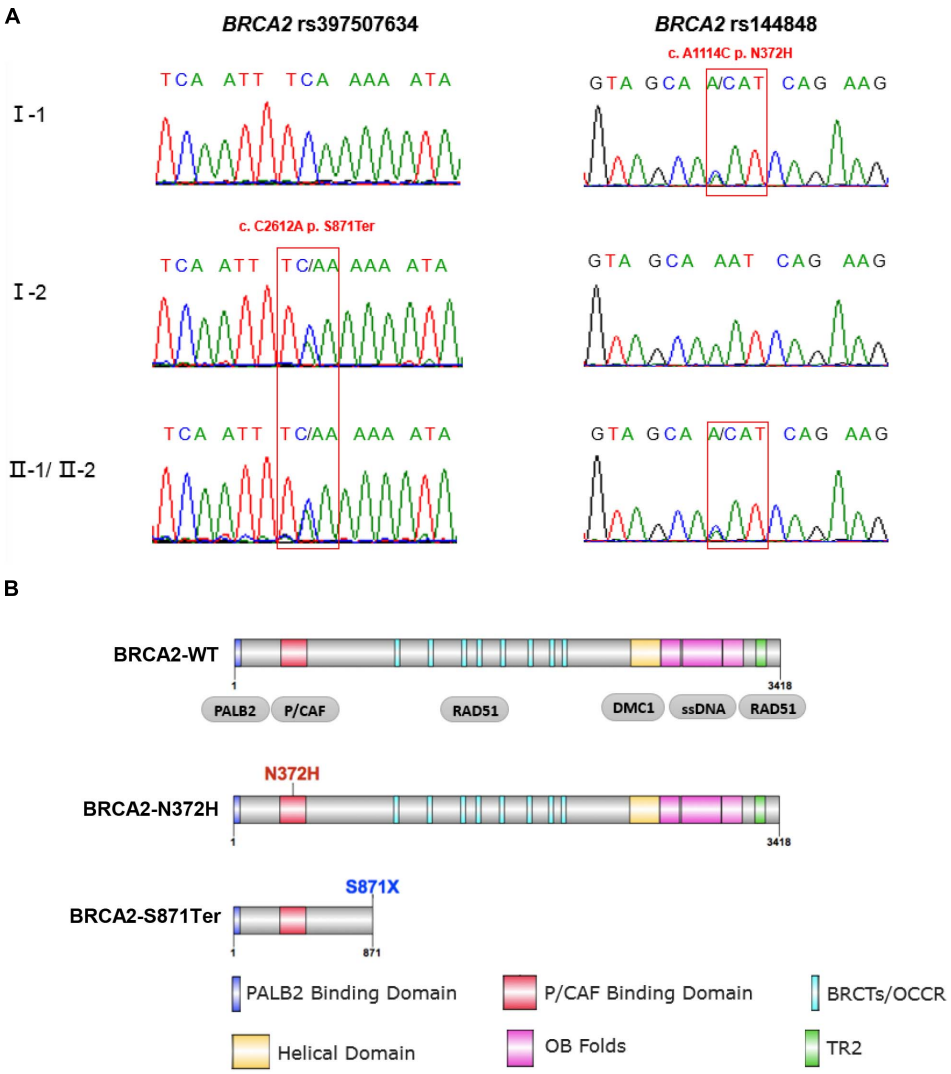

FIGURE 3 | Sanger sequencing validation and functional domains of full-length and mutant BRCA2 protein. **(A)** Sanger sequencing confirmed the mutations of *BRCA2*. The two patients (II-1 and II-2) carried both of these two variants. The father carried BRCA2-N372H, while the mother carried BRCA2-S871Ter. **(B)** The full-length BRCA2 protein harbors six functional domains, including PALB2 binding domain, P/CAF binding domain, BRCTs repeat region, Helical domain, OB folds, and TR2 domain. BRCA2-N372H mutant is in the P/CAF binding domain; while the BRCA2-S871Ter variant results in premature truncation at the 871st amino acid of BRCA2 protein, which only includes the PALB2 and P/CAF binding domains.

cancer, familial 2 (RCV000077282.4), and hereditary cancer-predisposing syndrome (RCV000213349.1).

By mining for the potential impact of stop-gain mutant BRCA2-S871Ter, differentially expressed genes were obtained from RNA-seq data of patients' tumor tissue (**Supplementary Table 4**) and the GEO database GSE62564. There are four samples carrying BRCA2-N372H without BRCA2-S871Ter in the GSE62564 database, including three samples with homozygous mutant (NB_380, NB_390, and NB_438) and one sample with heterozygous mutant (NB_459). KEGG pathway enrichment analysis based on differential genes was performed between siblings and these four samples from the GSE62564 database. The result showed that the significant enriched KEGG pathway included DNA replication, cell cycle, and HR, which were identified with the molecular functions of BRCA2 (**Figure 4A**).

Besides, BRCA2-S871Ter led to early termination during BRCA2 protein translation and produced a 1–871 amino acid truncation of BRCA2 or even led to degradation of BRCA2. To simulate the functions of BRCA2-S871Ter, the GSE62564 dataset (498 samples) was analyzed. Samples with the top 25% (BRCA2 high group) and the lowest 25% (BRCA2 low group) expression of *BRCA2* were taken separately. KEGG pathway enrichment analysis based on these two groups was performed (**Supplementary Figure 1**). The results also revealed that pathways of the cell cycle, DNA replication, and HR were enriched significantly.

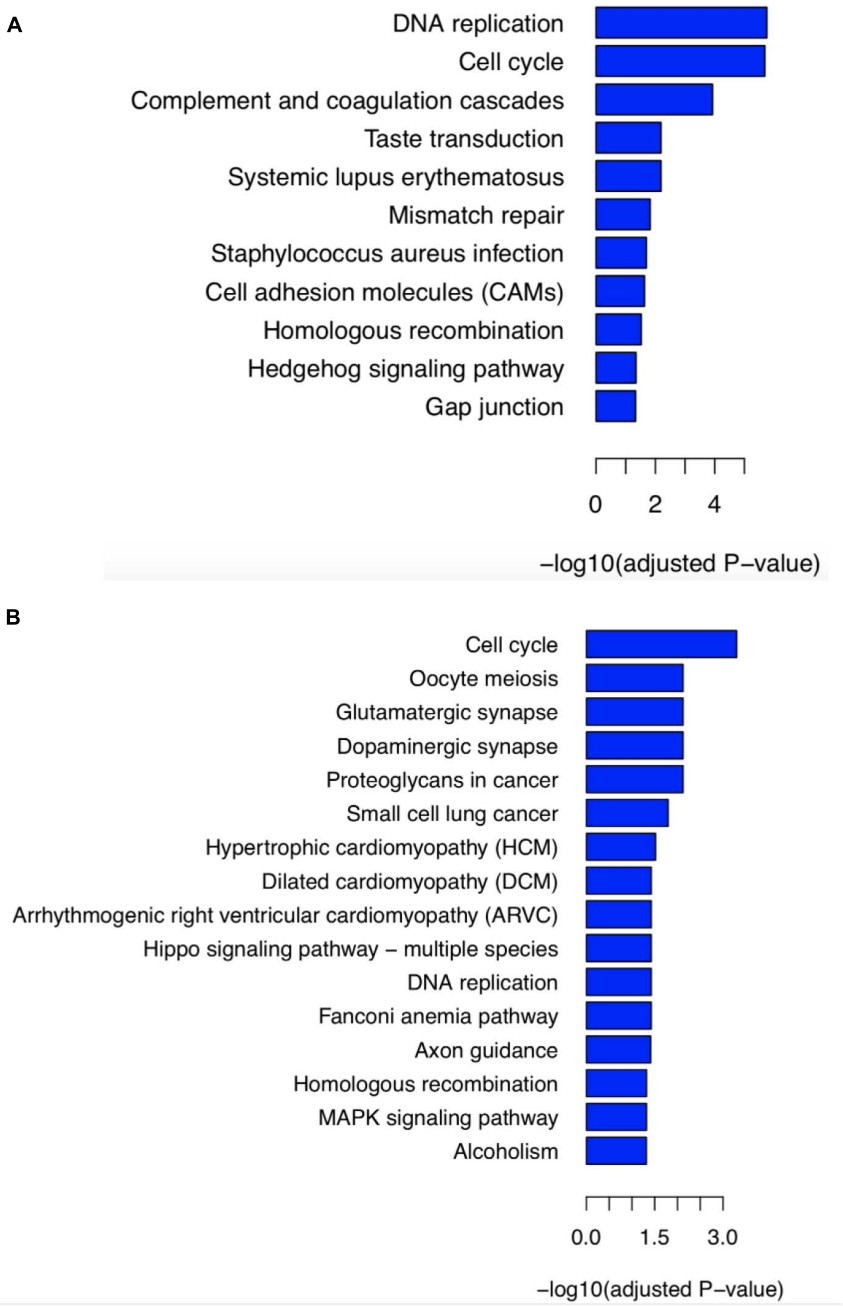

FIGURE 4 | KEGG pathway enrichment analysis of differentially expressed genes. **(A)** To analyze the contribution of stop-gain mutant BRCA2-S871Ter, analysis based on differential genes obtaining from RNA-seq data was performed between siblings and four samples with BRCA2-N372H variant belonging to the GSE62564 database. **(B)** To reveal whether the compound heterozygous mutant of *BRCA2* contributed to the PNTs, we compared the siblings' RNA-seq data, and 494 cases without either BRCA2-N372H or S871Ter mutant from the GSE62564 database.

BRCA2-N372H Mutant Mechanistic Association With PNTs

In addition to the BRCA2 premature truncating mutant, a missense variant rs144848 (N372H) of *BRCA2* inherited from proband's mother was considered a pathogenic variant. It has been reported to increase the risk of many kinds of cancers, especially breast and epithelial ovarian, and indicating that the variant could increase cancer susceptibility. However, there was no report that BRCA2-N372H was associated with neuroblastic tumors. This mutant locates in the P/CAF binding domain (residues 290–453) of BRCA2 (**Figure 3B**). This domain has been shown to mediate the interaction between BRCA2 and the histone acetyltransferase P/CAF. The formation of the BRCA2-P/CAF complex is essential for the transcriptional activation of

other genes (Fuks et al., 1998). Co-IP was performed and showed that BRCA2-N372H substitution did not affect the interaction between BRCA2 and P/CAF absenting chemotherapeutics drug (**Figure 5B**). To determine the optimum concentration and time of treatment with ADR, SH-SY5Y was treated with 1 μM ADR at a different time and then detected the expression of 53BP1 by WB. The results showed that 1 μM ADR treated for 6 h could induce the expression of 53BP1, which was enhanced significantly (**Figure 5A**). Remarkably, after the same condition of ADR treatment (1 μM, 6 h), the interaction between BRCA2 and P/CAF was reduced significantly due to the N372H mutant in the SH-SY5Y cell line (**Figure 5B**). The results suggest that the BRCA2-N372H mutant could destroy the function of P/CAF in the context of DNA damage.

Dysfunctional BRCA2 Due to *BRCA2* Mutants

To reveal the compound heterozygous of *BRCA2* was contributed to PNTs' carcinogenesis, KEGG pathway enrichment analysis was performed between the siblings' RNA-seq data and 494 cases without either BRCA2-S871Ter or BRCA2-N372H mutants from the GSE62564 database. It was shown that the pathways of the cell cycle, DNA replication, FA pathway, and HR were enriched. (**Figure 4B**). It indicated that the compound heterozygous of *BRCA2* resulted in the disordering of BRCA2 protein's function of DNA damage and response (DDR), which was the reason two child patients suffered PNTs.

To simulate the functions of the new compound heterozygous of *BRCA2,* the GSE62564 dataset (498 samples) was analyzed and the one intersection of samples with the top 25% expression of both *BRCA2* and *P/CAF* (BRCA2-P/CAF high group, 36 samples) was taken. Meanwhile, samples with the lowest 25% expression of both *BRCA2* and *P/CAF* were taken as the other intersection from GSE62564 (BRCA2-P/CAF low group, 39 samples). KEGG pathway enrichment analysis was performed between the two intersections (**Supplementary Figure 2**). The results also revealed that pathways of the cell cycle, DNA replication, FA pathway, and HR pathways were enriched significantly, which was similar to our previous results (**Figure 4B**).

DISCUSSION

BRCA2, a principal tumor suppressor gene, is involved in DNA damage repair and related to many types of cancer susceptibility, especially breast cancer, and ovarian cancer in adults (Rousset-Jablonski and Gompel, 2017; Flaum et al., 2019). Meanwhile, as one of the FA DNA repair-related genes, some specific *BRCA2* mutants lead to a severe subset FA accompanying the early onset of cancer, including acute myeloid leukemia, brain tumors, Wilms tumor, and so on (Alter, 2014). In pediatrics, BRCA2 p. W2830_E20splice and p. Y2215fs* are the only two *BRCA2* germline mutations to have been reported to associate with NB (Zhang et al., 2015; Cai et al., 2017).

This study revealed a new compound heterozygotes mutant of *BRCA2* by employing WGS of germline DNA in two siblings with PNTs and their unaffected parents. A stop-gain mutant BRCA2-S871Ter inherited from the mother and a missense mutant BRCA2-N372H inherited from the father were validated by Sanger sequencing (**Figure 3A**). According to Knudson's "two-hit" theory (Knudson, 1996) and KEGG pathway enrichment analysis (**Figure 4B** and **Supplementary Figure 2**), our results provide useful information that this novel compound heterozygotes mutant of *BRCA2* could be the cause of carcinogenesis and a potential biomarker for PNTs.

For the stop-gain mutant, inherited from the mother of the siblings, BRCA2-S871Ter is a rare frequency variation, which has been found through qPCR-HRM based on 210 patients (Coulet et al., 2010). Moreover, this variation is included in some panels to detect breast cancer or ovarian cancer in clinical settings. It can lead to early termination during BRCA2 protein translation and produce a 1–871 amino acids truncation of BRCA2. In this study, western blotting was performed to detect BRCA2 expression in the siblings' tumor tissue. Unfortunately, there was no specific signal of BRCA2 in WB (data not shown). The reason for failing to detect BRCA2 might be protein degradation for its large molecular weight. Although it is not certain whether this truncation is expressed or degraded after translation, the fact is that the functions of BRCA2 should be impacted to a great extent. As **Figure 3B** shows, the majority of the important domains, including BRCTs domain, helical domain, OB folds, and the TR2 domain, are not in the BRCA2 stop-gain truncation. These domains all play key roles in DNA double strand breaks repair. The BRCTs and OB fold domain are essential to BRCA2 binding to RAD51. Then the BRCA2-RAD51 complex will further promote RAD51 assembly onto single-strand DNA to carry forward HR repair. KEGG pathway enrichment analysis also confirmed the above results about BRCA2-S871Ter (**Figure 4A** and **Supplementary Figure 1**).

In addition, BRCA2-N372H is a variant associated with breast cancer and epithelial ovarian cancer (Healey et al., 2000; Wenham et al., 2003). Recently, an increasing number of studies have been performed to prove that the BRCA2-N372H variant is related to susceptibility to many other cancers, including multiple lymphoma, prostate cancer, advanced esophageal squamous cell carcinoma, familial colorectal tumors, and so on. Moreover, previous research has shown that BRCA2-N372H is located in the P/CAF binding domain of BRCA2. The domain from residues of amino acid 290–453 in the N-terminus of BRCA2 specifically interacts with the histone acetyltransferase P/CAF (Fuks et al., 1998). P/CAF belongs to the type three family of lysine acetyl transferases (KAT3) (Goodman and Smolik, 2000). As a transcriptional co-activator, P/CAF binds to transcription factors and performs lysine acetyltransferases activity to acetylate these transcription factors and histones to lose condensed chromosomes. N terminus of BRCA2 is equipped with HAT activity when it bounds to P/CAF. The BRCA2-N372H variation weakens BRCA2 expression and BRCA2-P/CAF interaction and further reduces sensitivity to paclitaxel in breast cancer cells (Kwon et al., 2017). It indicates that BRCA2 372-His substitution induces sequential changes of BRCA2-P/CAF interaction and HAT activity, which is important in paclitaxel resistance (Kwon et al., 2017). ADR as the first-line chemotherapy drug to NB therapy could induce DSBs to NB cells. In our study, with

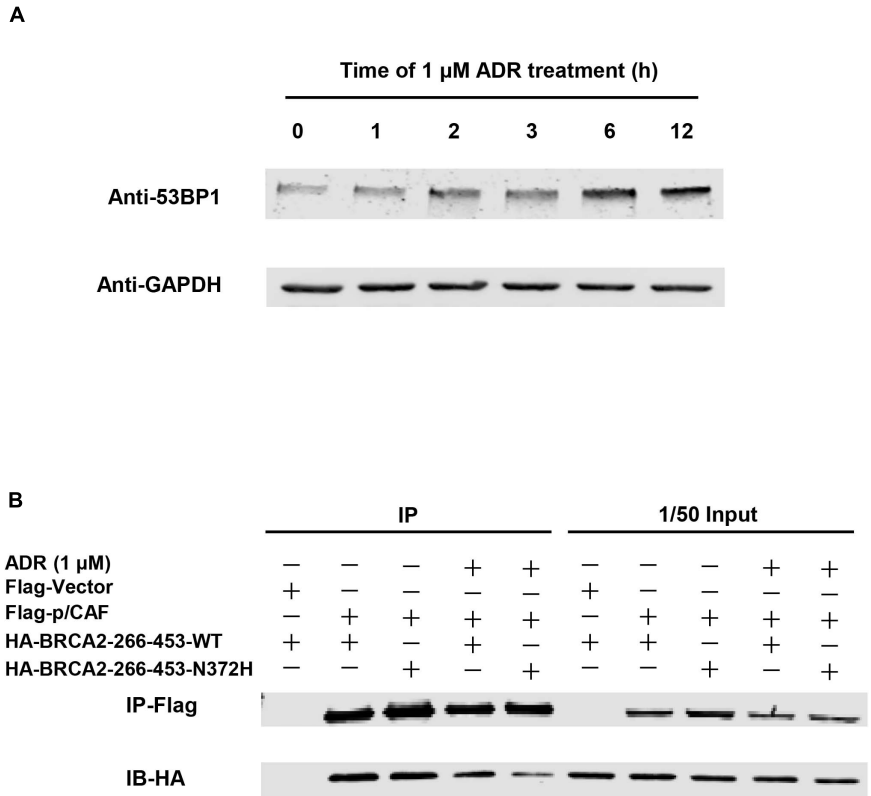

FIGURE 5 | Co-IP between BRCA2 and P/CAF. **(A)** The optimum concentration and time of treatment with ADR in SH-SY5Y to induce DNA double-strand breaks. 1 μM ADR treated with 6 h could induce the expression of 53 bp1 enhanced significantly. **(B)** Co-IP showed that BRCA2-N372H substitution reduced the interaction between BRCA2 and P/CAF with the treatment of ADR (1 μM and 6 h). SH-SY5Y cells were transfected with plasmids expressing Flag-p/CAF and HA-BRCA2-266-453-WT or HA-BRCA2-266-453-N372H and analyzed by co-IP with or without ADR treatment. Immunoprecipitates were analyzed by western blot with antibodies against Flag or HA.

ADR treatment, the interaction between BRCA2, and P/CAF was reduced significantly due to the N372H mutant in SH-SY5Y cells (**Figure 5**). Meanwhile, BRCA2-N372H substitution could not affect the interaction between BRCA2 and P/CAF without drugs. The results suggested that the BRCA2-N372H mutant could destroy the HAT activity of BRCA2-P/CAF in the context of DNA damage. Furthermore, like the paclitaxel resistance, this mutant could induce ADR resistance to NB cells due to lack of HAT activity. This indicates that the BRCA2-N372H mutant could be a new potential drug target in NB chemotherapy.

This study explored the reason for different malignant degrees between these two child patients. Firstly, the BRCA2-N372H variant was reported to affect fetal survival in a sex-dependent manner. With the same BRCA2-N372H variant, the survival rate of newborn females was significantly lower than males (Healey et al., 2000). Secondly, WGS of the two children's tumor tissue did not show enrichment of any distinct carcinogenesis related pathway in specific somatic variants in the siblings, respectively (**Supplementary Figure 2**). There were 721 individual mutant genes in the elder sister and 701 individual mutant genes in the brother. However, screening the IntOGen database revealed that the elder sister harbored more specific cancer driver genes than a brother (20 vs. 14 genes), indicating that the sister had a greater chance of developing a malignant tumor (**Supplementary Tables 6,7**). Furthermore, RNA-seq analysis based on the two siblings' tissue RNA revealed that the row counts of *BRCA2* mRNA expression were 91 in the sister and 898 in the brother (**Supplementary Table 5**). Real-time PCR was performed to confirm the relative expression of *BRCA2* in the siblings' tumor tissue. It was verified that the expression of *BRCA2* in the sister was significantly lower than in the younger brother (**Supplementary Figure 3**). Although the sample size was limited in this study, the siblings harbored similar genetic backgrounds and the same compound heterozygous mutation of *BRCA2*. This indicated that as tumor suppressor genes, a low amount of BRCA2 may lead to cancer susceptibility, and therefore the malignancy degree in the sister was higher than the younger brother. However, the mechanism of different malignant degrees between the siblings needs further study.

CONCLUSION

Our data revealed that a novel compound heterozygous mutation of the *BRCA2* gene is associated with PNTs by disordering

DDR signal pathway. One allele was stop-gain mutant BRCA2-S871Ter, inherited from their mother. The other was a missense mutant BRCA2-N372H inherited from their father, which was confirmed to impair the interaction between BRCA2 and P/CAF, and induced ADR resistance in NB. Moreover, these results provided potential laboratory evidence for the clinical and prenatal diagnosis of PNTs. These results could also potentially guide clinical precision medication. For further investigation of the genotype-phenotype relationship in PNTs, future efforts should be made to recruit more families affected by this disease, and despite its rarity.

LIMITATIONS

There were some limitations in the present study. Firstly, the sample size was small, due to the study object was only one family with two siblings harboring PNTs. Secondly, we needed to make a credible conclusion by incorporating prior knowledge, because of the rarity of samples. Therefore, to further investigate the genotype-phenotype relationship in PNTs, efforts should be made to recruit more families affected by this disease. Furthermore, to decipher the molecular mechanism of the compound heterozygous of *BRCA2* mutations, more studies, including cell level and animal models, and need to be performed in the future.

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: https://www.biosino.org/node/analysis/detail/OEZ005703, OEZ005703.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the Beijing Children's Hospital, Capital 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

YrY, JC, and HQ conducted the experiments, analyzed the data, and wrote the manuscript. YJ and LZ analyzed the data, reviewed, and edited the manuscript. SY, HW, LF, EH, YbY, JL, YC, and XN conducted the experiments. MX, TS, and YG designed the study, wrote, reviewed, and edited the manuscript, and administrated the project. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation of China (81702787, 81702463, and 31671377), the Special Fund of the Pediatric Medical Coordinated Development Center of Beijing Hospitals Authority (No. XTCX201806), Beijing Hospitals Authority Clinical Medicine Development of Special Funding Support (XMLX202121), Beijing Hospitals Authority' Ascent Plan (DFL20191201), and Shanghai Municipal Science and Technology Major Project (2017SHZDZX01).

ACKNOWLEDGMENTS

We are grateful to the child patients and their parents for their participation in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.652718/full#supplementary-material

Supplementary Figure 1 | KEGG pathway enrichment analysis of differentially expressed genes between BRCA2 high group and BRCA2 low group from GSE62564 dataset.

Supplementary Figure 2 | KEGG pathway enrichment analysis of differentially expressed genes between BRCA2-P/CAF high group and BRCA2-P/CAF low group from GSE62564 dataset.

Supplementary Figure 3 | KEGG pathway enrichment analysis of specific somatic variants. **(A)** KEGG pathway enrichment of 721 individual mutant genes in elder sister. **(B)** KEGG pathway enrichment of 701 individual mutant genes in brother.

Supplementary Figure 4 | Expression of *BRCA2* in the siblings. Real-time PCR to detect expression of *BRCA2* in the siblings' RNA from tumor tissue.

REFERENCES

Adzhubei, I., Jordan, D. M., and Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* Chapter 7:Unit7.20.

Alter, B. P. (2014). Fanconi anemia and the development of leukemia. *Best Pract. Res. Clin. Haematol.* 27, 214–221.

Bao, Y., Suo, L., Qian, P., Huang, H., Yang, Y., Tang, J., et al. (2019). Clinical and genetic analysis of Dent disease with nephrotic range albuminuria in Shaanxi, China. *Sci. China Life Sci.* 62, 1590–1593. doi: 10.1007/s11427-018-9829-0

Cai, J., Pan, C., Tang, Y., Chen, J., Zhou, M., Li, B., et al. (2017). Multivariate analysis of risk factors for patients with stage 4 neuroblastoma who were older than 18 months at diagnosis: a report from a single institute in Shanghai, China. *J. Cancer Res. Clin. Oncol.* 143, 1327–1335. doi: 10.1007/s00432-017-2379-5

Cheung, N. K., and Dyer, M. A. (2013). Neuroblastoma: developmental biology, cancer genomics and immunotherapy. *Nat. Rev. Cancer* 13, 397–411. doi: 10.1038/nrc3526

Coulet, F., Pires, F., Rouleau, E., Lefol, C., Martin, S., Colas, C., et al. (2010). A one-step prescreening for point mutations and large rearrangement in BRCA1

and BRCA2 genes using quantitative polymerase chain reaction and high-resolution melting curve analysis. *Genet. Test Mol. Biomark.* 14, 677–690. doi: 10.1089/gtmb.2009.0183

Devuyst, O. (2015). The 1000 genomes project: welcome to a new world. *Perit. Dial. Int.* 35, 676–677. doi: 10.3747/pdi.2015.00261

Flaum, N., Crosbie, E. J., Edmondson, R. J., Smith, M. J., and Evans, D. G. (2019). Epithelial ovarian cancer risk: a review of the current genetic landscape. *Clin. Genet.* 97, 54–63. doi: 10.1111/cge.13566

Fuks, F., Milner, J., and Kouzarides, T. (1998). BRCA2 associates with acetyltransferase activity when bound to P/CAF. *Oncogene* 17, 2531–2534. doi: 10.1038/sj.onc.1202475

Geng, J., Liu, Y., Guo, Y., Wang, H., Tai, J., Jin, Y., et al. (2019). Correlation between TERT C228T and clinic-pathological features in pediatric papillary thyroid carcinoma. *Sci. China Life Sci.* 62, 1563–1571. doi: 10.1007/s11427-018-9546-5

Goodman, R. H., and Smolik, S. (2000). CBP/p300 in cell growth, transformation, and development. *Genes Dev.* 14, 1553–1577.

Healey, C. S., Dunning, A. M., Teare, M. D., Chase, D., Parker, L., Burn, J., et al. (2000). A common variant in BRCA2 is associated with both breast cancer risk and prenatal viability. *Nat. Genet.* 26, 362–364. doi: 10.1038/81691

Jagadeesh, K. A., Wenger, A. M., Berger, M. J., Guturu, H., Stenson, P. D., Cooper, D. N., et al. (2016). M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. *Nat. Genet.* 48, 1581–1586. doi: 10.1038/ng.3703

Janoueix-Lerosey, I., Schleiermacher, G., and Delattre, O. (2010). Molecular pathogenesis of peripheral neuroblastic tumors. *Oncogene* 29, 1566–1579. doi: 10.1038/onc.2009.518

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Knudson, A. G. (1996). Hereditary cancer: two hits revisited. *J. Cancer Res. Clin. Oncol.* 122, 135–140. doi: 10.1007/bf01366952

Kwon, W. S., Rha, S. Y., Jeung, H. C., Kim, T. S., and Chung, H. C. (2017). Modulation of HAT activity by the BRCA2 N372H variation is a novel mechanism of paclitaxel resistance in breast cancer cell lines. *Biochem. Pharmacol.* 138, 163–173. doi: 10.1016/j.bcp.2017.04.015

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

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

Li, Z., Zhu, P., Huang, H., Pan, Y., Han, P., Cui, H., et al. (2019). Identification of a novel COL4A5 mutation in the proband initially diagnosed as IgAN from a Chinese family with X-linked Alport syndrome. *Sci. China Life Sci.* 62, 1572–1579. doi: 10.1007/s11427-018-9545-3

Lonergan, G. J., Schwab, C. M., Suarez, E. S., and Carlson, C. L. (2002). Neuroblastoma, ganglioneuroblastoma, and ganglioneuroma: radiologic-pathologic correlation. *Radiographics* 22, 911–934. doi: 10.1148/radiographics.22.4.g02jl15911

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550.

Luksch, R., Castellani, M. R., Collini, P., De Bernardi, B., Conte, M., Gambini, C., et al. (2016). Neuroblastoma (Peripheral neuroblastic tumours). *Crit. Rev. Oncol. Hematol.* 107, 163–181.

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

Mosse, Y. P., Laudenslager, M., Longo, L., Cole, K. A., Wood, A., Attiyeh, E. F., et al. (2008). Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature* 455, 930–935. doi: 10.1038/nature07261

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Ramus, S. J., Vierkant, R. A., Johnatty, S. E., Pike, M. C., Van Den Berg, D. J., Wu, A. H., et al. (2008). Consortium analysis of 7 candidate SNPs for ovarian cancer. *Int. J. Cancer.* 123, 380–388.

Ritenour, L. E., Randall, M. P., Bosse, K. R., and Diskin, S. J. (2018). Genetic susceptibility to neuroblastoma: current knowledge and future directions. *Cell Tissue Res.* 372, 287–307. doi: 10.1007/s00441-018-2820-3

Rousset-Jablonski, C., and Gompel, A. (2017). Screening for familial cancer risk: focus on breast cancer. *Maturitas* 105, 69–77. doi: 10.1016/j.maturitas.2017.08.004

Rudd, M. F., Sellick, G. S., Webb, E. L., Catovsky, D., and Houlston, R. S. (2006). Variants in the ATM-BRCA2-CHEK2 axis predispose to chronic lymphocytic leukemia. *Blood* 108, 638–644. doi: 10.1182/blood-2005-12-5022

Schwarz, J. M., Cooper, D. N., Schuelke, M., and Seelow, D. (2014). MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* 11, 361–362. doi: 10.1038/nmeth.2890

Trochet, D., Bourdeaut, F., Janoueix-Lerosey, I., Deville, A., de Pontual, L., Schleiermacher, G., et al. (2004). Germline mutations of the paired-like homeobox 2B (PHOX2B) gene in neuroblastoma. *Am. J. Hum. Genet.* 74, 761–764. doi: 10.1086/383253

Valencia, O. M., Samuel, S. E., Viscusi, R. K., Riall, T. S., and Neumayer, L. A. (2017). The role of genetic testing in patients with breast cancer: a review. *JAMA Surg.* 152, 589–594. doi: 10.1001/jamasurg.2017.0552

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

Wenham, R. M., Schildkraut, J. M., McLean, K., Calingaert, B., Bentley, R. C., Marks, J., et al. (2003). Polymorphisms in BRCA1 and BRCA2 and risk of epithelial ovarian cancer. *Clin. Cancer Res.* 9, 4396–4403.

Yang, Y., Liu, Z., Wang, F., Temviriyanukul, P., Ma, X., Tu, Y., et al. (2015). FANCD2 and REV1 cooperate in the protection of nascent DNA strands in response to replication stress. *Nucleic Acids Res.* 43, 8325–8339. doi: 10.1093/nar/gkv737

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

Zhang, J., Walsh, M. F., Wu, G., Edmonson, M. N., Gruber, T. A., Easton, J., et al. (2015). Germline mutations in predisposition genes in pediatric cancer. *N. Engl. J. Med.* 373, 2336–2346.

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.

frontiers in Pharmacology

MINI REVIEW
published: 23 August 2021
doi: 10.3389/fphar.2021.704586

Therapeutics Development for Alagille Syndrome

Phillip Sanchez[1], *Atena Farkhondeh*[1], *Ivan Pavlinov*[1], *Karsten Baumgaertel*[2], *Steven Rodems*[2] *and Wei Zheng*[1]*

[1]National Center for Advancing Translational Sciences, National Institutes of Health, Bethesda, MD, United States, [2]Travere Therapeutics, San Diego, CA, United States

Advancements in treatment for the rare genetic disorder known as Alagille Syndrome (ALGS) have been regrettably slow. The large variety of mutations to the JAG1 and NOTCH2 genes which lead to ALGS pose a unique challenge for developing targeted treatments. Due to the central role of the Notch signaling pathway in several cancers, traditional treatment modalities which compensate for the loss in activity caused by mutation are rightly excluded. Unfortunately, current treatment plans for ALGS focus on relieving symptoms of the disorder and do not address the underlying causes of disease. Here we review several of the current and potential key technologies and strategies which may yield a significant leap in developing targeted therapies for this disorder.

Keywords: alagille syndrome, JAG1, notch signaling pathway, liver, bile duct, drug development

OPEN ACCESS

Edited by:
Shraddha Thakkar, United States Food and Drug Administration, United States

Reviewed by:
Khalid Muhammad, United Arab Emirates University, United Arab Emirates
Ernesto Canalis, UCONN Health, United States

***Correspondence:**
Wei Zheng
wzheng@mail.nih.gov

Specialty section:
This article was submitted to Translational Pharmacology, a section of the journal Frontiers in Pharmacology

Received: *03 May 2021*
Accepted: *09 August 2021*
Published: *23 August 2021*

Citation:
Sanchez P, Farkhondeh A, Pavlinov I, Baumgaertel K, Rodems S and Zheng W (2021) Therapeutics Development for Alagille Syndrome. Front. Pharmacol. 12:704586. doi: 10.3389/fphar.2021.704586

INTRODUCTION

In 1975, a set of children suffering from cholestatic disease were identified to share several symptomatic similarities, distinct from others with comparable biliary dysfunctions. Characteristic facies as well as renal, vertebral and cardiac abnormalities allowed Dr. Daniel Alagille to append a single etiology, later named after him, to these children (Alagille et al., 1975). With an incidence of 1:30,000 to 1:50,000 births, the autosomal dominant Alagille Syndrome (ALGS) is a result of Notch signaling dysfunction caused by gene mutations mostly in JAG1 and NOTCH2 (Kamath et al., 2018). The most common and debilitating disease hallmark among ALGS patients is bile duct paucity, with almost all patients exhibiting cholestatic disease. Most patients experience intractable pruritis and the presence of xanthomas, localized cholesterol and fats deposits under the skin caused by this liver dysfunction. Due to the somatic nature of these mutations to Notch signaling however, several other organ systems are affected including kidney, heart, eye, nervous system, and bone. Because of the highly-variable nature of ALGS presentation, current therapeutic management is focused on addressing each patient's symptoms individually. The lack of effective and targeted drugs for ALGS treatment therefore constitutes an unmet medical need.

JAG1 (Jagged1) is a transmembrane ligand of the Notch signaling pathway. Partial loss of JAG1 protein due to a mutation in one allele of the JAG1 gene is sufficient to disrupt proper bile duct development, resulting in bile duct paucity seen in ALGS (Hofmann et al., 2010). Recent analysis of patients with JAG1 variants has identified mutations to all 26 exons. Haploinsufficiency caused by truncation or early transcriptional termination of JAG1 account for 83% of mutations seen in ALGS patients (Gilbert et al., 2019). NOTCH2 variants are far less often designated as the sole cause of ALGS, occurring in less than 3% of patients from a recent analysis. Comparatively, mutations in NOTCH2 are more likely to be missense (68%) than in JAG1 (15%) (Gilbert et al., 2019). Although haploinsufficiency might suggest a potential for introduction of exogenous JAG1, it is not without risk. The introduction of high-levels of JAG1 can dramatically increase the risk of hyperplasia and

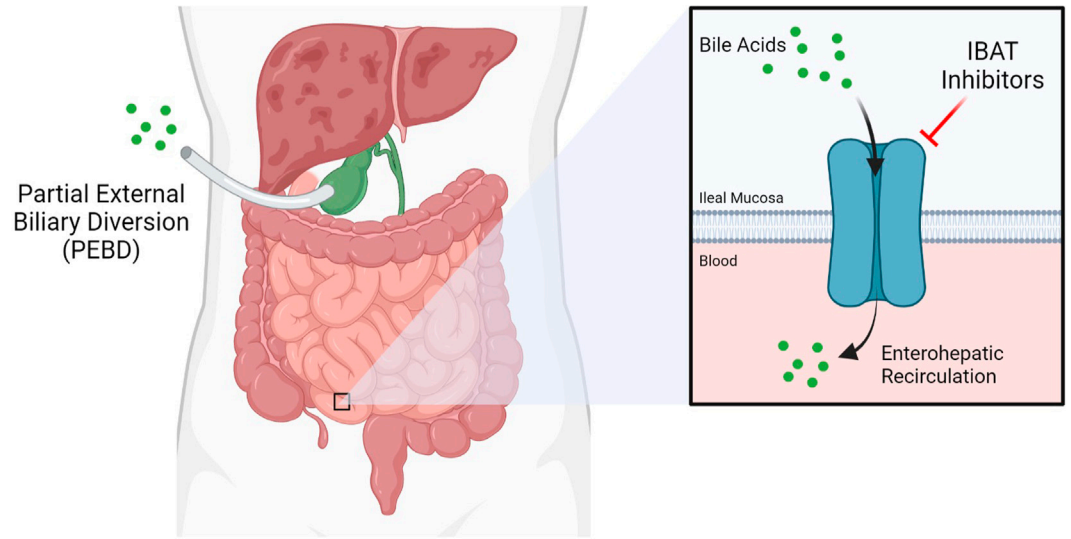

FIGURE 1 | Current therapeutic approaches for Alagille Syndrome include inhibitors of the Ileal Bile Acid Transporter (IBAT) such as maralixibat, which limit enterohepatic circulation, and Partial External Biliary Diversion (PEBD), a highly invasive procedure.

cancer since cell-cell based contact inhibition and proliferation is greatly influenced by Notch signaling (Aster et al., 2017). Strategies focused on developing targeted therapies, which restore physiological levels of Notch signaling, will yield treatments which address underlying causes of ALGS rather than just their symptoms.

Clinical Presentations, Treatment, and Prognosis

Alagille syndrome presents with several, often readily-observable, phenotypic traits. Distinctive facies with characteristic pointed chin, broad forehead and hypertelorism are included in diagnostic criteria. Pulmonary stenosis is frequently observed, leading to cardiac arrythmias in 63–98% of patients (Spinner et al., 1993; Kamath et al., 2018). Butterfly vertebrae are also observed, albeit less frequently and without significant symptoms. Posterior embryotoxon is the primary ophthalmological presentation for ALGS, exhibited in 78–89% of patients (Spinner et al., 1993). Renal anomalies are identified in significant (~58.9%) subset of ALGS patients, whether or not JAG1 mutations are observed (Kamath et al., 2012).

The aforementioned characteristics are important diagnostic criteria and include bile duct paucity which is manifested in the vast majority of patients with ALGS. Symptoms of liver dysfunction include pruritis, an intractable feeling of itchy skin, caused by elevated serum bile acids as seen in other cholestatic diseases like biliary atresia. Pruritis is a major complaint of children and young patients with ALGS. Conventional treatment options for pruritus are often not effective leading to reduced quality of life for these patients. A recent survey identified 42% of parents of children with ALGS are dissatisfied with current therapeutic approaches, with many stating that attempts to reduce pruritis are wholly ineffective (Patient Family Survey, 2019).

Presently, the only effective treatment options for liver disease in ALGS are highly invasive. Partial external biliary diversion (PEBD) is one such approach, aiming at reducing overall bile acids present in the blood (**Figure 1**) (Mattei et al., 2006). PEBD often has moderate success with reducing pruritis and xanthomas, though it is not necessarily sufficient to avoid later liver transplant (Wang et al., 2017a). End-stage liver disease is treated with liver transplantation, often required in cases of neonatal ALGS which are among the direst of cases. Unfortunately, unexpected and rapid worsening of overall health in patients who need a liver transplant accounts for significant mortality in ALGS, especially within the neonate population (Lykavieris et al., 2001). Moreover, such an invasive procedure is not without risk (Kamath et al., 2018). For instance, the long-term use of immunosuppressants designed to prevent the rejection of the donor liver can trigger or aggravate issues in the kidney, often necessitating a subsequent kidney transplant (Olyaei et al., 2001). Transplant-free survival is estimated at around 24% to the age of 18.5, demonstrating the need for new treatment paradigms (Kamath et al., 2020).

Notch Signaling and ALGS

The Notch signaling pathway is a highly conserved and indispensable component of cell-cell based signaling. Notch receptors on a signal receiving cell undergo regulated proteolysis after binding to a Notch ligand, like JAG1, from a signal transmitting cell. After several proteolytic steps on the inner membrane leaflet, the Notch intracellular domain (NICD) of the receptor is released and translocates to the nucleus. In Notch signal transduction only one signal, the NICD, is ever transmitted after proteolysis of the initial receptor. With no intermediate secondary messengers for amplification, expression of Notch target genes which are held in an "off" state by co-repressors is turned "on" by the NICD. Not

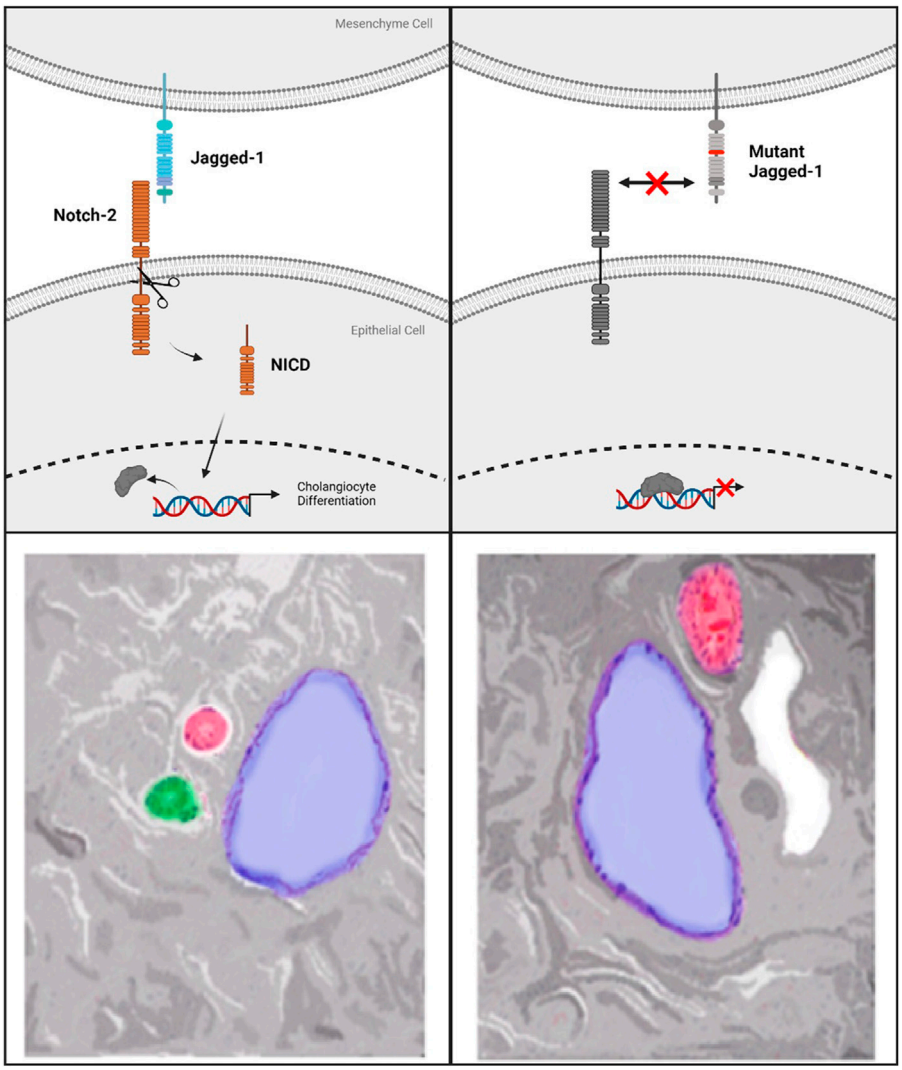

FIGURE 2 | Top left panel shows standard Notch-signaling *via* JAG1-NOTCH2 contact causing Notch-intracellular domain (NICD) cleavage and translocation to the nucleus. **Top right** panel indicates the lack of co-repressor displacement caused by a mutation to Jagged-1 which prevents Notch-2 activation and cleavage. **Bottom Left** panel shows normal hepatic triad with Hepatic Vein (blue), Hepatic Artery (red), Bile duct (green). On the **bottom right** is an example of the lack of bile ducts observed in Alagille Syndrome caused by mutations to Notch signaling components. ALGS livers also frequently present with cysts, in white, in place of functional bile ducts.

surprisingly, this tight regulation of transcription strongly influences a variety of important developmental and proliferation target genes (Kopan and Ilagan, 2009).

Notch signaling is a required feature of biliary genesis during liver development with Notch2 accounting for the majority of Notch receptor expressed in fetal livers. Coordination of contact signaling between JAG1 and Notch2 expressing cells results in terminal differentiation to cholangiocytes in the Notch2 expressing hepatoblasts (Freeburg and Goessling, 2020). In histopathological analysis of ALGS livers, a distinct dearth of staining for cholangiocyte-specific signaling factors is observed (Fabris et al., 2007). Further work involving 3D-spheroid hepatocyte co-cultures indicates a critical role for the JAG1 signal specifically from portal vein mesenchyme cells (PVC). This signal instructs the formation of relevant ductal structures by biliary epithelial cells (BEC) which stain positively for the marker Cytokeratin 19 (CK-19) (Hofmann et al., 2010). Additionally, mutant JAG1 in PVCs is responsible for several instructive phenotypes; namely improper spheroid formation, disrupted luminal development and loss of CK-19 staining. Due to the loss of functional JAG1 caused by ALGS-associated mutations, cholangiocyte specification is reduced, preventing proper biliary structural development (**Figure 2**) (Fabris et al., 2007; Hofmann et al., 2010). Therefore, recovering the reduced JAG1-NOTCH2 signaling in early stage ALGS patients is an important therapeutic consideration. To improve bile duct formation and regeneration for cholestasis in ALGS, a drug treatment which increases JAG1-NOTCH2 signaling may be used in a short-term manner until bile acid regulation is reestablished.

INVESTIGATIONAL THERAPIES AND APPROACHES

IBAT Inhibitors

Maralixibat is a therapy being evaluated as a treatment for children with rare cholestatic liver diseases, including ALGS and progressive familial intrahepatic cholestasis (PFIC) by Mirum Pharmaceuticals. Both of these cholestatic diseases involve an interruption in the flow of bile acid from the liver, resulting in elevated bile acid levels in the liver and serum. This leads to liver disease and a variety of severe and life-altering symptoms, including stunted growth and chronic and severe pruritis. Maralixibat works by inhibiting a carrier protein called the apical sodium dependent bile acid transporter (ASBT) also known as the ileal bile acid transporter (IBAT), as shown in **Figure 1**. ASBT is primarily responsible for mediating the uptake of bile acids from the small intestine to the liver. Approximately 95% of bile acids are recirculated *via* the ASBT to the liver. Accordingly, a product capable of inhibiting the ASBT could lead to a reduction in bile acids returning to the liver and may represent a promising approach for treating cholestatic liver diseases. Maralixibat is currently being evaluated in the Phase 3 study in children with PFIC. In the Phase 2b clinical trial of in ALGS patients, children taking Maralixibat had statistically significant reductions in pruritus and serum bile acids compared to placebo. At the 48-week measurement, pruritus and serum bile acid reductions were maintained and improvements in skin xanthomas and quality of life were also observed. This trial has also confirmed recent findings that pharmaceutical intervention to lower serum bile acids can significantly improve transplantation-free survival rates (van Wessel et al., 2020). The Food and Drug Administration (FDA) has granted Maralixibat Breakthrough Therapy and Orphan Drug designations for pruritus associated with Alagille syndrome in patients 1 year of age and older, which will accelerate clinical development on the path to approval by the FDA (Pharma, 2020a).

Odevixibat is currently being evaluated by Albireo Pharma, Inc. as a therapy for children with cholestatic liver disease and pruritus. In Phase 2 study in children with cholestatic liver disease and pruritus, odevixibat showed reductions in serum bile acids and pruritus in most patients and exhibited an overall tolerability profile. Albireo Pharma is developing odevixibat to treat patients with PFIC, biliary atresia and ALGS. Their Phase 3 trial for PFIC recently met its primary endpoints reducing serum bile acids and pruritis. FDA and EMA also have granted an Orphan Drug Designation to odevixibat for the treatment of biliary atresia, Alagille syndrome, PFIC and primary biliary cholangitis (Pharma, 2020b).

The aforementioned investigational drugs have advanced to late-stage clinical development with promising results and are likely to constitute an important treatment option for ALGS. Their mechanism of action however, is restricted to blocking bile acid reabsorption in the intestine. A therapy addressing the root cause of the disease, namely reduced JAG1-NOTCH2 signaling, is still required to improve overall therapeutic outcomes in ALGS. The following sections will cover several technologies that have been used to address other rare genetic diseases and could be applied to developing novel treatments for ALGS (**Table 1**).

RNA Targeted Therapies

Recently, several novel therapeutic modalities have been approved for rare diseases by the Food and Drug Administration. Among these therapies are eteplirsen and nusinersen for the treatment of Duchenne Muscular Dystrophy (DMD) and Spinal Muscular Atrophy (SMA), respectively (Finkel et al., 2016; Syed, 2016). Both drugs are anti-sense oligonucleotides (ASO) which modulate splicing of their disease-causing target transcripts: dystrophin for DMD and SMN2 for SMA. Specifically, eteplirsen targets exon 51 within dystrophin transcripts, allowing for restoration of an almost full-length transcript by skipping frameshift causing mutations. Although it is estimated that only 14% of DMD patients will benefit from skipping this particular exon, therapies targeting exon-specific mutations represent a bold new therapeutic strategy for the treatment of rare genetic diseases currently lacking effective treatment (Kole and Krieg, 2015).

TABLE 1 | Current treatments and potential future therapeutics for ALGS.

Current therapeutic modalities	
Surgical (available)	
	- Liver Transplant
	- Partial External Biliary Diversion (PEBD)
IBAT Inhibitors (in clinical trials)	
	- Maralixibat
	- Odevixibat
Potential Therapeutic Modalities	
RNA-Targeting	
	- Anti-sense Oligonucleotides (ASO)
	- Splice-switching small-molecules
	- CasRx Transcript Editing
Adenoviral Vector (AAV)-mediated gene therapy	
	- mRNA supplementation
	- CRISPR/Cas Gene Editing
	- Nuclease-based editing (ZFN's)

The application of RNA targeted therapy has notably progressed in n-of-1 clinical trials, showing great promise as a new therapeutic platform to treat rare genetic diseases using personalized exon-specific ASO's. Milasen, a tailor-made ASO for a single patient with the rare neurodegenerative Batten disease, went from proof-of-concept to intrathecal injection in the patient within a year. This Batten disease patient has a retrotransposon insertion within the CLN7 gene that disrupts a nominal splice-site between exons 6 and 7. By targeting the retrotransposon nestled between exons 6 and 7 of the CLN7 gene, the 22-nucleotide ASO was able to restore splice-site activity to exclude the deleterious insertion (Kim et al., 2019). The reason Milasen advanced so rapidly as a therapy was due in part to the fact that it shared similar backbone chemistry with nusinersen as well as having support from the regulatory agency (Woodcock and Marks, 2019).

One of the hallmarks of ASO therapeutics described above is the array of nucleoside chemistries which enhance oligonucleotide resistance to cleavage by endonucleases and enable ASO penetration into cells (Roberts et al., 2020). Both milasen and nusinersen are 2′O-methyl-phosphorothioate oligonucleotides (2′-OMePS), while eteplirsen is instead a phosphorodiamidate morpholino oligomer (PMO). Although there are many chemistries used to balance efficacy and pharmacokinetic parameters, the aforementioned ASO's are proof that shared nucleoside chemistry is a possibility for new therapeutics, indicating a pathway for ALGS RNA therapy development (Khvorova and Watts, 2017; Kim et al., 2019). Similar backbone structures from these approved drugs could be used as a template for the creation of ASOs which modulate the splicing of JAG1 or NOTCH2 for specific mutations to restore the reduced JAG1-NOTCH signaling in ALGS. Apart from ASOs, small molecules have yielded success in targeting mRNA splicing. One example approved by the FDA is risdiplam for SMA in August 2020. Similar to nusinersen, this drug prevents exclusion of exon 7 from SMN2 yielding a fully functional protein through the codon readthrough mechanism across a premature stop codon (Ratni et al., 2018).

Adverse events for ASO therapies consist of a wide range of categories. For example, inotersen which is used to treat hereditary transthyretin-mediated (hATTR) amyloidosis carries a label warning for glomerulonephritis, thrombocytopenia and stroke (TEGSEDI, 2019). More common side effects of ASO's as observed in nusinersen include injection site pain, respiratory congestion, headache, nausea and fever (SPINRAZA, 2020). Major side effects observed in ASO therapy also include renal toxicity and thrombocytopenia (Chi et al., 2017; Zaslavsky et al., 2021). Current research on expanding the types of chemical modifications to ASO backbones to reduce these side effects is ongoing (Echevarría et al., 2019; Roberts et al., 2020).

Alternative technologies include mRNA editing and mRNA replacement. Recently, CRISPR technology has made it possible to edit mRNA transcripts using a Cas13 system known as CasRx (Konermann et al., 2018). This system utilizes adenoviral vectors to deliver targeted Cas13 to relevant cell types. CasRx has been used to both silence and direct alternative splicing of targeted RNA transcripts (Konermann et al., 2018; Zhou et al., 2020). It may be possible to use this approach to correct many types of mutations within JAG1 and NOTCH2 transcripts in relevant patient tissues. Although this therapy requires additional layers of development such as adeno-associated viral (AAV) delivery, further discussed below, this approach is applicable to the highly specific nature of mutations for each patient as a personalized treatment.

In addition to editing nascent transcripts in the cell, additional transcripts can be delivered as mRNA replacement. This mode of therapy is being evaluated for several other disorders such as Cystic Fibrosis and Type 2 Diabetes. Similar to RNA editing, proper cell-type specific delivery is required. Once in the cell, the transcripts are subject to translational regulation which may aid in maintaining levels within physiological range. However, the level of JAG1 expression must be well-controlled, and constitutive expression must be avoided, as it is known that JAG1 overexpression has been linked to many cancer types (Capaccione and Pine, 2013; Dai et al., 2014). Careful delivery of JAG1 mRNA to supplement reduced Jag1 levels caused by haploinsufficiency may not be completely out of the question (Aster et al., 2017). The optimal dose would overcome haploinsufficiency while avoiding oncogenic effects. To lower nonspecific expression of JAG1, transcripts can be delivered through antibody labeling of lipid nanoparticles to specific cell types (Veiga et al., 2018). Alternatively, mRNA itself can be designed to include cell-type specific microRNA binding sites that could also lower nonspecific expression of JAG1 (Jain et al., 2018). More basic study is required to evaluate JAG1 mRNA supplementation in ALGS models, as well as other non-transformed cell types to determine the optimal therapeutic window for this approach.

Gene Therapy Using Adeno-Associated Virus

Adeno-Associated Virus (AAV) was first discovered and characterized in the 1960s and 70s and its ability to efficiently deliver genes *in vivo* made it a promising method for human gene therapy (Kaplitt et al., 1994; Carter, 2004). Wild-type AAVs contain a single-stranded ~4.7 kb genome containing the rep gene which encodes four proteins necessary for viral replication and the cap gene that codes for the three capsid proteins (Drouin and Agbandje-McKenna, 2013). The genome of AAVs are flanked by two T-shaped inverted terminal repeats (ITRs) which are responsible for assembly and serve as origins of replication. For gene therapy, recombinant AAV (rAAV) are produced in which viral genes are replaced with the therapeutic gene of interest. The leftover ITRs guide intra-molecular assembly into circularized episomes which allows for the long-term transcription of the inserted gene (Schultz and Chamberlain, 2008). Since rAAVs are devoid of viral replication genes, they do not actively integrate into the host genome and thus have a low cytotoxicity and carcinogenesis in contrast to other viral vectors. In addition, rAAVs capsid protein can be designed to target specific cells, with over a dozen well-characterized serotypes that have already paved the way in a variety of tissues (Wu et al., 2006). These advantages have led to the broad application of rAAVs in the treatment of rare genetic disorders of which there are currently 200 ongoing clinical trials (Goswami et al., 2019).

To date two successful rAAV-based drugs have gained approval from the FDA to treat inherited genetic diseases. The first is Luxturna, which treats retinal dystrophy due to Leber Congenital Amaurosis RPE65 deficiency, a severe form of incurable childhood blindness (Cideciyan et al., 2008). The second drug is Zolgensma, which was approved to treat spinal muscular atrophy type 1, characterized by mutations to SMN1. In addition to demonstrating the safety and long-term efficacy of rAAVs for these specific diseases, the development of these two drugs also presented protocols for the large-scale production and evaluation of rAAVs that can be applied to other genetic diseases such as ALGS. Both drugs earned approval because they showed long-term improvement in their respective disease phenotypes without significant side effects (Russell, 2007; Verdera et al., 2020).

Although the virus is nonpathogenic, a significant amount of preexisting immunity is present against wild-type and recombinant capsid protein, potentially affecting the efficacy of this method across patients (Verdera et al., 2020). Some studies have also linked rAAV treatment with carcinogenesis due to insertional mutagenesis, in particular at higher doses (Russell, 2007; Nault et al., 2015). In addition to these universal limitations, the future use of rAAVs to treat ALGS is dependent on the resolution of two major challenges specific to this disease. rAAVs typically cannot deliver genes that are larger than ~5.0 kb and the major genes involved in ALGS pathology are outside this range; JAG1 is slightly larger at 5.5 kb and NOTCH2 is 7.4 kb. One workaround would be to design a shortened, but functional, isoform of JAG1 or NOTCH2. Such a microgene approach has shown promise in preclinical models of Duchenne muscular dystrophy, which involves mutations in the 11.5 kb dystrophin gene, but has yet to achieve clinical proof-of-concept in patients (Duan, 2018). Alternatively, either of the genes can be split between two co-administered rAAV vectors with the reconstitution of the full gene facilitated by ITRs that can be removed during endogenous splicing events (Nakai et al., 2000). The second challenge is the oncogenic nature of Notch signaling components which inherently requires a more targeted approach, as previously mentioned (Xiu et al., 2020). Currently, many rAAV serotypes can efficiently transduce hepatocytes and have been used as delivery vectors for other liver diseases such as Crigler-Najjar. Safe and specific JAG1 expression may require vectors which infect a multitude of cell types in the liver as well as the introduction of regulatory elements that confer tissue- and cell type-specific expression (Manno et al., 2006; Nathwani et al., 2014; Collaud et al., 2019). Cell type specificity may be of particular importance since the intracellular domain of Jag1 (JICD) has been reported to prevent NICD dependent transcriptional regulation (Kim et al., 2011).

Nuclease-Based Genome Editing

Permanent alteration of the genome *in vivo* is accomplished using four main types of nucleases: zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), meganucleases and CRISPR/Cas-based nucleases (Silva et al., 2011; Gaj et al., 2013). These nucleases induce double stranded breaks within a specific region of DNA, which then allows for the tandem incorporation of an exogenously introduced gene at that cleavage site through a process called homology-directed repair (HDR) (Silva et al., 2011; Gaj et al., 2013). While ZFNs, TALENs and meganucleases contain DNA binding domains that can be engineered to bind to specific regions of DNA, CRISPR/Cas-based nucleases are unique in that they are guided by an RNA oligomer that is specific to a particular DNA sequence (Gaj et al., 2013). Practical delivery of these nucleases and the oligo targeting the gene of interest is usually mediated by viral vectors, such as rAAVs, though non-viral methods also exist (Wang et al., 2017b). The advent of CRISPR has greatly accelerated our ability to perturb the expression of most genes due to the simplicity of its targeting mechanism. The major advantage of using this technology to treat ALGS would be the permanence of the therapy relative to other techniques, since the mutation in the genomic DNA is corrected and would persist through cell division.

Though there are currently no FDA-approved *in vivo* gene editing therapeutics, many promising candidates are undergoing clinical trials (Li et al., 2020). ZFNs have only seen moderate success in the generation of genome edited cells *ex vivo*, while TALENs have only shown efficacy in animal models (Li et al., 2020). EDIT-101 made history in 2018 as the first ever CRISPR-Cas9-based genetic therapy to enter Phase I/II clinical trials. This therapy seeks to correct a point mutation in intron 26 of the CEP290 gene which introduces an alternative splice site resulting in a premature stop codon and is the most common mutation underlying Leber Congenital Amaurosis (Maeder et al., 2019). In this treatment, AAV5 delivers the coding sequence for the Cas9 protein along with two guide RNAs. These guide RNAs flank the sequence around the mutation in intron, IVS26, and lead to its deletion or inversion, abolishing the splice site. This method can be applied to the mutations present in JAG1 or NOTCH2 in a similar fashion retaining the ability to package an entire treatment payload in just one rAAV. Other methods, like ZFNs, require up to three separate vectors to deliver the coding sequences for the nucleases and the correct transcript. This has limited the efficacy of using ZFNs, as was the case with SB-915 which failed in Phase I/II clinical trials since the uptake of all three of its vectors in a single cell was needed for any successful gene editing to occur (Sheridan, 2018). Recently, the CRISPR toolbox has expanded further to allow for the editing of single base mutations directly by coupling an inactive Cas9 nuclease with cytidine deaminase (Gaudelli et al., 2017). While this method has not been applied in the clinic yet, it has been validated in murine models showing the ability to permanently correct genes in a broad variety of tissues, including the liver (Levy et al., 2020).

Nuclease-based treatments often fail in the clinic because of off-target effects, the low rate of gene integration by HDR and host immune response (Li et al., 2020). Continued development has addressed some of these limitations, particularly for Cas-based nucleases, improving integration efficiency while limiting the induction of double stranded breaks in other regions of the genome (Xie et al., 2014; Song et al., 2016). Since the CRISPR/Cas system is derived from a bacterial source however, overcoming the high innate immunity present in human

populations to these nucleases is likely to continue to impact efficacy (Charlesworth et al., 2019). Gene editing could aim to correct single mutations or replace larger segments covering multiple mutations found in patients. Theoretically, gene editing could also aim to correct the entire mutant copy of the JAG1 gene already present in patients with Alagille Syndrome. While ZFNs, TALENs and meganucleases can be designed to target any sequence for cleavage, Cas nucleases require the presence of a PAM sequence near the targeting site, potentially limiting what parts of the JAG1 gene can be corrected. The emergence of Cas-based point mutation editing is particularly promising for ALGS as it potentially alleviates the largest limitations of nuclease-based gene editing, offering further options for the development of gene editing-based therapeutics.

PERSPECTIVES FOR ALGS THERAPEUTICS DEVELOPMENT

Recent advances in organoid technology and patient-derived induced pluripotent stem cells (iPSC) have brought new modeling systems to the hunt for therapies of rare diseases. For Alagille Syndrome in particular, the difficulty of identifying therapeutics which modulate the Notch signaling pathway with a desired functional outcome would require such technologies. Improvements in iPSC liver cell differentiation have allowed modeling of human liver diseases with patient-derived cells. Human iPSCs can be generated from patient cell samples such as peripheral blood, dermal fibroblasts, urine, hair follicles and keratinocytes (Takahashi et al., 2007; Farkhondeh et al., 2019). The use of iPSC-derived hepatocytes, cholangiocytes and stellate cells will assist in advancing compound screening as well as drug efficacy evaluation. Furthermore, patient-derived iPSCs carry the unique genetic background of the donor which allows to precisely model disease-specific pathophysiology and phenotypes observed in complex genetic disorders such as ALGS. Liver cells derived from iPSCs can be produced in a large quantity and would be a new direction for the broader context of liver disease drug discovery.

Patient-derived organoids are a newly established platform for disease modeling and development of precision medicine (van de Wetering et al., 2015). Organoid models generated from mutation-specific iPSC lines can be used to establish disease-relevant phenotypes to evaluate and predict clinical efficacies for lead compounds. ALGS patient-derived iPSCs and their differentiated organoids represent a more relevant disease model system as they share the same genetic background as the patients they come from. Furthermore, these liver organoids can share similar disease phenotypes as observed in patients, providing further biological context for drug development. Intestinal organoids for example, have been used as a model for evaluating cystic fibrosis variants for drug development (Chen et al., 2019). Single-cell RNA-sequencing of organoids can yield additional valuable information about the transcriptional landscape of interacting cells. Using a combination of scRNA-Seq and immunostaining it may be possible to identify new

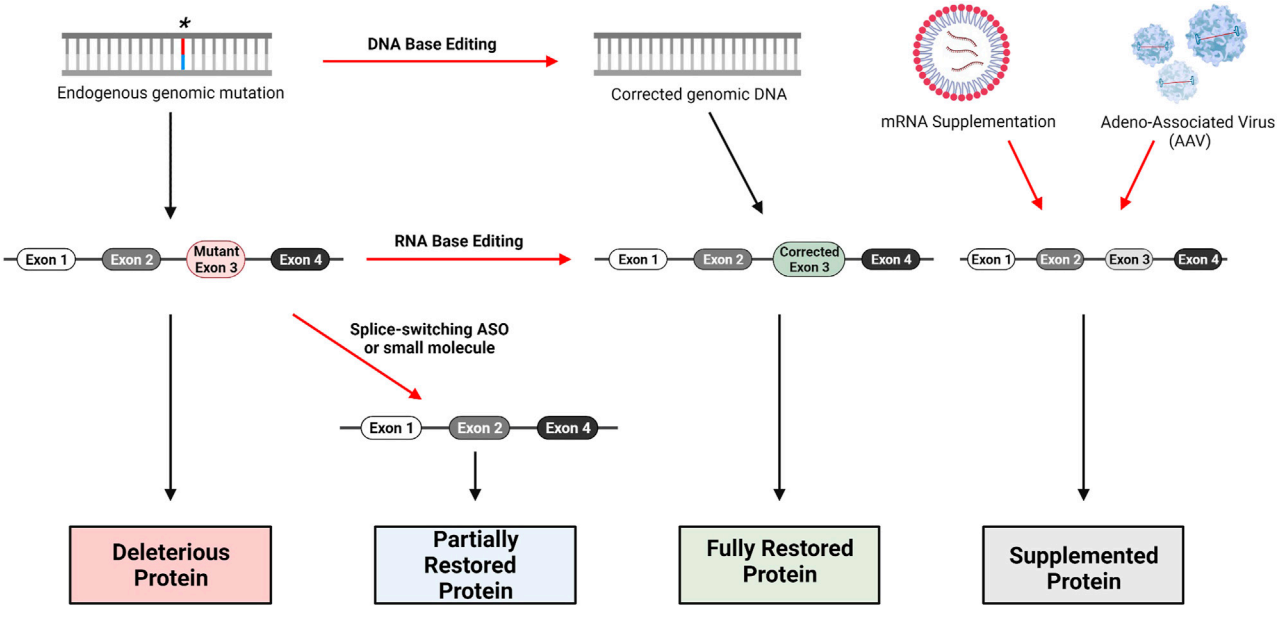

FIGURE 3 | Gene therapy approaches for loss-of-function genetic disorders. DNA base editing technologies include CRISPR/Cas9, Zinc-Finger Nucleases (ZFN) and Transcription Activator-Like Effector Nucleases (TALENS). RNA base editing of pre-mRNA can be performed with the CasRx system, utilizing Cas13. Skipping of mutant-containing exons during mRNA processing can be performed with Antisense Oligonucleotides (ASO) specifically designed to bind relevant intron-exon junctions; small-molecules such as risdiplam are also capable of altering splicing. Traditional replacement therapy including mRNA supplementation and AAV delivery are possible as well, though several factors complicate this approach for Alagille Syndrome as we have discussed.

targets for drug development. By having a patient-specific context for disease, as with ALGS, these technologies can be leveraged to enable the discovery of novel small-molecule and RNA therapies (Akbari et al., 2019; Sakabe et al., 2020). Although compound screening has expanded into the realm of ultra high-throughput technologies, i. e 1536-well screening of >1 million molecule libraries, incorporation of high-content imaging and machine learning for image analysis and modeling will accelerate drug development. Phenotypic screening in this context has already been developed for several orphan diseases like Batten's disease and Niemann-Pick Type C (NPC) disease (Chandrachud et al., 2015; Sun et al., 2017; Pugach et al., 2018).

The technologies presented here have the potential for development of therapeutics to treat all types of mutations in JAG1 or NOTCH2 that have been identified as pathologic variants for ALGS (**Figure 3**). The design of truly personalized RNA therapies as with Milasen is conceivable for ALGS. However, given the publicly available data surrounding ALGS mutations in both JAG1 and NOTCH2, it should be possible to generate RNA or gene editing therapies for subpopulations of patients that carry mutations in the same region of either of these genes. In addition to targeting specific mutations represented in the JAG1 and NOTCH2 mutational database, it may also be possible to develop long-term or even permanent treatments that can be applied to all patients using rAAVs or nuclease-based gene editing. The development of a targeted ALGS therapy would be a much-needed improvement over current approaches which only ameliorate symptoms of disease.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This work was supported by the Intramural Research Programs of the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (ZIA TR000018-06 to WZ). The open access fee will be paid with the above fund as well.

ACKNOWLEDGMENTS

This work was supported by the Intramural Research Programs of the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health, and was a CRADA collaboration between NCATS, Alagille Syndrome Alliance (alagille.org), and Travere Therapeutics. Figures were generated using BioRender.com.

REFERENCES

Akbari, S., Arslan, N., Senturk, S., and Erdal, E. (2019). Next-Generation Liver Medicine Using Organoid Models. *Front Cel Dev Biol* 7, 345. doi:10.3389/fcell.2019.00345

Alagille, D., Odièvre, M., Gautier, M., and Dommergues, J. P. (1975). Hepatic Ductular Hypoplasia Associated with Characteristic Facies, Vertebral Malformations, Retarded Physical, Mental, and Sexual Development, and Cardiac Murmur. *J. Pediatr.* 86 (1), 63–71. doi:10.1016/s0022-3476(75)80706-2

Aster, J. C., Pear, W. S., and Blacklow, S. C. (2017). The Varied Roles of Notch in Cancer. *Annu. Rev. Pathol.* 12, 245–275. doi:10.1146/annurev-pathol-052016-100127

Capaccione, K. M., and Pine, S. R. (2013). The Notch Signaling Pathway as a Mediator of Tumor Survival. *Carcinogenesis* 34 (7), 1420–1430. doi:10.1093/carcin/bgt127

Carter, B. J. (2004). Adeno-associated Virus and the Development of Adeno-Associated Virus Vectors: a Historical Perspective. *Mol. Ther.* 10 (6), 981–989. doi:10.1016/j.ymthe.2004.09.011

Chandrachud, U., Walker, M. W., Simas, A. M., Heetveld, S., Petcherski, A., Klein, M., et al. (2015). Unbiased Cell-Based Screening in a Neuronal Cell Model of Batten Disease Highlights an Interaction between Ca2+ Homeostasis, Autophagy, and CLN3 Protein Function. *J. Biol. Chem.* 290 (23), 14361–14380. doi:10.1074/jbc.M114.621706

Charlesworth, C. T., Deshpande, P. S., Dever, D. P., Camarena, J., Lemgart, V. T., Cromer, M. K., et al. (2019). Identification of Preexisting Adaptive Immunity to Cas9 Proteins in Humans. *Nat. Med.* 25 (2), 249–254. doi:10.1038/s41591-018-0326-x

Chen, K. G., Zhong, P., Zheng, W., and Beekman, J. M. (2019). Pharmacological Analysis of CFTR Variants of Cystic Fibrosis Using Stem Cell-Derived Organoids. *Drug Discov. Today* 24 (11), 2126–2138. doi:10.1016/j.drudis.2019.05.029

Chi, X., Gatti, P., and Papoian, T. (2017). Safety of Antisense Oligonucleotide and siRNA-Based Therapeutics. *Drug Discov. Today* 22 (5), 823–833. doi:10.1016/j.drudis.2017.01.013

Cideciyan, A. V., Aleman, T. S., Boye, S. L., Schwartz, S. B., Kaushal, S., Roman, A. J., et al. (2008). Human Gene Therapy for RPE65 Isomerase Deficiency Activates the Retinoid Cycle of Vision but with Slow Rod Kinetics. *Proc. Natl. Acad. Sci. U S A.* 105 (39), 15112–15117. doi:10.1073/pnas.0807027105

Collaud, F., Bortolussi, G., Guianvarc'h, L., Aronson, S. J., Bordet, T., Veron, P., et al. (2019). Preclinical Development of an AAV8-hUGT1A1 Vector for the Treatment of Crigler-Najjar Syndrome. *Mol. Ther. Methods Clin. Dev.* 12, 157–174. doi:10.1016/j.omtm.2018.12.011

Dai, Y., Wilson, G., Huang, B., Peng, M., Teng, G., Zhang, D., et al. (2014). Silencing of Jagged1 Inhibits Cell Growth and Invasion in Colorectal Cancer. *Cell Death Dis* 5, e1170. doi:10.1038/cddis.2014.137

Drouin, L. M., and Agbandje-McKenna, M. (2013). Adeno-associated Virus Structural Biology as a Tool in Vector Development. *Future Virol.* 8 (12), 1183–1199. doi:10.2217/fvl.13.112

Duan, D. (2018). Systemic AAV Micro-dystrophin Gene Therapy for Duchenne Muscular Dystrophy. *Mol. Ther.* 26 (10), 2337–2356. doi:10.1016/j.ymthe.2018.07.011

Echevarría, L., Aupy, P., Relizani, K., Bestetti, T., Griffith, G., Blandel, F., et al. (2019). Evaluating the Impact of Variable Phosphorothioate Content in Tricyclo-DNA Antisense Oligonucleotides in a Duchenne Muscular Dystrophy Mouse Model. *Nucleic Acid Ther.* 29 (3), 148–160. doi:10.1089/nat.2018.0773

Fabris, L., Cadamuro, M., Guido, M., Spirli, C., Fiorotto, R., Colledan, M., et al. (2007). Analysis of Liver Repair Mechanisms in Alagille Syndrome and Biliary Atresia Reveals a Role for Notch Signaling. *Am. J. Pathol.* 171 (2), 641–653. doi:10.2353/ajpath.2007.070073

Farkhondeh, A., Li, R., Gorshkov, K., Chen, K. G., Might, M., Rodems, S., et al. (2019). Induced Pluripotent Stem Cells for Neural Drug Discovery. *Drug Discov. Today* 24 (4), 992–999. doi:10.1016/j.drudis.2019.01.007

Finkel, R. S., Chiriboga, C. A., Vajsar, J., Day, J. W., Montes, J., De Vivo, D. C., et al. (2016). Treatment of Infantile-Onset Spinal Muscular Atrophy with

Nusinersen: a Phase 2, Open-Label, Dose-Escalation Study. *Lancet* 388 (10063), 3017–3026. doi:10.1016/S0140-6736(16)31408-8

Freeburg, S. H., and Goessling, W. (2020). Hepatobiliary Differentiation: Principles from Embryonic Liver Development. *Semin. Liver Dis.* 40 (4), 365–372. doi:10.1055/s-0040-1709679

Gaj, T., Gersbach, C. A., and Barbas, C. F. Z. F. N., 3rd (2013). ZFN, TALEN, and CRISPR/Cas-based Methods for Genome Engineering. *Trends Biotechnol.* 31 (7), 397–405. doi:10.1016/j.tibtech.2013.04.004

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., et al. (2017). Programmable Base Editing of A●T to G●C in Genomic DNA without DNA Cleavage. *Nature* 551 (7681), 464–471. doi:10.1038/nature24644

Gilbert, M. A., Bauer, R. C., Rajagopalan, R., Grochowski, C. M., Chao, G., McEldrew, D., et al. (2019). Alagille Syndrome Mutation Update: Comprehensive Overview of JAG1 and NOTCH2 Mutation Frequencies and Insight into Missense Variant Classification. *Hum. Mutat.* 40 (12), 2197–2220. doi:10.1002/humu.23879

Goswami, R., Subramanian, G., Silayeva, L., Newkirk, I., Doctor, D., Chawla, K., et al. (2019). Gene Therapy Leaves a Vicious Cycle. *Front. Oncol.* 9, 297. doi:10.3389/fonc.2019.00297

Hofmann, J. J., Zovein, A. C., Koh, H., Radtke, F., Weinmaster, G., and Iruela-Arispe, M. L. (2010). Jagged1 in the portal Vein Mesenchyme Regulates Intrahepatic Bile Duct Development: Insights into Alagille Syndrome. *Development* 137 (23), 4061–4072. doi:10.1242/dev.052118

Jain, R., Frederick, J. P., Huang, E. Y., Burke, K. E., Mauger, D. M., Andrianova, E. A., et al. (2018). MicroRNAs Enable mRNA Therapeutics to Selectively Program Cancer Cells to Self-Destruct. *Nucleic Acid Ther.* 28 (5), 285–296. doi:10.1089/nat.2018.0734

Kamath, B. M., Baker, A., Houwen, R., Todorova, L., and Kerkar, N. (2018). Systematic Review: The Epidemiology, Natural History, and Burden of Alagille Syndrome. *J. Pediatr. Gastroenterol. Nutr.* 67 (2), 148–156. doi:10.1097/MPG.0000000000001958

Kamath, B. M., Podkameni, G., Hutchinson, A. L., Leonard, L. D., Gerfen, J., Krantz, I. D., et al. (2012). Renal Anomalies in Alagille Syndrome: a Disease-Defining Feature. *Am. J. Med. Genet. A.* 158A (1), 85–89. doi:10.1002/ajmg.a.34369

Kamath, B. M., Ye, W., Goodrich, N. P., Loomes, K. M., Romero, R., Heubi, J. E., et al. (2020). Outcomes of Childhood Cholestasis in Alagille Syndrome: Results of a Multicenter Observational Study. *Hepatol. Commun.* 4 (3), 387–398. doi:10.1002/hep4.1468

Kaplitt, M. G., Leone, P., Samulski, R. J., Xiao, X., Pfaff, D. W., O'Malley, K. L., et al. (1994). Long-term Gene Expression and Phenotypic Correction Using Adeno-Associated Virus Vectors in the Mammalian Brain. *Nat. Genet.* 8 (2), 148–154. doi:10.1038/ng1094-148

Khvorova, A., and Watts, J. K. (2017). The Chemical Evolution of Oligonucleotide Therapies of Clinical Utility. *Nat. Biotechnol.* 35 (3), 238–248. doi:10.1038/nbt.3765

Kim, J., Hu, C., Moufawad El Achkar, C., Black, L. E., Douville, J., Larson, A., et al. (2019). Patient-Customized Oligonucleotide Therapy for a Rare Genetic Disease. *N. Engl. J. Med.* 381 (17), 1644–1652. doi:10.1056/NEJMoa1813279

Kim, M. Y., Jung, J., Mo, J. S., Ann, E. J., Ahn, J. S., Yoon, J. H., et al. (2011). The Intracellular Domain of Jagged-1 Interacts with Notch1 Intracellular Domain and Promotes its Degradation through Fbw7 E3 Ligase. *Exp. Cel Res* 317 (17), 2438–2446. doi:10.1016/j.yexcr.2011.07.014

Kole, R., and Krieg, A. M. (2015). Exon Skipping Therapy for Duchenne Muscular Dystrophy. *Adv. Drug Deliv. Rev.* 87, 104–107. doi:10.1016/j.addr.2015.05.008

Konermann, S., Lotfy, P., Brideau, N. J., Oki, J., Shokhirev, M. N., and Hsu, P. D. (2018). Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. *Cell* 173 (3), 665–e14. e14. doi:10.1016/j.cell.2018.02.033

Kopan, R., and Ilagan, M. X. (2009). The Canonical Notch Signaling Pathway: Unfolding the Activation Mechanism. *Cell* 137 (2), 216–233. doi:10.1016/j.cell.2009.03.045

Levy, J. M., Yeh, W. H., Pendse, N., Davis, J. R., Hennessey, E., Butcher, R., et al. (2020). Cytosine and Adenine Base Editing of the Brain, Liver, Retina, Heart and Skeletal Muscle of Mice via Adeno-Associated Viruses. *Nat. Biomed. Eng.* 4 (1), 97–110. doi:10.1038/s41551-019-0501-5

Li, H., Yang, Y., Hong, W., Huang, M., Wu, M., and Zhao, X. (2020). Applications of Genome Editing Technology in the Targeted Therapy of Human Diseases: Mechanisms, Advances and Prospects. *Signal. Transduct Target. Ther.* 5 (1), 1. doi:10.1038/s41392-019-0089-y

Lykavieris, P., Hadchouel, M., Chardot, C., and Bernard, O. (2001). Outcome of Liver Disease in Children with Alagille Syndrome: a Study of 163 Patients. *Gut* 49 (3), 431–435. doi:10.1136/gut.49.3.431

Maeder, M. L., Stefanidakis, M., Wilson, C. J., Baral, R., Barrera, L. A., Bounoutas, G. S., et al. (2019). Development of a Gene-Editing Approach to Restore Vision Loss in Leber Congenital Amaurosis Type 10. *Nat. Med.* 25 (2), 229–233. doi:10.1038/s41591-018-0327-9

Manno, C. S., Pierce, G. F., Arruda, V. R., Glader, B., Ragni, M., Rasko, J. J., et al. (2006). Successful Transduction of Liver in Hemophilia by AAV-Factor IX and Limitations Imposed by the Host Immune Response. *Nat. Med.* 12 (3), 342–347. doi:10.1038/nm1358

Mattei, P., von Allmen, D., Piccoli, D., and Rand, E. (2006). Relief of Intractable Pruritus in Alagille Syndrome by Partial External Biliary Diversion. *J. Pediatr. Surg.* 41 (1), 104–107. discussion 104-7. doi:10.1016/j.jpedsurg.2005.10.014

Nakai, H., Storm, T. A., and Kay, M. A. (2000). Increasing the Size of rAAV-Mediated Expression Cassettes *In Vivo* by Intermolecular Joining of Two Complementary Vectors. *Nat. Biotechnol.* 18 (5), 527–532. doi:10.1038/75390

Nathwani, A. C., Reiss, U. M., Tuddenham, E. G., Rosales, C., Chowdary, P., McIntosh, J., et al. (2014). Long-term Safety and Efficacy of Factor IX Gene Therapy in Hemophilia B. *N. Engl. J. Med.* 371 (21), 1994–2004. doi:10.1056/NEJMoa1407309

Nault, J. C., Datta, S., Imbeaud, S., Franconi, A., Mallet, M., Couchy, G., et al. (2015). Recurrent AAV2-Related Insertional Mutagenesis in Human Hepatocellular Carcinomas. *Nat. Genet.* 47 (10), 1187–1193. doi:10.1038/ng.3389

Olyaei, A. J., de Mattos, A. M., and Bennett, W. M. (2001). Nephrotoxicity of Immunosuppressive Drugs: New Insight and Preventive Strategies. *Curr. Opin. Crit. Care* 7 (6), 384–389. doi:10.1097/00075198-200112000-00003

Patient Family Survey (2019). *In Alagille Syndrome International Symposium*. Ohio: Cincinnati.

Pharma, A.. Odevixibat. 2020a; Available from: https://www.albireopharma.com/programs/odevixibat.[Accessed 06--11-2020].

Pharma, M.. Maralixibat. 2020b; Available from: https://mirumpharma.com/programs/#section_maralixibat. [Accessed 06-11-2020].

Pugach, E. K., Feltes, M., Kaufman, R. J., Ory, D. S., and Bang, A. G. (2018). High-content Screen for Modifiers of Niemann-Pick Type C Disease in Patient Cells. *Hum. Mol. Genet.* 27 (12), 2101–2112. doi:10.1093/hmg/ddy117

Ratni, H., Ebeling, M., Baird, J., Bendels, S., Bylund, J., Chen, K. S., et al. (2018). Discovery of Risdiplam, a Selective Survival of Motor Neuron-2 (SMN2) Gene Splicing Modifier for the Treatment of Spinal Muscular Atrophy (SMA). *J. Med. Chem.* 61 (15), 6501–6517. doi:10.1021/acs.jmedchem.8b00741

Roberts, T. C., Langer, R., and Wood, M. J. A. (2020). Advances in Oligonucleotide Drug Delivery. *Nat. Rev. Drug Discov.* 19 (10), 673–694. doi:10.1038/s41573-020-0075-7

Russell, D. W. (2007). AAV Vectors, Insertional Mutagenesis, and Cancer. *Mol. Ther.* 15 (10), 1740–1743. doi:10.1038/sj.mt.6300299

Sakabe, K., Takebe, T., and Asai, A. (2020). Organoid Medicine in Hepatology. *Clin. Liver Dis. (Hoboken)* 15 (1), 3–8. doi:10.1002/cld.855

Schultz, B. R., and Chamberlain, J. S. (2008). Recombinant Adeno-Associated Virus Transduction and Integration. *Mol. Ther.* 16 (7), 1189–1199. doi:10.1038/mt.2008.103

Sheridan, C. (2018). Sangamo's Landmark Genome Editing Trial Gets Mixed Reception. *Nat. Biotechnol.* 36 (10), 907–908. doi:10.1038/nbt1018-907

Silva, G., Poirot, L., Galetto, R., Smith, J., Montoya, G., Duchateau, P., et al. (2011). Meganucleases and Other Tools for Targeted Genome Engineering: Perspectives and Challenges for Gene Therapy. *Curr. Gene Ther.* 11 (1), 11–27. doi:10.2174/156652311794520111

Song, J., Yang, D., Xu, J., Zhu, T., Chen, Y. E., and Zhang, J. (2016). RS-1 Enhances CRISPR/Cas9- and TALEN-Mediated Knock-In Efficiency. *Nat. Commun.* 7, 10548. doi:10.1038/ncomms10548

Spinner, N. B., Gilbert, M. A., Loomes, K., and Mand Krantz, L. D. (1993). in *Alagille Syndrome, in GeneReviews((R))*. Editor M. P. Adam (WA: Seattle).

SPINRAZA (2020). *Prescribing Information*, 2. Cambridge, MA: Biogen. Biogen Data on File [as of 12/20].

Sun, W., Zheng, W., and Simeonov, A. (2017). Drug Discovery and Development for Rare Genetic Disorders. *Am. J. Med. Genet. A.* 173 (9), 2307–2322. doi:10.1002/ajmg.a.38326

Syed, Y. Y. (2016). Eteplirsen: First Global Approval. *Drugs* 76 (17), 1699–1704. doi:10.1007/s40265-016-0657-1

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., et al. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* 131 (5), 861–872. doi:10.1016/j.cell.2007.11.019

TEGSEDI (2019). *Package Insert*. Boston, MA: Akcea Therapeutics, Inc.

van de Wetering, M., Francies, H. E., Francis, J. M., Bounova, G., Iorio, F., Pronk, A., et al. (2015). Prospective Derivation of a Living Organoid Biobank of Colorectal Cancer Patients. *Cell* 161 (4), 933–945. doi:10.1016/j.cell.2015.03.053

van Wessel, D. B. E., Thompson, R. J., Gonzales, E., Jankowska, I., Sokal, E., Grammatikopoulos, T., et al. (2020). Genotype Correlates with the Natural History of Severe Bile Salt export Pump Deficiency. *J. Hepatol.* 73 (1), 84–93. doi:10.1016/j.jhep.2020.02.007

Veiga, N., Goldsmith, M., Granot, Y., Rosenblum, D., Dammes, N., Kedmi, R., et al. (2018). Cell Specific Delivery of Modified mRNA Expressing Therapeutic Proteins to Leukocytes. *Nat. Commun.* 9 (1), 4493. doi:10.1038/s41467-018-06936-1

Verdera, H. C., Kuranda, K., and Mingozzi, F. (2020). AAV Vector Immunogenicity in Humans: A Long Journey to Successful Gene Transfer. *Mol. Ther.* 28 (3), 723–746. doi:10.1016/j.ymthe.2019.12.010

Wang, K. S., Tiao, G., Bass, L. M., Hertel, P. M., Mogul, D., Kerkar, N., et al. (2017). Analysis of Surgical Interruption of the Enterohepatic Circulation as a Treatment for Pediatric Cholestasis. *Hepatology* 65 (5), 1645–1654. doi:10.1002/hep.29019

Wang, M., Glass, Z. A., and Xu, Q. (2017). Non-viral Delivery of Genome-Editing Nucleases for Gene Therapy. *Gene Ther.* 24 (3), 144–150. doi:10.1038/gt.2016.72

Woodcock, J., and Marks, P. (2019). Drug Regulation in the Era of Individualized Therapies. *N. Engl. J. Med.* 381 (17), 1678–1680. doi:10.1056/NEJMe1911295

Wu, Z., Asokan, A., and Samulski, R. J. (2006). Adeno-associated Virus Serotypes: Vector Toolkit for Human Gene Therapy. *Mol. Ther.* 14 (3), 316–327. doi:10.1016/j.ymthe.2006.05.009

Xie, S., Shen, B., Zhang, C., Huang, X., and Zhang, Y. (2014). sgRNAcas9: a Software Package for Designing CRISPR sgRNA and Evaluating Potential Off-Target Cleavage Sites. *PLoS One* 9 (6), e100448. doi:10.1371/journal.pone.0100448

Xiu, M. X., Liu, Y. M., and Kuang, B. H. (2020). The Oncogenic Role of Jagged1/Notch Signaling in Cancer. *Biomed. Pharmacother.* 129, 110416. doi:10.1016/j.biopha.2020.110416

Zaslavsky, A., Adams, M., Cao, X., Yamaguchi, A., Henderson, J., Busch-Østergren, P., et al. (2021). Antisense Oligonucleotides and Nucleic Acids Generate Hypersensitive Platelets. *Thromb. Res.* 200, 64–71. doi:10.1016/j.thromres.2021.01.006

Zhou, H., Su, J., Hu, X., Zhou, C., Li, H., Chen, Z., et al. (2020). Glia-to-Neuron Conversion by CRISPR-CasRx Alleviates Symptoms of Neurological Disease in Mice. *Cell* 181 (3), 590–e16. e16. doi:10.1016/j.cell.2020.03.024

Conflict of Interest: Authors KB and SR were employed by the company Travere Therapeutics.

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 24 August 2021
doi: 10.3389/fgene.2021.630650

Transmission of a Novel Imprinting Center Deletion Associated With Prader–Willi Syndrome Through Three Generations of a Chinese Family: Case Presentation, Differential Diagnosis, and a Lesson Worth Thinking About

Kaihui Zhang[1,2,3†], Shu Liu[4†], Wenjun Gu[3†], Yuqiang Lv[2], Haihua Yu[5], Min Gao[2], Dong Wang[2], Jianyuan Zhao[3], Xiaoying Li[2], Zhongtao Gai[2], Shimin Zhao[1,3,6], Yi Liu[2]* and Yiyuan Yuan[1,3]**

[1] Obstetrics and Gynecology Hospital of Fudan University, Fudan University, Shanghai, China, [2] Pediatric Research Institute, Qilu Children's Hospital of Shandong University, Jinan, China, [3] State Key Laboratory of Genetic Engineering and School of Life Sciences, Fudan University, Shanghai, China, [4] Children Inherited Metabolism and Endocrine Department, Guangdong Women and Children Hospital, Guangzhou, China, [5] Neonatal Intensive Care Unit, Qilu Children's Hospital of Shandong University, Jinan, China, [6] Key Laboratory of Reproduction Regulation of NPFPC, Collaborative Innovation Center of Genetics and Development, Institutes of Biomedical Sciences, Fudan University, Shanghai, China

Prader–Willi syndrome (PWS) is a complex genetic syndrome caused by the loss of function of genes in 15q11-q13 that are subject to regulation by genomic imprinting and expressed from the paternal allele only. The main clinical features of PWS patients are hypotonia during the neonatal and infantile stages, accompanied by delayed neuropsychomotor development, hyperphagia, obesity, hypogonadism, short stature, small hands and feet, mental disabilities, and behavioral problems. However, PWS has a clinical overlap with other disorders, especially those with other gene variations or chromosomal imbalances but sharing part of the similar clinical manifestations with PWS, which are sometimes referred to as Prader–Willi syndrome-like (PWS-like) disorders. Furthermore, it is worth mentioning that significant obesity as a consequence of hyperphagia in PWS usually develops between the ages of 1 and 6 years, which makes early diagnosis difficult. Thus, PWS is often not clinically recognized in infants and, on the other hand, may be wrongly suspected in obese and intellectually disabled patients. Therefore, an accurate investigation is necessary to differentiate classical PWS from PWS-like phenotypes, which is imperative for further treatment. For PWS, it is usually sporadic, and very rare family history and affected siblings have been described. Here, we report the clinical and molecular findings in a three-generation family with a novel 550-kb microdeletion affecting the chromosome 15 imprinting

OPEN ACCESS

Edited by:
Tieliu Shi,
East China Normal University, China

Reviewed by:
Theresa V. Strong,
Foundation for Prader–Willi Research, United States
Stormy J. Chamberlain,
University of Connecticut Health Center, United States

***Correspondence:**
Shimin Zhao
zhaosm@fudan.edu.cn
Yi Liu
liuyi-ly@126.com
Yiyuan Yuan
yiyuanyuan@fudan.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: *03 December 2020*
Accepted: *19 July 2021*
Published: *24 August 2021*

Citation:
Zhang K, Liu S, Gu W, Lv Y, Yu H, Gao M, Wang D, Zhao J, Li X, Gai Z, Zhao S, Liu Y and Yuan Y (2021) Transmission of a Novel Imprinting Center Deletion Associated With Prader–Willi Syndrome Through Three Generations of a Chinese Family: Case Presentation, Differential Diagnosis, and a Lesson Worth Thinking About. Front. Genet. 12:630650. doi: 10.3389/fgene.2021.630650

center (IC). Overall, the present study finds that the symptoms of our patient are somewhat different from those of typical PWS cases diagnosed and given treatment in our hospital. The familial occurrence and clinical features were challenging to our diagnostic strategy. The microdeletion included a region within the complex small nuclear ribonucleoprotein polypeptide protein N (*SNRPN*) gene locus encompassing the PWS IC and was identified by using a variety of techniques. Haplotype studies suggest that the IC microdeletion was vertically transmitted from an unaffected paternal grandmother to an unaffected father and then caused PWS in two sibling grandchildren when the IC microdeletion was inherited paternally. Based on the results of our study, preimplantation genetic diagnosis (PGD) was applied successfully to exclude imprinting deficiency in preimplantation embryos before transfer into the mother's uterus. Our study may be especially instructive regarding accurate diagnosis, differential diagnosis, genetic counseling, and PGD for familial PWS patients.

Keywords: Prader–Willi syndrome, Prader–Willi-like syndrome, imprinting center, microdeletion, familial transmission

INTRODUCTION

The Prader–Willi syndrome (PWS) is characterized by hypotonia and feeding problems in early infancy, as well as hypogonadism, small hands and feet, craniofacial signs, and hyperphagia leading to profound obesity. It is triggered by the loss of function of genes in 15q11-q13 that are regulated by genomic imprinting and expressed from the paternal allele only. In approximately 70% of the cases, PWS is the result of deletion of 5–7 Mb in the paternal 15q11-13 region; nearly 28% attributes to maternal uniparental disomy (mUPD); and in < 2%, it is developed as the consequence of mutation, microdeletion, or translocation disrupting the imprinting center (IC) (Buiting et al., 1995). In PWS IC microdeletion cases, the smallest region of microdeletion overlap (IC PWS-SRO) is about 4.3 kb and spans the promoter and exon 1 region of small nuclear ribonucleoprotein polypeptide N (SNRPN), and this region appears to be necessary for erasure of the maternal imprint and establishment and maintenance of the paternal imprint (Ohta et al., 1999). The vast majority of PWS patients, typically manifested as sporadic cases, are characterized by a low recurrence risk, whereas for PWS resulting from a familial microdeletion in the IC carried by the father, the recurrence risk could be as high as 50%. To the best of our knowledge, familial transmission of the IC microdeletion or multiple affected siblings have been proven to be very rare, and only four familial PWS cases with IC microdeletion transmitted through three generations have been reported so far (Hartin et al., 2018). This finding suggests that in families with an IC microdeletion, several generations may be unaffected and asymptomatic before some individual develops PWS. Even more complicated is that, recently, research has shown that phenotypic features typical of PWS can also be caused by other genetic variations that are associated with disorders defined as the so-called Prader–Willi syndrome-like (PWS-like) disorders (Gillessen-Kaesbach and Horsthemke, 1994; Lukusa and Fryns, 2000; Stalker et al., 2003; D'Angelo et al., 2006; Bonaglia et al., 2008a; Hosoki et al., 2009; Izumi et al., 2013; Schaaf et al., 2013; Desch et al., 2015; Dello Russo et al., 2016; Fountain et al., 2016; Geets et al., 2016; Linhares et al., 2016; Berger et al., 2017; McCarthy et al., 2018; Negishi et al., 2019). As a result, attempting to make a definite diagnosis of PWS as well as the differential diagnosing has become extremely challenging. Here, we reported two neonatal siblings with atypical but may be more severe phenotype of PWS triggered by a microdeletion in the PWS IC that was transmitted through three generations—unaffected paternal grandmother, unaffected father, and then these two affected sibling grandchildren. In the present study, we aimed to share our case presentation and explore some useful methodology for detecting the silent transmission of PWS IC microdeletion through the female germline that may cause considerable difficulties in diagnostic testing and genetic counseling in affected families.

MATERIALS AND METHODS

Patients

The 1-day-old male proposita was the second child of healthy non-consanguineous Chinese parents. Parental ages at birth of the child were 31 years (father) and 30 years (mother). During the pregnancy, ultrasound screening did not reveal any structural abnormalities except for decreased fetal activity. The first child of the couple was a baby girl who died 4 days after birth for unknown causes. Considering the family history of unexplained neonatal death, interventional prenatal diagnosis was recommended by the genetic counselor but refused by the couple. At 40 weeks + 5 days' gestation, the patient was hospitalized with fetal distress. An elective cesarean section was performed on the day of admission due to breech presentation. The neonatal evaluation recorded 2,950 g in birth weight (17.8th centile), 50 cm in length (41.7th centile), and 35 cm in head circumference (65.6th centile), with an Apgar score of 10, 7, and 5 at 1, 5, and 10 min, respectively. The boy was transferred to the neonatal intensive care unit directly after birth because of severe respiratory distress, poor suck, and

hypotonia. A thin male baby with poor response to external stimuli was noticed from a first glimpse at the patient. Physical examination revealed weak cry, flat nose, big nostrils, bilateral epicanthus, wide-spaced nipples, long slender fingers, and slight maldescent of the left testis (**Figure 1**). Neurological exam showed significant hypotonia. Ventilator and gastric tube feeding were applied on account of dyspnea and poor sucking. The boy died from respiratory infections (neonatal pneumonia), hypoventilation, and respiratory distress at the age of 10 days. Since the first child of the couple (the elder sister of the proposita) died 4 days after birth, retrospective analysis was conducted to compare the boy with his elder sister, with the results indicating more severe manifestation and rapid progression, particularly dyspnea and feeding difficulties, in the baby girl. In addition, as no similar findings were observed or reported from any of the other family members, neonatal lethal monogenic disorders such as fatty acid oxidation disorders and urea cycle disorder, as well as organic acid metabolism diseases including mitochondrial disease and chromosomal disease, were regarded as the overriding considerations. Also, PWS and PWS-like disorders could not be ruled out due to part of the clinical characteristics shared by the two siblings.

Genetic Testing

Karyotyping

Peripheral blood samples from the patient and his family members were obtained for karyotyping, with their information anonymized prior to submission. Standard G-banded chromosome analysis at a 550-band resolution was performed using phytohemagglutinin (PHA)-stimulated peripheral blood lymphocytes prepared from the samples according to standard procedures. Chromosomal abnormalities were designated based on the International System for Human Cytogenetic Nomenclature guidelines (ISCN 2016).

DNA Collection and Extraction

To isolate genomic DNA, the QIAamp DNA blood mini kit (cat. no. 51104; Qiagen GmbH) was used according to the manufacturer's protocol.

FIGURE 1 | Clinical features of our patient with Prader–Willi syndrome (PWS) (2 days after birth). **(A)** Non-invasive respiratory support with positive end expiratory pressure was used for the treatment of respiratory distress. Overall appearance of the boy indicated no significant difference between the patient and normal infants. Note muscular hypotonia and abnormal position of the hands and feet. **(B)** Facial features of the boy (narrow forehead, fair eyebrow, bilateral epicanthus, flat nose, big nostrils, and thin upper lip). **(C)** Overall appearance of abnormal position of the hands and feet. **(D)** Abnormal position of the fingers with thumbs adducted under index and middle fingers, flexed hands and wrists, persistently clenched hands and arachnodactyly. **(E)** Small feet and toes. **(F)** Genital hypoplasia with slight maldescent of the left testis.

Whole-Exome Sequencing and Variant Calling

Proband DNA was sequenced to identify the causal gene. The DNA was isolated from peripheral blood with CWE9600 Automated Nucleic Acid Extraction System using CWE2100 Blood DNA Kit V2 (CWBiotech, China, CW2553). Here, 750 ng of genomic DNA was fragmented into 200–300 bp by Scientz08-III Ultrasonic Homogenizer (SCIENTZ, China). The DNA fragments were then processed by end-repairing, A-tailing, and adaptor ligation using KAPA Library Preparation Kit (Illumina, KR0453, v3.13), followed by an eight-cycle pre-capture PCR amplification. Then, the amplified DNA sample was captured in the Agilent SureSelect XT2 Target Enrichment System (Agilent Technologies, Inc., United States), with the captured DNA fragments purified by Dynabeads MyOne Streptavidin T1 (Invitrogen, Thermo Fisher Scientific, United States) and amplified by 13 cycles of post-capture PCR. The final products were further purified by Agencourt AMPure XP (Beckman Coulter, Inc., United States) and quantitated with Life Invitrogen Qubit 3.0 using Qubit dsDNA HS Assay Kit (Invitrogen, Thermo Fisher Scientific, United States). Eventually, quantified DNA was sequenced with 150-bp paired-end reads on Illumina Novaseq 6000 platform (Illumina, Inc., United States) according to the standard manual. The coverage contained the exon regions and adjacent intron regions (50 bp) of all human genes by SeqCap EZ Choice XL Library (Roche NimbleGen). The average sequencing depth of the target region for the proband was 123.02×, among which 96.49% of the target sequence had a sequencing depth of more than 20×. The father was 140.85×, of which 97.90% of the target sequence had a sequencing depth of more than 20×. The mother was 160.89×, of which 97.83% of the target sequence had a sequencing depth of more than 20×.

The raw data produced on Novaseq platform were filtered and aligned against the human reference genome (hg19) using Burrows–Wheeler Aligner (BWA)[1] after being evaluated with Illumina Sequence Control Software (SCS). The single-nucleotide polymorphisms (SNPs) were called by using GATK software (Genome Analysis ToolKit)[2]. Variants were annotated using ANNOVAR[3], and the effects of single-nucleotide variants (SNVs) were predicted by SIFT, Polyphen-2, and Mutation_Taster programs (Li J. et al., 2019; Li Z. et al., 2019). All variants were interpreted according to ACMG standards and then categorized to be pathogenic, likely pathogenic, variants of unknown clinical significance (VUS), likely benign, and benign.

Copy Number Variation Calls by Whole-Genome Sequencing

Here, 750 ng of genomic DNA was fragmented to an average size of 200–300 bp, and DNA libraries were constructed using KAPA Library Preparation Kit, with reagent employed for one of the libraries. The constructed DNA library samples were then taken for high-throughput sequencing with Illumina Nova Seq 6000. The sequencing depth is 0.6×, whole-genome low-depth sequencing. High-quality double-ended sequencing reads were aligned to the human reference genome sequence from the UCSC database using the BWA tool. The window width was preset at 50 Kb, with an adjustment amount of 5 Kb. A two-step calibration of guanine-cytosine (GC) and population model was performed across each of the windows. After removing the abnormal windows, the standard deviation between the copy ratio and the reference set of each window was calculated. A standard deviation of less than 0.15 determined by the software was considered to be in accordance with the quality control. The size and copy ratio of the final copy number variation (CNV) segments were calculated by identifying the break point. Afterward, identified and mapped CNVs were interrogated against publicly available databases, including Decipher, Database of Genomic Variants (DGV), 1,000 Genomes, and Online Mendelian Inheritance in Man (OMIM).

Chromosome Microarray Analysis

DNA from the patient, his father, and paternal grandmother was genotyped using InfiniumOmniZhongHua-8 array (Illumina, San Diego, CA, United States). In addition to a genome-wide functional resolution of approximately 20 kb for deletions and 50 kb for duplications, the array also had a higher density coverage of the 15q11-q13 region. The experiments were carried out under the manufacturer's instructions. Genotype calling, quality control, and identification of CNV were performed using Illumina KaryoStudio software and cnvPartition algorithm, with various databases employed for array data evaluation and genotype–phenotype correlation analysis, including OMIM[4], DECIPHER[5], DGV[6], and ISCA[7].

Methylation-Specific Multiplex Ligation-Dependent Probe Amplification

Methylation-Specific Multiplex Ligation-dependent Probe Amplification (MS-MLPA®) reagents and kits obtained from MRC-Holland (MRC, Amsterdam, Netherlands) were used to verify the methylation status of chromosomes 15, including ME028-B2 kit containing sequence-specific probes that was applied for testing along the length of the 15q11.2-q13 region for the patient, his father, and paternal grandmother. In the presence of methylation-sensitive restriction enzymes, the B2 kit equipped with 48 MLPA probes was employed for copy number detection and methylation status verification. Approximately 50 ng of genomic DNA was introduced for each of the MS-MLPA reactions according to manufacturer's instructions. The PCR products were analyzed by capillary electrophoresis on an ABI 3100 sequencer (Applied Biosystems, CA, United States) using GeneScan 500 LIZ dye Size Standard and formamide (Applied Biosystems, CA, United States), and GeneMarker version 2.64 (SoftGenetics, LLC) was used to determine the copy number and methylation status associated with the critical region of Prader–Willi syndrome/Angelman syndrome (PWS/AS) patients. The fluorescent signals from the copy number probes, by comparing with the normal controls, showed the ratios of

[1]http://bio-bwa.sourceforge.net/

[2]www.broadinstitute.org/gatk

[3]annovar.openbioinformatics.org/en/latest/

[4]http://www.ncbi.nlm.nih.gov/omim

[5]http://decipher.sanger.ac.uk/

[6]http://projects.tcag.ca/variation

[7]http://dbsearch.clinicalgenome.org/search/

0.5 for deletions and 1.5 for duplications. Since the methylation probes were maternally imprinted (maternal allele methylated), the ratio of methylated probes to normal controls would increase accordingly in the presence of additional copies from maternal alleles but not paternal alleles.

RESULTS

Karyotyping

All cases presented normal karyotypes (not presented).

Whole-Exome Sequencing and Copy Number Variation

Copy number variation analysis of the proband identified a 500-kb interstitial microdeletion of 15q11.2 with the first breakpoint located at 24,932,524 bp and the last breakpoint at 25,482,598 bp on the distal region. There are five genes reported in OMIM (*SNRPN*, *SNHG14*, *PWAR6*, *SNORD115*, and *SNORD116*) according to UCSC Genome Browser on Human February 2009 Assembly. We extended the genetic analysis to the family members of the patient (father, mother, and paternal grandmother). The same deletion was detected in the paternal grandmother and the father of the patient but not in the mother, supporting the paternal origin of the deletion (**Figure 2**).

Whole-exome sequencing (WES) did not show findings of variants with pathogenic significance in clinically relevant candidate genes causing PWS phenotype in this patient and other family members (**Supplementary Table 1**).

Since all of the other family members are in good health, the results suggested vertical transmission of the IC microdeletion from the unaffected paternal grandmother to the unaffected father, which consequently resulted in PWS in these two sibling grandchildren when the IC microdeletion was inherited paternally.

Methylation-Specific Multiplex Ligation-Dependent Probe Amplification

Analysis of the patient and other family members was performed using an MS-MLPA kit specific for 15q11-q13 genomic region. Copy number changes were detected in 15 probes in the patient, indicating complete deletion of the *SNRPN* gene. The same deletion was also verified in the father and the paternal grandmother, but negative in the mother, which was corresponding to the paternal origin of the deletion as well. Methylation patterns within this region of this family revealed that the paternal grandmother and the father displayed an abnormal hypomethylation pattern in the *SNRPN* region due to complete loss on the maternal chromosome 15, whereas the patient presented with the typical PWS hypermethylation pattern in four *SNRPN* probes as a result of complete loss on his paternal chromosome 15 (**Figure 3**). In addition, a normal methylation pattern was observed in the mother. The findings from MS-MLPA confirmed the deletion of *SNURF/SNRPN* exon 1 and further identified that this deletion was paternal in origin.

Chromosome Microarray Analysis

High-resolution microarray analysis of this patient and other family members confirmed the results from CNV analysis data that demonstrated an interstitial microdeletion of 15q11.2. Meanwhile, trio analysis of SNP loci on chromosome 15 of the patient and his parent was also in accordance with the paternal inheritance. Chromosome microarray analysis (CMA) mapping revealed a 398-kb region of the chromosome 15 microdeletion, with the proximal breakpoint at 24,966,348 bp and the distal one at 25,364,551 bp (**Figure 4**). No additional aberrations were detected. As is known, *SNRPN*, *SNURF*, *PWRN2*, *SNORD116*, and *OCA2* are considered pathogenetic in the OMIM database. This deletion overlapped with upstream exons of the *SNURF-SNRPN* gene, thus verifying the findings from the MS-MLPA test.

In view of the clinical manifestation, karyotype, WES, CNV, CMA, and MS-MLPA, a diagnosis of PWS resulting from ∼550-kb loss on his paternal chromosome 15, which the breakpoint refers to the CNV calling for WGS, can eventually be confirmed for this patient [the sequencing reads for variant calling and the data for CMA had been deposited with NODE Bioproject (OEP001280 and OEP001281)]. The microdeletion of PWS IC was transmitted silently through two generations prior to being expressed in the third generation *via* the female germline—the paternal grandmother, the father, and then the two affected sibling grandchildren. With the aid of these genetic tools, PGD was therefore applied successfully to exclude imprinting deficiency in preimplantation embryos before transfer into the mother's uterus (Reproductive Hospital of Shandong University). As expected, the mother has given birth to a healthy boy.

DISCUSSION

Prader–Willi syndrome is typically featured by hypotonia during the neonatal and infantile stage, accompanied by delayed neuropsychomotor development, hyperphagia, obesity, hypogonadism, short stature, small hands and feet, as well as mental disabilities and behavioral problems (Cassidy and Driscoll, 2009; Cassidy et al., 2012). However, it is not uncommon that its clinical phenotype may be confounded by disorders caused by other genetic variations, which were defined as PWS-like, and may present with similar manifestations (Rocha and Paiva, 2014; Cheon, 2016). Though PWS-like disorders share features of PWS phenotype, the genetic basis of these rare disorders differs. As is well known, PWS is usually triggered by a paternal deletion, maternal uniparental disomy, or imprinting defect of the chromosome region 15q11-q13, while the genetic etiologies of PWS-like disorders are more diverse, including variations in *MAGEL2* gene (Schaaf-Yang syndrome) (Schaaf et al., 2013) and *RAI1* gene (Smith–Magenis syndrome) (Alaimo et al., 2015), 1p36 monosomy (Tsuyusaki et al., 2010), 2pter deletion (Doco-Fenzy et al., 2014), deletion of 3p26.3 (Geuns et al., 2003), deletion of 6q (Bonaglia et al., 2008b; Izumi et al., 2013), 10q26 deletion (Lukusa and Fryns, 2000), 12q subtelomere deletions (Niyazov et al., 2007), chromosome 14 maternal uniparental disomy (Hosoki et al., 2009), paracentric inversion (X)(q26q28) (Florez et al., 2003),

FIGURE 2 | Copy number variation (CNV) sequencing reveals a *de novo* 500-kb heterozygous deletion of 15q11.2 (chr15: 24,932,524–25,482,598) in the paternal grandmother, the father, and the patient. The deletion encompasses five Online Mendelian Inheritance in Man (OMIM) genes including *SNRPN*, *SNHG14*, *PWAR6*, *SNORD115*, and *SNORD116* (highlighted in pink box).

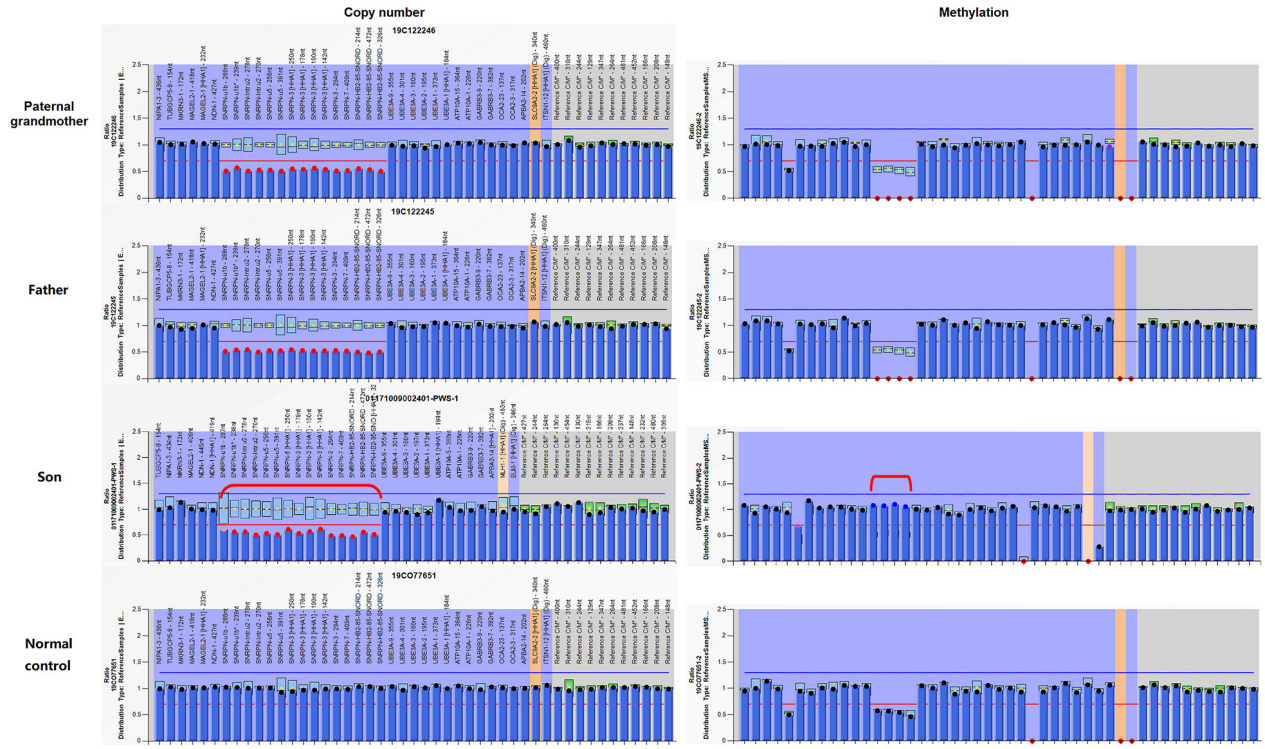

FIGURE 3 | Copy number and methylation patterns generated using Methylation-Specific Multiplex Ligation-dependent Probe Amplification B2 kit (MS-MLPA-B2). The Prader–Willi syndrome/Angelman syndrome (PWS/AS) kit contains 48 probes for copy number detection and methylation status analysis that are specific to regions in or near the PWS/AS critical region on chromosome 15q11-q13. The left and right columns display the results of copy number and methylation pattern, respectively. In the left column, copy number peak ratios are determined by comparing patient with normal control (2 copies/2 copies = 1.0). The figure reveals deletion (showing a copy number of 1) in 15 probes in and around the *SNRPN* region (highlighted in red dots) in the paternal grandmother, the father, and the patient. Meanwhile, the normal control has a normal copy number of 2 for all analyzed gene fragments of chromosome 15. The methylation probes were designed to hybridize to maternally imprinted loci, shown in the right column. When compared to the normal control, each of the four probes within the deleted region has a ratio of around 0.5. In this family, the paternal grandmother and the father display an abnormal methylation pattern (ratio = 0) in the *SNRPN* region due to complete loss on the maternal chromosome 15, while the patient displays the typical PWS methylation pattern (ratio = 1) in the four *SNRPN* probes due to complete loss on his paternal chromosome 15. *UBE3A* exon 1 and one other digestion control probe were used during the methylation analysis. These results indicate that in the paternal grandmother and the father, the paternal allele is present, and the deletion is maternal in origin, which explains their absence of clinical symptoms. In contrast, the child displays an abnormal methylation pattern in the four methylation-sensitive fragments digested in the *SNRPN* region due to loss on his paternal chromosome 15.

and duplication of X(q21.1-q21.31) (Pramyothin et al., 2010). Therefore, the variation of these genes in PWS-like disorders, though different from those of PWS, may be associated with phenotypes that are difficult to be differentiated from PWS due to the clinical overlap, especially atypical PWS, thus could consequently challenge the final diagnosis. The challenge for clinicians exists not only in accurate differentiation of the clinical manifestations between PWS and PWS-like disorders but also in the effort to provide conclusive genetic explanations for the phenotypes in order to offer uncompromised genetic counseling and treatment. Absence of correct diagnosis is highly likely to worsen the prognosis of the individuals due to the endocrine–metabolic malfunctioning associated with PWS. Given the phenotypic overlap between PWS and PWS-like disorders, tests aiming at other genetic variations should be considered in the context of PWS-like phenotypes but with negative results from PWS methylation analysis. Therefore, an appropriate and accurate genetic investigation strategy would be necessary and indispensable.

As a sporadic genetic disorder with remarkable developmental consequences, PWS is usually triggered by a paternal deletion, maternal uniparental disomy, or imprinting defect of the chromosome region 15q11-q13, which could be diagnosed using the standard methylation tests (MS-MLPA) with higher accuracy. The current genetic epidemiology indicates that approximately 65–75% PWS cases have a detectable deletion in this region, 20–30% cases are caused by mUPD, and imprinting errors have been observed in 1% (among 15% of the cases of either a sporadic or inherited microdeletion in the IC, there was a paternal chromosomal translocation in less than 1%) (Amos-Landgraf et al., 1999; Chai et al., 2003; Cassidy and Driscoll, 2009; Cassidy et al., 2012). Since the PWS patients with imprinting defect were rarely reported, no phenotypic feature is known to correlate exclusively with any of the three major molecular mechanisms that result in PWS. Previous research merely focused on the statistical differences in the frequency or severity of certain features between the two largest molecular classes (deletion and mUPD). To the best of our knowledge,

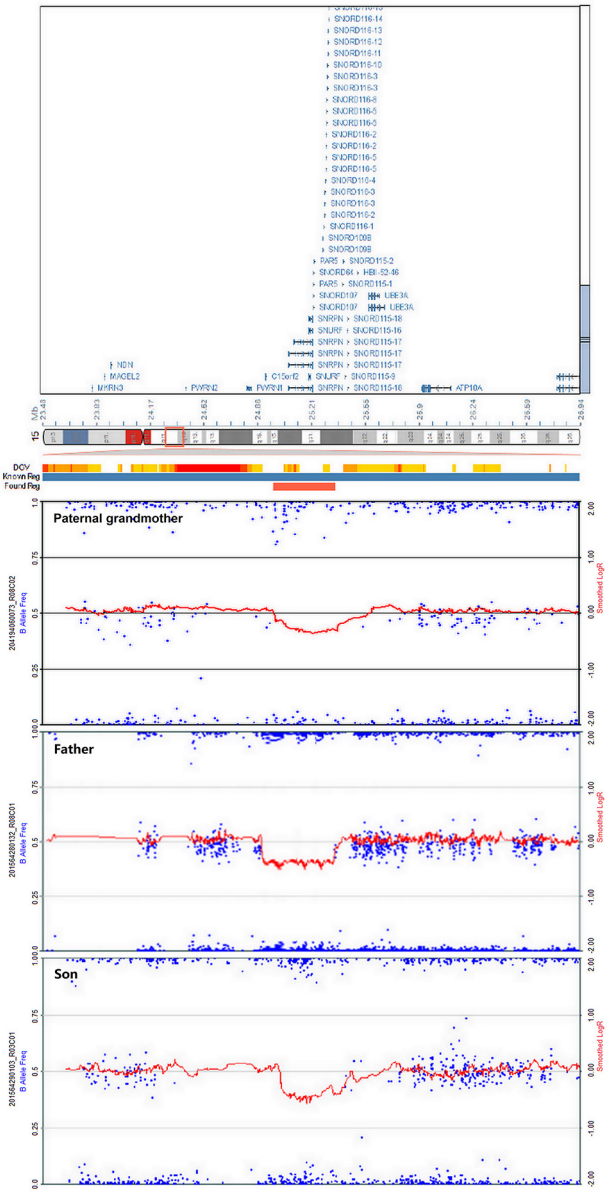

FIGURE 4 | Chromosome microarray analysis (CMA) demonstrating deletion of 15q11.2 in the paternal grandmother, the father, and the patient. The novel copy number variation (CNV) at the loci (chr15: 24,966,348–25,364,551) is of approximately 398 kb in heterozygous state, surrounding the *SNRPN* gene.

comparison among the three classes did reveal discrepancies in the phenotype, typically demonstrated as less features in IC PWS cases, including decreased fetal movement, typical facial phenotypes, excessive or rapid weight gain, hyperphagia, hypopigmentation, small hands/feet, and thick saliva (Hartin et al., 2018). Moreover, imprinting center deletions can be inherited, result in an increased risk of recurrence, and therefore it is important to diagnose them in a timely matter to enable preconception counseling or PGD in families carrying this type of genetic anomaly.

In our research, we described a rare familial occurrence and atypical clinical features of PWS with a 550-kb microdeletion at 15q11.2: a typical IC deletion spanning *SNRPN*, *SNHG14*, *PWAR6*, *SNORD115*, and *SNORD116* on the maternal chromosome 15 of the paternal grandmother and the father. The IC deletion caused atypical PWS-like phenotype in the proposita and his elder sister. The microdeletion was transmitted silently through the female germline (the paternal grandmother and the father) but impaired the erasure of the maternal imprint and/or the establishment of a paternal imprint in the male germline (the proposita). So, the proposita inherited a paternal chromosome with a missing paternal imprint, leading to the development of PWS. The proposita and his elder sister exhibited part of the major clinical manifestations of PWS, including neonatal and infantile hypotonia, weak cry, feeding difficulties, and hypoplastic external genitalia, followed by recurrent respiratory infections. Interestingly, additional features such as low birth weight, characteristic facial features (narrow forehead, almond-shaped eyes, thin upper lip, downturned corner of the mouth), and hypopigmentation (fair skin and hair) that are common clinical features of PWS were not seen in this patient. These findings, accompanied by early neonatal death, once misled us to the consideration of neonatal lethal monogenic diseases (such as fatty acid oxidation disorders, urea cycle disorder, and organic acid metabolism disease), mitochondrial diseases, or chromosomal diseases in the first place. Therefore, the familial history of neonatal death and atypical clinical features, though presented as rare events in PWS patients, could lead clinical judgment astray, thus should arouse vigilance among pediatricians.

The minimal critical region for PWS is proposed to be approximately 95 kb in size (at chr15:25280000-25375000, genome build hg 19) and contains two C/D box snoRNAs—the SNORD116 cluster and SNORD109A—as the only putative functional genes (**Figure 5**; Butler, 1990; Sahoo et al., 2008; de Smith et al., 2009; Duker et al., 2010; Bieth et al., 2015; Hassan and Butler, 2016; Tan et al., 2020). In review of the regions of deletion in the present and previously described cases that exhibit the key characteristics of the PWS phenotype, the SNORD116 cluster, SNORD109A, and the Imprinted in Prader–Willi (IPW) exons were found to be consistently deleted. In addition, by contrasting against PubMed, DECIPHER, and ClinVar database, the microdeletion in our patient was found to share a great similarity to the case reported in DECIPHER (patient number 288417). The deletion in patient 288417 of DECIPHER database, detected as 516 kb in size, encompassed five OMIM genes: *NPAP1*, *PWRN1*, *SNHG14*, *SNORD116*, and *SNRPN*. Also, patient 288417 showed a core phenotype characterized by obesity, aggressive behavior, intellectual disability, and psychosis (information of infant period was not provided), pointing toward a causative role of the genes in the minimal critical region in the broader phenotype of typical PWS. Furthermore, in animal models of PWS, knockout Snord116 mice displayed cognitive deficits (Adhikari et al., 2019), growth retardation (Ding et al., 2008), hyperphagia, and marked obesity (Polex-Wolf et al., 2018; Yang et al., 2019). So, all these findings indicate that PWS with microdeletion disrupting the IC should be considered in patients

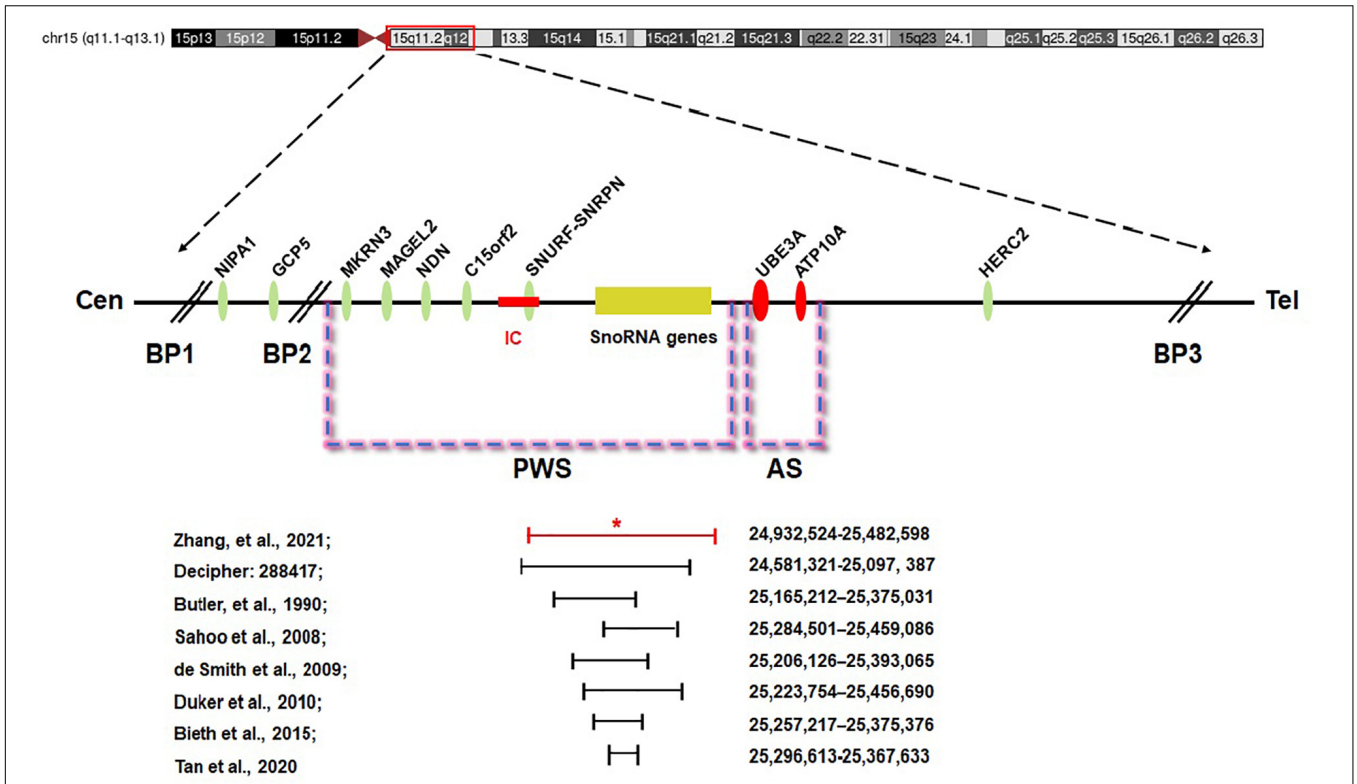

FIGURE 5 | Prader–Willi syndrome (PWS) and Angelman syndrome (AS) domain in proximal chromosome 15q11-q13. The positions of genes (oval) in PWS, AS, and the IC are illustrated. CEN, centromere; Tel, telomere; IC, PWS imprinting center (red rectangle); BP, breakpoint; *the breakpoint of the proband.

with hypotonia and developmental delay, even in the absence of the striking facial features. Furthermore, our research has provided further evidence that deletion of the SNORD116 region is sufficient to cause the key characteristics of PWS; therefore, suspicion of PWS should be aroused despite atypical physical features and rapid progression of the disorder. Meanwhile, the silent transmission of PWS IC microdeletion through the female germline has been recognized to be highly confounding for diagnostic testing and genetic counseling in affected patients and families. These results also suggest that other genes in the region may make specific phenotypic contributions, which necessitate further research and exploration to better understand the role of genes in the IC.

CONCLUSION

Overall, the present study finds that the symptoms of our patient are distinct from, and may be more severe than, those of typical PWS cases. The familial occurrence and atypical clinical features were challenging to our diagnostic strategy. Based on the results of our study, PGD was applied successfully to exclude imprinting deficiency in preimplantation embryos before transfer into the mother's uterus. As expected, the mother gave birth to a healthy boy. Our study may be especially instructive regarding accurate diagnosis, differential diagnosis, genetic counseling, treatment, and PGD for familial PWS patients.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the NODE BioProject repository, accession numbers are OEP001280 and OEP001281.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Qilu Children's Hospital of Shandong 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

YLiu, YY, and SZ supervised the project. KZ and SL wrote the manuscript. WG, KZ, YLiu, and YLv were involved in the clinical diagnosis and whole-exome sequencing and bioinformatics analysis. DW and MG performed the karyotype analysis and chromosome microarray analysis.

XL and ZG participated in case follow-up. All authors were involved in the conception, experiment design, and data analysis and approved the final manuscript.

FUNDING

This work was financially supported by the National Natural Science Foundation of China (Grant Numbers: 81971449 and 81671362), Shanghai Pujiang Program (Grant Number: 19PJ1402100), and Natural Science Training Foundation of Shandong Province (ZR2014HP051).

ACKNOWLEDGMENTS

The authors would like to thank the patients and their parents for their contribution to the project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.630650/full#supplementary-material

Supplementary Table 1 | The variants identified in WES for the family.

REFERENCES

Adhikari, A., Copping, N. A., Onaga, B., Pride, M. C., Coulson, R. L., Yang, M., et al. (2019). Cognitive deficits in the Snord116 deletion mouse model for Prader-Willi syndrome. *Neurobiol. Learn. Mem.* 165:106874. doi: 10.1016/j.nlm.2018.05.011

Alaimo, J. T., Barton, L. V., Mullegama, S. V., Wills, R. D., Foster, R. H., and Elsea, S. H. (2015). Individuals with Smith-Magenis syndrome display profound neurodevelopmental behavioral deficiencies and exhibit food-related behaviors equivalent to Prader-Willi syndrome. *Res. Dev. Disabil.* 47, 27–38. doi: 10.1016/j.ridd.2015.08.011

Amos-Landgraf, J. M., Ji, Y., Gottlieb, W., Depinet, T., Wandstrat, A. E., Cassidy, S. B., et al. (1999). Chromosome breakage in the Prader-Willi and Angelman syndromes involves recombination between large, transcribed repeats at proximal and distal breakpoints. *Am. J. Hum. Genet.* 65, 370–386. doi: 10.1086/302510

Berger, S. I., Ciccone, C., Simon, K. L., Malicdan, M. C., Vilboux, T., Billington, C., et al. (2017). Exome analysis of Smith-Magenis-like syndrome cohort identifies de novo likely pathogenic variants. *Hum. Genet.* 136, 409–420. doi: 10.1007/s00439-017-1767-x

Bieth, E., Eddiry, S., Gaston, V., Lorenzini, F., Buffet, A., Conte Auriol, F., et al. (2015). Highly restricted deletion of the SNORD116 region is implicated in Prader-Willi syndrome. *Eur. J. Hum. Genet.* 23, 252–255. doi: 10.1038/ejhg.2014.103

Bonaglia, M. C., Ciccone, R., Gimelli, G., Gimelli, S., Marelli, S., Verheij, J., et al. (2008a). Detailed phenotype-genotype study in five patients with chromosome 6q16 deletion: narrowing the critical region for Prader-Willi-like phenotype. *Eur. J. Hum. Genet.* 16, 1443–1449. doi: 10.1038/ejhg.2008.119

Bonaglia, M. C., Ciccone, R., Gimelli, G., Gimelli, S., Marelli, S., Verheij, J., et al. (2008b). Detailed phenotype-genotype study in five patients with chromosome 6q16 deletion: narrowing the critical region for Prader-Willi-like phenotype. *Eur. J. Hum. Genet.* 16, 1443–1449.

Buiting, K., Saitoh, S., Gross, S., Dittrich, B., Schwartz, S., Nicholls, R. D., et al. (1995). Inherited microdeletions in the Angelman and Prader-Willi syndromes define an imprinting centre on human chromosome 15. *Nat. Genet.* 9, 395–400. doi: 10.1038/ng0495-395

Butler, M. G. (1990). Prader-Willi syndrome - current understanding of cause and diagnosis. *Am. J. Med. Genet.* 35, 319–332. doi: 10.1002/ajmg.1320350306

Cassidy, S. B., and Driscoll, D. J. (2009). Prader-Willi syndrome. *Eur. J. Hum. Genet.* 17, 3–13. doi: 10.1038/ejhg.2008.165

Cassidy, S. B., Schwartz, S., Miller, J. L., and Driscoll, D. J. (2012). Prader-Willi syndrome. *Genet. Med.* 14, 10–26. doi: 10.1038/gim.0b013e31822bead0

Chai, J. H., Locke, D. P., Greally, J. M., Knoll, J. H., Ohta, T., Dunai, J., et al. (2003). Identification of four highly conserved genes between breakpoint hotspots BP1 and BP2 of the Prader-Willi/Angelman syndromes deletion region that have undergone evolutionary transposition mediated by flanking duplicons. *Am. J. Hum. Genet.* 73, 898–925. doi: 10.1086/378816

Cheon, C. K. (2016). Genetics of Prader-Willi syndrome and Prader-Will-Like syndrome. *Ann. Pediatr. Endocrinol. Metab.* 21, 126–135. doi: 10.6065/apem.2016.21.3.126

D'Angelo, C. S., Da Paz, J. A., Kim, C. A., Bertola, D. R., Castro, C. I. E., Varela, M. C., et al. (2006). Prader-Willi-like phenotype: investigation of 1p36 deletion in 41 patients with delayed psychomotor development hypotonia, obesity and/or hyperphagia, learning disabilities and behavioral problems. *Eur. J. Med. Genet.* 49, 451–460. doi: 10.1016/j.ejmg.2006.02.001

de Smith, A. J., Purmann, C., Walters, R. G., Ellis, R. J., Holder, S. E., Van Haelst, M. M., et al. (2009). A deletion of the HBII-85 class of small nucleolar RNAs (snoRNAs) is associated with hyperphagia, obesity and hypogonadism. *Hum. Mol. Genet.* 18, 3257–3265. doi: 10.1093/hmg/ddp263

Dello Russo, P., Demori, E., Sechi, A., Passon, N., Romagno, D., Gnan, C., et al. (2016). Microdeletion 15q26.2qter and Microduplication 18q23 in a Patient with Prader-Willi-Like Syndrome: Clinical Findings. *Cytogenet. Genome Res.* 148, 14–18. doi: 10.1159/000445923

Desch, L., Marle, N., Mosca-Boidron, A. L., Faivre, L., Eliade, M., Payet, M., et al. (2015). 6q16.3q23.3 duplication associated with Prader-Willi-like syndrome. *Mol. Cytogenet.* 8:42. doi: 10.1186/s13039-015-0151-6

Ding, F., Li, H. H., Zhang, S., Solomon, N. M., Camper, S. A., Cohen, P., et al. (2008). SnoRNA Snord116 (Pwcr1/MBII-85) deletion causes growth deficiency and hyperphagia in mice. *PLoS One* 3:e1709. doi: 10.1371/journal.pone.0001709

Doco-Fenzy, M., Leroy, C., Schneider, A., Petit, F., Delrue, M. A., Andrieux, J., et al. (2014). Early-onset obesity and paternal 2pter deletion encompassing the ACP1. TMEM18, and MYT1L genes. *Eur. J. Hum. Genet.* 22, 471–479. doi: 10.1038/ejhg.2013.189

Duker, A. L., Ballif, B. C., Bawle, E. V., Person, R. E., Mahadevan, S., Alliman, S., et al. (2010). Paternally inherited microdeletion at 15q11.2 confirms a significant role for the SNORD116 C/D box snoRNA cluster in Prader-Willi syndrome. *Eur. J. Hum. Genet.* 18, 1196–1201. doi: 10.1038/ejhg.2010.102

Florez, L., Anderson, M., and Lacassie, Y. (2003). De novo paracentric inversion (X)(q26q28) with features mimicking Prader-Willi syndrome. *Am. J. Med. Genet. A* 121A, 60–64. doi: 10.1002/ajmg.a.20129

Fountain, M. D., Aten, E., Cho, M. T., Juusola, J., Walkiewicz, M. A., Ray, J. W., et al. (2016). The phenotypic spectrum of Schaaf-Yang syndrome: 18 new affected individuals from 14 families (vol 18, pg 961, 2016). *Genet. Med.* 18, 1066–1066. doi: 10.1038/gim.2016.114

Geets, E., Zegers, D., Beckers, S., Verrijken, A., Massa, G., Van Hoorenbeeck, K., et al. (2016). Copy number variation (CNV) analysis and mutation analysis of the 6q14.1-6q16.3 genes SIM1 and MRAP2 in Prader Willi like patients. *Mol. Genet. Metab.* 117, 383–388. doi: 10.1016/j.ymgme.2016.01.003

Geuns, E., De Rycke, M., Van Steirteghem, A., and Liebaers, I. (2003). Methylation imprints of the imprint control region of the SNRPN-gene in human gametes and preimplantation embryos. *Hum. Mol. Genet.* 12, 2873–2879. doi: 10.1093/hmg/ddg315

Gillessen-Kaesbach, G., and Horsthemke, B. (1994). Clinical and molecular studies in fragile X patients with a Prader-Willi-like phenotype. *J. Med. Genet.* 31, 260–261. doi: 10.1136/jmg.31.3.260-b

Hartin, S. N., Hossain, W. A., Weisensel, N., and Butler, M. G. (2018). Three siblings with Prader-Willi syndrome caused by imprinting center microdeletions and review. *Am. J. Med. Genet. A* 176, 886–895. doi: 10.1002/ajmg.a.38627

Hassan, M., and Butler, M. G. (2016). Prader-Willi syndrome and atypical submicroscopic 15q11-q13 deletions with or without imprinting defects. *Eur. J. Med. Genet.* 59, 584–589. doi: 10.1016/j.ejmg.2016.09.017

Hosoki, K., Kagami, M., Tanaka, T., Kubota, M., Kurosawa, K., Kato, M., et al. (2009). Maternal uniparental disomy 14 syndrome demonstrates prader-willi syndrome-like phenotype. *J. Pediatr.* 155, 900–903.e1. doi: 10.1016/j.jpeds.2009.06.045

Izumi, K., Housam, R., Kapadia, C., Stallings, V. A., Medne, L., Shaikh, T. H., et al. (2013). Endocrine phenotype of 6q16.1-q21 deletion involving SIM1 and Prader-Willi syndrome-like features. *Am. J. Med. Genet. A* 161, 3137–3143. doi: 10.1002/ajmg.a.36149

Li, J., Lu, C., Wu, W., Liu, Y., Wang, R., Si, N., et al. (2019). Application of next-generation sequencing to screen for pathogenic mutations in 123 unrelated Chinese patients with Marfan syndrome or a related disease. *Sci. China Life Sci.* 62, 1630–1637. doi: 10.1007/s11427-018-9491-8

Li, Z., Zhu, P., Huang, H., Pan, Y., Han, P., Cui, H., et al. (2019). Identification of a novel COL4A5 mutation in the proband initially diagnosed as IgAN from a Chinese family with X-linked Alport syndrome. *Sci. China Life Sci.* 62, 1572–1579. doi: 10.1007/s11427-018-9545-3

Linhares, N. D., Valadares, E. R., da Costa, S. S., Arantes, R. R., de Oliveira, L. R., Rosenberg, C., et al. (2016). Inherited Xq13.2-q21.31 duplication in a boy with recurrent seizures and pubertal gynecomastia: clinical, chromosomal and aCGH characterization. *Meta Gene* 9, 185–190. doi: 10.1016/j.mgene.2016.07.004

Lukusa, T., and Fryns, J. P. (2000). Pure distal monosomy 10q26 in a patient displaying clinical features of Prader-Willi syndrome during infancy and distinct behavioural phenotype in adolescence. *Genet. Counsel.* 11, 119–126.

McCarthy, J. M., McCann-Crosby, B. M., Rech, M. E., Yin, J. N., Chen, C. A., Ali, M. A., et al. (2018). Hormonal, metabolic and skeletal phenotype of Schaaf-Yang syndrome: a comparison to Prader-Willi syndrome. *J. Med. Genet.* 55, 307–315. doi: 10.1136/jmedgenet-2017-105024

Negishi, Y., Ieda, D., Hori, I., Nozaki, Y., Yamagata, T., Komaki, H., et al. (2019). Schaaf-Yang syndrome shows a Prader-Willi syndrome-like phenotype during infancy. *Orphanet J. Rare Dis.* 14:277. doi: 10.1186/s13023-019-1249-4

Niyazov, D. M., Nawaz, Z., Justice, A. N., Toriello, H. V., Martin, C. L., and Adam, M. P. (2007). Genotype/phenotype correlations in two patients with 12q subtelomere deletions. *Am. J. Med. Genet. A* 143A, 2700–2705. doi: 10.1002/ajmg.a.32005

Ohta, T., Gray, T. A., Rogan, P. K., Buiting, K., Gabriel, J. M., Saitoh, S., et al. (1999). Imprinting-mutation mechanisms in Prader-Willi syndrome. *Am. J. Hum. Genet.* 64, 397–413. doi: 10.1086/302233

Polex-Wolf, J., Lam, B. Y., Larder, R., Tadross, J., Rimmington, D., Bosch, F., et al. (2018). Hypothalamic loss of Snord116 recapitulates the hyperphagia of Prader-Willi syndrome. *J. Clin. Invest.* 128, 960–969. doi: 10.1172/JCI97007

Pramyothin, P., Pithukpakorn, M., and Arakaki, R. F. (2010). A 47, XXY patient and Xq21.31 duplication with features of Prader-Willi syndrome: results of array-based comparative genomic hybridization. *Endocrine* 37, 379–382. doi: 10.1007/s12020-010-9330-8

Rocha, C. F., and Paiva, C. L. (2014). Prader-Willi-like phenotypes: a systematic review of their chromosomal abnormalities. *Genet. Mol. Res.* 13, 2290–2298. doi: 10.4238/2014.March.31.9

Sahoo, T., del Gaudio, D., German, J. R., Shinawi, M., Peters, S. U., Person, R. E., et al. (2008). Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. *Nat. Genet.* 40, 719–721. doi: 10.1038/ng.158

Schaaf, C. P., Gonzalez-Garay, M. L., Xia, F., Potocki, L., Gripp, K. W., Zhang, B. L., et al. (2013). Truncating mutations of MAGEL2 cause Prader-Willi phenotypes and autism. *Nat. Genet.* 45, 1405–1408. doi: 10.1038/ng.2776

Stalker, H. J., Keller, K. L., Gray, B. A., and Zori, R. T. (2003). Concurrence of fragile X syndrome and 47, XYY in an individual with a Prader-Willi-Like phenotype. *Am. J. Med. Genet. A* 116A, 176–178. doi: 10.1002/ajmg.a.10001

Tan, Q., Potter, K. J., Burnett, L. C., Orsso, C. E., Inman, M., Ryman, D. C., et al. (2020). Prader-Willi-Like Phenotype Caused by an Atypical 15q11.2 Microdeletion. *Genes* 11:128. doi: 10.3390/genes11020128

Tsuyusaki, Y., Yoshihashi, H., Furuya, N., Adachi, M., Osaka, H., Yamamoto, K., et al. (2010). 1p36 deletion syndrome associated with Prader-Willi-like phenotype. *Pediatr. Int.* 52, 547–550. doi: 10.1111/j.1442-200X.2010.03090.x

Yang, S., Mei, H., Mei, H., Yang, Y., Li, N., Tan, Y., et al. (2019). Risks of maternal prepregnancy overweight/obesity, excessive gestational weight gain, and bottle-feeding in infancy rapid weight gain: evidence from a cohort study in China. *Sci. China Life Sci.* 62, 1580–1589. doi: 10.1007/s11427-018-9831-5

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.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 06 October 2021
doi: 10.3389/fphar.2021.669516

Clinical Attributes and Electroencephalogram Analysis of Patients With Varying Alpers' Syndrome Genotypes

Hua Li[1†], Wei Wang[2,3], Xiaodi Han[1], Yujia Zhang[1], Lifang Dai[1], Manting Xu[1], Jie Deng[1], Changhong Ding[1], Xiaohui Wang[1], Chunhong Chen[1], Xiaofeng Yang[3†] and Fang Fang[1*†]*

[1]Department of Neurology, Beijing Children's Hospital, Capital Medical University, National Center For Children's Health, Beijing, China, [2]Laboratory of Brain Disorders, Ministry of Science and Technology, Collaborative Innovation Center for Brain Disorders, Beijing Institute of Brain Disorders, Capital Medical University, Beijing, China, [3]Bioland Laboratory Guangzhou Regenerative Medicine and Health Guangdong Laboratory, Guangzhou, China

OPEN ACCESS

Edited by:
Tieliu Shi,
East China Normal University, China

Reviewed by:
Bo Liu,
Sichuan University, China
Sun Dan,
Huazhong University of Science and Technology, China

***Correspondence:**
Xiaofeng Yang
xiaofengyang@yahoo.com
Fang Fang
fangfang@bch.com.cn

†ORCID:
Hua Li
orcid.org/0000-0001-7334-534X
Xiaofeng Yang
orcid.org/0000-0002-7847-0665
Fang Fang
orcid.org/0000-0001-6362-7896

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: 19 February 2021
Accepted: 30 August 2021
Published: 06 October 2021

Citation:
Li H, Wang W, Han X, Zhang Y, Dai L, Xu M, Deng J, Ding C, Wang X, Chen C, Yang X and Fang F (2021) Clinical Attributes and Electroencephalogram Analysis of Patients With Varying Alpers' Syndrome Genotypes. Front. Pharmacol. 12:669516. doi: 10.3389/fphar.2021.669516

Alpers' syndrome is an early inceptive neurodegenerative disorder with a poor prognosis, characterized by developmental regression, intractable epilepsy, and hepatic dysfunction. Candidate genes, such as *POLG, PARS2, CARS2, FARS2, NARS2, and GABRB2* are distinguished and registered following research on large cohorts that portray the clinical phenotype in such patients using expanded access to whole-exome sequencing (WES). In this study, we aimed to better understand the electroencephalogram (EEG) characteristics and clinical phenotype of different genotypes of the Alpers' syndrome, which are currently insufficiently studied. We conducted a study on seven patients with Alpers' syndrome who received treatment in Beijing Children's Hospital and had a detailed clinical EEG. Furthermore, a substantial literature search of the Chinese Biomedical Literature Database, PubMed, and Cochrane Central Register of Controlled Trials EMBASE was also conducted, which revealed a total of 22 reported cases between January 2008 to January 2021. We analyzed 29 cases of Alpers' syndrome caused by different gene variants, of which 22 cases were related to *POLG* gene mutation and 7 cases were related to *PARS2, CARS2, FARS2, NARS2, and GABRB2* gene mutation, and found that patients with distinctive pathogenic variants exhibited comparable phenotypes and similar EEG patterns. And we defined EEG characteristics found specifically in Alpers' syndrome. Rhythmic high-amplitude delta with superimposed (poly) spikes (RHADS) is a characteristic EEG finding in the early stages of Alpers' syndrome and is a kind of epileptic phenomenon, which can provide clues for the early diagnosis of the disease.

Keywords: Alpers, syndrome, POLG, electroencephalogram, RHADS, diagnosis

Abbreviations: Bg, background; BS, burst suppression pattern; D, damaging; DC, Disease causing; DFP, disease-associated polymorphism with supporting functional evidence; DM, disease-causing mutation; EEG, Electroencephalogram; EPC, epilepsia partialis continua; FED, focal epileptiform discharges; FS, focal seizure; MFCS, Multifocal clonic seizures; MFED, multifocal epileptiform discharges; MS, myoclonic seizure; GTCS, generalized tonic-clonic seizures; MLH, migraine-like headache; GSW, generalised spike-wave or poly spike-wave discharges; Hyps, hypsarrhythmia; GS, generalized seizure; MFMS, multifocal myoclonus seizure; TS, tonic seizure; IS, infantile spasms; LP, likely pathogenic; P, pathogenic; NA, not available; PD, probably damaging; RHADS, rhythmic high-amplitude delta with superimposed (poly) spikes; SE, status epilepticus; VUS, uncertain significance.

INTRODUCTION

Alpers' syndrome is an early inceptive neurodegenerative disorder with a poor prognosis, characterized by the triad of developmental regression, intractable epilepsy, and hepatic dysfunction (Sofou et al., 2012; Tarka et al., 2020). The patients with Alpers' syndrome are generally less than 4 years old after the initial presentation of symptoms (Hannah and Yasir, 2021). It usually occurs in infancy or early childhood, and rarely in juveniles. In the early stages of the disease, epileptic seizures, as an early symptom, are the most common. The epileptic seizures develop in the form of focal, multifocal or myoclonic seizures evolving into epilepsia partialis continua (EPC) or myoclonic status epilepticus (MSE). The sudden outbreak of therapeutically refractory seizures in infancy or adolescence is considered an average clinical indicator of Alpers' syndrome. The electroencephalograms (EEG) finds have shown a specific EEG phenomenon, known as rhythmic high-amplitude delta with superimposed (poly) spikes (RHADS) (Wolf et al., 2009; van Westrhenen et al., 2018). The epileptic seizures may be have an evolution of EEG in the development of Alpers' syndrome (Hannah and Yasir, 2021). EEG might provide in the potential diagnostic clue (Wolf et al., 2009; van Westrhenen et al., 2018).

Recent studies have identified molecular genetic causes of Alpers' syndrome, including pathogenic mutations in the gene encoding the catalytic subunit of polymerase gamma (*POLG*), *PARS2* which encodes prolyl-tRNA synthetase, *FARS2* which encodes phenylalanyl-tRNA synthetase, and *NARS2* which encodes asparaginyl-tRNA synthetase, etc, *POLG* gene variation is the most common (Sofou et al., 2015; Walker et al., 2016; Samanta et al., 2018; Sofou et al., 2021).

Alpers' syndrome is one of the most serious phenotypes of mitochondrial disease caused by *POLG* gene mutation. More than 90% of Alpers' syndrome cases happen due to autosomal recessive mutations in the nuclear-encoded catalytic subunit (Bicknese et al., 1992; Hance et al., 2005; Chrysostomou et al., 2016). The results of *POLG* flaws are mtDNA depletion/accumulation of mtDNA deletions. The depletion of mtDNA causes consequent cellular dysfunction, mitochondrial respiratory chain impairments, and apoptosis in the brain and liver (Anagnostou et al., 2016; Naviaux RK et al., 1999; Tzoulis C et al., 2014).

With the development of genetic detection and analysis techniques, the pathogenic genes of Alpers' syndrome have been discovered. The candidate genes such as *POLG, PARS2, CARS2, FARS2, NARS2,* and *GABRB2* have been distinguished and registered using large cohorts that portray the clinical phenotype in such patients using expanded access to WES. However, at present, no systematic studies have been performed on the clinical phenotypes and the EEG findings that occur in patients with different genetic pathogenic variation mutations. In this study, we aimed to characterize the EEG features and clinical phenotypes of patients with different genotypes of the Alpers' syndrome and assist in future studies.

METHODS

Patients

This study was conducted on seven patients with Alpers' syndrome who underwent a thorough medical treatment from May 2013 to December 2020 in the Beijing Children's Hospital. They were examined for detailed clinical, EEG, and pathogenic information.

In addition to these, a substantial literature search was carried out in EMBASE, Cochrane Central Register of Controlled Trials, and PubMed. Prior to an extensive literature search, an eligibility criterion was taken into consideration. It included EEG case studies based on patient's medical history with Alpers' syndrome. Publications in any language other than English, anthropoidal studies, limited resources of literature review, and unauthenticated abstracts were considered as criteria for exclusion, and comprehensive analysis was reported. After applying the inclusion and exclusion criteria, we shortlisted 22 cases that were reported in the various databases from January 2008 to January 2021. Thus, we examined a total of 29 patients whose details may be found in **Supplementary Table S1**.

The clinical features that were evaluated for these 29 patients with Alpers' syndrome were as follows: sex of the patient, age at disease onset, age at seizure onset, categories of seizures, seizure types at disease onset, development, EEG, pathogenic information, and other clinical features. Seizures and epilepsy syndromes were categorized as per the rules and regulations of the International League Against Epilepsy (ILAE) (Belousova et al., 2017).

Genetic Analysis

In the present study, all samples were captured by whole-exome reagent hybridization and using Illumina instruments for next-generation sequencing. Sequence data aligned with the human genome reference (hg19) and variants-calling were carried out with NextGene V2.3.4 software (Softgenetics LLC., State College, PA, United States). Variants were screened as follows:1) preference to the disease-related variants, small insertions and deletions (INDEL), canonical splice sites, and nonsense variants; 2) minor allele frequency (MAF) in normal populations <5% (except for known MAF ≥5% pathogenicity); 3) preference to variants in the Human Gene Mutation Database (HGMD), ClinVar; and 4) preference to variants in the Online Mendelian Inheritance in Man database. Pathogenic variants were defined according to the standards and guidelines for the interpretation of sequence variants published by the American College of Medical Genetics in 2015 with Human Genome Variation Society nomenclature (Richards et al., 2015). We performed Sanger sequencing of all samples to validate the identified *POLG* mutations and test the parental origin of available family members. Of these, we also used quantitative PCR technology to verify the deletion of large fragments of the genome in ample No. 6.

EEG Recording and Analysis

Video-EEG was documented and digitized at the sample rate of 1000 Hz using the standard international 10–20 system (Natus Medical Incorporated, Pleasanton, CA, United States). Impedances were kept below 10 KΩ. A low-cut filter at 1 Hz was used before digital sampling. At least a prolonged continuous EEG monitoring (more than 24-h-long video-EEG recording). All EEG data were interpreted and reviewed by qualified

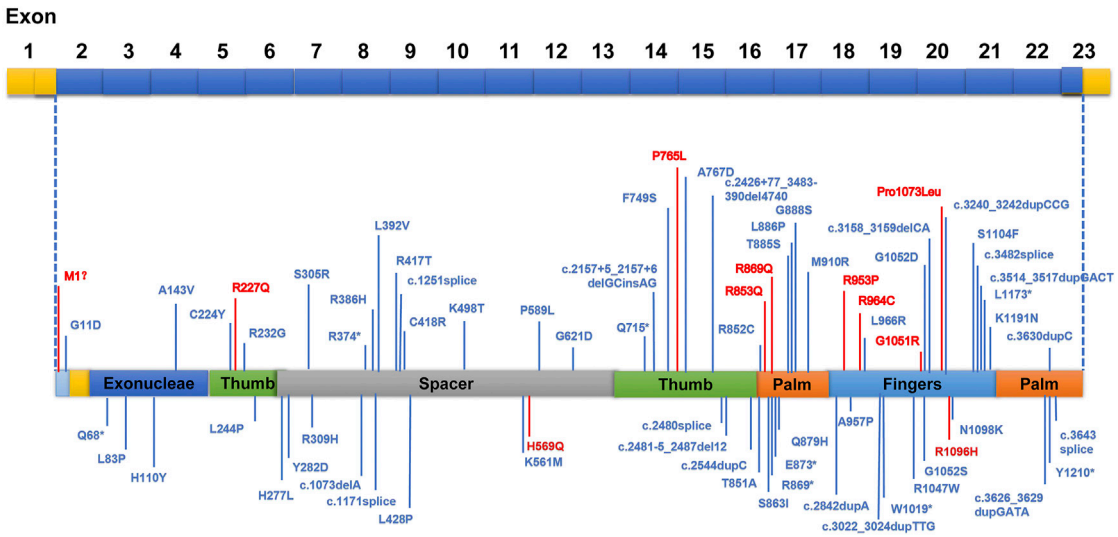

FIGURE 1 | Areas of distinguished variants in *POLG* mutation in Alpers' syndrome. Mutations presented in red were identified in this current study, whereas those in blue have been reported by past investigations.

neurophysiologists adopting the following criteria for RHADS: 1) occipital strength; 2) a moderate movement (<1 Hz) of high amplitude (200–1,000 uV); 3) frequent occurrence; and 4) a set of superimposed polyspikes. EEG data that showed results as 'inexplicit' were excluded for additional analysis.

In EEGs with RHADS, gamma oscillations (30–80 Hz) were marked visually via time-frequency analyses for EEG. First, the EEG data were reviewed visually in bipolar 10–20-Montage, and stage-II NREM sleep data (the delta band was estimated higher than 25% of all delta bands in a 30-s epoch by visual inspection) was selected (Bagshaw et al., 2009; Jacobs et al., 2016). Then, after band-pass filtering the raw data (30–80 Hz), time-frequency analysis techniques were used in each segment of data. Finally, the gamma oscillations were identified and compared with the original EEG data. The EEG data were visually examined by two EEG professional specialists, and channels with significant artifacts were excluded from the analysis.

Statistical Analysis

A total of 29 data samples were collected in this study, including 22 cases in the *POLG* group and 7 cases in the NON-*POLG* group. The differences of various indicators in the two groups were compared. For continuous variables that did not conform to the normal distribution, the median description was adopted, and a non-parametric test was used to compare the differences. For the comparison of classification variables, the composition ratio description was adopted, and the differences were compared by Fisher exact test was used to compare proportions of patients with microcephaly, Infantile spasm, Hepatic dysfunction, and EEG feature. For survival analysis, the endpoint was time to death which was defined as the time in months from the date of disease onset to the date of death. The survival rate of the two groups was compared by Kaplan-Meier survival curve analysis. The software selected for this analysis was SPSS26 and Graphpad7.0, and the significance level was 0.05.

Ethics Statement

The present study has been approved by the Ethics Committee of Beijing Children's Hospital and written informed consent was obtained from the guardians/families of the patients for their participation in the study.

RESULT

Mutation Outcomes

Our study revealed a total of 12 different variations within the *POLG* gene in seven patients (patient 1 to 5, previously reported (Dai et al., 2019) (**Figure 1**). 10 mutations were reported and two mutations were novel (**Table 1**, patient 6: ex.2-23del and patient 7: p. H569N). Patient 6 was compound heterozygous for the c.3218C>T mutation and a novel variation, a deletion in exon 2–23 (chr15:89859622–89876990). Quantitative PCR technology was used to verified the deletion, the proband and his mother were heterozygous deletions.

We identified pathogenic/likely pathogenic/VUS variants in seven patients in our cohort. One patient (patient5) in our study shared the same homozygous variant c.3218C>T (p.P1073L), which was also reported in another two Chinese studies (Qian et al., 2015). Three different recurrent variants (p.W748S,p.E1143G, and p. A467T) were reported (Uusimaa et al., 2008; Wolf et al., 2009). Two unrelated patients shared the same recurrent homozygous variant c.3218C>T (p.P1073L), this recommends a foundation effect in the Chinese population. The other five patients harbored compound heterozygous variants, including 2 variants (Exon 2–23 del and c.1705C>A,p.H569N) newly reported in this study (**Table 1**).

TABLE 1 | Summary of the Genetic mutation variants of 29 cases with Alpers' syndrome.

Patient	Ref	Gene	Nucleotide	Protein	Parental origin	ACMG-guideline	ACMG/ hgmd CLASS	Mutation Taster	Polyphen-2_HVAR	SIFT
1	Present case	*POLG* (NM_002693.2)	c.2858G > C	p.R953P	paternal	PM1 + PM5 + PM2_Supporting + PP3	LP	DC	PD	D
			c.2294C > T	p.P765L	maternal	PM1 + PM5 + PM2_Supporting + PP3	LP	DC	PD	D
2	Present case	*POLG* (NM_002693.2)	c.3151G > A	p.G1051R	maternal	PM1 + PM5 + PM2_Supporting + PP3	LP	DC	PD	D
			c.2606G > A	p.R869Q	paternal	PM3_Strong + PM1 + PM2_Supporting + PP3	VUS	DC	PD	D
3	Present case	*POLG* (NM_002693.2)	c.680G > A	p.R227Q	maternal	PM3 + PM5_Strong + PM2_Supporting + PP3	LP	DC	PD	D
			c.2T > G	p.M1?	paternal	PVS1_Supporting + PM3 + PM2_Supporting	VUS	DC	PD	D
4	Present case	*POLG* (NM_002693.2)	c.2558G > A	p.R853Q	*de novo**	PM1 + PM3 + PM5 + PM2_Supporting + PP3	P	DC	PD	D
			c.2890C > T	p.R964C	maternal	PM1 + PM3_Strong + PM2_Supporting + PS3_Spporting + PP1 + PP3 + PP4	VUS	DC	PD	D
5	Present case	*POLG* (NM_002693.2)	c.3218C > T	p.P1073L	paternal	PM3_Strong + PS3_Moderate + PM2_Supporting + PP3 + PP4	P	DC	PD	D
			c.3218C > T	p.P1073L	maternal	PM3_Strong + PS3_Moderate + PM2_Supporting + PP3 + PP4	P	DC	PD	D
6	Present case	*POLG* (NM_002693.2)	c.3218C > T	p.P1073L	paternal	PM3_Strong + PS3_Moderate + PM2_Supporting + PP3 + PP4	P	DC	PD	D
			ex.2-23del	p.?	maternal	PVS1 + PM3 + PM2_Supporting	LP	NA	NA	NA
7	Present case	*POLG* (NM_002693.2)	c.1705C > A	p.H569N	maternal	PM3 + PM5 + PM2_Supporting + PP3	VUS	DC	PD	D
			c.3287G > A	p.R1096H	paternal	PM5_Strong + PM3 + PM2_Supporting + PP3 + PS3_Supporting	LP	DC	PD	D
8	Wiltshire et al. (2008)	*POLG* (NM_002693.3)	c.2551A > G	p.T851A	NA	DM	DM	NA	NA	NA
			c.3139C > T	p.R1047W	NA	DM	DM	NA	NA	NA
9	Wolf et al. (2009)	*POLG* (NM_002693.3)	c.2243G > C	p.W748S	NA	DM	DM	NA	NA	NA
			c.2542G > A	p.G848S	NA	DM	DM	NA	NA	NA
10	Wolf et al. (2009)	*POLG* (NM_002693.3)	c.2243G > C	p.W748S	NA	DM	DM	NA	NA	NA
			c.2542G > A	p.G848S	NA	DM	DM	NA	NA	NA
11	Wolf et al. (2009)	*POLG* (NM_002693.3)	c.1399G > A	p.A467T	NA	DM	DM	NA	NA	NA
12	Wolf et al. (2009)	*POLG* (NM_002693.3)	c.2243G > C	p.W748S	NA	DM	DM	NA	NA	NA
			novel splice mutation	novel splice mutation	NA	NA	NA	NA	NA	NA
13	Wolf et al. (2009)	*POLG* (NM_002693.3)	c.1399G > A	p.A467T	NA	DM	DM	NA	NA	NA
14	Uusimaa et al. (2008)	*POLG* (NM_002693.3)	c.2243G > C	p.W748S	NA	DM	DM	NA	NA	NA
			c.3428A > G	p.E1143G	NA	DFP	DFP	NA	NA	NA
15	Uusimaa et al. (2008)	*POLG* (NM_002693.3)	c.2243G > C	p.W748S	NA	DM	DM	NA	NA	NA
			c.3428A > G	p.E1143G	NA	DFP	DFP	NA	NA	NA
16	Uusimaa et al. (2008)	*POLG* (NM_002693.3)	c.2243G > C	p.W748S	NA	NA	DM	NA	NA	NA
			c.3428A > G	p.E1143G	NA	DFP	DFP	NA	NA	NA
17	Mccoy et al. (2011)	*POLG* (NM_002693.3)	c.1681G > A	p.A467T	NA	NA	NA	NA	NA	NA
			c.2897T > G	p.L966R	NA	DM	NA	NA	NA	NA
18	Mccoy et al. (2011)	*POLG*	c.2824G > A	p.A467T	NA	NA	NA	NA	NA	NA
		POLG (NM_002693.3)	c.2542G > A	p.G848S	NA	DM	NA	NA	NA	NA

(Continued on following page)

TABLE 1 | (*Continued*) Summary of the Genetic mutation variants of 29 cases with Alpers' syndrome.

Patient	Ref	Gene	Nucleotide	Protein	Parental origin	ACMG-guideline	ACMG/ hgmd CLASS	Mutation Taster	Polyphen-2_HVAR	SIFT
19	Mccoy et al. (2011)	*POLG*	c.2836C > T	p.A467T	NA	NA	NA	NA	NA	NA
		POLG (NM_002693.3)	c.2554C > T	p.R852C	NA	DM	NA	NA	NA	NA
20	Mccoy et al. (2011)	*POLG* (NM_002693.3)	c.1399G > A	p.A467T	NA	DM	NA	NA	NA	NA
			c.1252T > C	p.C418R	NA	DM	NA	NA	NA	NA
21	Allen et al. (2014)	*POLG* (NM_002693.3)	c.2836C > T	p.A467T	NA	NA	NA	NA	NA	NA
			c.2740A > C	p.T914P	NA	DM	NA	NA	NA	NA
22	London et al. (2017)	*POLG* (NM_002693.3)	c.2836C > T	p.A467T	NA	NA	NA	NA	NA	NA
			c.2243G > C	p.W748S	NA	DM	NA	NA	NA	NA
23	sofou et al. (2015)	NARS2 (NM_024678.6)	c.641C > T	p.P214L	NA	NA	NA	NA	NA	NA
24	sofou et al. (2015)	PARS2 (NM_152,268.4)	c.1130dupC	p.K328fs*1	NA	DM	NA	NA	NA	NA
					NA	NA	NA	NA	NA	NA
25	Samanta et al. (2018)	CARS2 (NM_024537.4)	c.155T > G	p.V52G	NA	DM	NA	NA	NA	NA
			c.563C > T	p.T188M	NA	DM	NA	NA	NA	NA
26	Nishikawa et al. (2020)	GABRB2 (NM_021911.3)	c.784G > T	p.V262P	NA	DM	NA	NA	NA	NA
27	sofou et al. (2021)	NARS2 (NM_024678.6)	c.641C > T	p.P214L	NA	DM	NA	NA	NA	NA
28	sofou et al. (2021)	NARS2 (NM_024678.6)	c.641C > T	p.P214L	NA	DM	NA	NA	NA	NA
29	Walker et al. (2016)	FARS2 (NM_006567.5)	c.253C > G	p.P85A	NA	DM	NA	NA	NA	NA
			c.403C > G	p.H135D	NA	DM	NA	NA	NA	NA

Abbreviations: LP, likely pathogenic; P, pathogenic; VUS, uncertain significance; DC, disease causing; NA, not available; PD, probably damaging; D, damaging; DM, disease-causing mutation; DFP, disease-associated polymorphism with supporting functional evidence.

**we not verified that weather patient'father is chimerism or not biological father due to ethical and sample acquisition reasons.*

Genetic (the classification of the variants according to the American College of Medical Genetics and Genomics (ACMG), parental origin, etc) details of these patients are shown in **Table 1**. Among the 29 Alpers' syndrome patients, 79.31% (23/29) carried at least 1 missense variant. Only 1 truncation variant in *POLG* was identified (Wolf et al., 2009). Variants in *PARS2*, *CARS2*, *GABRB2*, *FARS2,* and *NARS2* were also reported in Alpers' syndrome patients (Sofou et al., 2015; Walker et al., 2016; Samanta et al., 2018; Sofou et al., 2021). All the variations were consistent with the autosomal recessive inheritance pattern, except for that in *GABRB2.* Aiko et al. reported a *de novo* missense variant in *GABRB2* (c.784G>T (p.Val262Phe)). A recurrent missense variant, c.641C>T (p.P214L), was identified in *NARS2* and 1 truncation variant was identified in *PARS2* (Sofou et al., 2015). All together 5 recurrent missense variants in 2 genes were identified in Alpers' syndrome patients, and only 2 truncation variants were observed in 2 genes.

Clinical Characteristics of Patients With Pathogenic *POLG* Variations

Our data showed that patients with pathogenic *POLG* variations showed the characteristics of mitochondrial encephalopathy, with refractory epilepsy, progressive liver function impairment, as the main clinical symptoms, acute liver failure occured after valproic acid application. For the analysis of *POLG*-related Alpers' syndrome, 22 out of 29 patients were included (8 females, 14 males, The median age at onset was 2.33 years, the range of age-onset was from 2 months to 17 years). The clinical characteristics of *POLG* mutation-related Alpers' syndrome were analyzed as follows.

Seizures

All 22 patients (100%) with *POLG* mutation-related Alpers' syndrome experienced epileptic seizures. A variety of refractory epileptic seizures were exhibited in 21 patients (95.5%). Seizures were the most common initial symptom (63.6%), and began at a median age of 2.67 years (ranging from 3 months to 17 years). The main type of seizures were focal/myoclonic seizures that lead to focal motor status or epileptic status. A total of 14 (63.6%) patients showed the symptoms of focal seizures, 12 (54.5%) patients had focal motor status or epileptic status, whereas myoclonic seizures occurred in 3 of 22 patients (13.6%) and generalized tonic-clonic seizures occurred in 2 of 22 patients (9.1%). Only 1 (4.5%) patient presented with the symptoms of generalized seizure. Eight of the 22 patients presented with the symptoms of epilepsia partialis continua (EPC; 36.4%) (**Supplementary Table S1**).

Development

A notable developmental delay with the signs of *POLG* mutation was found in all patients who suffered from Alpers' syndrome in life. A total of 6 out of 22 patients (27.3%) showed the symptoms

of developmental delay that were evident before the onset of seizures. Nine patients exhibited normal development until the beginning of intractable seizures. One patient who was mentioned in the literature review was having the symptoms of severe-to-profound intellectual disability.

Other Features

A series of features such as visual symptoms and visual disturbance, ataxia (including congenital ataxia), hypotonia, vomiting, language delayed/lost speech, and different levels of liver failure were found in all patients. Four patients had severe liver failure after valproic acid application and one patient showed multiorgan failure.

Clinical Characteristics of Patients Without Pathogenic *POLG* Variants

In addition, Our data showed that patients with pathogenic NON-*POLG* (*NARS2, FARS2, PARS2, CARS2*, and *GABRB2*) variations showed the characteristics of age at onset was earlier, mostly in the early infantile period, with prominent infantile spasms, less liver damage, and microcephaly, most of the survival time after treatment is longer. The characteristics are significantly different from *POLG*-related Alpers' syndrome. For the analysis of NON-*POLG* mutation-related Alpers' syndrome, 7 out of 29 patients were included (2 females, 5 males, The age at onset was the early infantile period). The clinical characteristics of NON-*POLG* mutation-related Alpers' syndrome were analyzed as follows.

Seizures

All 7 patients (100%) with NON-*POLG* mutation-related Alpers' syndrome experienced epileptic seizures. A variety of refractory epileptic seizures were exhibited in 7 patients (100%). Epileptic seizures were initial symptom in 3 patients. In terms of seizure type, two (43.1%) patients showed the symptoms of focal seizures, two (28.6%) exhibited focal motor status or epileptic status, whereas myoclonic seizures occurred in three of the seven patients (42.9%), generalized tonic-clonic seizures in three of the seven patients (42.9%), myoclonic seizures in 3 of 7 patients (42.9%). Only one (14.3%) patient presented with the symptoms of infantile spasm, one (14.3%) with the symptoms of atypical absence seizure, one (14.3%) with the symptoms of a tonic seizure, and three with the symptoms of EPC.

Development

A notable developmental delay with signs of NON-*POLG* mutation was found in six of the seven patients who suffered from Alpers' syndrome in life. Five out of seven patients (71.4%) showed symptoms of developmental delay that were evident before the seizure outbreak. One patient in the literature review showed symptoms of severe intellectual disability. Only one patient exhibited normal development until the beginning of intractable seizures.

Other Features

The other series of features such as hypotonia (85.7%, 6/7), visual symptoms (42.9%, 3/7), feeding difficulties, vomiting (42.9%, 3/7), progressive microcephaly (42.9%,3/7), language delayed/lost speech (14.3%, 1/7), mild liver failure (28.6%, 2/7), and autism spectrum (14.3%, 1/7) were observed in the patients in this category. A detailed description of the clinical features of each patient is provided in **Supplementary Table S1**.

EEG Results

Abnormal EEGs with epileptiform discharges and slow activity were detected at the early stage in all patients. A total of 23 out of 29 (79.3%) patients showed a reducing nonspecific background in the EEG. Of the 29 patients, focal epileptiform discharges on EEG were found in 16 patients (55.2%), multifocal epileptiform discharges were found in 2 (6.9%), and hypsarrhythmia occurred only in 1 (3.4%). Focal slowing and epileptiform discharges were seen relatively equally in the early and late stages of the disease. The incidence of RHADS was described in 18 out of 29 patients (62.1%) (**Figure 2**). RHADS were the most common EEG patterns and were based on focal abnormality along with epileptic discharges, and were described in four patients in our series with the gamma oscillations (**Figure 3**). RHADS and slowing were the most common factors identified in the occipital region (57.1%), followed by the impairment in the frontal lobe (28.6%) and temporal (14.3%) regions. In the late stage of the disease, six cases of RHADS on EEG performed focal changes (focal slowing, epileptiform discharges), multifocal epileptiform discharges, and burst suppression pattern.

A Comparative Relation Between Genetic and Clinical Phenotypes of Patients With Alpers' Syndrome

As demonstrated in **Table 2**, Age at onset, Age at seizure onset, Time of death showed a significant relationship with the *POLG* or NON-*POLG* group,$p < 0.05$; As demonstrated in **Table 2**, Age at onset, Age at seizure onset, Time of death showed a significant relationship with the *POLG or* NON-*POLG* group,$p < 0.05$; microcephaly, Infantile spasm, Hepaticdys function, EEG feature showed a significant relationship in *POLG* or NON-*POLG* group, $p < 0.05$ (**Table 3**); As can be seen from the above table, there was a significant difference in survival rate between the NON-*POLG* group and the *POLG* group ($p < 0.05$). From the above table, there are significant differences in survival between NON-*POLG* group and *POLG*, $p < 0.05$. Among them, the median survival of Group *POLG* was 1.75 years and NON-*POLG* 14.083, and the survival rate of Group NON-*POLG* was signnificantly higher than that of Group *POLG*, with the following KM survival curve (**Figure 4**).

DISCUSSION

Alpers' syndrome is a rare mitochondrial encephalopathy, which is thought to be caused by autosomal recessive mutations in the nuclear-encoded catalytic subunit of *POLG* (Hannah and Yasir, 2021). Pathogenic variants in *POLG* are common causes of multiple mtDNA deletions and/or mtDNA depletion. The brain and liver are susceptible to the disease because of their

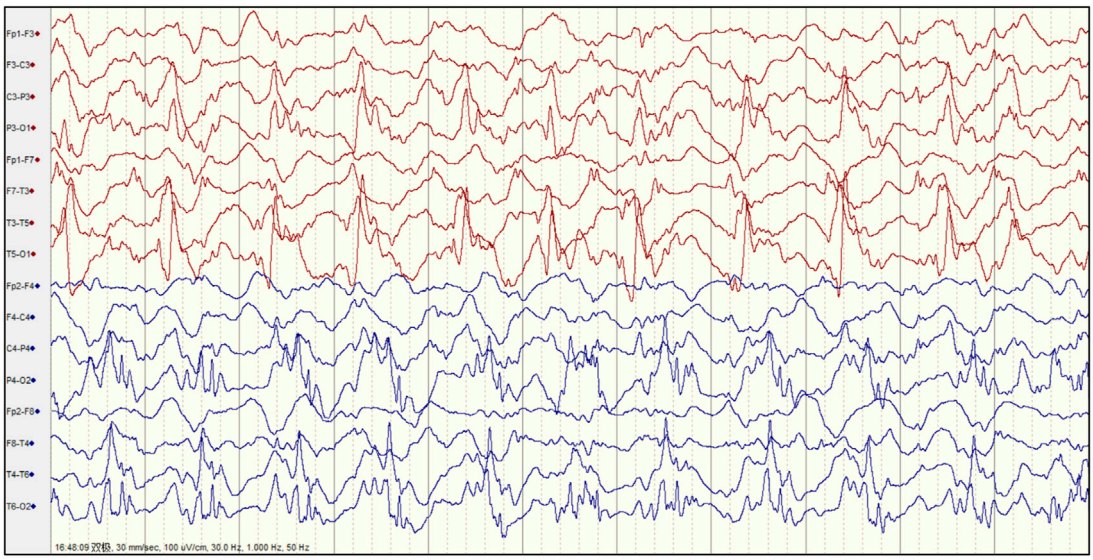

FIGURE 2 | The electroencephalogram (EEG) of patient 4 with Alpers' syndrome in our study showed typical rhythmic high amplitude delta waves with superimposed (poly) spikes (RHADS) in the bilateral posterior head region.

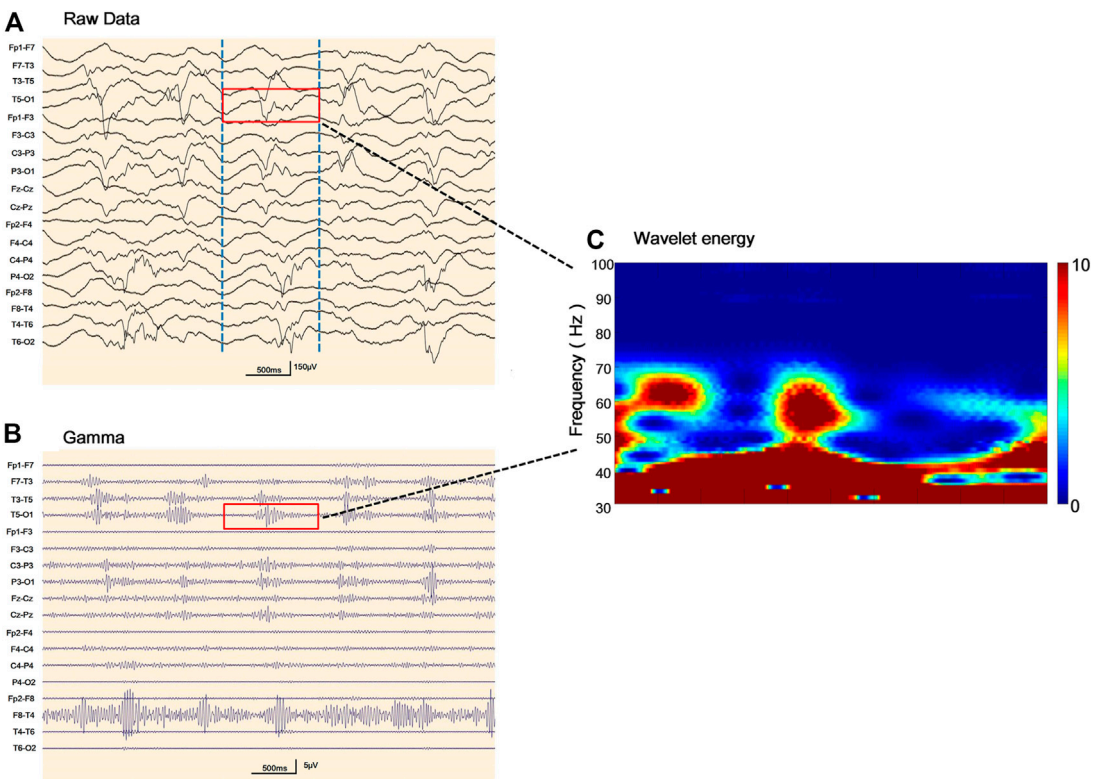

FIGURE 3 | Representative electroencephalogram (EEG) follow-ups and corresponding gamma oscillations **(A)** Raw EEG with the pattern of rhythmic high amplitude delta waves with superimposed (poly) spikes (RHADS) in sleep during stage II non-rapid eye movement period, the red layout marks 1 s data **(B)** EEG corresponding to **Figure 3A** after 30–80 Hz filtering, the red line addresses the gamma oscillations visually identified **(C)** Verification of gamma oscillations. Utilizing the wavelet transform method, the time-frequency spectra exhibit spectral blobs at around 30–80 Hz in the region T5-O1 in association with the corresponding spikes.

TABLE 2 | Comparison of mutations and clinical features of 29 cases with Alpers' syndrome by nonparametric test.

	Mutation involved		Z	P
	POLG	NON-*POLG*		
Age at onset	2.33 (0.65~9.6)	0.21 (0.08~0.92)	−2.271	0.023
Age at seizure onset	2.67 (0.83~9.6)	0.5 (0.33~8)	−1.963	0.049
Age at death	2.67 (1~11.5)	13.33 (4.08~15.5)	−1.487	0.137
Time of death	0.83 (0.17~1.75)	5.92 (1.15~15)	−2.053	0.04

TABLE 3 | Comparison of mutations and clinical features of 29 cases with Alpers, syndrome by Fisher exact test.

		Mutation involved		p
		POLG	NON-POLG	
sex	M	8 (36.4%)	5 (71.4%)	0.192
	F	14 (63.6%)	2 (28.6%)	
microcephaly	no	22 (100%)	4 (57.1%)	0.01
	yes	0 (0%)	3 (42.9%)	
Infantile spasm	no	22 (100%)	4 (57.1%)	0.01
	yes	0 (0%)	3 (42.9%)	
Hepatic dysfunction	no	6 (27.3%)	6 (85.7%)	0.011
	yes	16 (72.7%)	1 (14.3%)	
EEG feature	no	6 (27.3%)	6 (85.7%)	0.011
	yes	16 (72.7%)	1 (14.3%)	

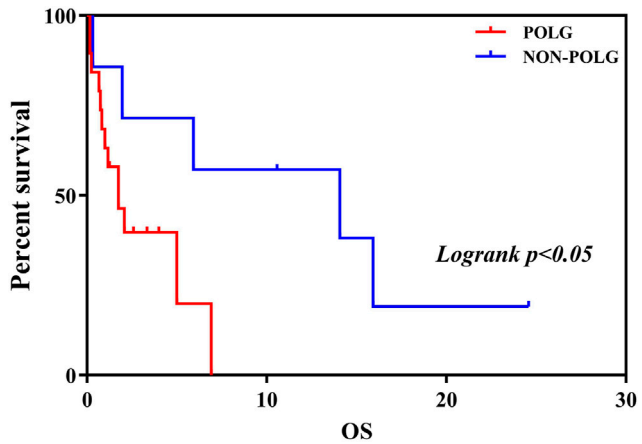

FIGURE 4 | Kaplan-Meier curve comparing survival of patients with the *POLG* or non-*POLG* group and patients with *POLG* Group showed carried significantly worse survival.

high demand for energy and proportionate need for mitochondria. The loss of mitochondria in these organ systems can lead to various symptoms, the most common being seizures and liver failure. Mitochondrial diseases affect cellular energy production through oxidative phosphorylation (OxPHOS) (Dimauro et al., 2013). The assembly and function of mitochondrial OXPHOS are dependent on a great number of proteins, mainly encoded by nuclear DNA and by mitochondrial DNA. Genetic defects in some of these proteins lead to various phenotypes with different ages of onset and course of the disease.

Next-generation sequencing has significantly broadened our understanding of the genetic causes of mitochondrial disease (Jin et al., 2019; Li et al., 2019; Shen and Shi, 2019). In recent years, an increasing number of genes have been related to Alpers' syndrome, many pathogenic mutations in the mitochondrial aminoacyl-tRNA synthetases (mt-aaRSs), such as *FARS2*, *CARS2*, *NARS2*, and *PARS2*, etc, are a novel cause of mitochondrial translation disorder, leading to Alpers' syndrome (Elo et al., 2012; Sofou et al., 2015; Walker et al., 2016; Samanta et al., 2018; Sofou et al., 2021). This study demonstrated the pathogenic mechanism of gene variants resulting in the loss of receptor function (Nishikawa et al., 2020). The patients in this study exhibited different genotypes of Alpers' syndrome, which suggests different genetic variations resulted in significant changes in mitochondrial function. In 2020, WES revealed a *de novo* missense variant of *GABRB2* (NM_021911.2: c.784G>T, p[Val262Phe]) in a child presenting with daily MSE and RHADS, that was associated with Alpers' syndrome phenotype. However, the *GABRB2* gene encodes γ-GABA receptors. At present, the relationship between γ-GABA receptors and mitochondrial function remains unclear, and further research will be of importance.

The relationship between clinical phenotypes and genotypes has been studied thoroughly based on the collected results of genetic mutation variants. The clinical phenotypes of Alpers' syndrome are variable, even in those with identical mutations, and this includes different degrees of neurological symptoms and hepatic dysfunction (Wu et al., 2018). The Alpers' syndrome caused by *POLG* gene mutations usually affects children aged less than 4 years (Hannah and Yasir, 2021), and inevitably leads to death. Refractory epilepsy is a major phenotypic feature leading to the poor prognostic range and rapidly progressive disease. In the early stages of the disease, epileptic seizures are the most common. The epileptic seizures develop in the form of focal, multifocal or myoclonic seizures evolving into EPC or MSE. In this study, all the patients had exhibited epileptic seizures, most had a focal motor seizure and epileptic status (**Supplementary Table S1**). In this study, most seizures captured by EEG monitoring originated from the occipital lobe, which appears to be the main site for seizure genesis in Alpers' syndrome. Numerous clinical and neuropathological studies revealed that the brain lobe is the key spot of the main area of neuronal degeneration and late post-stroke seizures in patients (Montine et al., 1995; Engelsen et al., 2008; Khan et al., 2012; Anagnostou et al., 2016; Wu et al., 2018). Recently, a complete study suggested interneuron and Purkinje cell pathology may be the mechanisms underlying seizure generation and maintenance (Hayhurst et al.,

2019). Development is another common neurological feature, this can occur at any time in childhood and may reflect the nature of the process (**Supplementary Table S1**). In the course of the disease, the individuals with Alpers' syndrome show different degrees of developmental delay, these findings in this study were consistent with those in previous studies (**Supplementary Table S1**). Liver involvement is the main characteristic of Alpers' syndrome, which is rarely the presenting symptom and can rapidly progress to end-stage liver failure. Sodium valproate accelerates liver dysfunction and invariably ends in a fatality (Park et al., 2017). In the current study, four patients presented with progressive liver failure (**Supplementary Table S1**). In addition, sometimes viral prodrome symptoms, headaches, and vomiting can be observed, which may give rise to clinical suspicion of encephalitis (Rahman and Copeland, 2019), expanding the phenotypic spectrum of Alpers' syndrome (Lim and Thomas, 2020). In recent years, with the advancement of genetic testing technology, Alpers syndrome caused by NON-*POLG* genes has also been continuously reported, including mitochondrial aminoacyl transfer RNA synthetase related genes (*PARS2, CARS2, FARS2 and NARS2*) and γ-GABA receptor related genes Gene (*GABRB2*). This study retrospectively compared the Alpers' syndrome caused by NON-*POLG* and *POLG* gene mutations, and found that NON-*POLG* onset was earlier, mostly in the early infantile period, with prominent infantile spasms, less liver damage, and some microcephaly, most of the survival time after treatment is longer (**Table 2**; **Figure 4**), the phenotype is significantly different from the Alpers' syndrome caused by *POLG* gene mutation (Sofou et al., 2015), so It may be better for these patients with mutations in NON-*POLG* to be diagnosed as Alpers-like syndrome. Differences in clinical phenotypes between the two groups may be related to the genes themselves, this aspect requires further research for clarification (Bao et al., 2019).

Our study systematically describes the EEG findings of various genotypes of Alpers' syndrome. In this study, background slowing was more commonly seen in Alpers' syndrome. Eighteen patients exhibited RHADS in the early stage of the disease, although six cases of RHADS on EEG performed nonspecific abnormalities during the late stage of the disease. RHADS may be associated with the staging phase (Boyd et al., 1986; Wolf et al., 2009; van Westrhenen et al., 2018). Our study is in agreement with the previous research that RHADS are highly distinctive in patients with Alpers' syndrome. RHADS were all in the early stage of the disease with the association of epileptic status and no traces of the influence of drug treatment (Boyd et al., 1986; Wolf et al., 2009; van Westrhenen et al., 2018). In addition, a notable finding was that the patient suffered from RHADS with gamma oscillations (**Figure 3**). This spatial overlap suggests between the brain areas generating spikes and fast oscillations involve similar "hyperactive" neuronal networks. In this study, RHADS were based on focal abnormality along with epileptic discharges, however, the gamma oscillations were described in four patients in our series, this finding demonstrated that spikes are more sensitive. Yet, it is important to note that gamma activity more specific and accurate to determine the regions of the brain participating in seizure generation, and the occurrence of the fast oscillations directly reflects the degree of epileptic activity of the EEG (Melani et al., 2013). This finding supported the idea that RHADS is a kind of epileptic phenomenon (Zijlmans et al., 2012). We also found that patients with Alpers' syndrome may have a high frequency of epileptic status. All patients experienced the same episodes of seizures and an indistinguishable EEG pattern during the early phase of the syndrome. EEG is an important tool for the examination of the diagnosis and prognosis of Alpers' syndrome because the EEG patterns are consistent with fluctuations in the electrical properties of the brain. The EEG correlations at different stages of the disease not only help to recognize the clinical course of the disease, but also provide a platform for better utilization of EEG in the diagnosis of clinical disease development. The explicit EEG pattern remains the foundation for the diagnosis and may help to comprehend the mechanism of Alpers' syndrome. However, it's worth noting that Alpers' syndrome caused by different genotypes showed similar EEG patterns in our study, which may be due to the involvement of a common electrophysiological mechanism. Therefore, future studies may be focused on applying digital video electroencephalography with high sampling rates and series of EEG patterns and the development of an automatic gamma oscillation analysis system.

To sum up, in this article, we the first integrated analysis of the EEG characteristics and clinical phenotype of patients with Alpers' syndrome of various genotypes. We found that the patients having different pathogenic variants exhibited comparable phenotypes. There are certain differences in the clinical phenotype of Alpers' syndrome caused by *POLG* gene mutations and NON-*POLG* gene mutations. and the differences in clinical phenotypes may be related to the genes themselves. So we believe that the clinical phenotype of NON-*POLG* gene mutations may be more appropriate to define Alpers-like syndrome. Due to our limited sample size, in the future, we will continue to collect samples and expand the sample size for further research. In addition, we defined EEG characteristics found specifically in Alpers' syndrome. RHADS is a characteristic EEG finding in the early stage of Alpers' syndrome, which is a kind of epileptic phenomenon, can provide clues for the early diagnosis of the disease. EEG characteristics of Alpers' syndrome with various genotypes show similar EEG patterns, the reason may be related to the common electrophysiological mechanism. We only have a few cases and have yet to draw a definitive conclusion in terms of the relationship between pathogenic genes and Alpers syndrome. This is a summative retrospective study, but it will provide useful information for a deeper understanding of the pathophysiology of Alpers' syndrome.

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 in the article/**Supplementary Material**.

ETHICS STATEMENT

The present study has been approved by the Ethics Committee of Beijing Children's Hospital and written informed consent was obtained from the guardians/families of the patients for their participation in the study.

AUTHOR CONTRIBUTIONS

All authors contributed to the study design, critically reviewed the manuscript, and approved the final version. HL performed literature search and analysis and wrote the manuscript. WW, XH, YZ, LD, MX, JD, CD, XW, and CC performed literature search and analysis. FF and XY revised the manuscript.

FUNDING

This work was supported by the China Association Against Epilepsy Scientific Research Project (CX-B -2021-09), Beijing Children's Hospital for Children Medication Foundation (YZZD202001), and the Capital Health Development Research Foundation Project (2018-2-2096).

ACKNOWLEDGMENTS

We thank the patients and their families who participated in this study. We also thank Berry Genomics Co. and Beijing Fulgent Co., for the technical support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.669516/full#supplementary-material

REFERENCES

Allen, N. M., Winter, T., Shahwan, A., and King, M. D. (2014). Explosive Onset Non-epileptic Jerks and Profound Hypotonia in an Infant with Alpers-Huttenlocher Syndrome. *Seizure* 23 (3), 237–239. doi:10.1016/j.seizure.2013.07.005

Anagnostou, M. E., Ng, Y. S., Taylor, R. W., and Mcfarland, R. (2016). Epilepsy Due to Mutations in the Mitochondrial Polymerase Gamma (*POLG*) Gene: A Clinical and Molecular Genetic Review. *Epilepsia* 57 (10), 1531–1545. doi:10.1111/epi.13508

Bagshaw, A. P., Jacobs, J., Levan, P., Dubeau, F., and Gotman, J. (2009). Effect of Sleep Stage on Interictal High-Frequency Oscillations Recorded from Depth Macroelectrodes in Patients with Focal Epilepsy. *Epilepsia* 50 (4), 617–628. doi:10.1111/j.1528-1167.2008.01784.x

Bao, Y., Suo, L., Qian, P., Huang, H., Yang, Y., Tang, J., et al. (2019). Clinical and Genetic Analysis of Dent Disease with Nephrotic Range Albuminuria in Shaanxi, China. *Sci. China Life Sci.* 62 (12), 1590–1593. doi:10.1007/s11427-018-9829-0

Belousova, E. D., Zavadenko, N. N., Kholin, A. A., and Sharkov, A. A. (2017). Psychiatry of the Future: an Overview of Foreign Scientists Opinions of the Position of Psychiatry in the Modern World. *Z. Nevrol. Psikhiatr. Im. S.S. Korsakova* 117 (7), 99. doi:10.17116/jnevro20171177199-106

Bicknese, A. R., May, W., Hickey, W. F., and Dodson, W. E. (1992). Early Childhood Hepatocerebral Degeneration Misdiagnosed as Valproate Hepatotoxicity. *Ann. Neurol.* 32 (6), 767–775. doi:10.1002/ana.410320610

Boyd, S. G., Harden, A., Egger, J., and Pampiglione, G. (1986). Progressive Neuronal Degeneration of Childhood with Liver Disease ("Alpers' Disease"): Characteristic Neurophysiological Features. *Neuropediatrics* 17 (02), 75–80. doi:10.1055/s-2008-1052505

Chrysostomou, A., Grady, J. P., Laude, A., Taylor, R. W., Turnbull, D. M., and Lax, N. Z. (2016). Investigating Complex I Deficiency in Purkinje Cells and Synapses in Patients with Mitochondrial Disease. *Neuropathol. Appl. Neurobiol.* 42 (5), 477–492. doi:10.1111/nan.12282

Dai, L. F., Fang, F., Liu, Z. M., Shen, D. M., Ding, C. H., Li, J. W., et al. (2019). [Phenotype and Genotype of Twelve Chinese Children with Mitochondrial DNA Depletion Syndromes]. *Zhonghua Er Ke Za Zhi* 57 (3), 211–216. doi:10.3760/cma.j.issn.0578-1310.2019.03.011

Dimauro, S., Schon, E. A., Carelli, V., and Hirano, M. (2013). The Clinical Maze of Mitochondrial Neurology. *Nat. Rev. Neurol.* 9 (8), 429–444. doi:10.1038/nrneurol.2013.126

Elo, J. M., Yadavalli, S. S., Euro, L., Isohanni, P., Götz, A., Carroll, C. J., et al. (2012). Mitochondrial Phenylalanyl-tRNA Synthetase Mutations Underlie Fatal Infantile Alpers Encephalopathy. *Hum. Mol. Genet.* 21 (20), 4521–4529. doi:10.1093/hmg/dds294

Engelsen, B. A., Tzoulis, C., Karlsen, B., Lillebø, A., Laegreid, L. M., Aasly, J., et al. (2008). POLG1 Mutations Cause a Syndromic Epilepsy with Occipital Lobe Predilection. *Brain* 131 (3), 818–828. doi:10.1093/brain/awn007

Hance, N., Ekstrand, M. I., and Trifunovic, A. (2005). Mitochondrial DNA Polymerase Gamma Is Essential for Mammalian Embryogenesis. *Hum. Mol. Genet.* 14 (13), 1775–1783. doi:10.1093/hmg/ddi184

Hannah, R. R., and Yasir, A. K. (2021). Alpers-Huttenlochen Syndrome (AHS, Alper Disease). *StatPearls* 2021, 01.

Hayhurst, H., Anagnostou, M. E., Bogle, H. J., Grady, J. P., Taylor, R. W., Bindoff, L. A., et al. (2019). Dissecting the Neuronal Vulnerability Underpinning Alpers' Syndrome: a Clinical and Neuropathological Study. *Brain Pathol.* 29 (1), 97–113. doi:10.1111/bpa.12640

Jacobs, J., Vogt, C., LeVan, P., Zelmann, R., Gotman, J., and Kobayashi, K. (2016). The Identification of Distinct High-Frequency Oscillations during Spikes Delineates the Seizure Onset Zone Better Than High-Frequency Spectral Power Changes. *Clin. Neurophysiol.* 127 (1), 129–142. doi:10.1016/j.clinph.2015.04.053

Jin, Y., Chen, G., Xiao, W., Hong, H., Xu, J., Guo, Y., et al. (2019). Sequencing XMET Genes to Promote Genotype-Guided Risk Assessment and Precision Medicine. *Sci. China Life Sci.* 62 (7), 895–904. doi:10.1007/s11427-018-9479-5

Khan, A., Trevenen, C., Wei, X. C., Sarnat, H. B., Payne, E., and Kirton, A. (2012). Alpers Syndrome: the Natural History of a Case Highlighting Neuroimaging, Neuropathology, and Fat Metabolism. *J. Child. Neurol.* 27 (5), 636–640. doi:10.1177/0883073811423973

Li, Z., Zhu, P., Huang, H., Pan, Y., Han, P., Cui, H., et al. (2019). Identification of a Novel COL4A5 Mutation in the Proband Initially Diagnosed as IgAN from a Chinese Family with X-Linked Alport Syndrome. *Sci. China Life Sci.* 62 (12), 1572–1579. doi:10.1007/s11427-018-9545-3

Lim, A., and Thomas, R. H. (2020). The Mitochondrial Epilepsies. *Eur. J. Paediatr. Neurol.* 24, 47–52. doi:10.1016/j.ejpn.2019.12.021

London, F., Hadhoum, N., Outteryck, O., Vermersch, P., and Zéphir, H. (2017). Late-onset of Alpers-Huttenlocher Syndrome: an Unusual Cause of Refractory Epilepsy and Liver Failure. *Acta Neurol. Belg.* 117 (1), 399–401. doi:10.1007/s13760-016-0672-8

Mccoy, B., Owens, C., Howley, R., Ryan, S., King, M., Farrell, M. A., et al. (2011). Partial Status Epilepticus - Rapid Genetic Diagnosis of Alpers' Disease. *Eur. J. Paediatr. Neurol.* 15 (6), 558–562. doi:10.1016/j.ejpn.2011.05.012

Melani, F., Zelmann, R., Dubeau, F., and Gotman, J. (2013). Occurrence of Scalp-Fast Oscillations Among Patients with Different Spiking Rate and Their Role as Epileptogenicity Marker. *Epilepsy Res.* 106 (3), 345–356. doi:10.1016/j.eplepsyres.2013.06.003

Montine, T. J., Powers, J. M., Vogel, F. S., and Radtke, R. A. (1995). Alpers' Syndrome Presenting with Seizures and Multiple Stroke-like Episodes in a 17-Year-Old Male. *Clin. Neuropathol.* 14, 322–326.

Naviaux, R. K., Nyhan, W. L., Barshop, B. A., Poulton, J., Markusic, D., Karpinski, N. C., et al. (1999). Mitochondrial DNA Polymerase Gamma Deficiency and mtDNA Depletion in a Child with Alpers' Syndrome. *Ann. Neurol.* 45 (1), 54–58. doi:10.1002/1531-8249(199901)45:1<54:aid-art10>3.0.co;2-b

Nishikawa, A., Otani, Y., Ito, S., Nagata, S., Shiota, M., Takanashi, J. I., et al. (2020). A De Novo GABRB2 Variant Associated with Myoclonic Status Epilepticus and Rhythmic High-Amplitude delta with Superimposed (Poly) Spikes (RHADS). *Epileptic Disord.* 22 (4), 476–481. doi:10.1684/epd.2020.1183

Park, S., Kang, H. C., Lee, J. S., Park, Y. N., Kim, S., and Koh, H. (2017). Alpers-Huttenlocher Syndrome First Presented with Hepatic Failure: Can Liver Transplantation Be Considered as Treatment Option? *Pediatr. Gastroenterol. Hepatol. Nutr.* 20 (4), 259–262. doi:10.5223/pghn.2017.20.4.259

Qian, Y., Ziehr, J. L., and Johnson, K. A. (2015). Alpers Disease Mutations in Human DNA Polymerase Gamma Cause Catalytic Defects in Mitochondrial DNA Replication by Distinct Mechanisms. *Front. Genet.* 6, 135. doi:10.3389/fgene.2015.00135

Rahman, S., and Copeland, W. C. (2019). *POLG*-related Disorders and Their Neurological Manifestations. *Nat. Rev. Neurol.* 15 (1), 40–52. doi:10.1038/s41582-018-0101-0

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

Samanta, D., Gokden, M., and Willis, E. (2018). Clinicopathologic Findings of CARS2 Mutation. *Pediatr. Neurol.* 87, 65–69. doi:10.1016/j.pediatrneurol.2018.06.009

Shen, Y., and Shi, T. (2019). Innovation for Better Health of Children. *Sci. China Life Sci.* 62 (12), 1561–1562. doi:10.1007/s11427-019-1589-2

Sofou, K., Kollberg, G., Holmström, M., Dávila, M., Darin, N., Gustafsson, C. M., et al. (2015). Whole Exome Sequencing Reveals Mutations in NARS2 and PARS2, Encoding the Mitochondrial Asparaginyl-tRNA Synthetase and Prolyl-tRNA Synthetase, in Patients with Alpers Syndrome. *Mol. Genet. Genomic Med.* 3 (1), 59–68. doi:10.1002/mgg3.115

Sofou, K., Moslemi, A. R., Kollberg, G., Bjarnadóttir, I., Oldfors, A., Nennesmo, I., et al. (2012). Phenotypic and Genotypic Variability in Alpers Syndrome. *Eur. J. Paediatr. Neurol.* 16 (4), 379–389. doi:10.1016/j.ejpn.2011.12.006

Sofou, K., Kollberg, G., Hedberg-Oldfors, C., and Oldfors, A. (2021). The Phenotypic Variability and Natural History of NARS2 Associated Disease. *Eur. J. Paediatric Neurol.* 31, 31–37. doi:10.1016/j.ejpn.2021.01.012

Tarka, S., Laure-Kamionowska, M., Wierzba-Bobrowicz, T., Witulska, K., Ciara, E., Szymańska, K., et al. (2020). *POLG* Gene Mutation. Clinico-Neuropathological Study. *Folia Neuropathol.* 58 (4), 386–392. doi:10.5114/fn.2020.102441

Tzoulis, C., Tran, G. T., Coxhead, J., Bertelsen, B., Lilleng, P. K., Balafkan, N., et al. (2014). Molecular Pathogenesis of Polymerase γ-related Neurodegeneration. *Ann. Neurol.* 76 (1), 66–81. doi:10.1002/ana.24185

Uusimaa, J., Hinttala, R., Rantala, H., Päivärinta, M., Herva, R., Röyttä, M., et al. (2008). Homozygous W748S Mutation in the POLG1 Gene in Patients with Juvenile-Onset Alpers Syndrome and Status Epilepticus. *Epilepsia* 49 (6), 1038–1045. doi:10.1111/j.1528-1167.2008.01544.x

van Westrhenen, A., Cats, E. A., van den Munckhof, B., van der Salm, S. M. A., Teunissen, N. W., Ferrier, C. H., et al. (2018). Specific EEG Markers in POLG1 Alpers' Syndrome. *Clin. Neurophysiol.* 129 (10), 2127–2131. doi:10.1016/j.clinph.2018.07.016

Walker, M. A., Mohler, K. P., Hopkins, K. W., Oakley, D. H., Sweetser, D. A., Ibba, M., et al. (2016). Novel Compound Heterozygous Mutations Expand the Recognized Phenotypes of FARS2-Linked Disease. *J. Child. Neurol.* 31 (9), 1127–1137. doi:10.1177/0883073816643402

Wiltshire, E., Davidzon, G., DiMauro, S., Akman, H. O., Sadleir, L., Haas, L., et al. (2008). Juvenile Alpers Disease. *Arch. Neurol.* 65 (1), 121–124. doi:10.1001/archneurol.2007.14

Wolf, N. I., Rahman, S., Schmitt, B., Taanman, J. W., Duncan, A. J., Harting, I., et al. (2009). Status Epilepticus in Children with Alpers' Disease Caused by POLG1 Mutations: EEG and MRI Features. *Epilepsia* 50 (6), 1596–1607. doi:10.1111/j.1528-1167.2008.01877.x

Wu, J., Yang, C., Collins, J., and Ginat, D. T. (2018). Case 250: Alpers-Huttenlocher Syndrome. *Radiology* 286 (2), 720–725. doi:10.1148/radiol.2017151748

Zijlmans, M., Jiruska, P., Zelmann, R., Leijten, F. S., Jefferys, J. G., and Gotman, J. (2012). High-frequency Oscillations as a New Biomarker in Epilepsy. *Ann. Neurol.* 71 (2), 169–178. doi:10.1002/ana.22548

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 11 October 2021
doi: 10.3389/fgene.2021.715437

Genetic Analysis of 28 Chinese Families With Tyrosinase-Positive Oculocutaneous Albinism

*Linya Ma[1†], Jianjian Zhu[1†], Jing Wang[2], Yazhou Huang[1], Jibo Zhang[1], Chao Wang[1], Yuan Zhou[1] and Dan Peng[1,3]**

[1] Department of Medical Genetics, Changde First People's Hospital, Changde, China, [2] Changsha Kingmed Center for Clinical Laboratory, Changsha, China, [3] Affiliated Hospital of Changde City, University of South China, Hengyang, China

Background: Tyrosinase-positive oculocutaneous albinism (OCA, type II, OCA2) is an autosomal recessive genetic disease in which the biosynthesis of melanin decreases in the skin, hair, and eyes. OCA2 disease is caused by mutations in *OCA2* gene. The gene product plays a role in regulating the pH of melanosomes. Up to now, hundreds of *OCA2* mutations have been reported and novel variants are still being discovered.

Methods: In this study, we reviewed the records of OCA2 patients who had conducted albinism genetic testing, and then analyzed the clinical and genetic information of 28 OCA2 patients who had been genetically diagnosed by using Sanger sequencing and next-generation sequencing.

Results: In this study, we reported 31 variants screened from 28 Chinese OCA2 families, and characterized the detailed molecular and clinical presentations. There were 12 novel variants among all detected variants, including 3 missense variants (p.G393V, p.T482A, and p.R720P), 4 frameshift variants (p.R53Gfs*49, p.N279Kfs*17, p.I469Lfs*4, p.I655Nfs*12), 2 splicing variants (c.1637-2A > G, c.1951 + 1G > C), 2 stopgain variants (p.L278X, p.W652X) and 1 insertion variants (p.P315LinsT). One potential cluster of missense variants was implicated indicating the important roles of the underlying domains in OCA2 pathogenesis.

Conclusion: Our results were beneficial for diagnosis and precision clinical management for *OCA2*-related disorder, and this study expanded the mutation spectrum of oculocutaneous albinism.

Keywords: *OCA2 gene*, missense variants, oculocutaneous albinism, next-generation sequencing, novel variants

OPEN ACCESS

Edited by:
Tieliu Shi,
East China Normal University, China

Reviewed by:
Lu Xie,
Shanghai Center For Bioinformation Technology, China
Yaqiong Jin,
Capital Medical University, China

****Correspondence:***
Dan Peng
dasan0988@126.com

[†]These authors 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: *28 May 2021*
Accepted: *30 July 2021*
Published: *11 October 2021*

Citation:
Ma L, Zhu J, Wang J, Huang Y, Zhang J, Wang C, Zhou Y and Peng D (2021) Genetic Analysis of 28 Chinese Families With Tyrosinase-Positive Oculocutaneous Albinism. Front. Genet. 12:715437. doi: 10.3389/fgene.2021.715437

INTRODUCTION

Albinism is a heterogeneous genetic disorder in which a group of genes related to pigment synthesis have mutated, leading to melanin deficiency. According to its clinical symptoms, it can be divided into oculocutaneous albinism (OCA) and ocular albinism (OA). Oculocutaneous albinism is a crowd of inherited disorders of melanin biosynthesis characterized by a generalized reduction

in pigmentation of hair, skin and eyes (Gronskov et al., 2007), whereas in OA patients only the ocular pigment is deficient. The degree of hypopigmentation of skin and hair varies with the types of OCA. Oculocutaneous albinism is classified into non-syndromic oculocutaneous albinism and syndrome oculocutaneous albinism based on symptoms such as the presence of bleeding diathesis, immunodeficiency and neurological dysfunction (Simeonov et al., 2013). Mutations in *TYR*, *OCA2*, *TYRP1*, *SLC45A2*, *SLC24A5,* and *LRMDA* lead to six different non-syndromic OCA subtypes, they are OCA1, OCA2, OCA3, OCA4, OCA6, and OCA7, respectively (Arveiler et al., 2017). The prevalence of all forms of albinism varies considerably worldwide, estimated at approximately 1/17000, the mutation carrier rate is about 1 in 65 (Montoliu et al., 2014).

Tyrosinase-positive oculocutaneous albinism (OCA, type II, OMIM 203200) is an autosomal recessive genetic disease that reduces the biosynthesis of melanin in the skin, hair and eyes. Patients with OCA2 have characteristic visual abnormalities associated with albinism, including decreased vision and nystagmus. These symptoms are usually milder than those with OCA1 (Lee et al., 1994). OCA2 was first identified in 1993, in a patient with tyrosinase-positive oculocutaneous albinism. According to epidemiological survey, OCA2 is the most prevalent subtype in Africa and accounts for 30% of OCA cases worldwide. Studies have revealed that OCA2 is caused by mutations in *OCA2*, the human cDNA contains 3143 bases which encodes a protein that corresponds to the "pink-eyed dilution" (p) mouse mutant. In addition, the gene product plays a role in regulating the pH of melanosomes (Yuasa et al., 2007). Up to now, lots of *OCA2* mutations have been reported and novel variants are still being discovered (Maurano et al., 2012; Martinez-Garcia and Montoliu, 2013; Wang et al., 2016; Qiu et al., 2018; Lin et al., 2019; Luo et al., 2019; Yang et al., 2019; Zhong et al., 2019; Chuan et al., 2021; Xu et al., 2021). In recent years, multiple groups of researchers have conducted epidemiological surveys on Chinese albinism patients. Although the proportions of each type are different, OCA1 is recognized as the main type of albinism, and exons 1 and 2 are mutation hotspots. The proportions of OCA2, OCA4, HPS1 and unknown mutations are behind (Lin et al., 2019; Luo et al., 2019; Chuan et al., 2021; Xu et al., 2021).

In this study, we performed mutational screening of OCA2 in 28 Chinese OCA2 patients. After literature search and database retrieval, 12 novel variants in OCA2 were identified. Our results provide new insight into the underlying mechanisms of OCA2 and at the same time, provide important information for genetic testing and counseling.

MATERIALS AND METHODS

Patients

We reviewed the records of patients undergoing genetic testing for albinism in the Changde First People's Hospital and Changsha Kingmed Center for Clinical Laboratory in recent years. We recruited 28 OCA2 families from Hunan Province, China. Among the 28 families, there are 28 patients, including 15 females and 13 males. All patients denied family history of consanguinity. These patients all have obvious albinism phenotype, but due to the incompletely medical records, the clinical data did not record their skin color, hair color, iris color and ophthalmic phenotype in detail.

Ethics Statement

A written informed consent was obtained from each subject or their guardians to participate in this study. The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Changde First People's Hospital (protocol code 2020-014-01) and approved on January 23, 2020.

Genomic DNA Preparation and Next-Generation Sequencing

Genomic DNA was extracted using whole blood DNA extraction kit (Tiangen Biotechnology, Beijing, China). The library was prepared by randomly fragmenting 2 μg of genomic DNA into fragments by ultrasonic shearing. According to the standard Illumina protocol, a HiSeq2500 sequencer was used for sample sequencing. The amplified DNA was captured with a Genodermatosis-related Gene Panel, which can capture all exons and splicing sites of more than 100 related genes. The raw reads were filtered to obtain high-quality clean reads, and mapped to the human reference genome (UCSC hg19), and the sention software suite was used to call Single Nucleotide Variants (SNVs) and small insertions or deletions (InDels).

Validation of Variants and Inheritance Analysis

All candidate pathogenic variants were verified by Sanger sequencing in the reported families to verify the heritability of the variants. We designed specific primers (**Supplementary Table 1**) to amplify the region containing the corresponding variation by polymerase chain reaction (PCR). The PCR products were then sequenced on ABI 3730XL Genetic Analyzer (Applied Biosystems Life Technologies) according to manufacturer s protocols.

Variant Classification and *in silico* Analysis

The data obtained by sequencing was screened according to the following criteria. First, the variants with minor allele frequency (MAF) >0.01 in the following three SNP databases were excluded. Including gnomAD,[1] 1000 Genome project,[2] dbSNP[3] and ESP6500.[4] Second, we used the ClinVar database,[5] the albinism database,[6] the human gene mutation database (HGMD[7]) and OMIM[8] to annotate variants. Third, in order

[1]https://gnomad.broadinstitute.org/
[2]http://www.internationalgenome.org
[3]http://www.ncbi.nlm.nih.gov/dbvar
[4]http://evs.gs.washington.edu/EVS
[5]http://www.ncbi.nlm.nih.gov/clinvar
[6]http://www.ifpcs.org/albinism
[7]http://www.biobase-international.com/product/hgmd
[8]https://omim.org/

to determine the pathogenicity of novel mutations, we *in silico* analyzed the pathogenicity of novel mutations with various tools, which included the programs of Sorting Intolerant From Tolerant (SIFT[9]), Polymorphism Phenotyping v2 (PolyPhen2[10]), Mutation Assessor,[11] Protein Variation Effect Analyzer (PROVEAN[12]) and CADD.[13] Subsequently, we used I Mutant2.0[14] to evaluate the protein stability changes upon novel variants and clustalX2 for the protein conservation analysis. The structure changes of protein caused by amino acid substitutions were simulated by I-TASSER.[15] Finally, according to the American Medical Genetics and Genomics (ACMG) guidelines, all detected novel variants were classified into pathogenic, likely pathogenic, uncertain significance, and likely benign or benign.

[9]http://sift.jcvi.org
[10]http://genetics.bwh.harvard.edu/pph2
[11]http://mutationassessor.org/
[12]http://provean.jcvi.org/index.php
[13]http://cadd.gs.washington.edu/download
[14]http://folding.biofold.org/cgi-bin/i-mutant2.0.cgi
[15]https://zhanglab.ccmb.med.umich.edu/I-TASSER/

RESULTS

Clinical Manifestation

By reviewing the records of patients who have performed albinism genetic testing in the First People's Hospital of Changde and Changsha Kingmed Center for Clinical Laboratory in recent years, the clinical information of 28 patients who had been diagnosed as OCA2 through genetic testing was collected. All patients had typical OCA symptoms, including varying degrees of hypopigmentation on the skin, hair and iris. Moreover, patients did not show any other symptoms involving other systems. Among 28 patients, only 7 parents had tested for albinism-related genes (**Supplementary Table 1**). The clinical characteristics and mutant alleles of these 28 patients were shown in **Table 1**.

Mutation Pattern and Potential Missense Variant Clusters in *OCA2*

Based on the analysis and statistics of the sequencing results, 2 of the 28 patients were single homozygous (Patient 1 and 3 in **Table 1**), and the rest 26 probands were compound heterozygous.

TABLE 1 | Clinical characteristics and genotypes of the 28 patients.

Patient ID	Gender	Age	Molecular diagnosis	Mutations			
				Allele 1	Parental	Allele 2	Parental
1	M	25	OCA2	c.1178G > T	Ma	c.1178G > T	Pa
2	F	5	OCA2	c.1963dupA	Ma	c.1444A > G	Pa
3	F	9	OCA2	c.1255C > T	Ma	c.1255C > T	Pa
4	M	3	OCA2	c.833T > G	NA	c.406C > T	NA
5	F	1	OCA2	c.1349C > T	NA	c.1349C > T	NA
6	M	10	OCA2	c.2228C > T	NA	c.593C > T	NA
7	F	9	OCA2	c.1441G > A	NA	c.2159G > C	NA
8	F	14	OCA2	c.2344G > A	NA	c.1844A > G	NA
9	F	6	OCA2	c.406C > T	NA	c.1955G > A	NA
10	M	4	OCA2	c.1663C > T	NA	c.2330G > A	NA
11	M	4	OCA2	c.1349C > T	NA	c.1441G > A	NA
12	M	24	OCA2	c.1255C > T	Ma	c.2180T > C	Pa
13	F	5	OCA2	c.1182 + 1G > A	NA	c.1405_1406delATinsC	NA
14	F	8	OCA2	c.1255C > T	NA	c.2323G > C	NA
15	F	26	OCA2	c.1182 + 1G > A	NA	c.1405_1406delATinsC	NA
16	F	25	OCA2	c.156delC	NA	c.1441G > A	NA
17	F	17	OCA2	c.1255C > T	NA	c.1349C > T	NA
18	F	3	OCA2	c.1255C > T	NA	c.2180T > C	NA
19	M	13	OCA2	c.163delG	NA	c.1441G > A	NA
20	F	22	OCA2	c.1441G > A	NA	c.2344G > A	NA
21	M	5	OCA2	c.808-3C > G	NA	c.1441G > A	NA
22	M	21	OCA2	c.1951 + 1G > C	NA	c.1423A > G	NA
23	F	9	OCA2	c.2359G > A	NA	c.1255C > T	NA
24	M	2	OCA2	c.1363A > G	NA	c.1637-2A > G	NA
25	F	7	OCA2	c.593C > T	NA	c.2228C > T	NA
26	M	4	OCA2	c.406C > T	Ma	c.593C > T	Pa
27	M	2	OCA2	c.2323G > A	Ma	c.830_836dup	Pa
28	M	1	OCA2	c.944_945insCAC	Ma	c.1139_1141del	Pa

Moreover, missense variants (54.8%, 17/31) were the most prevalent mutation. To investigate the mutation pattern of *OCA2*, we aggregated the mutations information of all patients. There were 31 variants in 28 patients, including twelve likely gene-disrupting (LGD) variants (6 frameshift, 2 stopgain, 4 splice-site), 1 deletion,1 insertion and seventeen missense variants (**Table 2** and **Figure 1**). We observed one potential missense cluster, 16 out of 17 missense variants were clustered in P-permease domain (**Figure 1**).

Pathogenic Novel Variant Pattern of *OCA2*-Related OCA2 Genes

There were 12 novel variants among all variants, including 3 missense variants. Of the remaining 19 reported mutations, six were reported for the first time in the Chinese population, including p.A55Lfs*47, p.L727P, p.P743L, p.G775S, p.C777Y, and p.G782R. By comparing the alignments of OCA2 orthologous peptide sequences of the three novel missense variants (p.G393V, p.T482A, p.R720P), most variants show evolutionary conservation (**Figure 2A**) and predicted to be damaging or probably damaging by PolyPhen2, PROVEAN, SIFT and had a high CADD and Mutation Assessor score (**Table 3**). The p.T482A and p.R720P were assessed to decrease the OCA2 protein stability by I Mutant2.0, which was consistent with the simulated result by I-TASSER that amino acid substitutions resulted in the structure changes of protein (**Figure 2B**).

DISCUSSION

Oculocutaneous albinism is a group of autosomal recessive genetic diseases caused by the reduction or lack of melanin synthesis in melanocytes that affects the pigmentation of hair, skin and eyes (Yang et al., 2019). Different types of OCA cannot be completely distinguished by clinical phenotype, therefore, molecular diagnosis have become a useful tool and a necessary condition for genetic consultation (Gronskov et al., 2007). OCA2 is one of the main subtypes of OCA, which is caused by mutations in *OCA2* (Yuasa et al., 2007). The protein encoded by this gene has 12 transmembrane domains and is an intact membrane

TABLE 2 | Summary of OCA2 variants identified in this study.

Nt change	gDNA change(chr15)	Exon No.	AA change	Function	Minor Allele frequency (gnomAD)	Mutation frequency
c.156delC	g.28326865delG	Exon2	p.R53Gfs*49	frameshift	0.00001217	1/56
c.163delG	g.28326858delC	Exon2	p.A55Lfs*47	frameshift	0.00001079	1/56
c.406C > T	g.28273126G > A	Exon4	p.R136X	stopgain	0.00001591	3/56
c.593C > T	g.28267700G > A	Exon6	p.P198L	missense	0.0001238	3/56
c.808-3C > G	g.28261332-3G > C	Intron7	–	splicing	0.000003977	1/56
c.833T > G	g.28261307A > C	Exon8	p.L278X	frameshift	–	1/56
c.830_836dup	g.28261304_g.28261310dup	Exon8	p.N279Kfs*17	frameshift	–	1/56
c.944_945insCAC	g.28260021_g.28260022insGTG	Exon9	p.P315LinsT	insertion	–	1/56
c.1139_1141del	g.28234788_g.28234790delCCA	Exon11	p.V380del	deletion	–	1/56
c.1178G > T	g.28234751C > A	Exon11	p.G393V	missense	0.000003979	1/56
c.1182 + 1G > A	g.28230320 + 1C > T	Intron11	–	splicing	–	2/56
c.1255C > T	g.28230319G > A	Exon13	p.R419W	missense	0.0002659	6/56
c.1349C > T	g.28230225G > A	Exon13	p.T450M	missense	0.00001776	3/56
c.1363A > G	g.28230211T > C	Exon13	p.R455G	missense	0.0002453	1/56
c.1405_1406delATinsC	g.28228588_g.28228589delinsG	Exon14	p.I469Lfs*4	frameshift	–	2/56
c.1423A > G	g.28228571T > C	Exon14	p.T475A	missense	–	1/56
c.1441G > A	g.28228553C > T	Exon14	p.A481T	missense	0.008427	6/56
c.1444A > G	g.28228550T > C	Exon14	p.T482A	missense	–	1/56
c.1637-2A > G	g.28202881-2T > C	Intron15	–	splicing	–	1/56
c.1663C > T	g.28202855G > A	Exon16	p.R555C	missense	0.0001186	1/56
c.1844A > G	g.28197037T > C	Exon18	p.H615R	missense	0.04345	1/56
c.1951 + 1G > C	g.28196930 + 1C > G	Intron18	–	splicing	0.000003984	1/56
c.1955G > A	g.28171397C > 7	Exon19	p.W652X	stopgain	–	1/56
c.1963dupA	g.28171390dupT	Exon14	p.I655Nfs*12	frameshift	–	1/56
c.2159G > C	g.28116385C > G	Exon21	p.R720P	missense	–	1/56
c.2180T > C	g.28116364A > G	Exon21	p.L727P	missense	–	2/56
c.2228C > T	g.28116316G > A	Exon21	p.P743L	missense	0.0001344	2/56
c.2323G > A	g.28096543C > T	Exon22	p.G775S	missense	0.000008028	2/56
c.2330G > A	g.28096536C > T	Exon22	p.C777Y	missense	0.00003277	1/56
c.2344G > A	g.28090193C > T	Exon23	p.G782R	missense	0.000007071	2/56
c.2359G > A	g.28090178C > T	Exon23	p.A787T	missense	0.00004374	1/56

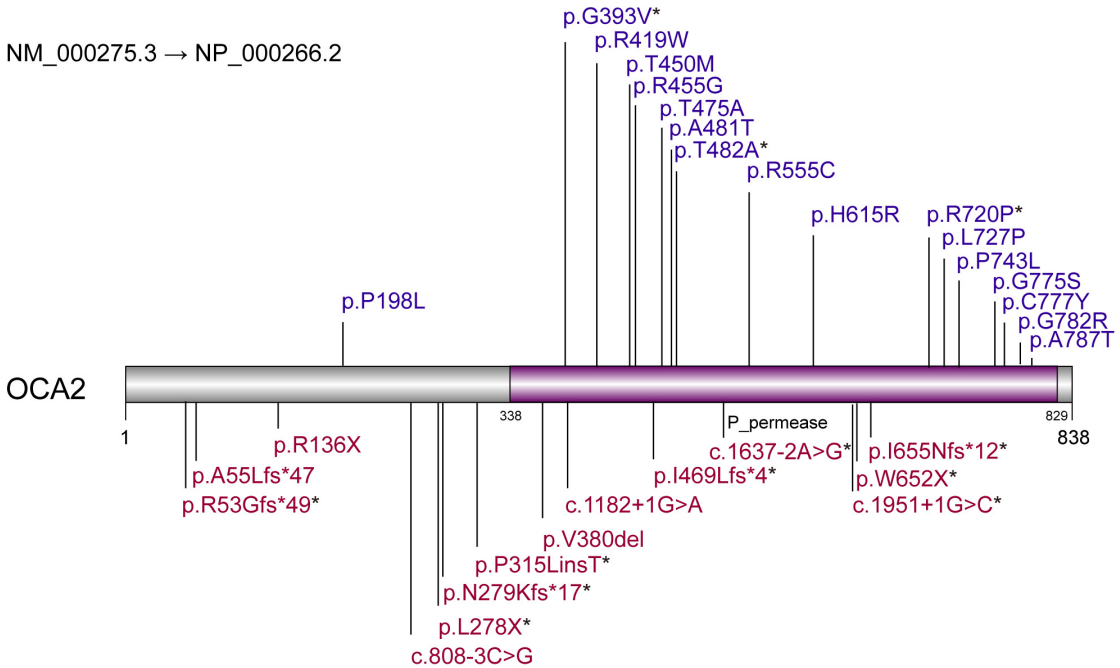

FIGURE 1 | Variants pattern of *OCA2*. Loss-of-function variants are labeled above the diagram. Missense variants are labeled below the diagram. Purple region represents P-permease domain. * represents the novel variants.

protein involved in the transport of small molecules, especially tyrosine, which is the precursor of melanin synthesis (Rinchik et al., 1993; Gronskov et al., 2007).

In this study, we collected 28 oculocutaneous albinism patients with compound heterozygous or homozygous *OCA2* variants identified by sequencing. All patients have typical clinical phenotypes of oculocutaneous albinism (**Table 1**). Thirty one variants have been identified in all patients, including 19 reported variants and 12 novel variants. Among all the variants, missense variants (54.8%, 17/31) were the most common mutation (Chumakov and Kronrod, 1990; Spritz et al., 1997; Sviderskaya et al., 1997; Passmore et al., 1999; Suzuki et al., 2003; Duan et al., 2006; Hongyi et al., 2007; Yuasa et al., 2007; Chiang et al., 2008; Dai et al., 2008; Rooryck et al., 2008; Gronskov et al., 2009; Wei et al., 2010, 2011, 2013; Gargiulo et al., 2011). The missense variants, p.R419W and p.A481T were the most general OCA2 variants, six patients have inherited these variants, respectively. R A Spritz (Spritz et al., 1997) have found p.R419W in OCA2 patients, but the variant classification of the HGMD database and ClinVar database are inconsistent. According to the comprehensive analysis of ACMG, the variant was consistent with uncertain significance. The high frequency of p.R419W in our study could provide more evidence for its pathogenicity. In addition, the pathogenicity of p.A481T has not been determined, it was recorded as benign in the ClinVar database, but reported to be an Asian-specific hypopigmentation allele in 2007 (Yuasa et al., 2007). The p.P743L variant in the OCA2 gene has been reported in the homozygous state or in the compound heterozygous state in multiple unrelated individuals with OCA2 (Lee et al., 1994; King et al., 2003; Hutton and Spritz, 2008; Sengupta et al., 2010; Jaworek et al., 2012; Richards et al., 2015; Shahzad et al., 2017). Variant p.P743L is a semi-conservative amino acid substitution. Due to the differences in physiochemical properties between the two residues, the switch of proline to leucine residue will have significant impact on protein structure, indicating the pathogenic effect of the mutation. In our study, 2 patients inherited this missense substitution, most of the previous reports for this substitution were in European, American and African, this was the first report in Chinese. The splice variant c.1182 + 1G > A was also found in two patients, in 2019, Dan Luo et al., found four patients carrying this mutation, and one of them was homozygous (Luo et al., 2019). In addition, a research showed that the variant was expected to affect splicing following exon 11, thereby leading to abnormal splicing of the transcript (Montoliu et al., 2014).

Among the 31 different mutations in the OCA2 gene, there are 12 novel mutations (4 frameshift, 2 stopgain, 2 splice-site,1 insertion and 3 missense). Frameshift mutations p.R53Gfs*49, p.N279Kfs*17, p.I469Lfs*4, p.I655Nfs*12 and stopgain mutations p.Leu278X, p.W652X were expected to affect the original protein OCA2 function by changing the protein sequence or leading to early termination of protein translation. The truncated protein lacks intact transmembrane domain, which could cause the dysfunction of tyrosine transport and precursor melanin synthesis, and lead to the location of the protein in the nucleus. Splicing mutations c.1637-2A > G and c.1951 + 1G > C were expected to affect splicing of the transcript, thereby leading to abnormal protein function. The literature research showed that most missense mutations occur in the

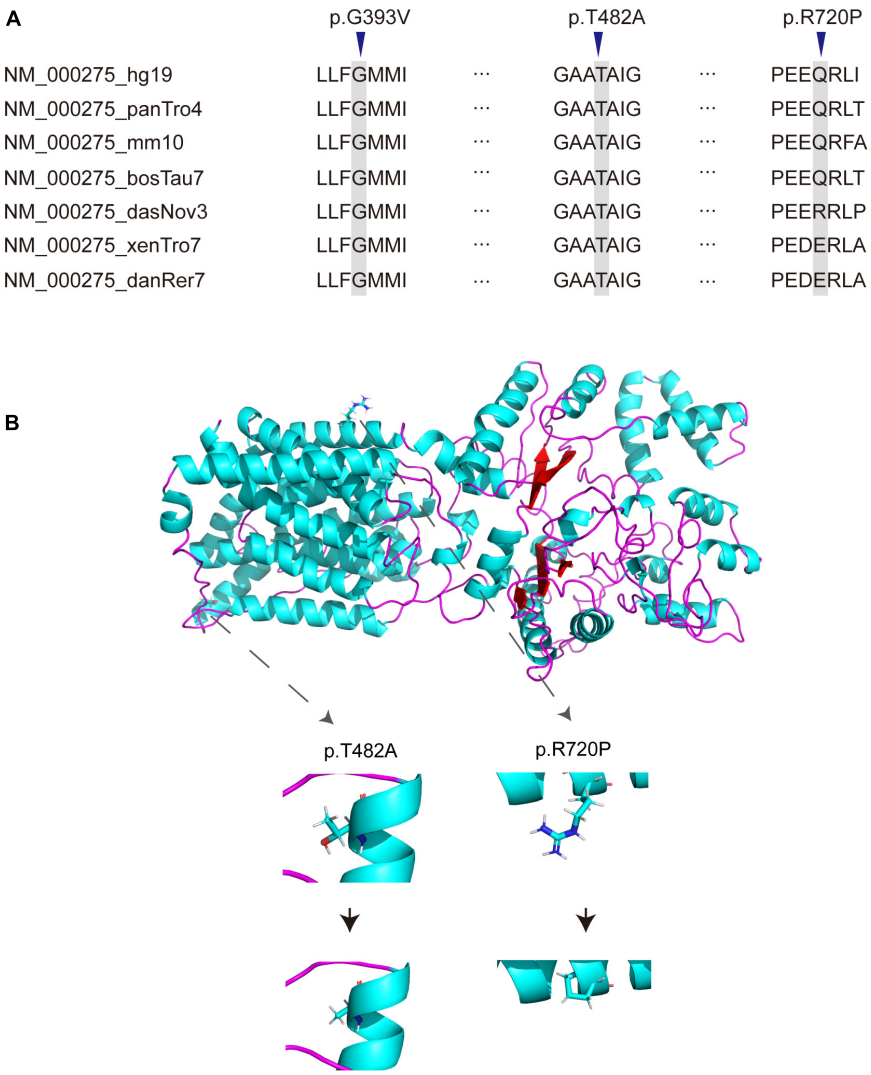

FIGURE 2 | Novel missense variants of *OCA2*. **(A)** Conservation analysis of *OCA2* missense variants. **(B)** Simulation of the amino acids conformation changes by I-TASSER.

TABLE 3 | Pathogenicity analysis of novel variants.

Nt change	AA change	SIFT	Polyphen2	Mutation assessor	PROVEAN	CADD score	Variant classification	Evidence criterion
c.156delC	p.R53Gfs*49	–	–	–	–	–	Pathogenic	PVS1 + PM2 + PP3
c.830_836dup	p.N279Kfs*17	–	–	–	–	–	Pathogenic	PVS1 + PM2 + PP3
c.1405_1406delATinsC	p.I469Lfs*4	–	–	–	–	–	Pathogenic	PVS1 + PM2 + PP3
c.1963dupA	p.I655Nfs*12	–	–	–	–	–	Likely pathogenic	PVS1 + PM2
c.1637-2A > G	c.1637-2A > G	–	–	–	–	–	Pathogenic	PVS1 + PM2 + PP3
c.1951 + 1G > C	c.1951 + 1G > C	–	–	–	–	–	Pathogenic	PVS1 + PM2 + PP3
c.833T > G	p.Leu278X	–	–	–	–	–	Pathogenic	PVS1 + PM2 + PP4
c.1955G > A	p.W652X	–	–	–	–	–	Pathogenic	PVS1 + PM2 + PP3
c.944_945insCAC	p.P315LinsT	–	–	–	D	–	Uncertain significance	PM2 + PP3
c.1178G > T	p.G393V	D	PD	High	D	33	Uncertain significance	PM2 + PP3
c.1444A > G	p.T482A	D	PD	High	D	25.5	Uncertain significance	PM2 + PM3 + PP3
c.2159G > C	p.R720P	D	PD	Medium	D	33	Uncertain significance	PM2 + PP3

PD, Probably Damaging; D, Deleterious. The type of evidence refers to. ACMG/AMP 2015 guideline (http://wintervar.wglab.org/).

loop between the transmembrane domains (Spritz, 1994). In addition, most of the missense variants found in this study were located in the P_permease domain, which was linked to human melanosomal P gene. Variants in P gene were responsible for classic phenotype of OCA2 (Lee et al., 1995; Puri et al., 2000; Brilliant, 2001; Staleva et al., 2002). Missense variants p.G393V occurred in the fourth transmembrane domain of P protein, which would lead to inactivity of the P protein. Three novel missense variants p.G393V, p.T482A and p.R720P were evolutionary conservation and *in silico* predicted to be damaging or probably damaging. The T482A and R720P are assessed to decrease the OCA2 protein stability, thereby affect to protein function possibly. According to the ACMG, these missense variants are recorded as uncertain significance, therefore, the pathogenicity of these missense mutations needs in-depth exploration.

At present, due to the lack of effective treatment for albinism, prenatal diagnosis is particularly important to prevent the birth of patients. On the basis of prenatal genetic testing and diagnosis, genetic counseling and inspection guidance can effectively prevent the birth of severely patients (Bao et al., 2019; Geng et al., 2019; Li et al., 2019).

In summary, we identified 31 distinct variants of *OCA2* by next-generation sequencing, in addition, 12 variants were novel ones. We characterized the molecular and phenotypic data for patients with OCA2 variants and revealed one potential missense variant cluster by curating published data. Our findings will benefit not only for the genetic diagnosis and counseling but also provide motivation for further functional characterizations of *OCA2* variants.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the China National GeneBank with the accession code of CNP0002229.

ETHICS STATEMENT

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

LM and JZhu analyzed the data and prepared the manuscript. JW and YH collected the data. JZha and CW rechecked the data. DP and YZ provided theoretical guidance. DP revised the manuscript and read and approved the final manuscript. All authors approved the final manuscript.

FUNDING

This work was supported by the 2020 Hunan Provincial Clinical Medical Technology Innovation Guidance Project (2020SSK50201).

ACKNOWLEDGMENTS

We thank all the health care technicians in the Department of Medical Genetics, Affiliated Hospital of Changde city of South China University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.715437/full#supplementary-material

Supplementary Figure 1 | Pedigree drawings of the seven families.

REFERENCES

Arveiler, B., Lasseaux, E., and Morice-Picard, F. (2017). [Clinical and genetic aspects of albinism]. *Presse Med.* 46, 648–654. doi: 10.1016/j.lpm.2017.05.020

Bao, Y., Suo, L., Qian, P., Huang, H., Yang, Y., Tang, J., et al. (2019). Clinical and genetic analysis of Dent disease with nephrotic range albuminuria in Shaanxi. China. *Sci. China Life Sci.* 62, 1590–1593.

Brilliant, M. H. (2001). The mouse p (pink-eyed dilution) and human P genes, oculocutaneous albinism type 2 (OCA2), and melanosomal pH. *Pigment Cell Res.* 14, 86–93. doi: 10.1034/j.1600-0749.2001.140203.x

Chiang, P. W., Fulton, A. B., Spector, E., and Hisama, F. M. (2008). Synergistic interaction of the OCA2 and OCA3 genes in a family. *Am. J. Med. Genet. A* 146A, 2427–2430. doi: 10.1002/ajmg.a.32453

Chuan, Z., Yan, Y., Hao, S., Zhang, Q., Zhou, B., Feng, X., et al. (2021). Mutation Analysis of 63 Northwest Chinese Probands with Oculocutaneous Albinism. *Curr. Eye Res.* 46, 140–143. doi: 10.1080/02713683.2020.1781192

Chumakov, A. A., and Kronrod, B. A. (1990). [The status and outlook of the training of pathologists in the RSFSR]. *Arkh. Patol.* 52, 50–52.

Dai, C., Li, W., Gao, B., Li, L. Y., and Lu, G. X. (2008). [Mutation screening of the TYR and P gene in three patients with oculocutaneous albinism]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 25, 373–377.

Duan, H. L., Li, H. Y., Wu, W. Q., Zheng, H., and Chen, Z. (2006). [A novel P gene mutation in a Chinese family with oculocutaneous albinism]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 23, 614–617.

Gargiulo, A., Testa, F., Rossi, S., Di Iorio, V., Fecarotta, S., de Berardinis, T., et al. (2011). Molecular and clinical characterization of albinism in a large cohort of Italian patients. *Invest. Ophthalmol. Vis. Sci.* 52, 1281–1289. doi: 10.1167/iovs.10-6091

Geng, J., Liu, Y., Guo, Y., Wang, H., Tai, J., Jin, Y., et al. (2019). Correlation between TERT C228T and clinicpathological features in pediatric papillary thyroid carcinoma. *Sci. China Life Sci.* 62, 1563–1571.

Gronskov, K., Ek, J., and Brondum-Nielsen, K. (2007). Oculocutaneous albinism. *Orphanet. J. Rare Dis.* 2:43. doi: 10.1186/1750-1172-2-43

Gronskov, K., Ek, J., Sand, A., Scheller, R., Bygum, A., Brixen, K., et al. (2009). Birth prevalence and mutation spectrum in danish patients with autosomal recessive albinism. *Invest. Ophthalmol. Vis. Sci.* 50, 1058–1064. doi: 10.1167/iovs.08-2639

Hongyi, L., Haiyun, W., Hui, Z., Qing, W., Honglei, D., Shu, M., et al. (2007). Prenatal diagnosis of oculocutaneous albinism type II and novel mutations in two Chinese families. *Prenat. Diagn.* 27, 502–506. doi: 10.1002/pd.1713

Hutton, S. M., and Spritz, R. A. (2008). Comprehensive analysis of oculocutaneous albinism among non-Hispanic caucasians shows that OCA1 is the most prevalent OCA type. *J. Invest. Dermatol.* 128, 2442–2450. doi: 10.1038/jid.2008.109

Jaworek, T. J., Kausar, T., Bell, S. M., Tariq, N., Maqsood, M. I., Sohail, A., et al. (2012). Molecular genetic studies and delineation of the oculocutaneous albinism phenotype in the Pakistani population. *Orphanet. J. Rare Dis.* 7:44. doi: 10.1186/1750-1172-7-44

King, R. A., Willaert, R. K., Schmidt, R. M., Pietsch, J., Savage, S., Brott, M. J., et al. (2003). MC1R mutations modify the classic phenotype of oculocutaneous albinism type 2 (OCA2). *Am. J. Hum. Genet.* 73, 638–645. doi: 10.1086/377569

Lee, S. T., Nicholls, R. D., Bundey, S., Laxova, R., Musarella, M., and Spritz, R. A. (1994). Mutations of the P gene in oculocutaneous albinism, ocular albinism, and Prader-Willi syndrome plus albinism. *N. Engl. J. Med.* 330, 529–534. doi: 10.1056/NEJM199402243300803

Lee, S. T., Nicholls, R. D., Jong, M. T., Fukai, K., and Spritz, R. A. (1995). Organization and sequence of the human P gene and identification of a new family of transport proteins. *Genomics* 26, 354–363. doi: 10.1016/0888-7543(95)80220-g

Li, Z., Zhu, P., Huang, H., Pan, Y., Han, P., Cui, H., et al. (2019). Identification of a novel COL4A5 mutation in the proband initially diagnosed as IgAN from a Chinese family with X-linked Alport syndrome. *Sci. China Life Sci.* 62, 1572–1579.

Lin, Y., Chen, X., Yang, Y., Che, F., Zhang, S., Yuan, L., et al. (2019). Mutational analysis of TYR, OCA2, and SLC45A2 genes in chinese families with oculocutaneous albinism. *Mol. Genet. Genomic Med.* 7:e00687. doi: 10.1002/mgg3.687

Luo, D., Linpeng, S., Zeng, L., Tan, H., Li, Z., and Wu, L. (2019). Molecular genetic study of 59 Chinese Oculocutaneous albinism families. *Eur. J. Med. Genet.* 62:103709. doi: 10.1016/j.ejmg.2019.103709

Martinez-Garcia, M., and Montoliu, L. (2013). Albinism in Europe. *J. Dermatol.* 40, 319–324. doi: 10.1111/1346-8138.12170

Maurano, M. T., Wang, H., Kutyavin, T., and Stamatoyannopoulos, J. A. (2012). Widespread site-dependent buffering of human regulatory polymorphism. *PLoS Genet.* 8:e1002599. doi: 10.1371/journal.pgen.1002599

Montoliu, L., Gronskov, K., Wei, A. H., Martinez-Garcia, M., Fernandez, A., Arveiler, B., et al. (2014). Increasing the complexity: new genes and new types of albinism. *Pigment Cell Melanoma Res.* 27, 11–18. doi: 10.1111/pcmr.12167

Passmore, L. A., Kaesmann-Kellner, B., and Weber, B. H. (1999). Novel and recurrent mutations in the tyrosinase gene and the P gene in the German albino population. *Hum. Genet.* 105, 200–210. doi: 10.1007/s004390051090

Puri, N., Gardner, J. M., and Brilliant, M. H. (2000). Aberrant pH of melanosomes in pink-eyed dilution (p) mutant melanocytes. *J. Invest. Dermatol.* 115, 607–613. doi: 10.1046/j.1523-1747.2000.00108.x

Qiu, B., Ma, T., Peng, C., Zheng, X., and Yang, J. (2018). Identification of five novel variants in chinese oculocutaneous albinism by targeted next-generation sequencing. *Genet. Test. Mol. Biomarkers* 22, 252–258. doi: 10.1089/gtmb.2017.0211

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

Rinchik, E. M., Bultman, S. J., Horsthemke, B., Lee, S. T., Strunk, K. M., Spritz, R. A., et al. (1993). A gene for the mouse pink-eyed dilution locus and for human type II oculocutaneous albinism. *Nature* 361, 72–76. doi: 10.1038/361072a0

Rooryck, C., Morice-Picard, F., Elcioglu, N. H., Lacombe, D., Taieb, A., and Arveiler, B. (2008). Molecular diagnosis of oculocutaneous albinism: new mutations in the OCA1-4 genes and practical aspects. *Pigment Cell Melanoma Res.* 21, 583–587. doi: 10.1111/j.1755-148X.2008.00496.x

Sengupta, M., Mondal, M., Jaiswal, P., Sinha, S., Chaki, M., Samanta, S., et al. (2010). Comprehensive analysis of the molecular basis of oculocutaneous albinism in Indian patients lacking a mutation in the tyrosinase gene. *Br. J. Dermatol.* 163, 487–494. doi: 10.1111/j.1365-2133.2010.09830.x

Shahzad, M., Yousaf, S., Waryah, Y. M., Gul, H., Kausar, T., Tariq, N., et al. (2017). Molecular outcomes, clinical consequences, and genetic diagnosis of Oculocutaneous Albinism in Pakistani population. *Sci. Rep.* 7:44185. doi: 10.1038/srep44185

Simeonov, D. R., Wang, X., Wang, C., Sergeev, Y., Dolinska, M., Bower, M., et al. (2013). DNA variations in oculocutaneous albinism: an updated mutation list and current outstanding issues in molecular diagnostics. *Hum. Mutat.* 34, 827–835. doi: 10.1002/humu.22315

Spritz, R. A. (1994). Molecular genetics of oculocutaneous albinism. *Hum. Mol. Genet.* 3, 1469–1475. doi: 10.1093/hmg/3.suppl_1.1469

Spritz, R. A., Lee, S. T., Fukai, K., Brondum-Nielsen, K., Chitayat, D., Lipson, M. H., et al. (1997). Novel mutations of the P gene in type II oculocutaneous albinism (OCA2). *Hum. Mutat.* 10, 175–177.

Staleva, L., Manga, P., and Orlow, S. J. (2002). Pink-eyed dilution protein modulates arsenic sensitivity and intracellular glutathione metabolism. *Mol. Biol. Cell* 13, 4206–4220. doi: 10.1091/mbc.e02-05-0282

Suzuki, T., Miyamura, Y., Matsunaga, J., Shimizu, H., Kawachi, Y., Ohyama, N., et al. (2003). Six novel P gene mutations and oculocutaneous albinism type 2 frequency in Japanese albino patients. *J. Invest. Dermatol.* 120, 781–783. doi: 10.1046/j.1523-1747.2003.12127.x

Sviderskaya, E. V., Bennett, D. C., Ho, L., Bailin, T., Lee, S. T., and Spritz, R. A. (1997). Complementation of hypopigmentation in p-mutant (pink-eyed dilution) mouse melanocytes by normal human P cDNA, and defective complementation by OCA2 mutant sequences. *J. Invest. Dermatol.* 108, 30–34. doi: 10.1111/1523-1747.ep12285621

Wang, X., Zhu, Y., Shen, N., Peng, J., Wang, C., Liu, H., et al. (2016). Mutation analysis of a Chinese family with oculocutaneous albinism. *Oncotarget* 7, 84981–84988. doi: 10.18632/oncotarget.13109

Wei, A., Wang, Y., Long, Y., Wang, Y., Guo, X., Zhou, Z., et al. (2010). A comprehensive analysis reveals mutational spectra and common alleles in Chinese patients with oculocutaneous albinism. *J. Invest. Dermatol.* 130, 716–724. doi: 10.1038/jid.2009.339

Wei, A., Yang, X., Lian, S., and Li, W. (2011). Implementation of an optimized strategy for genetic testing of the Chinese patients with oculocutaneous albinism. *J. Dermatol. Sci.* 62, 124–127. doi: 10.1016/j.jdermsci.2011.02.009

Wei, A. H., Yang, X. M., Lian, S., and Li, W. (2013). Genetic analyses of Chinese patients with digenic oculocutaneous albinism. *Chin. Med. J.* 126, 226–230.

Xu, C., Xiang, Y., Li, H., Xu, Y., Mao, Y., Zhou, L., et al. (2021). Genetic analysis and prenatal diagnosis of 20 Chinese families with oculocutaneous albinism. *J. Clin. Lab. Anal.* 35:e23647. doi: 10.1002/jcla.23647

Yang, Q., Yi, S., Li, M., Xie, B., Luo, J., Wang, J., et al. (2019). Genetic analyses of oculocutaneous albinism types 1 and 2 with four novel mutations. *BMC Med. Genet.* 20:106. doi: 10.1186/s12881-019-0842-7

Yuasa, I., Umetsu, K., Harihara, S., Miyoshi, A., Saitou, N., Park, K. S., et al. (2007). OCA2 481Thr, a hypofunctional allele in pigmentation, is characteristic of northeastern Asian populations. *J. Hum. Genet.* 52, 690–693. doi: 10.1007/s10038-007-0167-9

Zhong, Z., Gu, L., Zheng, X., Ma, N., Wu, Z., Duan, J., et al. (2019). Comprehensive analysis of spectral distribution of a large cohort of Chinese patients with non-syndromic oculocutaneous albinism facilitates genetic diagnosis. *Pigment Cell Melanoma Res.* 32, 672–686. doi: 10.1111/pcmr.12790

Conflict of Interest: JW was employed by Changsha Kingmed Center for Clinical Laboratory.

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.

frontiers in Pediatrics

CASE REPORT
published: 21 October 2021
doi: 10.3389/fped.2021.689901

Case Report: Identification of Polygenic Mutations by Exome Sequencing

Yanfeng Liu[1], *Zhongshi Zheng*[1] *and Qingling Zhu*[2]*

[1] Department of Endocrinology, Quanzhou Women and Children's Hospital, Quanzhou, China, [2] Department of Children Health Care, Quanzhou Women and Children's Hospital, Quanzhou, China

The discovery of rare genetic variation through different gene sequencing methods is a very challenging subject in the field of human genetics. A case of a 1-year-old boy with metabolic acidosis and hypokalemia, a small penis, growth retardation, and G-6PD deficiency was reported. Since the clinical symptoms are complex and seem uncorrelated, the authors hypothesized that the child had chromosome or gene problems, and exome sequencing (ES) was applied to samples from him and his parents. Three main locus mutations in three genes were found in the proband, including *SLC4A1*, *FGFR1*, and *G6PD* genes. A missense mutation (c.1766G>T, p.R589 L) was found in exon 14 of *SLC4A1* gene, which was a *de novo* mutation. Another missense mutation (c.1028 A>G, p.H343R) was found in exon 9 of *FGFR1* gene, which was also a *de novo* mutation. These findings further demonstrate the utility of ES in the diagnosis of rare diseases.

Keywords: genetic diseases, gene mutation, distal renal tubular acidosis, *SLC4A1*, *FGFR1*, *KAT6B*

OPEN ACCESS

Edited by:
Zhichao Liu,
National Center for Toxicological Research (FDA), United States

Reviewed by:
Tingting Yu,
Shanghai Children's Medical Center, China
Yifan Zhang,
University of Arkansas at Little Rock, United States

***Correspondence:**
Qingling Zhu
zhu_qingling2012@163.com

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

Received: *01 April 2021*
Accepted: *09 September 2021*
Published: *21 October 2021*

Citation:
Liu Y, Zheng Z and Zhu Q (2021) Case Report: Identification of Polygenic Mutations by Exome Sequencing. Front. Pediatr. 9:689901. doi: 10.3389/fped.2021.689901

INTRODUCTION

It is a challenging task to study the pathogenesis of genetic diseases. Exome sequencing (ES), also known as target exome capture, is a genomic analysis method that uses sequence capture technology to capture and enrich whole genome exon region DNA and conduct high-throughput sequencing. Previous studies (1) have shown that ES has high clinical practicability and is cost-effective. It has the highest diagnostic rate for children with genetic heterogeneity or characteristics that overlap multiple diseases (2). Our study reports a male boy who was found to have G-6PD deficiency and a small penis on his 15-day visit and showed growth retardation, metabolic acidosis, and hypokalemia on his 1-year second visit. The proband had no elder brother or sister; his parents were not blood relatives, and both of their phenotypes were normal. To further study the proband's disease, we sequenced the whole exome of the proband and his parents after obtaining consent from the proband's parents.

CASE DESCRIPTION

A male boy with normal birth weight and length was found to have G-6PD deficiency and a small penis on his 15-day first visit. Hearing screening passed in both ears. G2P1, the first fetus, was embryo stop pregnancy, with no special maternal pregnancy. At the time of his birth, his mother was 30 years old, and the father was 34 years old. There was no regular physical examination after his birth. The second visit was at 1 year of age with a slow increase in body weight, and the size

and length of the penis did not change. His body length was 69.3 cm (−2.5 SD), and his weight was 6.0 kg (−4.1). Clinical evaluation showed a thin and small stature with normal appearance; penile length was 1.5 * 0.5 cm with right testicular volume test approximately 1 ml (the left was not touched), pubic hair Tanner I stage. Blood gas analysis in the outpatient department showed pH 7.226, PCO_2 21.60 mmHg, PO_2 63.50 mmHg, BE^- 16.70 mmol/L, $BEECF^-$ 18.8 mmol/L, HCO^{3-} 8.80 mmol/L, K^+ 2.72 mmol/L, Na^+ 137.5 mmol/L, and Cl^- 109.5 mmol/L (hypokalebicarbonate acid correction, which was related to complete examinations). After repeated treatment with acidosis and oral potassium supplementation for 7 days, the blood gas basically returned to normal.

A hormonal basal study revealed low values of follicle-stimulating hormone (FSH), luteinizing hormone (LH), PREG, TESTOST, and estradiol for the proband's sex and chronological age. The gonadotropin-releasing hormone (GnRH) stimulation test elicited a prepubertal LH response with LH and FSH peaks of 0.45 and 7.10 mUI/ml, respectively. The basal value of LH was <0.07 mIU/ml. The human chorionic gonadotropin (hCG) excitation test showed a TESTOST peak of 12.36 ng/dl; the extended hCG excitation test showed that the level of TESTOST was 119.93 ng/dl, both of which suggested no obvious increase in testosterone. The results of inhibin B (INH-B) and anti-Müllerian hormone (AMH) were 21.05 pg/ml and 14.24 ng/ml, respectively, which showed poor function of testicular Sertoli cells. Drostenedione <0.44 nmol/L; dehydroepiandrosterone sulfate <0.1 μg/dl. IGF-1 was 25.0 ng/ml, and IGFBP-3 was 1.7 μg/ml, both of which were lower than in patients of the same race, sex, and age group. Plasma ammonia was elevated at 73.0 μmol/L. Citrulline was elevated at 42.71 μmol/L. A child neuropsychological development assessment suggested a low development quotient. The levels of 17-hydroxyprogesterone and hydrocortisone (COR), adrenocorticotropic hormone (ACTH), alpha-fetoprotein (AFP), carcinoembryonic antigen (CEA), and βHCG were normal. The ACTH cortisol rhythm and thyroid function were normal. The reference ranges of the test results are shown in **Table 1**.

The karyotype was 46, XY. A pituitary MRI scan and enhancement showed that the pituitary gland was small and the pituitary stalk was thin, which implicated pituitary hypoplasia, and the position of the cerebellar tonsil was low. Olfactory bulb MRI showed that both olfactory tracts were asymmetric and that the olfactory sulcus was not clear. Color Doppler ultrasound of the urinary system showed that bilateral renal medulla echoes were significantly enhanced, and calcareous deposition was considered, with spermatic cord effusion on the right region and cryptorchidism on the left region. The child had normal ACTH cortisol rhythm and thyroid function.

The initial diagnoses were as follows: (1) distal renal tubular acidosis (dRTA; type I); (2) growth retardation; (3) G-6PD enzyme deficiency; and (4) congenital small penis.

After the informed consent of family members and the consent of the hospital ethics committee was obtained, the detection of trio-ES in the child and his parents was performed to detect potential variants. Sanger sequencing technology was used to verify the suspected pathogenic mutations.

TABLE 1 | Laboratory test results and reference range of the child.

Index	Test value	Reference range
pH	7.226	7.35–7.45
PCO_2	21.60 mmHg	34.8–44.9 mmHg
PO_2	63.50 mmHg	79.8–100 mmHg
BE^-	16.70 mmol/L	−3.0 to +3.0 mmol/L
$BEECF^-$	18.8 mmol/L	−3.0 to +3.0 mmol/L
HCO^{3-}	8.80 mmol/L	21.4–27.3 mmol/L
K^+	2.72 mmol/L	3.5–5.3 mmol/L
Na^+	137.5 mmol/L	136–145 mmol/L
Cl^-	109.5 mmol/L	96–108 mmol/L
Peak of LH	0.45 mIU/ml	1.5–9.3 mIU/ml
Basal value of LH	<0.07 mIU/ml	1.5–9.3 mIU/ml
Peak of FSH	7.10 mUI/ml	1.4–18.1 mUI/ml
Peak of TESTOST	12.36 ng/dl	123.06–813.86 ng/dl
TESTOST*	119.93 ng/dl	94–327 pg/ml
INH-B	21.05 pg/ml	94–327 pg/ml
AMH	14.24 ng/ml	2.0–6.8 ng/ml
Drostenedione	<0.44 nmol/L	2.0–4.6 nmol/L
Dehydroepiandrosterone sulfate	<0.1 μg/dl	0.47–19.4 μg/dl
IGF-1	25.0 ng/ml	55–327 ng/ml
IGFBP-3	1.7 μg/ml	0.7–0.36 ng/ml
Plasma ammonia	73.0 μmol/L	10–47 μmol/L
Citrulline	42.71 μmol/L	5–35 μmol/L

LH, luteinizing hormone; FSH, follicle-stimulating hormone; INH-B, inhibin B; AMH, anti-Müllerian hormone.

**The level of TESTOST in the extended human chorionic gonadotropin (hCG) excitation test.*

METHODS FOR TRIO-EXOME SEQUENCING

DNA was extracted from peripheral blood, and ES was carried out on the proband and his parents. The coding exons were captured using the xGen® Exome Research Panel v1.0 (IDT). The captured fragments were sequenced using a NovaSeq 6000 sequencer (Illumina, San Diego, CA, USA) to an average depth of 100 reads per target base. The clean data were aligned to the National Center for Biotechnology Information (NCBI) human reference genome (HG19) using the Burrows–Wheeler Aligner (BWA), and variants were determined using GATK. Samtools and Pindel were used to determine single-nucleotide polymorphisms (SNPs) and indels, respectively. The clean data were filtered for further analysis according to the quality of the sequencing. For variant annotation and prediction, non-synonymous substitutions and SNPs with minor allele frequencies (MAFs) lower than 5% were filtered using SIFT. The function of mutated genes and their pathogenicity were analyzed referencing the dbSNP, 1,000 Genomes Project, ExAC, ESP, OMIM, Swiss-var, HGMD, ClinVar, and other disease databases. The variants with unknown pathogenicity of single bases were screened using Provean, SIFT, Polyphen2-HVAR, Polyphen2-HDIV, Mutationtster, and other protein structure prediction software. MaxEntScan was used to screen potential splice sites.

TREATMENT

After admission, the child was treated with sodium bicarbonate 20 ml qd by intravenous injection for 7 days and 10% KCl 15.5 ml qd for 4 days, later changed to potassium citrate 2 g tid p.o. The medicine after discharge was potassium citrate 2 g tid p.o. The child has regular follow-up visits and taking medications on time without adverse drug events occurring, symptoms are well-controlled, and body length and weight are increasing steadily. The timeline for the treatment and follow-up is shown in **Table 2**. During the recent follow-up, parents said that the child was more active and his appetite improved significantly.

RESULTS

Seven locus mutations in seven genes were found in the proband: *SLC4A1*, *FGFR1*, *G6PD*, *GLI3*, *BCOR*, *FAT4*, and *KAT6B*.

A missense mutation (c.1766G>T, p.R589 L) was found in exon 14 of *SLC4A1* gene of the proband, at base 1,766, in which G was mutated to T, resulting in the mutation of the amino acid residue at position 589 from arginine to leucine. The mutation was not found in either of his parents and was a *de novo* mutation. After searching the SNP website (https://www.ncbi.nlm.nih.gov/snp/, retrieval date July 2020), we found that the R589 L mutation in the proband had not been reported before, and we named it *SLC4A1* Quanzhou according to nomenclature by city (3, 4). In addition, a missense mutation (c.1028 A>G, p.H343R) was found in exon 9 of *FGFR1* gene in the proband, which was also a *de novo* mutation. A missense mutation (c.296G>A, p.R2329H) was found in exon 3 of *KAT6B* gene, which was observed in the father, so it was a hereditary mutation. Similarly, these two mutations have not been reported. We named them *FGFR1* Quanzhou and *KAT6B* Quanzhou. The other four gene mutation sites were noted in previous reports, as shown in **Table 3**. The inclusion status and frequency of the three "pathogenic" genes in the database are shown in **Table 4**.

According to the American College of Medical Genetics and Genomics (ACMG) guidelines (2015), *SLC4A1* gene mutation was mutated as "pathogenic (PS2 + PM1 + PM2 + PM5 + PP3)," *FGFR1* gene mutation was mutated as "likely pathogenic (PS2 + PM1 + PM2 + PP3)," and *G6PD* gene mutation was mutated as "likely pathogenic (PS1 + PM1 + PP3)." The evidence of the remaining four genes becoming pathogenic was insufficient, but possible pathogenic variation was not ruled out (**Table 3**). The verification results of gene mutations with high pathogenicity are shown in **Figures 1**–**3**.

DISCUSSION

SLC4A1 Gene Mutation: Pathogenic

The 17-kb *SLC4A1* gene is located on chromosome 17q21 and has 20 exons. It encodes ion exchange protein anion exchanger 1 (AE1; also called band 3 protein), first reported in 1972 (5). More than 10 mutations have been reported (6) that comprise multiple mutations of 589 loci, including R589 L, R589H, R589C, and R589S, and are considered hot spot mutations. In previous reports, *SLC4A1* gene mutation is observed with hereditary dRTA, hereditary spherocytosis, and Southeast Asian ovalocytosis.

SLC4A1, which encodes AE1, has two subtypes. The eAE1 subtype is expressed in red blood cells and plays a role in maintaining the normal morphology of red blood cells. The KAE1 subtype is expressed in the kidney and plays a role in transporting Cl^-/HCO_3^-. After gene mutation, the activity of Cl^-/HCO_3^- transport was affected, and the acidification function of distal renal tubules was impaired, which led to

TABLE 2 | Timeline for the treatment and follow-up.

Time	Length (cm)	Weight (kg)	pH (7.35–7.45)	PCO_2 (mmHg) (34.8–44.9)	PO_2 (mmHg) (79.80–100.00)	BE (mmol/L) (−3.0 to 3.0)	HCO^{3-} (mmol/L) (21.4–27.3)	K^+ (mmol/L) (3.5–5.3)	Cl^- (mmol/L) (96–108)	Treatment
Outpatient visits 2019.6.14 (15 days)	49.0	3.3	Visited due to failing to pass newborn disease screening; G-6PD deficiency and small penis were found							Advised that the child avoid exogenous drugs such as primaquine or fava beans
Hospitalized 2020.6.19 (1 year)	63.9	6.0	7.268	20.9	53.9	−15.3	9.3	2.75	110.2	Sodium bicarbonate 20 ml qd 7 days + 10% KCl 15.5 ml qd 4 days, later changed to potassium citrate 2 g tid p.o. Medicine after discharge: potassium citrate 2 g p.o. tid
Follow-up 2020.8 (1 year 2 months)	64.5	6.5	7.41	27.6	74.7	−5.3	17.5	3.9	105.1	Potassium citrate 2 g p.o. tid
Follow-up 2020.12 (1 year 6 months)	73.5	9.0	Due to the steady increase in the length and weight of the child, and there was no complaint of discomfort; the blood gas and electrolytes were not reviewed							Potassium citrate 2 g p.o. tid
Follow-up 2021.4.20 (1 year 10 months)	78.5	11.0	7.40	30.0	76.1	−5.0	18.7	3.8	103.2	Potassium citrate 2 g p.o. tid

TABLE 3 | Results of exome sequencing.

	Chromosomal location	Nucleic acid change (exon no.)	Amino acid changes	RS no.	ACMG pathogenicity grade	Proband (male)	Father (normal)	Mother (normal)	Related diseases (OMIM), genetic pattern
SLC4A1	Chr17:42333075	c.1766 (exon 14) G>T	p.R589L (NM_000342)	No items found	Pathogenic	Heterozygosity (40/75)	Wild type (0/40)	Wild type (0/630)	Renal tubular acidosis, distal (OMIM: 179800), AD
FGFR1	Chr8:38279344	c.1028 (exon 9) A>G	p.H343R (NM_001174064)	No items found	Likely pathogenic	Heterozygosity (32/65)	Wild type (0/41)	Wild type (0/54)	Hypogonadotropic hypogonadism 2 with or without anosmia (OMIM: 147950), AD Hartsfield syndrome (OMIM: 615465), AD
G6PD	ChrX:153760484	c.1466 (exon 12) G>T	p.R489L (NM_000402)	rs 72554665	Likely pathogenic	Hemizygote (49/49)	Wild type (0/52)	Heterozygosity (53/101)	Hemolytic anemia, *G6PD* deficient (OMIM: 300908), XLD
GLI3	Chr7:42116438	c.386 (exon 4) C>G	p.P129R (NM_000168)	rs 1276292491	Uncertain	Heterozygosity	Wild type	Heterozygosity	Pallister–Hall syndrome (OMIM: 146510), AD
BCOR	ChrX:39934000	c.599 (exon 4) C>T	p.T200M (NM_001123385)	rs 777945715	Uncertain	Hemizygote	Wild type	Heterozygosity	Microphthalmia, syndromic 2 (OMIM: 300166), XLD
FAT4	Chr4:126337745	c.6986 (exon 6) G>A	p.R2329H (NM_024582)	rs 754622270	Likely benign	Heterozygosity	Wild type	Heterozygosity	Hennekam lymphangiectasia-lymphedema syndrome 2 (OMIM: 616006), AR Van Maldergem syndrome 2 (OMIM: 615546), AR
KAT6B	Chr10:76602911	c.296 (exon 3) G>A	p.C99Y (NM_012330)	No items found	Uncertain	Heterozygosity	Heterozygosity	Wild type	Genitopatellar syndrome (OMIM: 606170), AD

ACMG, American College of Medical Genetics and Genomics; AD, autosomal dominant; XLD, X-linked dominant; AR, autosomal recessive.

TABLE 4 | The inclusion status and frequency of the three "pathogenic" genes in the database.

Gene	dbSNP	1,000 Genomes Project	Thousands of people in the South	Thousands of people in the North	Genome AD	Genome AD in East Asia
SLC4 A1	Not included	Not included	Not included	Not included	Not included	Not included
FGFR1	Not included	Not included	Not included	Not included	Not included	Not included
G6PD	0.00045	0.0031	0.00	0.0063	0.0083	0.0083

acidosis and increased urinary pH-value. Distal renal tubular acidosis caused by *SLC4A1* gene mutation is usually caused by autosomal dominant (AD) inheritance. However, in equatorial countries such as Thailand, Malaysia, the Philippines, and Papua New Guinea, autosomal recessive inheritance is observed (7). In addition, it was reported that patients with the A858D homozygous mutation of *SLC4A1* had manifestations of both hereditary spherocytosis and dRTA (8).

The main clinical feature of dRTA is that the disorder of H^+ secretion in the distal renal tubules causes renal acidification dysfunction, leading to high chlorine metabolic acidosis with normal anions, which is often accompanied by hypokalemia, hypercalciuria, nephrocalcinosis, and kidney stones. With the occurrence of acidosis, growth retardation and rickets can occur and have been observed in children with variants in *SLC4A1* p.R589C, p.R589H, and p.R589S. The proband presented with typical manifestations of acidosis, high blood chlorine, low blood potassium, renal calcinosis, and slow growth; and after repeated correction of acidosis and potassium supplementation, the blood gas returned to normal. Previous studies have reported that children with dRTA have a good prognosis after reasonable treatment. However, due to the possibility of deterioration of renal function in adulthood, especially after puberty, it is necessary to be vigilant and follow up regularly. In addition to *SLC4A1* gene described in this case, the hereditary pathogenic genes that have been identified to cause dRTA can also be found in *ATP6V1B1* and *ATP6V0A4* genes. Studies have found that patients with dRTA caused by

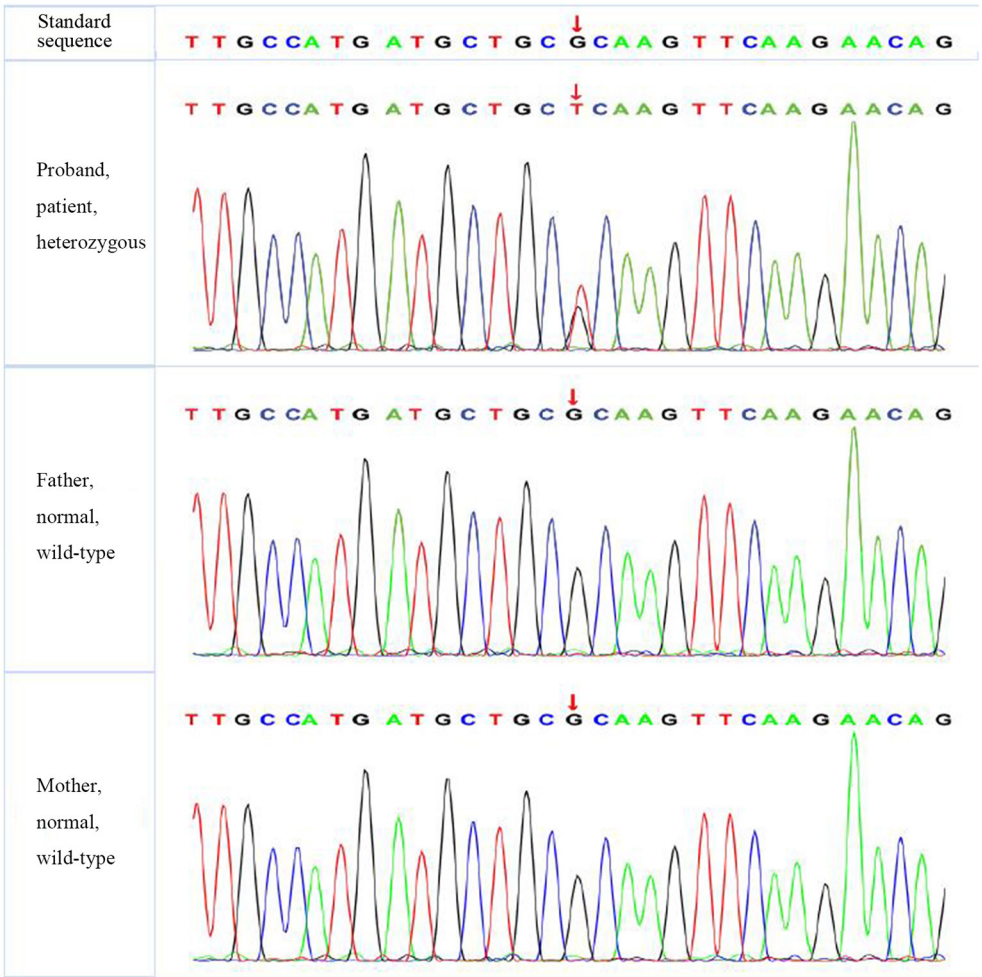

FIGURE 1 | The verification results of *SLC4A1*. *SLC4A1*:c.1766 (exon 14) G>T.

SLC4A1 gene mutations are mostly diagnosed in adolescence, while those caused by *ATP6V1B1* and *ATP6V0A4* gene mutations are more commonly identified in infancy and early childhood (9). The age of onset in this study (1 year) was significantly younger than that of *SLC4A1* gene mutations reported in previous reports.

FGFR1 Gene Mutation: Likely Pathogenic

Fibroblast growth factor receptor 1 (*FGFR1*), also known as acidic fibroblast growth factor receptor (aFGFR), is located on chromosome 8p12 and has 19 exons and a molecular weight of 78 kDa. Its mutation or enhanced expression can lead to changes in protein tyrosine kinase activity, thereby enhancing tyrosine kinase phosphorylation and its downstream effects. *FGFR1* is involved in the growth, differentiation, migration, apoptosis, angiogenesis, and drug resistance of tumor cells (10). Previous reports focused on the role of *FGFR1* in the field of cancer. *FGFR1* gene amplification seems to be one of the most common genetic changes in tumors and can be seen in gastric cancer, oral squamous cell carcinoma, ovarian cancer, bladder cancer, and breast cancer. *FGFR1* also plays an important role in the development of the reproductive and nervous systems. The role of nerve cell development in the embryonic stage may be related to the development of GnRH neurons and olfactory neurons. *FGFR1* protein defects can cause GnRH neuromigration and abnormal olfactory bulb development.

FGFR1 gene mutation was reported to be associated with Kallmann syndrome (KS), which is a disease that results in idiopathic hypogonadotropic hypogonadism (IHH) combined with anosmia or hyposmia (11). Some patients with hypogonadotropic hypogonadism 2 with or without anosmia (IHH) can present specific non-reproductive phenotypes, including cryptorchidism, hypogonadotropic hypogonadism, congenital penile, olfactory loss, osteopenia, developmental delay, agenesis of the corpus callosum, bimanual synkinesis, short stature, hearing impairment, unilateral (occasionally bilateral) renal hypoplasia, cleft lip and palate, loss of teeth, and obesity (12). At present, the common causes of IHH are *FGFR1* and KAL1 gene mutations; in addition, the *FGFR1* gene is associated with Hartsfield syndrome (13).

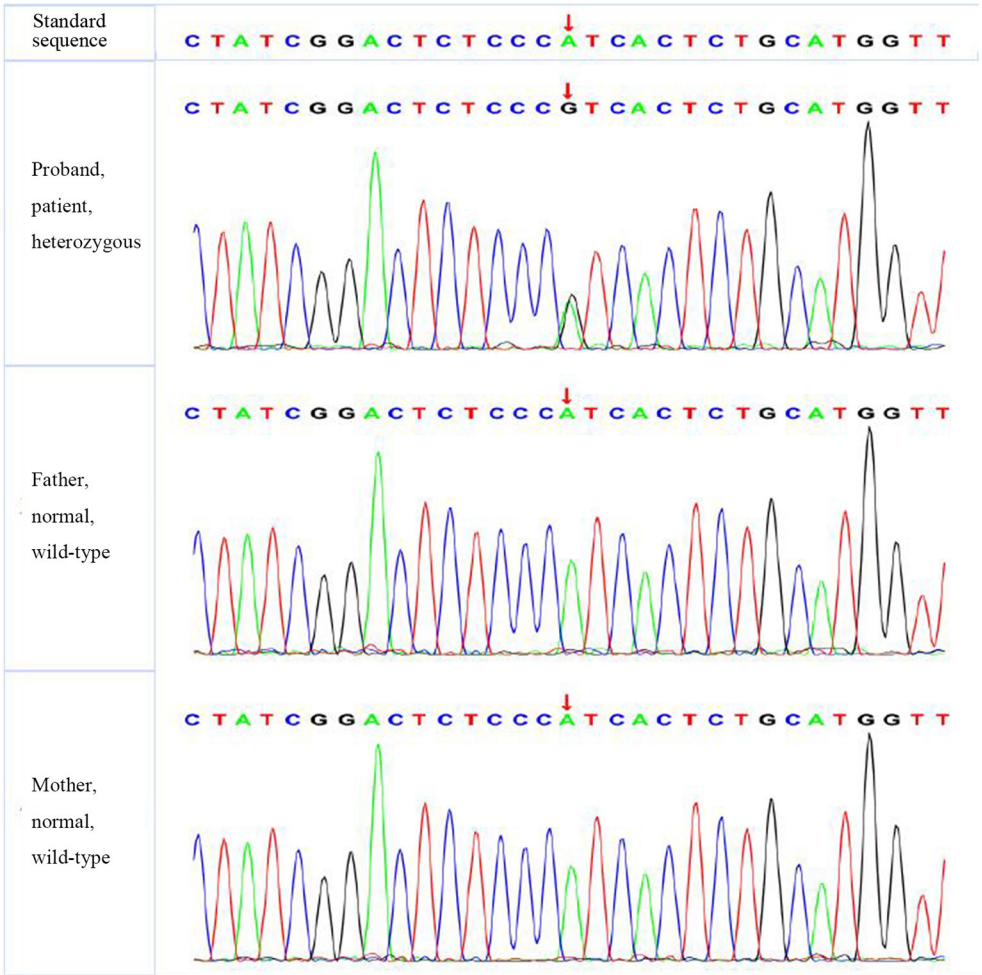

FIGURE 2 | The verification results of *FGFR1*. *FGFR1*:c.1028 (exon 9) A>G.

In this case, the proband had a congenital small penis (micropenis), which means that the stretched penile length (SPL) was lower than the average of the population of the same age or the same developmental status by 2.5 standard deviations or more, accompanied by deformities such as cryptorchidism, small-volume testes, and scrotal hypoplasia. The normal value of the penis length of newborns measured by Feldman in the United States is 28–42 mm (14). Some researchers in the West have diagnosed a small penis with a penis length of <2 cm in full-term newborns. The normal penile length of full-term newborns in Asia is 26–46 mm, and it is considered that penile length <26 mm is a small penis (15).

Most *FGFR1* gene mutations were observed in male patients and were usually found in adolescence, rarely in infancy. The proband had an *FGFR1* gene heterozygous *de novo* mutation, which is consistent with the pathogenesis of AD inheritance, with clinical manifestations of congenital penile, left cryptorchidism, and growth retardation and no significant increase in testosterone in the hCG provocation test, suggesting that the genotype is consistent with the phenotype. According to the ACMG guidelines (2015), the mutation is likely to be pathogenic. In addition, *FGFR1* gene is inherited in an AD manner, so heterozygous mutations can cause disease. In this case, *FGFR1* gene of the proband is a heterozygous mutation. His parents should be wary that there is a possibility of subsequent development of IHH or Hartsfield syndrome.

G6PD Gene Mutation: Likely Pathogenic

The human *G6PD* gene, located on chromosome Xq28, has 13 exons and spans 18 kb. *G6PD* gene plays a critical role in the production of ribose 5-phosphate and the generation of NADpH in the hexose monophosphate pathway. To date, there have been more than 140 mutations in *G6PD* gene found worldwide (16), and its genotype has the characteristics of regional mutations (17). Previous studies have confirmed that *G6PD* deficiency has a protective effect for malaria (18). With some environmental factors, such as infection and certain drugs and foods, *G6PD* deficiency may lead to hemolysis. It is the most common group of human enzyme deficiency

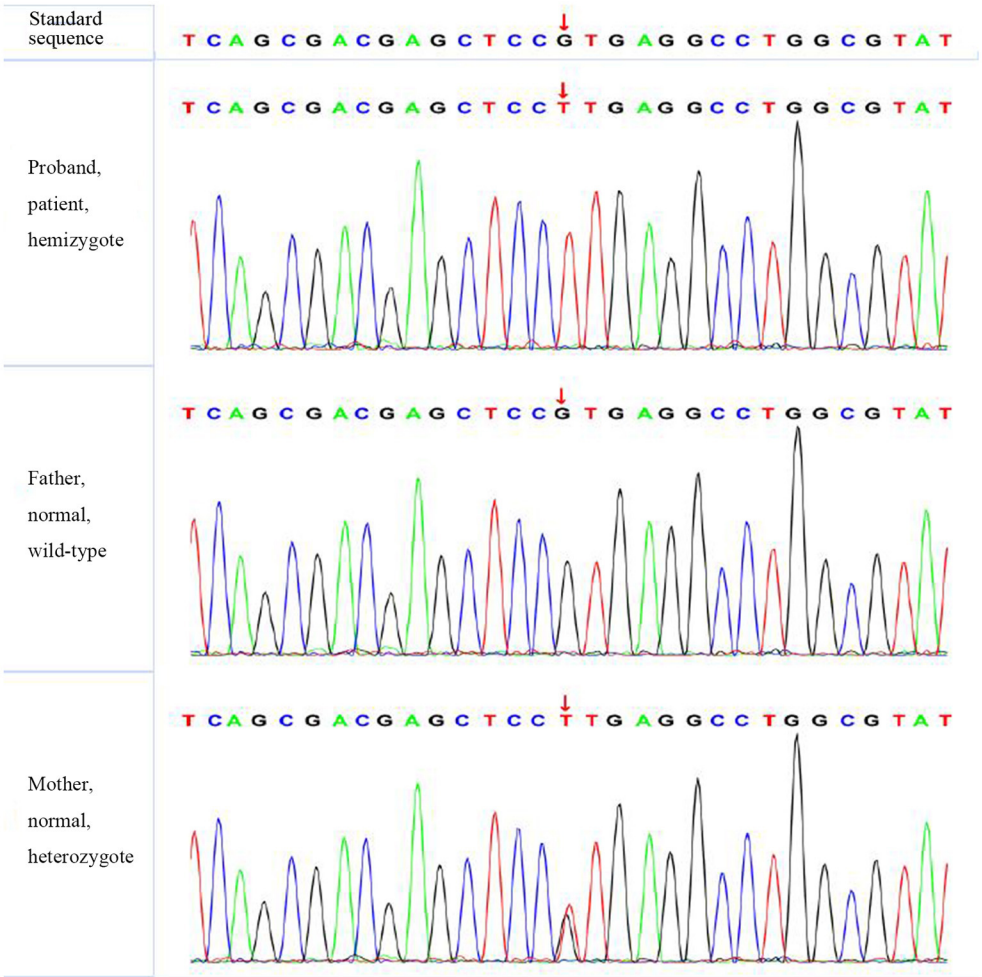

FIGURE 3 | The verification results of *G6PD*. *G6PD*:c.1466 (exon 12) G>T.

syndrome and follows an X-linked incomplete dominant inheritance model.

Most *G6PD* deficiency patients are asymptomatic for life, but it can also be life-threatening, especially in children (19). The most common clinical manifestations of *G6PD* deficiency are neonatal jaundice and acute hemolytic anemia caused by exogenous drugs such as primaquine or fava beans. No persistent jaundice or clinical manifestations such as obvious hemolytic anemia were found in the neonatal period of the proband. He was discovered due to "Chinese abnormal screening of neonatal diseases" upon his first visit and was confirmed by ES by chance. The mutation was not found in the proband's father, and a heterozygous mutation was found in his mother, so he was considered to be a hemizygote.

No Sufficient Evidence of Pathogenicity, but Variation With Pathogenicity Cannot Be Excluded

GLI3 gene mutation was reported to be associated with Pallister–Hall syndrome (pHS; MIM 146510) (20), which was characterized by polydactyly, micropenis with undescended or hydroxyplastic testes in males, and urinary dysplasia. *BCOR* gene mutation is related to oculofaciocardiodental (OFCD) syndrome (OMIM 300166) (21), with clinical manifestations of microphthalmia, congenital cataracts, radiculomegaly, and cardiac and digital abnormalities, as well as mild developmental delay, delayed motor development, and short stature. *FAT4* gene mutation can lead to Hennekam lymphangiectasia-lymphedema syndrome 2 (OMIM 616006), and Van Maldergem syndrome 2 (OMIM 615546) (22). The former is mainly manifested as lymphedema, claudication dementia, and developmental delay, while the latter is characterized by intellectual disability, and some may suffer from renal hyperplasia. Both of these diseases may have intellectual disability and developmental delay. In addition, small kidney and genital abnormalities, including micropenis, cryptorchidism, and double scrotum, have been observed in Van Maldergem syndrome (23). *KAT6B* gene mutation is relevant to genitopatellar syndrome (OMIM 606170) (24), with main manifestations of patellar loss, congenital flexion contracture of lower limbs,

psychomotor disorders, and abnormalities of external genitalia and kidneys.

In our study, the proband had the clinical manifestations of delayed growth and development, slow height and weight gain, cryptorchidism, and micropenis, which may have a certain degree of crossover with the four gene-related diseases mentioned above. According to the ACMG guidelines, the four gene mutations were identified as having insufficient evidence of pathogenicity, but possible pathogenic variation was not excluded.

Polygenic Mutation

Many human diseases are likely caused by single-gene and/or multiple-gene mutations, which regulate the occurrence and development of diseases through gene interactions, thereby affecting human health (25). It is not uncommon to have polygenic mutations in the same patient. However, the mechanism of these diseases and the principle of gene interaction are not clear. Multiple mutations tend to have different effects on genes compared with single mutations.

Multiple-gene mutations associated with the same disease have been reported, with incidences of digenic and trigenic variants of 8/48 (16.7%) and 1/48 (2.1%), respectively (26). Various gene mutations may work together in the same pathway to affect the phenotype of the disease, while they may have no or only minimal effect on the phenotype when these heterozygous mutations exist alone.

Multiple mutations in the same gene, also known as compound mutations, have been found to be associated with various genetic diseases. Some studies have also reported that the combined effect of multiple mutations in the same gene leads to more serious cardiovascular disease than a single mutation (27). They believe that these multiple-gene mutations have additive effects on cardiovascular pathology and function and highlight the importance of basic molecular, cellular, and animal model studies in elucidating key pathogenic pathways. However, some studies suggest that multigene mutations can either aggravate the development of the disease or lighten it; the latter is known as a compensatory mutation (28).

Our case is different from the above two cases. In this case, seven gene mutations were detected by the whole-exome test, among which one was pathogenic and two were likely to be pathogenic. The other four genes with insufficient evidence need to be further confirmed. These gene mutations are not the cause of the same disease, and there is no correlation among them that has been reported. However, the cause of this patient's polygenic mutation is unclear. The existence of multiple mutations directly affects the strategy of gene diagnosis. It is important to study the combined effect of multiple mutations. Thus, we look forward to more research on the mechanism of multigene mutations.

Inevitably, there were some limitations to this study. First, we did not have the capacity to conduct functional research; thus, missense mutations could not be verified. In addition, this article focuses on the results of exon detection of this rare polygenic mutation; thus, only the details related to the clinical phenotype and treatment of this child were discussed, not all the relevant details of the related diseases.

The discovery of rare genetic variation through different gene sequencing methods is a very challenging subject in the field of human genetics. Our research confirms the role of ES in the diagnosis of dRTA. With the help of ES, case detection, and diagnosis become easier.

CONCLUSIONS

The child needs to take bismuth potassium citrate 2 g qd orally for life, and his parents should be wary of the possibility of subsequent development of IHH or Hartsfield syndrome. In addition, the child should avoid exogenous drugs such as primaquine or fava beans. Our case expanded the gene mutation spectrum and enriched the human gene mutation library. With the expansion of newborn screening programs, including for citrullinemia and *G6PD* deficiency, numerous asymptomatic infants have been identified, aiding in treatment decisions and genetic counseling. In addition, our study extends the mutation spectrum of dRTA and is helpful in early molecular diagnoses of dRTA. In clinical work, attention should be given to the follow-up of multiple-gene mutations, especially asymptomatic infants identified in infancy.

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 Quanzhou Women and Children's Hospital Ethics Committee. 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 participants' legal guardian/next of kin for the publication of this case report.

AUTHOR CONTRIBUTIONS

YL and ZZ were involved in two sessions with the patient. QZ analyzed the sequencing data and performed the literature research. QZ wrote the first draft of the report, which was edited by all authors. All authors reviewed and agreed to the final version of the article. QZ had the final responsibility for the decision to submit for publication.

ACKNOWLEDGMENTS

The authors thank Shengping Yang of Quanzhou Orthopedic-traumatological Hospital (Master of Medicine), the husband of the corresponding author (QZ), for providing assistance with comments on revising the manuscript.

REFERENCES

1. Tan TY, Dillon OJ, Stark Z, Schofield D, Alam K, Shrestha R, et al. Diagnostic impact and cost-effectiveness of whole-exome sequencing for ambulant children with suspected monogenic conditions. *JAMA Pediatr.* (2017) 171:855–62. doi: 10.1001/jamapediatrics.2017.1755
2. Theunissen TE, Sallevelt SC, Hellebrekers DM, De Koning B, Hendrickx AT, van den Bosch BJ, et al. Rapid resolution of blended or composite multigenic disease in infants by whole-exome sequencing. *J. Pediatrics.* (2017) 182:371.e2–4.e2. doi: 10.1016/j.jpeds.2016.12.032
3. Jomoui W, Panichchob P, Rujirachaivej P, Panyasai S, Tepakhan W. Coinheritance of Hb A 2-Melbourne (HBD: c.130G>A) and Hb E (HBB: c.79G>A) in Laos and simultaneous high resolution melt detection of Hb A 2-Melbourne and Hb A 2-Lampang (HBD: c.142G>A) in a single tube. *Hemoglobin.* (2019) 43: 214–7. doi: 10.1080/03630269.2019.1651332
4. Shao L, Yan X, Dong Q, Lang Y, Yue S, Miao Z, et al. novel *SLC4A1* variant in an autosomal dominant distal renal tubular acidosis family with a severe phenotype. *Endocrine.* (2010) 37:473–8. doi: 10.1007/s12020-010-9340-6
5. Richards P, Wrong OM. Dominant inheritance in a family with familial renal tubular acidosis. *Lancet.* (1972) 2:998–9. doi: 10.1016/s0140-6736(72)92406-3
6. Deejai N, Wisanuyotin S, Nettuwakul C, Khositseth S, Sawasdee N, Saetai K, et al. Molecular diagnosis of solute carrier family 4 member 1 (*SLC4A1*) mutation-related autosomal recessive distal renal tubular acidosis. *Lab Med.* (2019) 50:78–86. doi: 10.1093/labmed/lmy051
7. Khositseth S, Bruce LJ, Walsh SB, Bawazir WM, Ogle GD, Unwin RJ, et al. Tropical distal renal tubular acidosis: clinical and epidemiological studies in 78 patients. *QJM.* (2012) 105:861–77. doi: 10.1093/qjmed/hcs139
8. Shmukler BE, Kedar PS, Warang P, Desai M, Madkaikar M, Ghosh K, et al. Hemolytic anemia and distal renal tubular acidosis in two Indian patients homozygous for *SLC4A1*/AE1 mutation A858D. *Am J Hematol.* (2010) 85:824–8. doi: 10.1002/ajh.21836
9. Palazzo V, Provenzano A, Becherucci F, Sansavini G, Mazzinghi B, Orlandini V, et al. The genetic and clinical spectrum of a large cohort of patients with distal renal tubular acidosis. *Kidney Int.* (2017) 91:1243–55. doi: 10.1016/j.kint.2016.12.017
10. Yu T, Yang Y, Liu Y, Zhang Y, Xu H, Li M, et al. A *FGFR1* inhibitor patent review: progress since 2010. *Expert Opin Ther Pat.* (2017) 27:439–54. doi: 10.1080/13543776.2017.1272574
11. Choi JH, Oh A, Lee Y, Kim GH, Yoo HW. Functional characteristics of novel *FGFR1* mutations in patients with isolated gonadotropin-releasing hormone deficiency. *Exp Clin Endocrinol Diabetes.* (2020) 129:457–63. doi: 10.1055/a-1151-4800
12. Hardelin JP, Dodé C. The complex genetics of Kallmann syndrome: KAL1, *FGFR1*, FGF8, PROKR2, PROK2, et al. *Sex Dev.* (2008) 2:181–93. doi: 10.1159/000152034
13. Courage C, Jackson CB, Owczarek-Lipska M, Jamsheer A, Sowińska-Seidler A, Piotrowicz M, et al. Novel synonymous and missense variants in *FGFR1* causing Hartsfield syndrome. *Am J Med Genet A.* (2019) 179:2447–53. doi: 10.1002/ajmg.a.61354
14. Feldman KW, Smith DW. Fetal phallic growth and penile standards for newborn male infants. *J Pediatr.* (1975) 86:395–8. doi: 10.1016/s0022-3476(75)80969-3
15. Lian WB, Lee WR, Ho LY. Penile length of newborns in Singapore. *J Pediatr Endocrinol Metab.* (2000) 13:55–62. doi: 10.1515/jpem.2000.13.1.55
16. Cappellini MD, Fiorelli G. Glucose-6-phosphate dehydrogenase deficiency. *Lancet.* (2008) 371:64–74. doi: 10.1016/s0140-6736(08)60073-2
17. Lo E, Zhong D, Raya B, Pestana K, Koepfli C, Lee MC, et al. Prevalence and distribution of *G6PD* deficiency: implication for the use of primaquine in malaria treatment in Ethiopia. *Malar J.* (2019) 18:340. doi: 10.1186/s12936-019-2981-x
18. Min-Oo G, Gros P. Erythrocyte variants and the nature of their malaria protective effect. *Cell Microbiol.* (2005) 7:753–63. doi: 10.1111/j.1462-5822.2005.00524.x
19. Luzzatto L, Arese P. Favism and glucose-6-phosphate dehydrogenase deficiency. *N Engl J Med.* (2018) 378:60–71. doi: 10.1056/nejmra1708111
20. Kariminejad A, Ghaderi-Sohi S, Keshavarz E, Hashemi SA, Parsimehr E, Szenker-Ravi E, et al. A *GLI3* variant leading to polydactyly in heterozygotes and Pallister-Hall-like syndrome in a homozygote. *Clin Genet.* (2020) 97:915–9. doi: 10.1111/cge.13730
21. Ng D, Thakker N, Corcoran CM, Donnai D, Perveen R, Schneider A, et al. Oculofaciocardiodental and Lenz microphthalmia syndromes result from distinct classes of mutations in *BCOR*. *Nat Genet.* (2004) 36:411–6. doi: 10.1038/ng1321
22. Alders M, Al-Gazali L, Cordeiro I, Dallapiccola B, Garavelli L, Tuysuz B, et al. Hennekam syndrome can be caused by FAT4 mutations and be allelic to van Maldergem syndrome. *Hum Genet.* (2014) 133:1161–7. doi: 10.1007/s00439-014-1456-y
23. Neuhann TM, Müller D, Hackmann K, Holzinger S, Schrock E, Di Donato N, et al. further patient with van Maldergem syndrome. *Eur J Med Genet.* (2012) 55:423–8. doi: 10.1016/j.ejmg.2012.02.012
24. Okano S, Miyamoto A, Fukuda I, Tanaka H, Hata K, Kaname T, et al. Genitopatellar syndrome: the first reported case in Japan. *Hum Genome Var.* (2018) 5:8. doi: 10.1038/s41439-018-0010-1
25. Loeb KR, Loeb LA. Significance of multiple mutations in cancer. *Carcinogenesis.* (2000) 21:379–85. doi: 10.1093/carcin/21.3.379
26. Quaynor SD. Bosley ME2, Duckworth CG, Porter KR, Kim SH, Kim HG, et al. Targeted next generation sequencing approach identifies eighteen new candidate genes in normosmic hypogonadotropic hypogonadism and Kallmann syndrome. *Mol Cell Endocrinol.* (2016) 437:86–96. doi: 10.1016/j.mce.2016.08.007
27. Kelly M, Semsarian C. Multiple mutations in genetic cardiovascular disease: a marker of disease severity? *Circ Cardiovasc Genet.* (2009) 2:182–90. doi: 10.1161/CIRCGENETICS.108.836478
28. Liu M, Watson LT, Zhang L. Predicting the combined effect of multiple genetic variants. *Hum Genomics.* (2015) 9:18. doi: 10.1186/s40246-015-0040-4

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

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: 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