COPY NUMBER VARIATION IN RARE DISORDERS

EDITED BY: Judit Bene, Katalin Komlosi and Attila Gyenesei
PUBLISHED IN: Frontiers in Genetics and Frontiers in Pediatrics

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-018-3
DOI 10.3389/978-2-88976-018-3

About Frontiers

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

Frontiers Journal Series

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

Dedication to Quality

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

What are Frontiers Research Topics?

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

COPY NUMBER VARIATION IN RARE DISORDERS

Topic Editors:
Judit Bene, University of Pécs, Hungary
Katalin Komlosi, Medical Center University of Freiburg, Germany
Attila Gyenesei, University of Pécs, Hungary

Citation: Bene, J., Komlosi, K., Gyenesei, A., eds. (2022). Copy Number Variation in Rare Disorders. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-018-3

Table of Contents

frontiers | Frontiers in Genetics

EDITORIAL
published: 05 April 2022
doi: 10.3389/fgene.2022.898059

Editorial: Copy Number Variation in Rare Disorders

Katalin Komlósi[1], *Attila Gyenesei*[2] *and Judit Bene*[3]*

[1]Institute of Human Genetics, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany, [2]Bioinformatics Research Group, Genomics and Bioinformatics Core Facility, Szentagothai Research Center, University of Pécs, Pécs, Hungary, [3]Department of Medical Genetics, Clinical Centre, Medical School, University of Pécs, Pécs, Hungary

Keywords: copy number variation (CNV), rare disorders, genomic disorders, Mendelian disease, genomic rearrangement

Editorial on the Research Topic

Copy Number Variation in Rare Disorders

Copy number variation (CNV), encompassing losses or gains of relatively large genomic DNA segments, is one of the major sources of genetic diversity in humans (Zhang et al., 2009). Recent studies revealed that *de novo* locus-specific mutation rates appear much higher for CNVs than for SNPs (Lupski, 2007; Turner et al., 2008). CNVs comprise approximately 12–16% of the human genome and 3 to 7 rare CNVs can be found in an average genome (Harel and Lupski, 2018). The frequency of a CNV shows strong anticorrelation with its size and its gene density (Itsara et al., 2009). Although several CNVs are presumably benign, the role of CNVs in the pathogenesis of various diseases has increasingly gained attention nowadays thanks to the multiple sophisticated molecular laboratory technologies capable detecting various CNVs. Benign CNVs are frequently small, intergenic, or comprise genes that can tolerate copy number changes. Pathogenic CNVs are significantly enriched for genes involved in development and genes with constrained evolutionary patterns of gene duplication and loss (Rice and McLysaght, 2017). At the early era of CNV detection large CNVs (>500 kb) appeared to be associated with genomic disorders only; however, it is now clear that CNVs can also be involved in susceptibilities to complex traits, and nowadays there is an emerging evidence that CNVs may cause Mendelian diseases or sporadic traits as well (Zhang et al., 2009; Harel and Lupski, 2018).

The disease-causing genomic rearrangements can be either recurrent or non-recurrent. Recombination-based as well as replication-based mechanisms have been proposed to be responsible for the formation of CNVs such as nonallelic homologous recombination (NAHR), non-homologous end-joining (NHEJ), L1-mediated retrotransposition or Fork Stalling and Template Switching (FoSTeS) (Kazazian and Moran, 1998; Lupski and Stankiewicz, 2005; Korbel et al., 2007). There is a variety of molecular mechanisms by which CNVs can lead to abnormal phenotypes, encompassing dosage sensitivity of a gene within the CNV, gene interruption or gene fusion at the breakpoint junctions, deletion of a regulatory element, or unmasking of a recessive allele or functional polymorphism (Lupski and Stankiewicz, 2005). Moreover, CNVs can affect noncoding regulatory elements such as promoters or enhancers as well (Harel and Lupski, 2018).

The goal of this Research Topic was to provide the cutting-edge knowledge of CNVs leading to the development of rare disorders. Rare diseases are conditions that affect less than 5 in 10,000 people. To date more than 7,000 entities exist and the numbers are continuously increasing. Today little is known about the genetic background of still a significant portion of rare diseases, therefore their diagnostics is challenging. Furthermore, patients with undiagnosed genetic diseases often face a diagnostic odyssey, which lasts for 8 years on average. CNVs were initially proposed to represent a significant contribution to rare disease formation; however, there is now evidence from a recent study that CNV should be responsible for the disease phenotype in approx. 10% of cases (Truty et al., 2019).

OPEN ACCESS

Edited and reviewed by:
Stephen J. Bush,
University of Oxford, United Kingdom

***Correspondence:**
Judit Bene
bene.judit@pte.hu

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *16 March 2022*
Accepted: *18 March 2022*
Published: *05 April 2022*

Citation:
Komlósi K, Gyenesei A and Bene J (2022) Editorial: Copy Number Variation in Rare Disorders. Front. Genet. 13:898059. doi: 10.3389/fgene.2022.898059

The following topics were considered in this special issue:

- novel CNVs detected in rare disorders
- state-of-the-art technology for detection or evaluation of CNVs
- functional or animal studies related to the functional validation of CNVs
- comparative studies revealing phenotype–genotype correlation
- mechanism of non-coding CNVs in rare disorders and genotype-phenotype correlation

For our special topic we received five case reports, twelve original research reports and one review article, from which one review article, nine original research reports and three case reports were accepted for publication.

Pócza et al. provided a comprehensive review of the landscape of germline structural variation (SV) types and the various methodologies capable detecting SVs mainly focusing on cancer predisposition genes.

Xiao et al. investigated the phenotype of pediatric patients with 17q12 duplication syndrome. They demonstrated first in the literature that annular pancreas can be observed in approx. 20% of this patient cohort. Moreover, among the 15 genes encompassed within the 17q12 recurrent duplication/deletion region they verified the role of the *HNF1B* gene in pancreatic development using zebrafish studies.

Song et al. studied the association between a common prenatal ultrasound soft marker, echogenic intracardiac focus (EIF), and chromosomal abnormalities in pregnancies. No correlation was found between the appearance of isolated EIFs in early or mid-trimester and an increased risk of fetal chromosomal abnormalities. However, the persistence of EIFs in late trimester showed an association with a higher risk of pathology-related CNVs and may indicate heart development defects after birth.

Czakó et al. investigated the rare duplication of the Xp11.23p11.22 region in female patients with intellectual disability, epilepsy and minor anomalies. Based on their phenotypic and molecular cytogenetic data they concluded that Xp11.23p11.22 duplication can result in a neurodevelopmental disorder in females. A comparison of the studied patients with others reported so far clearly demonstrates that in addition to the breakpoints of the duplication and the role of the genes involved a number of other factors influencing gene expression may affect the symptoms observed in females with the Xp11.23p11.22 duplication.

Zodanu et al. investigated 22q11.2 copy number variations in pediatric and adult patients with congenital heart disease (CHD). Their data further confirmed previous findings that demonstrated high phenotypic diversity in 22q11.2 CNV carriers. Their results highlight the necessity for large-scale genetic screening of CHD-patients and the importance of early genetic diagnosis in their clinical management.

Cai et al. presented a family with Birt-Hogg-Dubé syndrome and a novel intragenic deletion spanning exons 10–14 in the *FLCN* gene detected by targeted next generation sequencing as a rare cause of the disease. Moreover, they demonstrated that the precision and accuracy of the applied NGS approach is similar to that of the MLPA (multiplex ligation-dependent probe amplification) technique.

Büki et al. performed genotype-phenotype association analyses in mostly paediatric patients with type-1, type-2, and atypical *NF1* microdeletions. In this study three novel atypical deletions were identified. The authors established that MLPA is a feasible, cost-effective technique for the identification and the classification of the *NF1* microdeletions.

Nevado et al. performed a deep phenotyping and genetic characterization of a large patient cohort with 5p minus syndrome. Within this clinically heterogeneous syndrome, around 39% of the studied patients carried clinically significant additional genomic rearrangements, mainly a duplication in other chromosomes which may explain part of the broad clinical spectrum.

Piras et al. investigated the genomic architecture of the *CFH-CFHR* region and characterized CNVs in a large cohort of patients with C3 glomerulopathy (C3G) and immune complex-mediated membranoproliferative glomerulonephritis (IC-MPGN). They identified novel CNVs leading to structural variants in 5 C3G and 2 IC-MPGN patients.

Nevado et al. performed a thorough clinical and genetic characterization in a cohort of 210 patients with Phelan-McDermid syndrome (PMS). Multiple variant types were observed among patients, including a significant number of small deletions and *SHANK3* sequence variants. Furthermore, multiple types of rearrangements were detected among microdeletion cases, including post-zygotic mosaicism, ring chromosome 22, unbalanced translocations and additional rearrangements at 22q13 as well as other copy number variations in other chromosomes. Their findings suggest that *SHANK3* plays an important role in this syndrome, but is probably not uniquely responsible for all the features in the PMS patients.

Pan and Fu reported the clinical characteristics of a patient carrying a large 10p deletion involving the 10p15.3–10p13 region as a second reported case in the literature. The patient had facial dysmorphism, swallowing dysfunction, hypoparathyroidism, hypocalcemia and neurological abnormalities.

Chen et al. reported two siblings with familial Beckwith-Wiedemann syndrome (BWS) due to a maternal deletion in *H19* and its upstream regulatory elements which can result in loss of function of the IGF2-H19 imprinting control element in the offspring and lead to BWS. Since studies on adult BWS patients are scarce the case report gives some insight on the presentation of BWS in adulthood and some of the potential reproductive issues such as subfertility or infertility in males.

Nagy et al. reported a family with distinct severity of expressive speech disorder, mild behavioral abnormality and dysmorphic features carrying a 7.87 Mb interstitial deletion of the 7q31.1q31.31 region involving the *FOXP2* gene. They concluded that the "phenotype first" then targeted diagnostic strategy can improve the diagnostic yield of speech disorders in the routine clinical practice.

In the 13 accepted manuscripts the contribution of CNVs to disease mechanism was investigated in a great variety of rare and ultra-rare diseases. 7/13 studies contributed to new clinical and

genetic insights into rare microdeletion/microduplication syndromes while in 5/13 contributions the role of CNVs in disease mechanism was addressed in classically monogenic diseases.

AUTHOR CONTRIBUTIONS

KK, AG, and JB organized the Research Topic as guest editors, supervised the reviewing of the manuscripts and contributed equally to this Editorial paper.

REFERENCES

Harel, T., and Lupski, J. R. (2018). Genomic Disorders 20 Years On-Mechanisms for Clinical Manifestations. *Clin. Genet.* 93 (3), 439–449. doi:10.1111/cge.13146

Itsara, A., Cooper, G. M., Baker, C., Girirajan, S., Li, J., Absher, D., et al. (2009). Population Analysis of Large Copy Number Variants and Hotspots of Human Genetic Disease. *Am. J. Hum. Genet.* 84 (2), 148–161. doi:10.1016/j.ajhg.2008.12.014

Kazazian, H. H., Jr., and Moran, J. V. (1998). The Impact of L1 Retrotransposons on the Human Genome. *Nat. Genet.* 19 (1), 19–24. doi:10.1038/ng0598-19

Korbel, J. O., Urban, A. E., Affourtit, J. P., Godwin, B., Grubert, F., Simons, J. F., et al. (2007). Paired-end Mapping Reveals Extensive Structural Variation in the Human Genome. *Science* 318 (5849), 420–426. doi:10.1126/science.1149504

Lupski, J. R. (2007). Genomic Rearrangements and Sporadic Disease. *Nat. Genet.* 39 (7 Suppl. l), S43–S47. doi:10.1038/ng2084

Lupski, J. R., and Stankiewicz, P. (2005). Genomic Disorders: Molecular Mechanisms for Rearrangements and Conveyed Phenotypes. *Plos Genet.* 1 (6), e49. doi:10.1371/journal.pgen.0010049

Rice, A. M., and McLysaght, A. (2017). Dosage Sensitivity Is a Major Determinant of Human Copy Number Variant Pathogenicity. *Nat. Commun.* 8, 14366. doi:10.1038/ncomms14366

Truty, R., Paul, J., Kennemer, M., Lincoln, S. E., Olivares, E., Nussbaum, R. L., et al. (2019). Prevalence and Properties of Intragenic Copy-Number Variation in Mendelian Disease Genes. *Genet. Med.* 21 (1), 114–123. doi:10.1038/s41436-018-0033-5

FUNDING

This work was supported by grant from the Medical School, University of Pécs (KA 2020-27).

ACKNOWLEDGMENTS

The guest editors are grateful to all the authors contributing to this special issue and thank all the reviewers who helped improve the manuscripts.

Turner, D. J., Miretti, M., Rajan, D., Fiegler, H., Carter, N. P., Blayney, M. L., et al. (2008). Germline Rates of De Novo Meiotic Deletions and Duplications Causing Several Genomic Disorders. *Nat. Genet.* 40 (1), 90–95. doi:10.1038/ng.2007.40

Zhang, F., Gu, W., Hurles, M. E., and Lupski, J. R. (2009). Copy Number Variation in Human Health, Disease, and Evolution. *Annu. Rev. Genom. Hum. Genet.* 10, 451–481. doi:10.1146/annurev.genom.9.081307.164217

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

CASE REPORT
published: 19 February 2021
doi: 10.3389/fgene.2021.621096

Case Report: A Novel Deletion in the 11p15 Region Causing a Familial Beckwith–Wiedemann Syndrome

*Juan Chen, Jian Xu, Yang Yu and Ling Sun**

Department of Assisted Reproductive Technology, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

Beckwith–Wiedemann syndrome (BWS; OMIM 130650) is a human overgrowth and cancer susceptibility disorder with a wide clinical spectrum, which cannot be predicted based on genomic variants alone. Most reports on BWS cases focus on childhood patients. Studies on adult BWS patients are scarce. Our study reports a BWS family in which the disorder appears to be caused by deletion of *H19* and its upstream regulatory elements. Genetic analysis showed a heterozygous microdeletion (~chr11:2009895-2070570 (GRCh37)) in the patients. Maternal deletion in *H19* can result in loss of function of the IGF2-H19 imprinting control element, which leads to BWS. The male proband in this family was affected by the testicular anomaly and cryptorchidism. Early orchidopexy did not rescue his azoospermia, which might be not the consequence of cryptorchidism, but due to genetic defects associated with H19 deletion. In summary, our study gives some insights on the presentation of BWS in adulthood.

Keywords: Beckwith–Wiedemann syndrome, 11p15, MLPA, azoospermia, imprinting

OPEN ACCESS

Edited by:
Attila Gyenesei, University of Pécs, Hungary

Reviewed by:
Maria Paola Lombardi, University of Amsterdam, Netherlands
Theresa V. Strong, Foundation for Prader-Willi Research, United States

****Correspondence:***
Ling Sun
sunling6299@163.com

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

Received: *25 October 2020*
Accepted: *18 January 2021*
Published: *19 February 2021*

Citation:
Chen J, Xu J, Yu Y and Sun L (2021) Case Report: A Novel Deletion in the 11p15 Region Causing a Familial Beckwith–Wiedemann Syndrome. Front. Genet. 12:621096. doi: 10.3389/fgene.2021.621096

INTRODUCTION

Beckwith–Wiedemann syndrome (BWS; OMIM 130650) is a human overgrowth disorder, characterized by macrosomia, hemihyperplasia, abdominal wall defects, macroglossia, and neonatal hypoglycemia (Choufani et al., 2010). Cryptorchidism is a common symptom in male BWS patients and adult male patients usually face subfertility problems (Cohen, 1971; Kosseff et al., 1976; Taylor, 1981; Watanabe and Yamanaka, 1990; Elliott et al., 1994; Gazzin et al., 2019).

BWS is associated with genetic and epigenetic changes on the chromosome 11p15 region (Choufani et al., 2010). There are two imprinting centers in the 11p15 region (**Figure 1A**). One includes a region that encodes a long noncoding RNA (lncRNA) *H19* and insulin-like growth factor 2 (*IGF2*) and is controlled by *H19/IGF2* intergenic differentially methylated region (*H19/IGF2*: IG DMR), which is also called imprinting control region 1 (IC1). Another includes *KCNQ1*, the regulatory lncRNA *KCNQ1OT1*, and *CDKN1C*. This region is controlled by the *KCNQ1OT1* transcription start site differentially methylated region (*KCNQ1OT1*: TSS DMR), which is called imprinting control region 2 (IC2) (Hark et al., 2000; Diaz-Meyer et al., 2003; Pandey et al., 2008).

Approximately 80% of BWS patients have a molecular defect in the 11p15 region, mostly due to abnormal DNA methylation (Choufani et al., 2010). Only 15% of these cases are inherited, and nearly half of them are associated with *CDKN1C* mutations (Algar et al., 2000; Li et al., 2001; Brioude et al., 2015).

Beyond that, chromosomal duplications, deletions, and translocations of the 11p15 region contribute to about 1% of BWS cases (Niemitz et al., 2004; Sparago et al., 2004; Prawitt et al., 2005b; Krzyzewska et al., 2019). Here, we report a BWS family with an uncommon DNA aberration, which results in deletion of *H19* and its upstream regulatory genes (**Figure 1B**). The male offspring underwent bilateral orchidopexy in childhood but still developed azoospermia.

CASE PRESENTATION

Patient #1 is a 24-year-old man who came to consult about fertility due to orchidopexy in his childhood (**Figure 2A**). He was diagnosed BWS when he was born based on the features of macroglossia, abdominal wall defects, and bilateral cryptorchidism. He had surgical correction for macroglossia at 6 months of age and orchidopexy at 18 months. Now he is 193 cm tall, and his testis were smaller than usual with volume <8 ml. Multiple-semen analyses showed azoospermia. There were no other abnormalities noted on the annual health examination. Patient #2 is the younger sister of Patient #1, who is 20 years old, diagnosed as BWS as a neonate. She was born with hypoglycemia, macroglossia, and abdominal wall defects and underwent surgical correction when she was 2 years old. Abdominal ultrasound demonstrates a structurally normal uterus and ovaries. There were no other abnormalities noted on the annual health examination. Patients' parents have no physical issues and reported no drugs or abnormal environmental exposures during the pregnancy.

MATERIALS AND METHODS

All the patients in this study had been informed and gave their informed consent prior to their inclusion.

DNA was extracted from peripheral blood using QIAamp DNA Blood Mini Kit (Qiagen, USA), following the manufacturer's instructions. All DNA samples were quantified by Qubit™ 1× dsDNA HS Assay Kit (Thermo Fisher, USA).

To detect the methylation of IC1 and IC2, methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) was performed using SALSA MS-MLPA Probemix (ME030 BWS/RSS, MRC HOLLAND, Netherlands) following the manufacturer's instructions. The data were analyzed by Coffalyser.NET software (MRC HOLLAND, Netherlands) for MS-MLPA analysis.

To explore whether other genetic mutations might have contributed to cryptorchidism and azoospermia, whole-exome sequencing (WES) (Berry Genomics, China) and detection of Y chromosome microdeletions (Y chromosome microdeletion detection kit, PCR-fluorescence probe method, TOGEN, China) were performed. SNP array (HumanCytoSNP-12 BeadChip, Illumina, USA) was performed to confirm the other aberrant copy number variation in the patient.

RESULTS

MS-MLPA results showed two abnormal findings: first, the molecular genetic analysis demonstrated a heterozygous deletion of *H19* and its upstream regulatory region [chr11:2016835-2025813 (GRCh37)] in the patients and their mother (**Figure 2B**). Secondly, the ratio of IC1 methylation was abnormal and associated with a lost copy of the gene. A methylation rate of 50% would be expected in a normal person, whereas it was zero in the mother and 100% in both patients (**Figure 2B**). These results indicated that the patients' mother likely lost the methylated copy of the IC1 region, and she passed the microdeletion to her children, therefore, the two offspring possess a single methylated copy of IC1 from their father. As the patient and his young sister also demonstrated the loss of the non-methylated copy of IC1, this silenced *H19* gene expression and eventually results in BWS.

The results of Y chromosome microdeletion detection showed no microdeletion in the Y chromosome. WES results showed a heterozygous mutation c.302G>A (p.R101Q) in *PROK2* (prokineticin 2, MIM607002). PROK2 is a newly identified molecular culprit in Kallmann syndrome (KS), and it can be inherited as an autosomal dominant or recessive trait. According to the American College of Medical Genetics and Genomics (ACMG) standards and guidelines, mutation c.302G>A is likely pathogenic. However, this mutation was also found in patients' father, who does not show any symptoms related to KS. Therefore, the association of this mutation with azoospermia is uncertain.

WES and SNP array results were reanalyzed to confirm the breakpoint of the deletion in the patient. The microdeletion appeared to be located around chr11:2009895–2070570 (GRCh37), but the precise breakpoint was not detectable.

DISCUSSION

Microdeletion and microduplication are uncommon phenomena in BWS patients, accounting for <9% in familial BWS. Sparago et al. (2004) first reported the microdeletion of *H19* DMR in BWS, and several articles demonstrated chromosomal microdeletion in the imprinting center region associated with BWS (Niemitz et al., 2004; Prawitt et al., 2005a,b; Zollino et al., 2010; De Crescenzo et al., 2011; Beygo et al., 2013; Baskin et al., 2014). Two cases in Baskin's study (Baskin et al., 2014) are similar to our patient. One of them is maternal deletion of H19 and IC1 in a patient leading to BWS. The patient is female and also had macrosomia, abdominal wall defects, macroglossia, and neonatal hypoglycemia. However, this girl patient suffered from Wilms tumor, mild nephrosis, nephromegaly, and polydactyly, which were not observed in our patients. The other case in their study is a male patient and had a *de novo* deletion of H19 and IC1. The clinical data is similar to our case except for chryptorchidism and azoospermia, which have been observed in our male patient

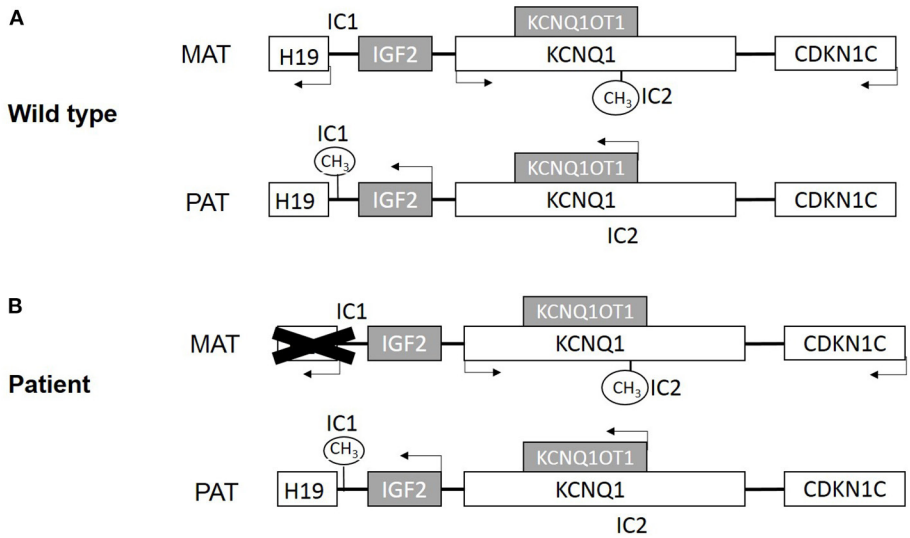


FIGURE 1 | (A) Scheme of imprinted gene cluster in wild type. **(B)** Scheme of imprinted gene cluster in patient. MAT, maternal derived; PAT, paternal derived. White boxes present maternal expression genes, and gray present paternal expression genes. IC1 is paternal methylation and IC2 is maternal methylation.

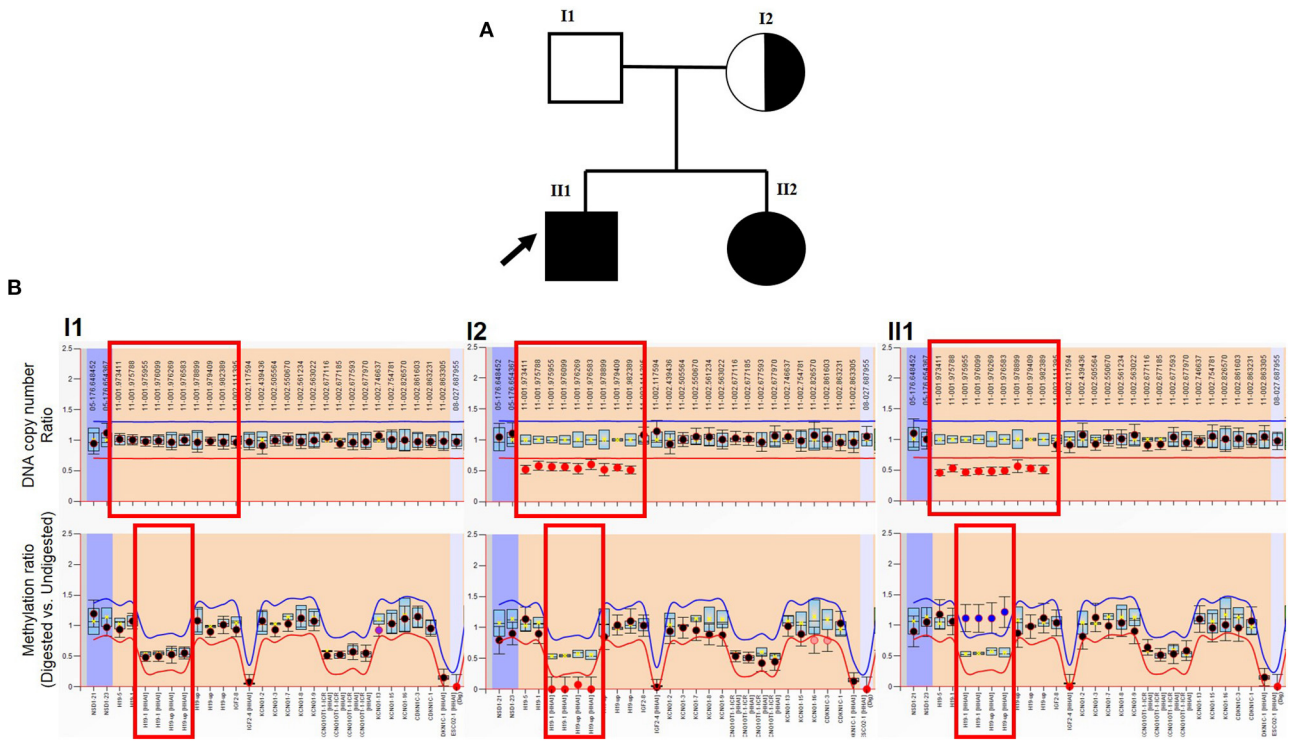


FIGURE 2 | MLPA result of patient. **(A)** Genogram of the BWS family. The arrow points out the male patient (proband) in our study. **(B)** Top row: DNA copy number ratio of patient vs. normal reference. The *X*-axis shows hg18 locations. The red box points out H19 and its upstream location. It shows normal ratio (∼1) in the patient's father, and about half (∼0.5) in the male patient, and his mother (also in his sister, which is not presented here). Bottom row: methylation ratio of H19. The *X*-axis shows gene exons. The red box points out *H19* and the upstream region. It shows normal ratio (∼0.5) in the patient's father. Because his mother lost the methylation copy of *H19*, the methylation level is about 0. The male patient lost the maternal copy of *H19* (also in his sister, which is not presented here), which should not be methylated, which leads to the methylation level increase (∼1).

but not reported in their case. The male patient in our case had chryptorchidism which was corrected by surgery before 2 years old. According to the literature (Feyles et al., 2014), when cryptorchid patients have surgery before age 2, more than 95% patients can reach normal sperm count and motility. However, BWS patients do not appear to experience such high fertility

rates with orchidopexy. Gazzin et al. (2019) followed four BWS males who suffered cryptorchidism. All of them had azoospermia after surgery, as in our case. Therefore, azoospermia of our male patient might be not the consequence of cryptorchidism, but due to genetic issues, which lead to dysfunction of the testis.

In recent years, more and more studies focus on the genetic and epigenetic factors in male infertility (Dong et al., 2017; Gunes et al., 2018; Lujan et al., 2019). Some studies showed that abnormal methylation of *H19* may be associated with male infertility. In our case, the male patient lost maternal H19, which thus may be related with his cryptorchidism and azoospermia.

Furthermore, the female patient carried the H19 gene microdeletion, which can be passed on to her children and would lead to BWS. Therefore, it may be reasonable to consider PGT (preimplantation genomic testing) to detect potential for BWS offspring.

In addition, we also reanalyzed WES and SNP array result to find the breakpoint of the DNA deletion for this family. Because of the limitation of these techniques, we need other methods (gap-PCR and Sanger sequencing) to confirm the precise breakpoint of this microdeletion.

CONCLUSION

In this study, we report a BWS family, which was due to maternal deletion in *H19* and its upstream regulatory genes. Now that the patients have reached childbearing age, it gives some insight on the presentation of BWS in adulthood and some of the potential reproductive issues. As BWS male patients could face subfertility, these may be associated with specific molecular subtypes. Azoospermia in these patients may not be the consequence of cryptorchidism, but due to genetic or epigenetic issues. Undescended testis greatly increases the risk of several serious complications like testicular torsion and testicular cancer. Therefore, the surgery is probably indicated regardless of its effect on fertility in males with BWS. This report is limited to a small family and thus presents only a portion of possible clinical scenarios. More studies on adult BWS patients are warranted in the future.

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 Reproductive Medical Ethics Committee of Guangzhou Women and Children's medical center. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JC designed the study, performed the experiments, analyzed the data, and edited the manuscript. JX and YY collected the clinical information and communicated with patients. LS designed the research, supervised the studies, analyzed the data, and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study is supported by the National Natural Science Foundation of China (81800110).

ACKNOWLEDGMENTS

We thank Dr. Abraham Nick Morse for his kind suggestions and language modification.

REFERENCES

Algar, E., Brickell, S., Deeble, G., Amor, D., and Smith, P. (2000). Analysis of CDKN1C in Beckwith Wiedemann syndrome. *Hum. Mutat.* 15, 497–508. doi: 10.1002/1098-1004(200006)15:6<497::AID-HUMU2>3.0.CO;2-F

Baskin, B., Choufani, S., Chen, Y. A., Shuman, C., Parkinson, N., Lemyre, E., et al. (2014). High frequency of copy number variations (CNVs) in the chromosome 11p15 region in patients with Beckwith-Wiedemann syndrome. *Hum. Genet.* 133, 321–330. doi: 10.1007/s00439-013-1379-z

Beygo, J., Citro, V., Sparago, A., De Crescenzo, A., Cerrato, F., Heitmann, M., et al. (2013). The molecular function and clinical phenotype of partial deletions of the IGF2/H19 imprinting control region depends on the spatial arrangement of the remaining CTCF-binding sites. *Hum. Mol. Genet.* 22, 544–557. doi: 10.1093/hmg/dds465

Brioude, F., Netchine, I., Praz, F., Le Jule, M., Calmel, C., Lacombe, D., et al. (2015). Mutations of the imprinted CDKN1C gene as a cause of the overgrowth Beckwith-Wiedemann Syndrome: clinical spectrum and functional characterization. *Hum. Mutat.* 36, 894–902. doi: 10.1002/humu.22824

Choufani, S., Shuman, C., and Weksberg, R. (2010). Beckwith-Wiedemann syndrome. *Am. J. Med. Genet. C Semin. Med. Genet.* 154C, 343–354. doi: 10.1002/ajmg.c.30267

Cohen, M. M. Jr. (1971). Macroglossia, omphalocele, visceromegaly, cytomegaly of the adrenal cortex and neonatal hypoglycemia. *Birth Defects Orig. Artic. Ser.* 7, 226–232.

De Crescenzo, A., Coppola, F., Falco, P., Bernardo, I., Ausanio, G., Cerrato, F., et al. (2011). A novel microdeletion in the IGF2/H19 imprinting centre region defines a recurrent mutation mechanism in familial Beckwith-Wiedemann syndrome. *Eur. J. Med. Genet.* 54, e451–e454. doi: 10.1016/j.ejmg.2011.04.009

Diaz-Meyer, N., Day, C. D., Khatod, K., Maher, E. R., Cooper, W., Reik, W., et al. (2003). Silencing of CDKN1C (p57KIP2) is associated with hypomethylation at KvDMR1 in Beckwith-Wiedemann syndrome. *J. Med. Genet.* 40, 797–801. doi: 10.1136/jmg.40.11.797

Dong, H., Wang, Y., Zou, Z., Chen, L., Shen, C., Xu, S., et al. (2017). Abnormal methylation of imprinted genes and cigarette smoking: assessment of their association with the risk of male infertility. *Reprod. Sci.* 24, 114–123. doi: 10.1177/1933719116650755

Elliott, M., Bayly, R., Cole, T., Temple, I. K., and Maher, E. R. (1994). Clinical features and natural history of Beckwith-Wiedemann syndrome: presentation of 74 new cases. *Clin. Genet.* 46, 168–174. doi: 10.1111/j.1399-0004.1994.tb04219.x

Feyles, F., Peiretti, V., Mussa, A., Manenti, M., Canavese, F., Cortese, M. G., et al. (2014). Improved sperm count and motility in young men surgically treated

for cryptorchidism in the first year of life. *Eur. J. Pediatr. Surg.* 24, 376–380. doi: 10.1055/s-0033-1349715

Gazzin, A., Carli, D., Sirchia, F., Molinatto, C., Cardaropoli, S., Palumbo, G., et al. (2019). Phenotype evolution and health issues of adults with Beckwith-Wiedemann syndrome. *Am. J. Med. Genet. A* 179, 1691–1702. doi: 10.1002/ajmg.a.61301

Gunes, S., Agarwal, A., Henkel, R., Mahmutoglu, A. M., Sharma, R., Esteves, S. C., et al. (2018). Association between promoter methylation of MLH1 and MSH2 and reactive oxygen species in oligozoospermic men-A pilot study. *Andrologia* 50:e12903. doi: 10.1111/and.12903

Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., and Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405, 486–489. doi: 10.1038/35013106

Kosseff, A. L., Herrmann, J., Gilbert, E. F., Viseskul, C., Lubinsky, M., and Opitz, J. M. (1976). Studies of malformation syndromes of man XXIX: the Wiedemann-Beckwith syndrome. Clinical, genetic and pathogenetic studies of 12 cases. *Eur J Pediatr.* 123, 139–66. doi: 10.1007/BF00452093

Krzyzewska, I. M., Alders, M., Maas, S. M., Bliek, J., Venema, A., Henneman, P., et al. (2019). Genome-wide methylation profiling of Beckwith-Wiedemann syndrome patients without molecular confirmation after routine diagnostics. *Clin. Epigenet.* 11, 53. doi: 10.1186/s13148-019-0649-6

Li, M., Squire, J., Shuman, C., Fei, Y. L., Atkin, J., Pauli, R., et al. (2001). Imprinting status of 11p15 genes in Beckwith-Wiedemann syndrome patients with CDKN1C mutations. *Genomics* 74, 370–376. doi: 10.1006/geno.2001.6549

Lujan, S., Caroppo, E., Niederberger, C., Arce, J. C., Sadler-Riggleman, I., Beck, D., et al. (2019). Sperm DNA methylation epimutation biomarkers for male infertility and FSH therapeutic responsiveness. *Sci. Rep.* 9, 16786. doi: 10.1038/s41598-019-52903-1

Niemitz, E. L., DeBaun, M. R., Fallon, J., Murakami, K., Kugoh, H., Oshimura, M., et al. (2004). Microdeletion of LIT1 in familial Beckwith-Wiedemann syndrome. *Am. J. Hum. Genet.* 75, 844–849. doi: 10.1086/425343

Pandey, R. R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., et al. (2008). Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell* 32, 232–246. doi: 10.1016/j.molcel.2008.08.022

Prawitt, D., Enklaar, T., Gartner-Rupprecht, B., Spangenberg, C., Lausch, E., Reutzel, D., et al. (2005a). Microdeletion and IGF2 loss of imprinting in a cascade causing Beckwith-Wiedemann syndrome with Wilms' tumor. *Nat Genet.* 37, 785–6; author reply 6–7. doi: 10.1038/ng0805-785

Prawitt, D., Enklaar, T., Gartner-Rupprecht, B., Spangenberg, C., Oswald, M., Lausch, E., et al. (2005b). Microdeletion of target sites for insulator protein CTCF in a chromosome 11p15 imprinting center in Beckwith-Wiedemann syndrome and Wilms' tumor. *Proc. Natl. Acad. Sci. U. S. A.* 102, 4085–4090. doi: 10.1073/pnas.0500037102

Sparago, A., Cerrato, F., Vernucci, M., Ferrero, G. B., Silengo, M. C., and Riccio, A. (2004). Microdeletions in the human H19 DMR result in loss of IGF2 imprinting and Beckwith-Wiedemann syndrome. *Nat. Genet.* 36, 958–960. doi: 10.1038/ng1410

Taylor, W. N. (1981). Urological implications of the beckwith-wiedemann syndrome. *J. Urol.* 125, 439–441. doi: 10.1016/S0022-5347(17)55068-5

Watanabe, H., and Yamanaka, T. (1990). A possible relationship between Beckwith-Wiedemann syndrome, urinary tract anomaly and prune belly syndrome. *Clin. Genet.* 38, 410–414. doi: 10.1111/j.1399-0004.1990.tb03605.x

Zollino, M., Orteschi, D., Marangi, G., De Crescenzo, A., Pecile, V., Riccio, A., et al. (2010). A case of Beckwith-Wiedemann syndrome caused by a cryptic 11p15 deletion encompassing the centromeric imprinted domain of the BWS locus. *J. Med. Genet.* 47, 429–432. doi: 10.1136/jmg.2009.071142

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: 25 February 2021
doi: 10.3389/fped.2021.603666

Case Report: Clinical Description of a Patient Carrying a 12.48 Mb Microdeletion Involving the 10p13–15.3 Region

*Yu-qing Pan and Jian-hua Fu**

Department of Pediatrics, Shengjing Hospital of China Medical University, Shenyang, China

Partial deletion of 10p chromosome is a rare chromosomal aberration. Submicroscopic deletion of 10p15.3 is mainly related to cognitive deficits, speech disorders, motor delay, and hypotonia with the deleted region ranging from 0.15 to 4 Mb. The clinical phenotype is mainly determined by the *ZMYND11* and *DIP2C* genes. Here, we report a rare case of feeding difficulties, hypocalcemia, and psychomotor retardation. Our patient has a 12.48 Mb deletion in 10p15.3–10p13, which is the second case of large 10p deletion among reported cases thus far.

Keywords: 10p deletion, copy number variation, feeding difficulty, hypocalcemia, psychomotor retardation

OPEN ACCESS

Edited by:
Katalin Komlosi, Medical Center University of Freiburg, Germany

Reviewed by:
Rita Selvatici, University of Ferrara, Italy
Judit Kárteszi, Zala Megyei Szent Raphael Kórház, Hungary

****Correspondence:***
Jian-hua Fu
fujh77@163.com

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

Received: *07 September 2020*
Accepted: *13 January 2021*
Published: *25 February 2021*

Citation:
Pan Y-q and Fu J-h (2021) Case Report: Clinical Description of a Patient Carrying a 12.48 Mb Microdeletion Involving the 10p13–15.3 Region. Front. Pediatr. 9:603666. doi: 10.3389/fped.2021.603666

INTRODUCTION

Partial deletion of chromosome 10p is a rare chromosomal aberration that is associated with different syndromes (1). It involves a known monogenic syndrome; the hypoparathyroidism, sensorineural deafness, and renal dysplasia (HDR or Barakat) syndrome (OMIM #146255); and DiGeorge syndrome 2 (DGS2), as well as other syndromes (2). HDR is a rare autosomal dominant disorder, which is characterized by varying degrees of HDR. It is caused by the dysfunction of the glutamyl-amidotransferase-subunit A (*GATA3*) gene located on 10p14 (3). *GATA3* plays an important role in the embryonic development of the parathyroid gland, inner ear, kidney, and central nervous system (4). However, haploinsufficiency of a more centromeric region on 10p13–10p14 was previously found to be related to DGS2 (velocardiofacial syndrome complex three), which also includes as a feature congenital heart defects and thymus hypoplasia/aplasia or T cell defects (5). Additional features include renal anomalies, eye anomalies, hypoparathyroidism, skeletal defects, and developmental delay (6). Importantly, hypocalcemia is one of the key characteristics present in DGS2 and HDR. Thus, this characteristic can aid in the diagnosis of the possible genetic disorders (7) (the causes of hypocalcemia are summarized in **Supplementary Table 1**).

In 10p13–10p15.3 microdeletions, zinc finger MYND-type containing 11 (*ZMYND11*), disco-interacting protein 2 homolog C (*DIP2C*), La ribonucleoprotein 4B (*LARP4B*), and other genes have been reported to be responsible for DGS2, HDR syndrome, or other similar phenotypes (8). The protein encoded by the *ZMYND11* gene (also called BS69 or BRAM1), a cellular nuclear protein containing PHD, Bromo, PWWP, and MYND domains, was originally identified as an adenovirus E1A-binding protein that inhibits the transactivation function of E1A (9). It is associated with autosomal dominant non-syndromic intellectual disability (or autosomal dominant mental retardation type 30, OMIM #616083) involving complex cognitive, behavioral, and developmental difficulties (10, 11). The *DIP2C* gene encodes a member of the disco-interacting

protein homolog two family with two other isoforms, DIP2A and DIP2B. It is a candidate for developmental dyslexia and autism (12). The *LARP4B* gene encodes a member of an evolutionarily conserved protein family implicated in RNA metabolism and translation (13).

Recently, Kim et al. reported the so far largest deletion (16 Mb) in chromosome 10p in a Korean neonate (8). The clinical manifestations included hypoparathyroidism, hearing loss, genitourinary and cardiac anomalies, thymus hypoplasia, and neural system abnormalities and limb deformities.

In the present study, a rare case of feeding difficulties, hypocalcemia, and psychomotor retardation is reported in which our patient harbors a 12.48 Mb deletion in 10p15.3–10p13, which is the second case of large 10p deletion among reported cases so far.

CLINICAL REPORT

Our patient was admitted when she was 6 days old to the Department of Neonatology, Shengjing Hospital of China Medical University. She was born at 39 weeks of gestation from a 34-year-old G4P2 mother without asphyxia. Her parents were non-consanguineous, and there was one healthy older female sibling in the family. The patient was spontaneously conceived, and her prenatal history was reportedly uneventful with no exposure to known teratogens. Birth weight, length, and head circumference were within normal ranges. Facial dysmorphism and feeding difficulties were noted soon after birth. Accordingly, the following features were noted: brachycephaly, round face, down-ward slanting palpebral fissures, hypertelorism, curled eye lashes, a broad and low nasal root, micrognathia, high arched palate, low-set ears, muscular hypertonia, irritability, cyanosis during feeding, weak sucking, and severe swallowing dysfunction (**Figure 1**). In addition, the patient had multiple daily episodes of apnea, desaturation, and cyanosis for a few seconds with spontaneous resolution.

She presented with hypoparathyroidism, and her persistent hypocalcemia was difficult to correct despite treatment with calcium gluconate solution and calcitriol. Additionally, parathyroid ultrasound did not demonstrate any anomalies, whereas a chest CT scan showed normal thymus but diffuse bilateral bronchopneumonia, and laryngotracheal CT was normal. The patient's 25-hydroxyvitamin D blood level was low (8.96 ng/ml, normal: 30–70 ng/ml), and immunological studies demonstrated normal counts and a ratio of T and B cells. Myocardial enzyme spectrum was abnormal: creatine kinase (CK) 1,335 U/L (normal: <171 U/L), CK-MB 67 U/L (normal: <24 U/L), lactate dehydrogenase (LDH) 612 U/L (normal: 80–285 U/L), and troponin I: 0.657 μg/L (normal: 0–0.04 μg/L). Myoglobin was normal. Serum creatinine and urine routine test were normal. Thus, laboratory results did not disclose renal functional abnormality. Abdominal ultrasound was planned to be performed on control examination to rule out any urogenital developmental disorder. Brain MRIs performed during hospitalization were found to be normal; however, the electroencephalogram (EEG) was severely abnormal: (1) the sleep–wake cycle was disturbed and not consistent with the corresponding gestational age; (2) EEG activity in the QS phase was abnormal with multifocal spikes and sharp waves released asynchronously; and (3) multiple θ rhythms were issued in the central midline and parietal regions. The infant's clinical manifestations and the abnormal EEG pointed toward a severe neurological dysfunction in the patient. Moreover, upper gastrointestinal radiography exhibited severe swallowing dysfunction of the epiglottis. Visual and brainstem auditory evoked potentials were observed to be normal. Fundoscopy, screening for inborn errors of metabolism, and echocardiography were found to be without a pathological sign.

MATERIALS AND METHODS

Karyotype Analysis

G-band karyotype analysis from the patient's peripheral blood was performed (14). Karyotypes were reviewed according to the International System for Human Cytogenetics Nomenclature (ISCN 2013).

Detection of Chromosome Copy Number Variation

Genomic DNA was extracted from the patient's peripheral blood using the DNeasy blood and tissue kit (Qiagen GmbH, Hilden, Germany). The quality and concentration of DNA were measured using the NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). Next-generation sequencing (NGS) and copy number variation (CNV) detection was performed by Berry Genomics Corporation (Beijing, China) according to previous studies (15). Briefly, 50 ng of amniocyte DNA was fragmented to an average size of 300 bp. DNA libraries were constructed by end filling, adapter ligation, and PCR amplification as previously described (16). DNA libraries were then subjected to massively parallel sequencing on the NextSeq 500 platform (Illumina, San Diego, CA, USA) in order to generate approximately five million 36-bp single-end reads, representing 0.06–0.1-fold genome coverage. All the sequences were aligned to the unmasked hg19 genome using the Burrows–Wheeler algorithm (17). Mapped reads were allocated progressively to 20 kb bin sizes from the p to q arms of the 24 chromosomes.

RESULTS

The routine karyotype analysis already pointed toward a deletion in one of the short arms of chromosome 10. Subsequent NGS and CNV analyses of the data found a heterozygous 10p15.3–10p13 deletion [arr 10p15.3–p13 (120,000–12,600,000) × 1], illustrating that the patient had an ∼12.48 Mb deletion in the short arm of chromosome 10 (**Figure 2**). The genes located in the deleted region were checked using the UCSC Genome Browser (www.genome.ucsc.edu) (**Figure 3**). The *ZMYND11*, *DIP2C*, *LARP4B*, and *GATA3* were included in the deleted region.

FIGURE 1 | Our patient at the age of 45 days showed brachycephaly, round face, down-ward slanting palpebral fissures, hypertelorism, curled eye lashes, a broad and low nasal root, micrognathia, high arched palate, low-set ears.

DISCUSSION

The clinical characteristics of reported patients with a microdeletion in the 10p13–15 region are summarized in **Table 1**. The present patient who has a 12.48 Mb deletion in 10p15.3–10p13 shows clinical features of facial dysmorphism, swallowing dysfunction, hypoparathyroidism, and neurological abnormalities. These phenotypes are not typical of HDR and DGS2, though the patient is only the second reported case with a large deletion.

Previously, Tremblay et al. described that the deletion of chromosome 10p15.3–10p15.2 was associated with ventricular septal defects/septal aneurysms among 18 family members in three generations (20). Moreover, 19 unrelated individuals

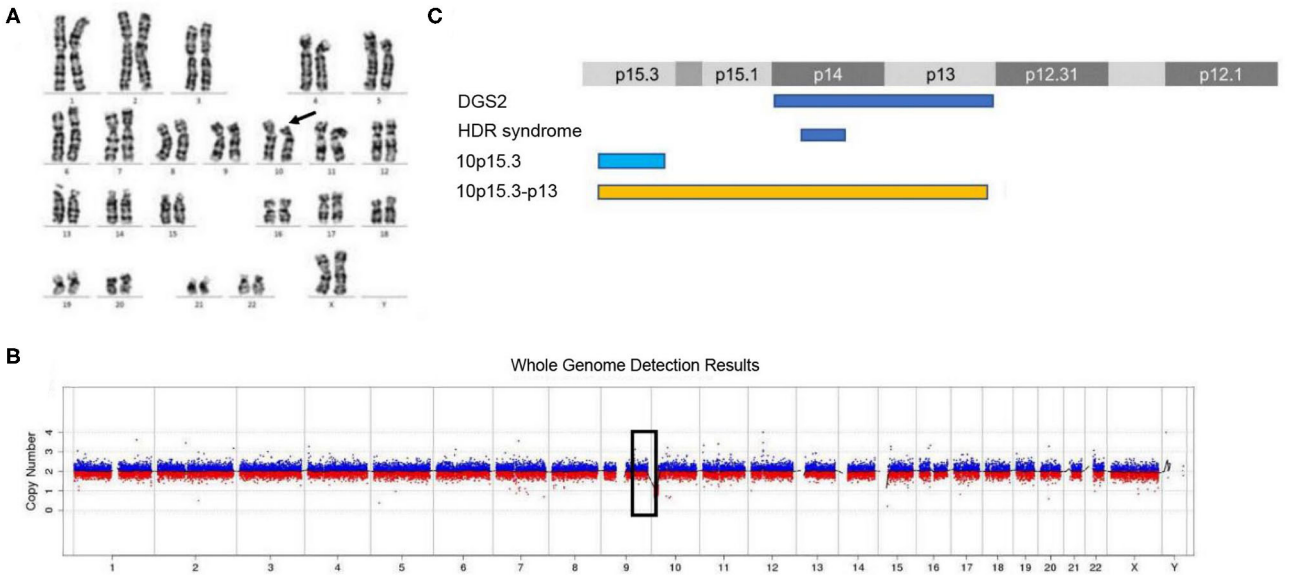


FIGURE 2 | Karyotype analysis and CNV analysis showed ∼12.48 Mb deletion at 10p15.3–10p13. **(A)** Karyotype analysis; **(B)** CNV sequence; **(C)** the published deletions in this region. **(C)** Showed the four deletion types in the 10p15–10p13 region. The deletion of 10p15.3 has been reported by Descipio et al. (11), Vargiami et al. (12), Eggert et al. (18), and Poluha et al. (19). The deletion of 10p15.3–p13 has been reported by Kim et al. (8) and our study.

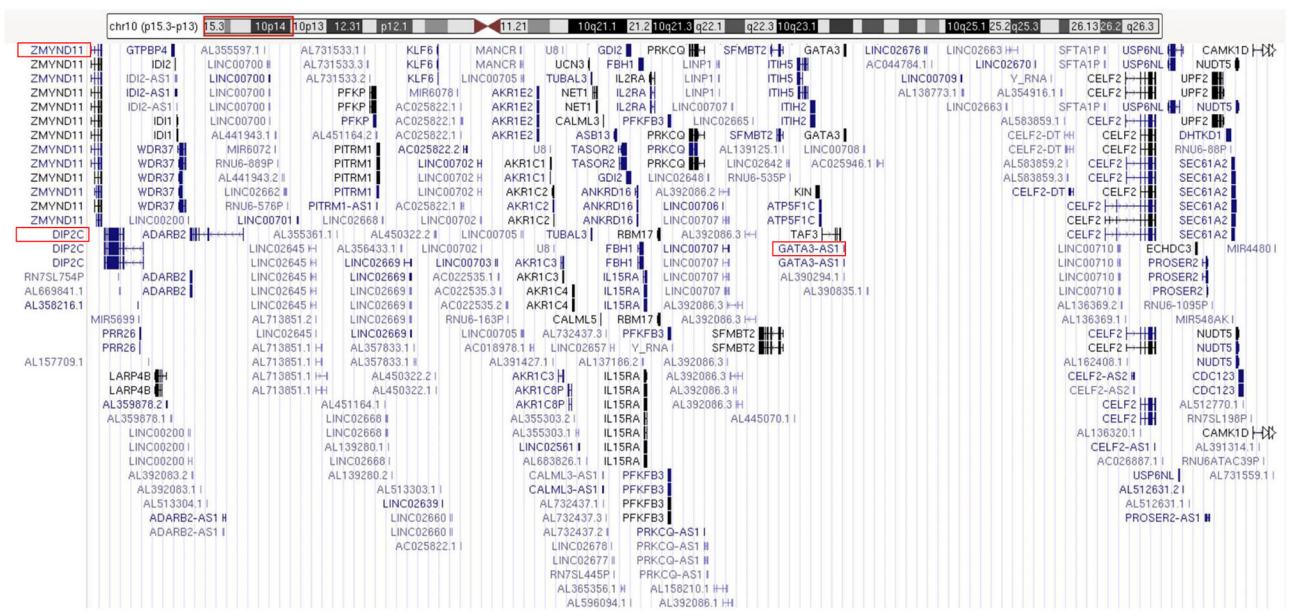


FIGURE 3 | The genes included in the deleted region (chr10: 120,000–12,600,000).

were characterized with submicroscopic deletions involving 10p15.3 (21). Their common clinical features (12 of the 19 individuals) included cognitive/behavioral/developmental difficulties, speech delay/language disorder, motor delay (10/10), craniofacial dysmorphism, hypotonia, brain anomalies, and seizures. A monozygotic female twin pair with a *de novo* 2.7 Mb deletion of 10p15.3 presented with severe developmental delay, including severe visual and sensorineural hearing impairment. Additionally, both showed generalized dystonia, microcephaly, complete absence of voluntary movements, and visual/auditory unresponsiveness (22). Submicroscopic deletion of 10p15.3 is mainly related to cognitive deficits, speech disorders, motor delay, and hypotonia with a deleted region from 0.15 to 4 Mb (21). However, Gamba et al. reported a male child

TABLE 1 | Clinical features of patients with a microdeletion/mutation involving the 10p15–10p13 region.

Microdeletion/mutation	10p13–p14		10p15.3				10p15.3–p13	
Patient characteristics (ref.)	**DGS2 syndrome**	**HDR syndrome**	**Descipio et al. (11) (11/19 enrolled patients*)**	**Vargiami et al. (12) (two patients)**	**Eggert et al. (18) (two patients)**	**Poluha et al. (19) (one patient)**	**Kim et al. (8) (one patient)**	**Our patient**
Hypoparathyroidism	n.a.	+	–	n.a.	–	n.a.	+	+
Hearing loss	n.a.	+	–	2/2	–	n.a.	+	–
Genitourinary anomalies/hypoplasia of the kidney	n.a.	+	≥1/11	–	–	n.a.	+	–
Cardiac anomalies	+	n.a.	≥2/11	–	1/2	+	+	–
Thymus hypoplasia	+	n.a.	n.a.	n.a.	n.a.	n.a.	+	–
ADHD	n.a.	n.a.	2/11	n.a.	n.a.	n.a.	n.d	n.d
Autism	n.a.	n.a.	1/11	n.a.	n.a.	n.a.	n.d	n.d
Cognitive/behavioral abnormality	n.a.	n.a.	≥10/11	2/2	2/2	+	n.d	n.d
Motor delay	n.a.	n.a.	≥10/11	2/2	2/2	+	n.d	+
Speech delay	n.a.	n.a.	≥10/11	2/2	2/2	+	n.d	n.d
Brain CT/MRI anomalies	n.a.	n.a.	≥3/11	2/2	n.a.	+	+	Normal
Facial dysmorphism	n.a.	n.a.	9/11	2/2	2/2	+	+	+
Hypertonia	n.a.	n.a.	–	+	–	+	+	+
Hypotonia	n.a.	n.a.	≥6/11	–	–	+	–	–
Seizures	n.a.	n.a.	≥3/11	–	–	+	+	+
Others	n.a.	n.a.	n.a.	n.a.	Hand/foot anomalies	n.a.	IUGR, hand/foot anomalies, hip dislocation/ subluxation	n.a.

*n.d meant that the symptom was not determined in the patients, n.a. meant no data available. *Descipio et al. reported 19 patients with the microdeletion, and only 11 patients have the available clinical data.*

with a 5.6 Mb deletion at 10p15.3–10p14 who exhibited short stature, cleft lip/palate, and feeding problems (23). In view of the aforementioned studies, the microdeletion of 10p15.3 was inferred to give rise to different features and is seemingly not related to the deletion size.

Tumiene et al. suggested that the clinical features of 10p15.3 microdeletion included neurodevelopmental disorders, characteristic dysmorphic features, and some other more frequent symptoms, and that the *ZMYND11* gene was responsible for the above phenotype (24). As a transcriptional repressor, mutations of *ZMYND11* have been associated with autosomal dominant mental retardation type 30 leading to intellectual disability, behavioral abnormalities, and seizures (25). Pathogenic single nucleotide variants (SNVs) of *ZMYND11* were also associated with Cornelia de Lange syndrome in a large study (26). In addition, the *DIP2C* gene is located on the minimal region of the overlap of the deletions (24). It is highly expressed in the brain and encompasses various neurological functions, such as "memory," "neuropeptide signaling pathway," and "response to amphetamine" according to a gene ontology analysis in the *DIP2C* gene knock-out mice (12). The deletion of *DIP2C* could induce the cell enlargement and growth retardation by stimulating DNA methylation (27). However, no pathogenic point mutations or gene deletions of the *DIP2C* gene have been described so far with a human phenotype. Although the *ZMYND11* and *DIP2C* genes are located on the minimal region of the overlap of the deletions (24), it seems that they serve as key genes for 10p15.3 microdeletion syndrome. However, the deletion is so large that it is impossible to correlate the entire manifestation with the function of single genes. The exact role of the *ZMYND11* and *DIP2C* genes in regard to their clinical features requires further investigation.

In summary, this study reports a patient with a 12.48 Mb deletion in 10p15.3–10p13. The phenotype was not found to be typical of HDR and DGS2, though the patient is only the second reported case having a large deletion thus far.

DATA AVAILABILITY STATEMENT

The original contributions generated for 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 Ethics Committee of Shengjing hospital of China 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

Y-qP performed the data collection and analysis and wrote the original draft. J-hF performed the supervision and edited the writing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fped.2021.603666/full#supplementary-material

REFERENCES

1. Melis D, Genesio R, Boemio P, Del Giudice E, Cappuccio G, Mormile A, et al. Clinical description of a patient carrying the smallest reported deletion involving 10p14 region. *Am J Med Genet A.* (2012) 158A:832–5. doi: 10.1002/ajmg.a.34133
2. Bilous RW, Murty G, Parkinson DB, Thakker RV, Coulthard MG, Burn J, et al. Brief report: autosomal dominant familial hypoparathyroidism, sensorineural deafness, and renal dysplasia. *N Engl J Med.* (1992) 327:1069–74. doi: 10.1056/NEJM199210083271506
3. Van Esch H, Groenen P, Nesbit MA, Schuffenhauer S, Lichtner P, Vanderlinden G, et al. GATA3 haplo-insufficiency causes human HDR syndrome. *Nature.* (2000) 406:419–22. doi: 10.1038/35019088
4. Nesbit MA, Bowl MR, Harding B, Ali A, Ayala A, Crowe C, et al. Characterization of GATA3 mutations in the hypoparathyroidism, deafness, and renal dysplasia (HDR) syndrome. *J Biol Chem.* (2004) 279:22624–34. doi: 10.1074/jbc.M401797200
5. Lichtner P, Konig R, Hasegawa T, Van Esch H, Meitinger T, Schuffenhauer S. An HDR (hypoparathyroidism, deafness, renal dysplasia) syndrome locus maps distal to the DiGeorge syndrome region on 10p13/14. *J Med Genet.* (2000) 37:33–7. doi: 10.1136/jmg.37.1.33
6. McDonald-McGinn DM, Sullivan KE. Chromosome 22q11.2 deletion syndrome (DiGeorge syndrome/velocardiofacial syndrome). *Medicine.* (2011) 90:1–18. doi: 10.1097/MD.0b013e3182060469
7. Kita M, Kuwata Y, Usui T. Familial congenital choanal atresia with GATA3 associated hypoparathyroidism-deafness-renal dysplasia syndrome unidentified on auditory brainstem response. *Auris Nasus Larynx.* (2019) 46:808–12. doi: 10.1016/j.anl.2018.10.005
8. Kim SB, Kim YE, Jung JM, Jin HY, Lim YJ, Chung ML. Clinical description of a neonate carrying the largest reported deletion involving the 10p15.3p13 region. *Clin Case Reports.* (2017) 5:1369–75. doi: 10.1002/ccr3.1070
9. Yang JP, Yang JK, Li C, Cui ZQ, Liu HJ, Sun XF, et al. Downregulation of ZMYND11 induced by miR-196a-5p promotes the progression and growth of GBM. *Biochem Biophys Res Commun.* (2017) 494:674–80. doi: 10.1016/j.bbrc.2017.10.098
10. Moskowitz AM, Belnap N, Siniard AL, Szelinger S, Claasen AM, Richholt RF, et al. A de novo missense mutation in ZMYND11 is associated with global developmental delay, seizures, and hypotonia. *Cold Spring Harbor Mol Case Stud.* (2016) 2:a000851. doi: 10.1101/mcs.a000851
11. Coe BP, Witherspoon K, Rosenfeld JA, van Bon BW, Vulto-van Silfhout AT, Bosco P, et al. Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat Genetics.* (2014) 46:1063–71. doi: 10.1038/ng.3092
12. Oo ZM, Adlat S, Sah RK, Myint MZZ, Hayel F, Chen Y, et al. Brain transcriptome study through CRISPR/Cas9 mediated mouse Dip2c gene knock-out. *Gene.* (2020) 758:144975. doi: 10.1016/j.gene.2020.144975
13. Kuspert M, Murakawa Y, Schaffler K, Vanselow JT, Wolf E, Juranek S, et al. LARP4B is an AU-rich sequence associated factor that promotes mRNA accumulation and translation. *RNA.* (2015) 21:1294–305. doi: 10.1261/rna.051441.115
14. Shimono J, Miyoshi H, Kiyasu J, Kamimura T, Eto T, Miyagishima T, et al. Clinicopathological analysis of polyploid diffuse large B-cell lymphoma. *PLoS ONE.* (2018) 13:e0194525. doi: 10.1371/journal.pone.0194525
15. Liang D, Peng Y, Lv W, Deng L, Zhang Y, Li H, et al. Copy number variation sequencing for comprehensive diagnosis of chromosome disease syndromes. *J Mol Diagn.* (2014) 16:519–26. doi: 10.1016/j.jmoldx.2014.05.002
16. Liang D, Lv W, Wang H, Xu L, Liu J, Li H, et al. Non-invasive prenatal testing of fetal whole chromosome aneuploidy by massively parallel sequencing. *Prenat Diagn.* (2013) 33:409–15. doi: 10.1002/pd.4033
17. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* (2009) 25:1754–60. doi: 10.1093/bioinformatics/btp324
18. Eggert M, Muller S, Heinrich U, Mehraein Y. A new familial case of microdeletion syndrome 10p15.3. *Eur J Med Genet.* (2016) 59:179–82. doi: 10.1016/j.ejmg.2016.02.008
19. Poluha A, Bernaciak J, Jaszczuk I, Kedzior M, Nowakowska BA. Molecular and clinical characterization of new patient with 1.08 Mb deletion in 10p15.3 region. *Mol Cytogenet.* (2017) 10:34. doi: 10.1186/s13039-017-0336-2
20. Tremblay N, Yang SW, Hitz MP, Asselin G, Ginns J, Riopel K, et al. Familial ventricular aneurysms and septal defects map to chromosome 10p15. *Eur Heart J.* (2011) 32:568–73. doi: 10.1093/eurheartj/ehq447
21. DeScipio C, Conlin L, Rosenfeld J, Tepperberg J, Pasion R, Patel A, et al. Subtelomeric deletion of chromosome 10p15.3: clinical findings and molecular cytogenetic characterization. *Am J Med Genet A.* (2012) 158A:2152–61. doi: 10.1002/ajmg.a.35574
22. Vargiami E, Ververi A, Kyriazi M, Papathanasiou E, Gioula G, Gerou S, et al. Severe clinical presentation in monozygotic twins with 10p15.3 microdeletion syndrome. *Am J Med Genet A.* (2014) 164A:764–8. doi: 10.1002/ajmg.a.36329
23. Gamba BF, Rosenberg C, Costa S, Richieri-Costa A, Ribeiro-Bicudo LA. Cleft lip/palate, short stature, and developmental delay in a boy with a 5.6-mb interstitial deletion involving 10p15.3p14. *Mol Syndromol.* (2015) 6:39–43. doi: 10.1159/000371404
24. Tumiene B, Ciuladaite Z, Preiksaitiene E, Mameniskiene R, Utkus A, Kucinskas V. Phenotype comparison confirms ZMYND11 as a critical gene for 10p15.3 microdeletion syndrome. *J Appl Genet.* (2017) 58:467–74. doi: 10.1007/s13353-017-0408-3
25. Yates TM, Drucker M, Barnicoat A, Low K, Gerkes EH, Fry AE, et al. ZMYND11-related syndromic intellectual disability: 16 patients delineating and expanding the phenotypic spectrum. *Hum Mutat.* (2020) 41:1042–50. doi: 10.1002/humu.24001
26. Aoi H, Mizuguchi T, Ceroni JR, Kim VEH, Furquim I, Honjo RS, et al. Comprehensive genetic analysis of 57 families with clinically suspected Cornelia de Lange syndrome. *J Hum Genet.* (2019) 64:967–78. doi: 10.1038/s10038-019-0643-z
27. Larsson C, Ali MA, Pandzic T, Lindroth AM, He L, Sjoblom T. Loss of DIP2C in RKO cells stimulates changes in DNA methylation

and epithelial-mesenchymal transition. *BMC Cancer*. (2017) 17:487. doi: 10.1186/s12885-017-3472-5

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 26 March 2021
doi: 10.3389/fgene.2021.626044

Application of Copy Number Variation Detection to Fetal Diagnosis of Echogenic Intracardiac Focus During Pregnancy

Yaxian Song[1†], Jingjing Xu[1†], Hongmiao Li[1], Jiong Gao[2], Limin Wu[1], Guoping He[1], Wen Liu[1], Yue Hu[1], Yaqin Peng[1], Fang Yang[3], Xiaohua Jiang[1] and Jing Wang[1]**

[1] Department of Obstetrics and Gynecology, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China, [2] Clinical Laboratory of Beijing Genomics Institute (BGI) Health, BGI-Shenzhen, Shenzhen, China, [3] Department of Obstetrics and Gynecology, The First Affiliated Hospital of Anhui University of Science and Technology, Anhui University of Science and Technology, Huainan, China

Echogenic intracardiac focus (EIF) is one of the most common ultrasound soft markers (USMs) in prenatal screening. However, the association of EIF with chromosomal abnormalities is still controversial. From January 2018 to April 2020, a total of 571 fetuses with USMs in our center were enrolled, among which 150 (26.27%) presented EIFs. We analyzed the karyotype anomalies and copy number variations (CNVs) in fetuses who presented EIFs by comparing their ultrasound indications, maternal ages and gestational stages. There were no statistically significant differences in the incidence of chromosomal abnormalities between fetuses with EIFs and the fetuses with USMs (4.00 vs. 7.71%, $p = 0.112$). Additionally, the incidence of chromosomal abnormalities was not related to maternal age (4.10% in maternal age below 35 yeas vs. 3.57% in maternal age above 35, $p = 1.000$). Interestingly, after 28 weeks of gestation, fetuses with EIFs showed more chromosomal abnormalities (20.00%) than that in the group before 28 weeks of gestation (2.22%, $p = 0.014$), and this result was attributed to the detection of pathogenic CNVs. After birth, 25 of children conducted cardiac development re-examination. Among them, 9 (36%, 9/25) were diagnosed with congenital heart disease, primarily patent foramen oval and ventricular septal defects (7/9, 77.77%). We concluded that the appearance of EIFs in early or mid-trimester would not indicate an increased risk of fetal chromosomal abnormalities. However, the persistence of EIFs in late trimester was associated with a higher risk of pathology-related CNVs and its persistent appearance may indicate heart development defects after birth. Thus, our results suggest that CNV detection has its advantages in prenatal diagnosis, especially for those with EIFs that persist in the third trimester.

Keywords: echogenic intracardiac focus, ultrasound soft markers, copy number variation, karyotype, congenital heart defects

OPEN ACCESS

Edited by:
Judit Bene,
University of Pécs, Hungary

Reviewed by:
Thomas Liehr,
Friedrich Schiller University Jena, Germany
Sandor Nagy,
Széchenyi István University, Hungary

***Correspondence:**
Xiaohua Jiang
biojxh@ustc.edu.cn
Jing Wang
ahwangjing1968@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: *04 November 2020*
Accepted: *10 March 2021*
Published: *26 March 2021*

Citation:
Song Y, Xu J, Li H, Gao J, Wu L, He G, Liu W, Hu Y, Peng Y, Yang F, Jiang X and Wang J (2021) Application of Copy Number Variation Detection to Fetal Diagnosis of Echogenic Intracardiac Focus During Pregnancy. Front. Genet. 12:626044. doi: 10.3389/fgene.2021.626044

INTRODUCTION

Ultrasound screening in the first and second trimesters of pregnancy is one of the most commonly performed genetic screenings. Besides sonographic structural defects, a group of findings classified as ultrasound soft markers (USMs) are often considered to indicate increased risk of underlying fetal aneuploidy, including echogenic intracardiac focus (EIF), thickening of the nuchal translucence, nasal bone dysplasia, echogenic bowel, single umbilical artery, short long bones, enlarged cisterna magna, cerebral ventriculomegaly, choroid plexus cyst, external left superior cavity, permanent right umbilical vein, right aortic arch, mild pyelectasis, and other conditions (Van Den Hof et al., 2005; Rembouskos et al., 2012; Choi et al., 2016; Lide et al., 2016).

EIF is defined as foci of echogenicity comparable to bone in the region of papillary muscle in either or both ventricles of fetal heart (Van Den Hof et al., 2005). EIF is one of the most common USMs in prenatal screening, with a prevalence ranging from less than 1% to 20% in different populations (Sepulveda and Romero, 1998; Wax et al., 2003). However, the association of EIF and aneuploidy is still controversial. For example, an increased incidence of trisomy 21 was found in fetuses who presented EIFs in high-risk pregnancies, however, some studies had failed to show this association (Dildy et al., 1996; Simpson et al., 1996; Achiron et al., 1997; Bromley et al., 1998; Manning et al., 1998; Winter et al., 2000). Additionally, previous results were mostly obtained based on traditional karyotyping, and their studies were always focused on trisomy syndromes, such as trisomy 21 syndrome. Recently, due to the availability of high-throughput sequencing, by measuring chromosomal microdeletions or microrepetitions, CNV-seq increases the detection efficiency of chromosomal abnormalities (Cohen et al., 2015; Committee on Genetics and the Society for Maternal-Fetal Medicine, 2016; Society for Maternal-Fetal Medicine et al., 2016; Clinical Genetics Group Of Medical Genetics Branch Chinese Medical Association et al., 2019). It was reported that 6.0% of fetuses presenting structural anomalies under ultrasound scanning had abnormal CNVs even karyotypes were normal (Wapner et al., 2012). Nevertheless, the CNV character in fetuses with EIFs is still seldom revealed.

It is generally accepted that most echogenic foci disappear with the progress of pregnancy and their constant presence may not imply poor pregnancy outcome (Simpson et al., 1996; Wolman et al., 2000; Wax et al., 2003; Chiu et al., 2019). Conflicting conclusions were elicited by the finding that euploid fetuses with EIFs showed cardiac diastolic dysfunction in the second trimester (Degani et al., 2001). Moreover, it was reported that the infants who showed fetal EIFs suffered more cardiac defects after birth than the general population (Goncalves et al., 2006).

Thus, in clinical consultation, although treatment of indications such as high risk in serological screening or in NIPT (Non-invasive Prenatal Testing) and fetal structural abnormalities is clear; perplexity arises when only EIF or USMs appear. In this investigation, to better estimate the risk of chromosomal abnormalities in fetuses with EIFs, we applied CNV-seq at a resolution of 100 kb simultaneously with conventional karyotyping in our prenatal diagnosis procedure and described the CNVs and karyotype abnormalities of fetuses with USMs and EIFs. The pregnancy outcomes from birth to 2 years of age were recorded and used to evaluate the association of EIFs with the presence of birth defects.

MATERIALS AND METHODS

Editorial Policies and Ethical Considerations

The study was conducted with the Ethics Committee of the First Affiliated Hospital of University of Science and Technology of China (USTC). All procedures were performed in accordance with the ethical standards set forth in the Helsinki Declaration of 1964 and its latest amendments or comparable ethical standards. The participating pregnant women signed informed consent forms and agreed to allow the sequencing data to be used in research after anonymization.

Patients and Study Design

This was a retrospective study. From January 2018 to April 2020, fetuses with USMs on prenatal diagnosis at the Prenatal Diagnosis Center of the First Affiliated Hospital of USTC were examined. Among them, fetuses who showed EIF were further investigated.

All of the pregnant women involved in this study received genetic counseling and provided written informed consent followed by amniocentesis or umbilical cord blood puncture. Using these samples, karyotype analysis and CNV-seq were conducted to identify fetal chromosomal abnormalities.

Ultrasound soft markers are based on sonographic findings, including EIF, thickening of the nuchal translucence, nasal bone dysplasia, echogenic bowel, single umbilical artery, short long bones, enlarged cisterna magna, hydrocystoma of neck, cerebral ventriculomegaly, choroid plexus cyst, external left superior cavity, permanent right umbilical vein, right aortic arch, mild pyelectasis, and other conditions.

Patients were excluded from this study when (1) the mother had previously delivered a child with a congenital defect; (2) the fetus displayed sonographic structural anomalies; (3) the couple had known genetic defects; or (4) the pregnancy was indicated to be high-risk based on serum screening or NIPT.

Karyotype Analysis

Fetal amniotic fluid cells were cultured in complete medium (Baorong, Hangzhou, China). Cordocentesis samples were cultured in Cell Preservation Medium (Sinochrome, Shanghai, China). Karyotype analysis was performed on G-band metaphases prepared from amniotic fluid or cord blood samples with resolution between 320 and 420 bands. At least thirty metaphases were counted, and five from each sample were analyzed. Karyotypes are described according to the ISCN 2016 nomenclature.

Copy Number Variation Sequencing (CNV-Seq)

Briefly, genomic DNAs were extracted using the Whole Blood DNA kit (BGI, Shenzhen, China) according to the manufacturer's protocol. The DNA quality and concentration were assessed using a Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, United States). Approximate 5 million sequencing reads per sample were mapped to the NCBI human reference genome (hg19/GRCh37) using the Burrows-Wheeler Aligner (BWA) tool and then allocated to 20 kb sequencing bins with 5 kb sliding to achieve higher resolution for CNV detection. The CNV-seq profiles of each chromosome were represented as log2 of the mean CNV of each sequencing bin along the length of the chromosome.

Sequence variants were annotated using population and literature databases including DECIPHER[1], DGV[2], 1000 Genomes Project[3], OMIM[4], ClinVar[5], ClinGen[6], and ISCA CNV[7]. Online software was used to analyze the structures of proteins, predict the conserved and functional domains and perform multiple sequence alignment. CNVs were classified into three categories (benign, uncertain clinical significance and pathogenic) according to the American College of Medical Genetics and Genomics (ACMG) standards and guidelines for the interpretation of CNVs.

Follow-Up

The outcomes of fetuses with EIF were followed by telephone. Follow-up was conducted when the children were 3–6 months, 6–12 months, and 1–2 years of age. The results of ultrasonography and physical examinations after birth were recorded.

Statistics

Identifiable personal information was removed from the data used in the analysis to protect individuals' privacy. The data are presented as mean with median and n (% or ‰). Statistical significance was evaluated using the Student's *t*-test, Chi-square test or Fisher's exact test, and $p < 0.05$ was considered statistically significant.

RESULTS

Characteristics of Patients

From January 2018 to April 2020, 3,377 invasive pregnancy diagnoses were performed in our center. Among them, 571 (571/3,377, 16.9%) fetuses with USMs (USMs) were enrolled in this study. EIFs, which were detected in 150 fetuses, comprised 4.44% (150/3,377) of the entire diagnosis population and 26.27% (150/571) of the fetuses with USMs. The mean ages of the pregnant women with EIF and USM groups were 29.72 and 29.67 years, respectively (the EIF group ranged from 21 to 44 years, median 29 years, *SD* = 4.91; the USMs group ranged from 18 to 44 years, median 29 years, *SD* = 4.67). Student's t-test revealed no significant differences in the ages of these two groups ($p = 0.912$). As shown in **Table 1**, 82 (82/150, 54.67%) cases had only one EIF, and 68 (68/150, 45.33%) cases showed multiple EIFs by ultrasound. The locations of the EIFs were recorded; most of them were in left ventricle (113/150, 75.33%), some were in both ventricles (31/150, 20.67%), and others were found in right ventricle (6/150, 4.00%).

Chromosomal Abnormalities in Fetuses With EIFs

The main chromosomal abnormalities detected through prenatal diagnosis were abnormal karyotypes and pCNVs (pathogenic CNVs and likely pathogenic CNVs). As shown in **Table 2**, among the 571 fetuses with USMs, a total of 44 (44/571, 7.71%) chromosomal abnormalities were identified; 28 (28/571, 4.90%) were detected by karyotype analysis, including 16 cases of trisomy 21 (16/28, 57.14% of karyotype abnormalities, one presented a Robertson translocation between chromosome 14 and 21), 5 cases of trisomy 18 (5/28, 17.86% of karyotype abnormalities), 6 cases of sex chromosome aneuploidies (6/28, 21.43% of karyotype abnormalities, of which 3 were mosaics), and one case with large fragment deletion in chromosome 18 (1/28, 3.57% of karyotype abnormalities). 16 (16/571, 2.80%) anomalies were pathogenic or likely pathogenic CNVs (pCNVs) which smaller than 5 Mb. In which, one case (1/16, 6.25% of pCNVs)detected microdeletion related to a X-linked ichthyosis, one case (1/16, 6.25% of pCNVs) was Y chromosome microdeletion related to spermatogenic failure, two were chromosome 22q11.2 microduplication syndromes (2/16, 12.50% of pCNVs), two cases detected chromosome 16p11.2 deletion syndromes (2/16, 12.50% of pCNVs), pCNVs in three cases were microdeletions on chromosome 15 (3/16, 18.75% of pCNVs), two pathological/likely pathological microdeletions were on chromosome 17 (2/16, 12.50% of pCNVs), two were microdeletions on chromosome 9 (2/16, 12.50% of pCNVs), and 3 pCNVs were found on chromosome2 10, and 20, respectively (**Table 3** and **Supplementary Table 1**).

For further investigation, we focused on the association of EIFs and chromosomal abnormalities. Compared to fetuses

TABLE 1 | Characteristics of EIFs in fetuses.

Types	No. of patients
Total	150
Focus number	
Single EIF	82 (54.67%)
Multi EIFs	68 (45.33%)
Focus position	
Left ventricle	113 (75.33%)
Right ventricle	6 (4.00%)
Both ventricles	31 (20.67%)

[1] https://decipher.sanger.ac.uk/
[2] http://dgv.tcag.ca/
[3] http://www.internationalgenome.org/
[4] http://omim.org/
[5] http://www.ncbi.nlm.nih.gov/clinvar
[6] https://www.clinicalgenome.org/
[7] https://www.iscaconsortium.org

TABLE 2 | Incidence of chromosomal abnormalities in fetuses with EIFs and USMs.

	No. of patients	No. of patients with chromosomal abnormalities (% of No. of patients)	*p*	No. of patients with abnormal karyotype (% of No. of patients)	*p*	No. of patients with < 5Mb pCNVs (% of No. of patients)	*p*
USMs	571	44 (7.71)		28 (4.90)		16 (2.80)	
EIFs	150	6 (4.00)	0.112[a]	2 (1.33)	0.051[a]	4 (2.67)	1.000[a]
Isolated EIFs	59	0		0		0	
With other USMs	91	6 (6.59)	0.082[b]	2 (2.20)	0.5200[b]	4 (4.40)	0.154[b]
Gestational weeks <28	135	3 (2.22)		2 (1.48)		1 (0.74)	
Gestational weeks ≥28	15	3 (20.00)	0.014[c]	0	1.000[c]	3 (20.00)	0.003[c]
Age <35	122	5 (4.10)		2 (1.64)		3 (2.46)	
Age ≥35	28	1 (3.57)	1.000[d]	0	1.000[d]	1 (3.57)	0.566[d]

[a]*Compared with USM group, by Chi-square (X^2) test.*
[b]*Compared with isolated EIF group, by fisher's exact test.*
[c]*Compared with gestational weeks <28, by fisher's exact test.*
[d]*Compared with maternal age <35, by fisher's exact test.*

TABLE 3 | Chromosomal abnormalities and pregnancy outcomes of fetuses withEIFs.

Case No.	Maternal age (years)	Ultrasound indications	CNV-seq results		Pathogenicity	Karyotype	Pregnancy outcome
			Segments of CNVs	Clinical phenotype for ultrasound findings			
18S4620963	30	Echogenic intracardiac focus, echogenic bowel, cerebral ventriculomegaly	Trisomy 21	Down's syndrome	Pathogenic	47,XY, + 21	Labor induction
18S4364136	23	Echogenic intracardiac focus, thicken of nuchal translucence	Trisomy 21	Down's syndrome	Pathogenic	47,XY, + 21	Labor induction
19S2775503	27	Echogenic intracardiac focus, nasal bone dysplasia,echogenic bowel	del(17p11.2p12).seq[GRCh37/hg19] (14,989,438–16,852,433) (1.86 Mb)	–	Pathogenic	46,XX	Labor induction
20S2508263	41	Echogenic intracardiac focus, mild pyelectasis	del(15q11.2).seq[GRCh37/hg19] (22,646,193–23,514,853) (868.66 Kb)	Congenital heart disease	Likely pathogenic	46,XY	Full-term delivery, develop normal at 3 months
19S3641497	33	Echogenic intracardiac focus, permanent right umbilical vein	46,XN,dup(22q11.21).seq[GRCh37/hg19] (18,765,311–21,630,621) (2.87 Mb)	Congenital heart disease	Pathogenic	46,XY	Refused to disclose
18S3921249	28	Echogenic intracardiac focus, mild pyelectasis	del(9p22.3).seq[GRCh37/hg19] (14,263,199–15,390,161) (1.13 Mb)	Renal hypoplasia	Likely pathogenic	46,XX	Refused to disclose

with USMs, 6 of the total 150 fetuses (6/150, 4.00%) with EIFs suffered from chromosomal abnormalities, but the difference was not significant (4.00 vs. 7.71%, $p = 0.112$). In detailed, the chromosomal abnormalities included 2 (2/150, 1.33%) karyotype abnormalities (both were trisomy 21) and 4 (4/150, 2.67%) pCNVs, similar to the USMs group (**Table 2**).

Subsequently, the fetuses with EIFs were divided into two groups based on existence of other USMs. Among the 91 fetuses who presented other USMs beside EIFs, chromosomal abnormalities were detected in 6 (6/91, 6.59%) fetuses, including 2 cases of trisomy 21 (2/91, 2.20%) and 4 cases of pCNVs (4/91, 4.40%); while no abnormal karyotype or pCNVs were found in 59 fetuses showed isolated EIFs.

Finally, we focused on the relationship between fetal EIFs and gestational stage/mother age. When scanned by ultrasound, 135 EIF cases were found and referred to prenatal diagnosis at the second trimester (gestational stage ranged from 18 to 26 weeks); the EIFs in the remaining 15 fetuses were detected until the late second trimester, and the invasive diagnosis was conducted over 28 weeks of gestation (from 28 to 31 weeks). The incidence of chromosomal abnormalities was dramatically increased in group over 28 weeks of gestation (2.22% in less than 28 weeks of gestation vs. 20.00% in over 28 weeks of

gestation, p = 0.014). This discrepancy could be attributed to the number of pCNVs, for example, only 1 out of 135 (0.74%) pregnancies before 28 weeks of gestation were found pCNVs, and this incidence increased to 20.00% (3 in 15) in the pregnancies over 28 weeks (p = 0.003). On the other hand, there was no difference in the observed incidence of chromosomal abnormalities in fetuses with EIFs when the maternal ages were below or above 35 years (4.10 vs. 3.57%, p = 1.000, **Table 2**).

Pregnancy Outcomes

Six chromosomal abnormalities in the group of fetuses with EIFs and the outcomes of these pregnancies were shown in **Table 3**. Three cases terminated pregnancy, 2 refused to disclose the pregnancy outcome, and 1 fetus obtained the same likely pCNV from the mother and was followed up to 3 months after birth. In the latter case, physical examination was normal, but ultrasound examination could not be conducted.

In addition, 144 fetuses without pCNVs were followed up from birth to 2 years of age. Of these, 127 were born full term, 5 were born prematurely at 33–36 W of gestation, 3 mothers chose labor induction, and 9 patients refused to disclose their pregnancy outcomes. As shown in **Table 4**, 25 children underwent re-examination by ultrasound after birth. Among them, 9 cases (6.25%, 9/144) of congenital heart disease were found, in which 1 patent foramen ovale self-cured at 9 months, 1 case had a valve defect and accepted cardiac surgery at 1 year old, 2 displayed persistent EIFs without obvious heart defects, and 14 showed no heart defects (the oldest was over 2 years of age at the last examination and developed normally). In addition, although no re-examination after birth, 22 pregnant women accepted ultrasound screening several weeks after invasive prenatal diagnosis, EIFs were still detected in 6 (6/22, 27.27%) of these fetuses, whereas disappeared in 16 (16/22, 72.73%) fetuses. The incidence of congenital heart disease after birth was not different in fetuses with a single EIF and those with more than one EIF (p = 1.000). In the group that received ultrasound before birth, the ratio of EIF disappearance on re-examination was similar in fetuses who showed only one EIF and those with multiple EIFs (p = 0.136).

TABLE 4 | Re-examination results by ultrasound.

	No. of patients	Single EIF	Multi EIFs	p*
Ultrasound after birth	25	10	15	
Congenital heart disease	9	4	5	1.000
Patent foramen oval	4	2	2	
Ventricular septal defect	3	1	2	
Valve defect	1	0	1	
Right aortic arch	1	1	0	
Echogenic intracardiac focus	2	0	2	
Normal	14	6	8	
Ultrasound before birth	22	12	10	
With EIF	6	3	3	0.136
Without EIF	16	9	7	

**Compared between fetuses with single EIF and multi EIFs.*

DISCUSSION

The current study investigated fetuses who presented EIFs in prenatal diagnosis. From year 2018 to 2020, among 3,377 cases of prenatal diagnosis in our center, the incidence of EIFs identified by sonographic screening was 4.44% (150 in 3,377); and the most common cardiac lesions were in the left ventricle (75.33%), in accordance to the previous literatures (How et al., 1994; Bromley et al., 1995, 1998; Wax et al., 2000; Nyberg et al., 2001; Wang et al., 2018). In our data, EIFs, comprised 26.27% of USMs (150 in 571), were one of the most prevalent USMs in fetuses. Moreover, in clinical consulting, it is confusing when EIFs present without other information available from serum screening or without association with a definite structural anomaly by ultrasound. Thus, this investigation focused on fetuses who presented EIFs while patients do not have other definite or high-risk indications for chromosomal abnormalities.

Here, we analyzed the risk of fetal chromosomal abnormalities by both karyotyping and CNV-seq. Among a total of 571 pregnancies that displayed USMs, beside 28 fetuses with karyotype abnormalities (28/571, 4.90%), 16 (16/571, 2.80%) extra pathogenic CNVs were identified. The pCNVs founded here can result in a wide range of syndromes including X-linked ichthyosis, spermatogenic failure, 22q11.2 microduplication syndromes, 16p11.2 deletion syndrome, etc. Notably, we did not detect pCNVs cause serious heart defects in fetuses showed EIFs, such as microdeletions on 22q11.2 for DiGeorge syndrome. This may due to our exclusion of samples who presented structure defects under ultrasound screening. To be noted, besides the USM indications, including EIFs, other phenotypes can be detect by CNV-seq. For example, the detection of two 16p11.2 microdeletion syndromes came from both fetuses displayed external left superior cavity, which has no relation with typical epilepsies or intellectual disability of 16p11.2 microdeletion syndromes; and in fetuses showed thickened nuchal translucence, a wide spectrum of pCNVs involved from spermatogenesis to mental development was found; in fetuses showed EIFs and other USMs simultaneously, two pCNVs had no correlation with their ultrasound results were detected. Thus, the application of CNVs detection can expand the detection area of syndromes in prenatal diagnosis, and the symptoms of underlying syndromes may not only restrict in ultrasound findings. Moreover, our data is in agreement with previous reports based on chromosomal microarray analyses, which demonstrated that fetuses with abnormal ultrasound findings included 2.8–3.5% pathogenic CNVs that were not detectable by karyotyping (Wapner et al., 2012; Society for Maternal-Fetal Medicine et al., 2016; Lostchuck et al., 2019). Therefore, although aneuploidy in EIF fetuses has been extensively studied, subchromosomal abnormalities still need to be evaluated.

In fetuses with EIFs, the incidence of chromosomal abnormalities was 4.00% which is not different from the

whole USM group (vs. 7.71%, $p = 0.125$), and is also similar to other soft markers that had been proposed previously (Nyberg et al., 2001). In EIF group, among 59 fetuses showed isolated EIFs, no chromosomal abnormality was found. Our data support the idea that EIFs alone is not indicative for an increased risk of chromosomal abnormalities (Bromley et al., 1998; Manning et al., 1998; Lamont et al., 2004; Bradley et al., 2005; Van Den Hof et al., 2005).

It is generally accepted that nearly half of EIFs observed in the second trimester may resolve with advancing gestation, although the genetic characteristics of the remaining EIF-positive fetuses are still not clear (Hurd and Nelson, 2009; Su et al., 2011). We analyzed the results of 15 invasive diagnosed cases conducted at late gestational stages in which echogenic foci persisted through the late second to third trimester. No aneuploidy was detected at gestational stages over 28 weeks, and this was not different from the earlier gestational age group. Interestingly, 20% incidence (3 in 15 cases) of pCNVs in the late gestational age group was significantly higher than that was found in earlier gestational age group (0.74%, $p = 0.003$). This result suggests that EIFs, especially appeared simultaneously with other USMs that persist to late gestational ages could indicate higher risk of pathogenic CNVs which may be missed by conventional karyotyping. In addition, because of its shorter reporting cycle compared to karyotyping, CNV-seq has more advantages for patients at late gestational stages. This finding emphasizes the efficiency of CNV-seq in a field in which conventional karyotyping was previously considered standard.

Maternal age is generally assumed as an important factor that associated with the aneuploidy rate in fetuses with EIFs (Bromley et al., 1998; Goncalves et al., 2006). However, when we analyzed the data on the basis of maternal age (younger vs. older age), no correlation of age was found. This may be due to the relatively low incidence of aneuploidy in our study, which is attributed to the exclusion of some high-risk samples by serum screening, NIPT, or fetal structural anomalies, etc. The incidence of pCNVs was also similar in old and young patients, in according with the idea that CNVs can occur in any pregnancy independent of maternal ages (Chau et al., 2019).

In our clinical management, the parents of fetuses who have been detected of carrying pCNVs were suggested to take the verification of CNVs. Meanwhile, the ultrasound results throughout pregnancy were considered comprehensively together with pCNVs during clinical consulting. Moreover, we have conducted telephone calls to follow up with the outcomes of fetuses with pCNVs. It was regrettable that the pregnancy terminations were not conducted in our hospital; for this reason, autopsies to verify the dysplasia could not be performed. In one case of full-term birth, though the infant possessed a likely pathogenic CNV that was inherited from the mother, the infant behaved normally and showed normal development according to gross physical examination. To be noted, cardiac ultrasound was not performed because the infant was only 3 months old at the time of follow-up, furthermore, the development of intelligence and language could not be assessed.

Previously, EIF was considered as a normal developmental variant that was not associated with congenital heart disease (Simpson et al., 1996; Wolman et al., 2000; Wax et al., 2003). However, in clinical consulting, cases with EIFs were found to be associated with heart defects, causing confusion and possibly leading to poor prognosis. A conflicting conclusion was reached that the prevalence of cardiac defects in fetuses with EIFs was twice than that in the general newborn population (Goncalves et al., 2006). Furthermore, some benign defects, such as ventricular septal defect and patent foramen ovale maybe self-cure prior to 3 years of age (Jortveit et al., 2016; Cho et al., 2017) and might not be observed at the time of re-examination, could cause an artificial decrease of the heart defect incidence. Here, we followed the outcomes of fetuses and found that among 25 neonates who underwent cardiac ultrasound prior to 6 months of age, 9 (36.00%) presented abnormalities, mostly with patent foramen ovale (44.44%) and ventricular septal defects (33.33%). This result was relatively higher than expected and may be that the population underwent re-examination after birth including cases with persisted EIFs throughout the pregnancy. It is reported that 8–75% of EIFs disappear as gestation progresses (Arda et al., 2007; Lorente et al., 2017); a similar percentage was found in our study, in which 72.73% of EIFs could not be detected in re-emanations after invasive prenatal diagnosis. Thus, we calculated that the percentage of heart defects in the total population in this study was 6.25%, higher than the 0.8% observed in the general population (Prefumo et al., 2003). This was consistent with previous reports, suggests that EIFs in fetuses may be associated with the presence of heart defects (Lorente et al., 2017; Guo et al., 2018).

In conclusion, by applying high-resolution sequencing simultaneously with conventional karyotyping, additional chromosomal anomalies were detected in our study. The still-controversial association between EIFs and chromosomal abnormalities was reevaluated using data on CNVs. Isolated EIFs appear not to indicate increased risk of chromosomal abnormalities. Interestingly, although most EIFs disappeared with advanced stages of gestation, EIFs detected with other USMs that persisted in late trimester may indicate the presence of pathogenic CNVs in the fetus. Thus, at late gestational ages, CNVs detection has its advantages over karyotyping. By analyzing the follow up data, more congenital heart disease was found after birth in fetuses with EIFs than in fetuses with normal conception and no ultrasound abnormality, a comprehensive sonographic examination throughout the entire pregnancy and after birth could be recommended. Our investigation provides detailed information from the perspective of genetic disorders and prognosis after birth of fetuses with EIFs for clinical consulting.

DATA AVAILABILITY STATEMENT

According to the national legislation/guidelines, specifically the Administrative Regulations of the People's Republic

of China on Human Genetic Resources (http://www.gov.cn/zhengce/content/2019-06/10/content_5398829.htm, http://english.www.gov.cn/policies/latest_releases/2019/06/10/content_281476708945462.htm), no additional raw data are available at this time. The data presented in the study are temporarily deposited in the BGI data delivery system (http://cdts.genomics.org.cn/). The accession code 20210310F12FHQHSSJ0877B and the password ZhaoYanF12345! should be included in the application, or you can email to the corresponding authors for details.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of First Affiliated Hospital of the University of Science and Technology of China (USTC). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YS and JW conceived and designed the study. YS, JX, and XJ wrote the manuscript. HL performed the ultrasound scan. JG and XJ contributed to manuscript revision. JW, LW, and GH provided clinical information of patients and conducted invasive diagnosis. YS, JX, WL, YH, and YP performed the experiments and genetic detections. YS and FY completed the follow ups. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Fundamental Research Funds for the Central Universities (WK9110000051) and Anhui Provincial Key R&D Program (202004j07020024).

ACKNOWLEDGMENTS

We gratefully acknowledge all of our patients for agreeing to allow their personal data to be used in our study and for allowing these data to be published, and we would like to thank Kang Zhang, and other members of the genetic analysis group and the high-throughput sequencing facilities of BGI.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.626044/full#supplementary-material

REFERENCES

Achiron, R., Lipitz, S., Gabbay, U., and Yagel, S. (1997). Prenatal ultrasonographic diagnosis of fetal heart echogenic foci: no correlation with Down syndrome. *Obstet. Gynecol.* 89, 945–948. doi: 10.1016/s0029-7844(97)00131-2

Arda, S., Sayin, N. C., Varol, F. G., and Sut, N. (2007). Isolated fetal intracardiac hyperechogenic focus associated with neonatal outcome and triple test results. *Arch. Gynecol. Obstet.* 276, 481–485. doi: 10.1007/s00404-007-0366-9

Bradley, K. E., Santulli, T. S., Gregory, K. D., Herbert, W., Carlson, D. E., and Platt, L. D. (2005). An isolated intracardiac echogenic focus as a marker for aneuploidy. *Am. J. Obstet. Gynecol.* 192, 2021–2026;discussion2026–2028.

Bromley, B., Lieberman, E., Laboda, L., and Benacerraf, B. R. (1995). Echogenic intracardiac focus: a sonographic sign for fetal Down syndrome. *Obstet. Gynecol.* 86, 998–1001. doi: 10.1016/0029-7844(95)00323-j

Bromley, B., Lieberman, E., Shipp, T. D., Richardson, M., and Benacerraf, B. R. (1998). Significance of an echogenic intracardiac focus in fetuses at high and low risk for aneuploidy. *J. Ultrasound Med.* 17, 127–131. doi: 10.7863/jum.1998.17.2.127

Chau, M. H. K., Cao, Y., Kwok, Y. K. Y., Chan, S., Chan, Y. M., Wang, H., et al. (2019). Characteristics and mode of inheritance of pathogenic copy number variants in prenatal diagnosis. *Am. J. Obstet. Gynecol.* 221, 493.e1–493.e11.

Chiu, G., Zhao, A., Zhang, B., and Zhang, T. (2019). Intracardiac echogenic focus and its location: association with congenital heart defects. *J. Matern. Fetal Neonatal. Med.* 32, 3074–3078. doi: 10.1080/14767058.2018.1558200

Cho, Y. S., Park, S. E., Hong, S. K., Jeong, N. Y., and Choi, E. Y. (2017). The natural history of fetal diagnosed isolated ventricular septal defect. *Prenat. Diagn.* 37, 889–893. doi: 10.1002/pd.5100

Choi, E. Y., Hong, S. K., and Jeong, N. Y. (2016). Clinical characteristics of prenatally diagnosed persistent left superior vena cava in low-risk pregnancies. *Prenat. Diagn.* 36, 444–448. doi: 10.1002/pd.4801

Clinical Genetics Group Of Medical Genetics Branch Chinese Medical Association, Professional Committee For Prenatal Diagnosis Of Genetic Diseases Medical Genetics Branch Of Chinese Medical Association, and Group Of Genetic Disease Prevention And Control Birth Defect Prevention And Control Committee Of Chinese Society Of Preventive Medicine (2019). [Expert consensus on the application of low-depth whole genome sequencing in prenatal diagnosis]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 36, 293–296.

Cohen, K., Tzika, A., Wood, H., Berri, S., Roberts, P., Mason, G., et al. (2015). Diagnosis of fetal submicroscopic chromosomal abnormalities in failed array CGH samples: copy number by sequencing as an alternative to microarrays for invasive fetal testing. *Ultrasound Obstet. Gynecol.* 45, 394–401. doi: 10.1002/uog.14767

Degani, S., Leibovitz, Z., Shapiro, I., Gonen, R., and Ohel, G. (2001). Cardiac function in fetuses with intracardiac echogenic foci. *Ultrasound Obstet. Gynecol.* 18, 131–134. doi: 10.1046/j.1469-0705.2001.00433.x

Dildy, G. A., Judd, V. E., and Clark, S. L. (1996). Prospective evaluation of the antenatal incidence and postnatal significance of the fetal echogenic cardiac focus: a case-control study. *Am. J. Obstet. Gynecol.* 175, 1008–1012. doi: 10.1016/s0002-9378(96)80043-3

Committee on Genetics and the Society for Maternal-Fetal Medicine (2016). Committee opinion no.682: microarrays and next-generation sequencing technology: the use of advanced genetic diagnostic tools in obstetrics and gynecology. *Obstet. Gynecol.* 128, e262–e268.

Goncalves, T. R., Zamith, M. M., Murta, C. G., Bussamra, L. C., Torloni, M. R., and Moron, A. F. (2006). Chromosomal and cardiac anomalies in fetuses with intracardiac echogenic foci. *Int. J. Gynaecol. Obstet.* 95, 132–137. doi: 10.1016/j.ijgo.2006.06.020

Guo, Y., He, Y., Gu, X., Zhang, Y., Sun, L., Liu, X., et al. (2018). Echogenic intracardiac foci and fetal cardiac anomalies: a review of cases from a tertiary care center in China. *J. Clin. Ultrasound* 46, 103–107. doi: 10.1002/jcu.22533

How, H. Y., Villafane, J., Parihus, R. R., and Spinnato, J. A. II (1994). Small hyperechoic ventricle: a benign foci of the fetal cardiac sonographic finding? *Ultrasound Obstet. Gynecol.* 4, 205–207. doi: 10.1046/j.1469-0705.1994.04030205.x

Hurd, P. J., and Nelson, C. J. (2009). Advantages of next-generation sequencing versus the microarray in epigenetic research. *Brief Funct. Genomic. Proteomic.* 8, 174–183. doi: 10.1093/bfgp/elp013

Jortveit, J., Leirgul, E., Eskedal, L., Greve, G., Fomina, T., Dohlen, G., et al. (2016). Mortality and complications in 3495 children with isolated ventricular septal defects. *Arch. Dis. Child.* 101, 808–813. doi: 10.1136/archdischild-2015-310154

Lamont, R. F., Havutcu, E., Salgia, S., Adinkra, P., and Nicholl, R. (2004). The association between isolated fetal echogenic cardiac foci on second-trimester ultrasound scan and trisomy 21 in low-risk unselected women. *Ultrasound Obstet. Gynecol.* 23, 346–351. doi: 10.1002/uog.1018

Lide, B., Lindsley, W., Foster, M. J., Hale, R., and Haeri, S. (2016). Intrahepatic persistent right umbilical vein and associated outcomes: a systematic review of the literature. *J. Ultrasound Med.* 35, 1–5. doi: 10.7863/ultra.15.01008

Lorente, A. M. R., Moreno-Cid, M., Rodriguez, M. J., Bueno, G., Tenias, J. M., Roman, C., et al. (2017). Meta-analysis of validity of echogenic intracardiac foci for calculating the risk of Down syndrome in the second trimester of pregnancy. *Taiwan J. Obstet. Gynecol.* 56, 16–22. doi: 10.1016/j.tjog.2016.11.002

Lostchuck, E., Poulton, A., Halliday, J., and Hui, L. (2019). Population-based trends in invasive prenatal diagnosis for ultrasound-based indications: two decades of change from 1994 to 2016. *Ultrasound Obstet. Gynecol.* 53, 503–511. doi: 10.1002/uog.19107

Manning, J. E., Ragavendra, N., Sayre, J., Laifer-Narin, S. L., Melany, M. L., Grant, E. G., et al. (1998). Significance of fetal intracardiac echogenic foci in relation to trisomy 21: a prospective sonographic study of high-risk pregnant women. *AJR Am. J. Roentgenol.* 170, 1083–1084. doi: 10.2214/ajr.170.4.9530064

Nyberg, D. A., Souter, V. L., El-Bastawissi, A., Young, S., Luthhardt, F., and Luthy, D. A. (2001). Isolated sonographic markers for detection of fetal Down syndrome in the second trimester of pregnancy. *J. Ultrasound Med.* 20, 1053–1063. doi: 10.7863/jum.2001.20.10.1053

Prefumo, F., Presti, F., Thilaganathan, B., and Carvalho, J. S. (2003). Association between increased nuchal translucency and second trimester cardiac echogenic foci. *Obstet. Gynecol.* 101, 899–904. doi: 10.1016/s0029-7844(02)03128-9

Rembouskos, G., Passamonti, U., De Robertis, V., Tempesta, A., Campobasso, G., Volpe, G., et al. (2012). Aberrant right subclavian artery (ARSA) in unselected population at first and second trimester ultrasonography. *Prenat. Diagn.* 32, 968–975. doi: 10.1002/pd.3942

Sepulveda, W., and Romero, D. (1998). Significance of echogenic foci in the fetal heart. *Ultrasound Obstet. Gynecol.* 12, 445–449. doi: 10.1046/j.1469-0705.1998.12060445.x

Simpson, J. M., Cook, A., and Sharland, G. (1996). The significance of echogenic foci in the fetal heart: a prospective study of 228 cases. *Ultrasound Obstet. Gynecol.* 8, 225–228. doi: 10.1046/j.1469-0705.1996.08040225.x

Society for Maternal-Fetal Medicine, Dugoff, L., Norton, M. E., and Kuller, J. A. (2016). The use of chromosomal microarray for prenatal diagnosis. *Am. J. Obstet. Gynecol.* 215, B2–B9.

Su, Z., Li, Z., Chen, T., Li, Q. Z., Fang, H., Ding, D., et al. (2011). Comparing next-generation sequencing and microarray technologies in a toxicological study of the effects of aristolochic acid on rat kidneys. *Chem. Res. Toxicol.* 24, 1486–1493. doi: 10.1021/tx200103b

Van Den Hof, M. C., Wilson, R. D., Diagnostic Imaging Committee, Society of Obstetricians and Gynaecologists of Canada, and Genetics Committee, Society of Obstetricians and Gynaecologists of Canada. (2005). Fetal soft markers in obstetric ultrasound. *J. Obstet. Gynaecol. Can.* 27, 592–636. doi: 10.1016/s1701-2163(16)30720-4

Wang, J., Chen, L., Zhou, C., Wang, L., Xie, H., Xiao, Y., et al. (2018). Identification of copy number variations among fetuses with ultrasound soft markers using next-generation sequencing. *Sci. Rep.* 8:8134.

Wapner, R. J., Martin, C. L., Levy, B., Ballif, B. C., Eng, C. M., Zachary, J. M., et al. (2012). Chromosomal microarray versus karyotyping for prenatal diagnosis. *N. Engl. J. Med.* 367, 2175–2184.

Wax, J. R., Donnelly, J., Carpenter, M., Chard, R., Pinette, M. G., Blackstone, J., et al. (2003). Childhood cardiac function after prenatal diagnosis of intracardiac echogenic foci. *J. Ultrasound Med.* 22, 783–787. doi: 10.7863/jum.2003.22.8.783

Wax, J. R., Royer, D., Mather, J., Chen, C., Aponte-Garcia, A., Steinfeld, J. D., et al. (2000). A preliminary study of sonographic grading of fetal intracardiac echogenic foci: feasibility, reliability and association with aneuploidy. *Ultrasound Obstet. Gynecol.* 16, 123–127. doi: 10.1046/j.1469-0705.2000.00206.x

Winter, T. C., Anderson, A. M., Cheng, E. Y., Komarniski, C. A., Souter, V. L., Uhrich, S. B., et al. (2000). Echogenic intracardiac focus in 2nd-trimester fetuses with trisomy 21: usefulness as a US marker. *Radiology* 216, 450–456. doi: 10.1148/radiology.216.2.r00au32450

Wolman, I., Jaffa, A., Geva, E., Diamant, S., Strauss, S., Lessing, J. B., et al. (2000). Intracardiac echogenic focus: no apparent association with structural cardiac abnormality. *Fetal Diagn. Ther.* 15, 216–218. doi: 10.1159/000021009

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: 01 April 2021
doi: 10.3389/fgene.2021.636900

A Novel *FLCN* Intragenic Deletion Identified by NGS in a BHDS Family and Literature Review

Minghui Cai[1†], Xinxin Zhang[2†], Lizhen Fan[3], Shuwen Cheng[3], Abdukahar Kiram[3], Shaoqin Cen[3], Baofu Chen[1], Minhua Ye[1], Qian Gao[3], Chengchu Zhu[1], Long Yi[1,3‡] and Dehua Ma[1*‡]*

[1] Department of Cardiothoracic Surgery, Taizhou Hospital of Zhejiang Province affiliated to Wenzhou Medical University, Linhai, China, [2] Department of Histology and Embryology, School of Medicine, Southeast University, Nanjing, China, [3] Jiangsu Key Laboratory of Molecular Medicine, School of Medicine, Nanjing University, Nanjing, China

OPEN ACCESS

Edited by:
Katalin Komlosi,
University Medical Center Freiburg, Germany

Reviewed by:
Laura Schmidt,
National Cancer Institute at Frederick, United States
Nikoletta Nagy,
University of Szeged, Hungary

***Correspondence:**
Dehua Ma
madh@enzemed.com
Long Yi
yilong@nju.edu.cn

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

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

Received: 02 December 2020
Accepted: 05 March 2021
Published: 01 April 2021

Citation:
Cai M, Zhang X, Fan L, Cheng S, Kiram A, Cen S, Chen B, Ye M, Gao Q, Zhu C, Yi L and Ma D (2021) A Novel FLCN Intragenic Deletion Identified by NGS in a BHDS Family and Literature Review. Front. Genet. 12:636900. doi: 10.3389/fgene.2021.636900

Birt–Hogg–Dubé syndrome (BHDS, MIM #135150), caused by germline mutations of *FLCN* gene, is a rare autosomal dominant inherited disorder characterized by skin fibrofolliculomas, renal cancer, pulmonary cysts and spontaneous pneumothorax. The syndrome is considered to be under-diagnosed due to variable and atypical manifestations. Herein we present a BHDS family. Targeted next generation sequencing (NGS) and multiplex ligation-dependent probe amplification (MLPA) revealed a novel *FLCN* intragenic deletion spanning exons 10-14 in four members including the proband with pulmonary cysts and spontaneous pneumothorax, one member with suspicious skin lesions and a few pulmonary cysts, as well as two asymptomatic family members. In addition, a linkage analysis further demonstrated one member with pulmonary bullae to be a BHDS-ruled-out case, whose bullae presented more likely as an aspect of paraseptal emphysema. Furthermore, the targeted NGS and MLPA data including our previous and present findings were reviewed and analyzed to compare the advantages and disadvantages of the two methods, and a brief review of the relevant literature is included. Considering the capability of the targeted NGS method to detect large intragenic deletions as well as determining deletion junctions, and the occasional false positives of MLPA, we highly recommend targeted NGS to be used for clinical molecular diagnosis in suspected BHDS patients.

Keywords: BHD syndrome, *FLCN* gene, large intragenic deletion, targeted NGS, MLPA

INTRODUCTION

Birt–Hogg–Dubé syndrome (BHDS, MIM #135150), caused by germline mutations of *FLCN* gene, is a rare autosomal dominant inherited disorder characterized by skin fibrofolliculomas, renal cancer, pulmonary cysts and spontaneous pneumothorax (Roth et al., 1993; Birt et al., 1977; Nickerson et al., 2002). It is often considered to be underdiagnosed due to variable and atypical

manifestations. Multiple and bilateral pulmonary cysts are the most common manifestation of BHDS and can be observed in more than 80% of BHDS patients (Zbar et al., 2002; Schmidt et al., 2005; Toro et al., 2007; Agarwal et al., 2011). They could exhibit a pneumothorax dominant phenotype with no or reduced penetrance of the skin or renal manifestations (Ren et al., 2008). It is in some degree challenging for clinicians or radiologists to distinguish these BHDS patients from patients with other cystic lung diseases such as lymphangioleiomyomatosis, lymphoid interstitial pneumonia and Langerhans cell histiocytosis (Raoof et al., 2016). However, various cystic lung diseases could have a characteristic CT appearance in terms of distribution, extent and morphology of cysts (Agarwal et al., 2011; Raoof et al., 2016), that allows their distinction, which could narrow the differential diagnosis considerably. Early and accurate diagnosis of BHDS is crucial, which will lead to early identification and treatment of renal cancer in patients and their family members. Due to the variability in the clinical manifestations and the complicacy of diagnostic criteria (Menko et al., 2009), a DNA-based diagnosis is necessary.

FLCN, currently the only gene known to be associated with BHDS, is located on chromosome 17p11.2, consists 14 exons, and encodes an evolutionarily conserved protein whose function has not yet been completely understood. As is shown in the online Locus-Specific Database for *FLCN*[1] (Lim et al., 2010), there are 286 unique public DNA variants (last accessed: 08/31/2020), in which small indels and nonsense mutations account for the majority of pathogenic variants detected by DNA sequencing. Whereas, splice-site mutation and large intragenic deletions/duplications are also vital because these pathogenic mutation can lead to premature protein truncation or haploinsufficiency (Kunogi et al., 2010; Furuya et al., 2016; Xu et al., 2016; Jensen et al., 2017).

At present, DNA-based diagnosis of BHDS mainly relies on Sanger sequencing and multiplex ligation-dependent probe amplification (MLPA), which makes the diagnosis more time-consuming and labor-intensive. Additionally, there are different kinds of complex mutations in *FLCN* that may not be detected by Sanger sequencing, such as large intragenic deletions or duplications and deep intronic mutations that may lead to abnormal splicing (Benhammou et al., 2011; Ding et al., 2015). Therefore, a new rapid next generation sequencing (NGS) strategy which could identify not only point mutations and indels but also copy number variations (CNV) in *FLCN* gene was developed to conduct the molecular diagnosis of BHDS (Zhang et al., 2016).

In the current study, six individuals including three patients with primary spontaneous pneumothorax (PSP) or lung cysts/bullae from a three-generation family were investigated. Linkage analysis and targeted NGS were both conducted to identify the potential mutation of *FLCN* gene.

[1]https://databases.lovd.nl/shared/genes/FLCN

CASE PRESENTATION

Clinical Report

The family (**Figure 1A**) was ascertained through a proband (II-2) with spontaneous pneumothorax, he came to the Taizhou Hospital at the age of 44 for his second attack, who had his first episode at age of 43. He was a non-smoker and he had neither skin lesion nor renal cancer. The chest computed tomography (CT) showed that he had a right-sided pneumothorax and multiple bilateral pulmonary cysts in different pulmonary segments, especially the basal lung region (**Figure 1B**). The CT scan was performed in each family member except the proband's mother (I-2) who died of colon cancer 20 years ago and revealed that the proband's father (I-1, 71 years old) and brother (II-1, 40 years old) had cysts or bullae (**Figures 1C,D**), the others (III-1, 24 years old, III-2, 14 years old and III-3, 8 years old) had no obvious abnormity in the lung. The CT images of I-1 who was a heavy smoker and had never suffered from spontaneous pneumothorax showed polygonal-shaped, emphysematous subpleural bullae in the upper lobe (**Figure 1C**). The CT images of II-1, who had skin lesions on his face and neck with multiple dome-shaped, white or skin-colored papules (**Figure 1E**), showed that he had a few 5–8 mm thin-walled cysts in the upper and middle lobes (**Figure 1D**).

Genetic Analysis

Mutation analysis of *FLCN* was performed and no pathogenic point or indel mutation was identified by Sanger sequence analysis. As phenotype analysis in this family could not exclude the possibility of BHDS, we performed a targeted NGS in twelve suspect BHDS patients without any pathogenic mutation of *FLCN* gene including the proband. Comprehensive analysis including CNV analysis was subsequently carried out. By introducing a normalized depth-based method (Zhang et al., 2016), we detected a novel large deletion spanning exons 10–14 in the proband (**Supplementary Figure 1** and **Supplementary Table 3**), who was diagnosed as BHDS consequently. By calculating the normalized depth of a set of 20 bp-interval reads covered the entire *FLCN* gene, we defined the approximate sites of the breakpoints (**Figure 2A**). One breakpoint is located near chr 17: 17113680, the other breakpoint is near chr 17: 17121880, the estimated deletion size is approximately 8.2 kb in length.

Multiplex ligation-dependent probe amplification assay was subsequently performed to validate and determine the large deletion in proband and family members mentioned above. The large deletion of exons 10–14 in the proband detected by NGS method was confirmed (**Figure 2B**), and MLPA analysis showed that the large deletion also occurred in II-1, III-1 and III-3.

To determine the precise breakpoint, the junction fragments adjacent to the deleted regions were amplified by Polymerase Chain Reaction (PCR) with specially designed primers. The 1.5 kb PCR product of the junction fragments were separated by agarose gel electrophoresis, while the 10.0 kb wild-type sequences were too long to be amplified (**Figure 2C**). Bidirectional sequencing of the PCR products indicated an

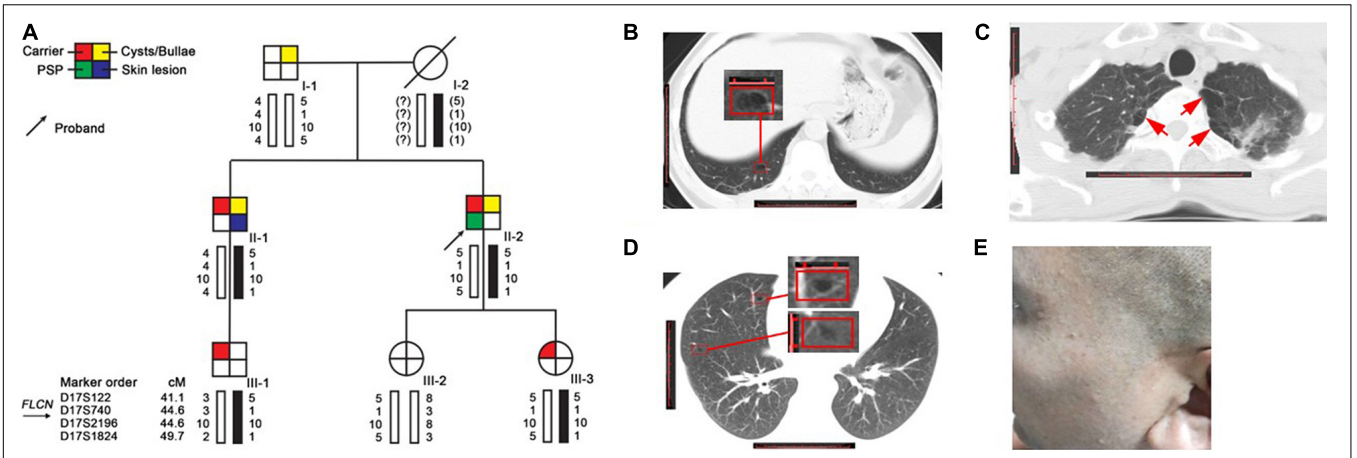


FIGURE 1 | Pedigree and clinical pictures of the family. **(A)** Pedigree and haplotype analysis. Numbers shown below each individual indicate the genotypes of the microsatellite markers in chromosomal order. The black bars represent the disease-associated haplotype. **(B–D)** Representative CT images of II-2 **(B)**, I-1 **(C)**, and II-1 **(D)**; red frame: cyst; red arrow: bullae. **(E)** Several dome-shaped, skin-colored papules were seen on the face of II-1.

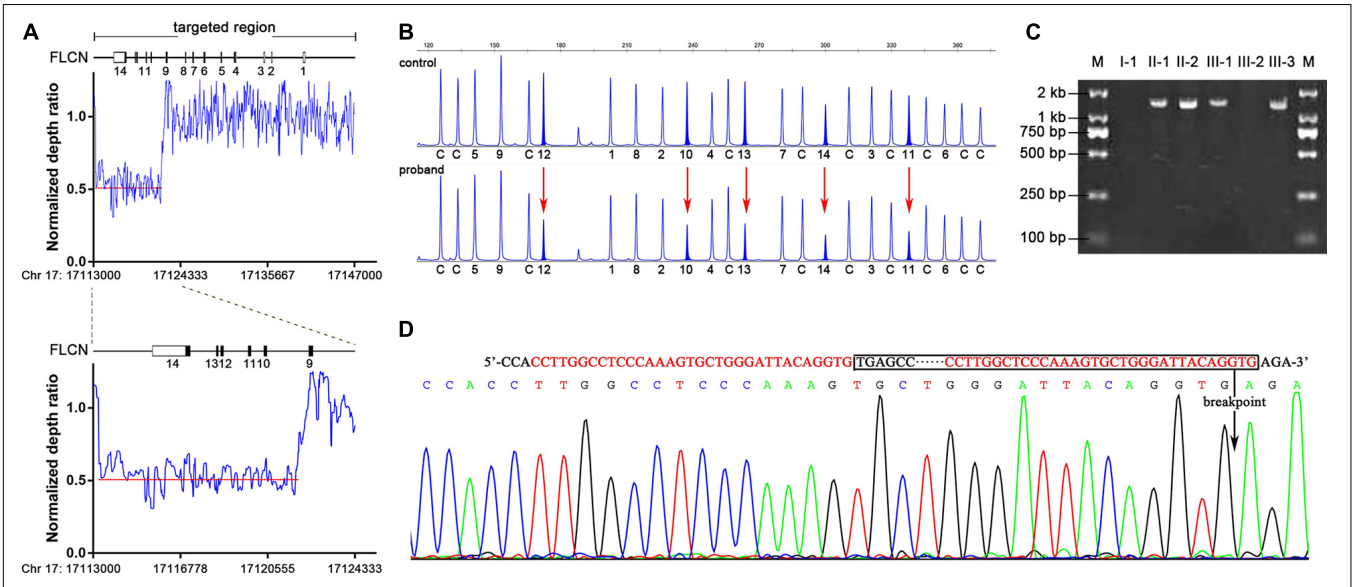


FIGURE 2 | Identification of *FLCN* intragenic deletions and breakpoint analysis in II-2. **(A)** Representation of the normalize depth ratio of 20-bp intervals in *FLCN* gene detected by NGS method. The red dotted lines represent the average ratio of the deletion area. **(B)** MLPA analysis showed heterozygous exonic deletions (red arrows) in II-2 compared with the control individual. The characters and numbers below the probes indicate exons and control probes. **(C)** PCR products of junction fragments were separated by 2% agarose gel electrophoresis. The numbers above the lanes corresponded to the subjects of the Pedigree in **Figure 1A**. **(D)** Bidirectional sequencing of the junction region. The sequences of microhomology between the up and downstream breakpoints were marked in red; boxed nucleotides indicate the deletion and an arrow indicate the location of the deletion.

8,384 bp deletion encompassing exons 10–14 (chr17: 17113559-17121942) (**Figure 2D**), the mutation was not recorded in the online Locus-Specific Database for *FLCN*[2] or ClinVar [3]. Further analysis illustrated that the up- and down-stream breakpoints of the deletion were flanked by Alu-Sg and Alu-Sx1 repeats, respectively, in which there were microhomology-mediated break-induced replication sequences (**Figure 2D**).

As results showed that I-1, who also exhibited multiple pulmonary bullae, did not carry the mutation as others (II-1, II-2, III-1, and III-3). We conducted a linkage analysis to validate the segregation of the mutation. Evidence of linkage was observed in family members with the large deletion mutation, and inferred that the father (I-1), did not pass on the affected haplotype (5-1-10-1) to the two affected children and this has ruled him out as the obligate carrier. The affected haplotype was transmitted from the mother (I-2) (**Figure 1A**), suggesting that I-2 should be regarded as an obligate carrier.

[2]https://databases.lovd.nl/shared/genes/FLCN

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

Furthermore, to evaluate the precision and accuracy of the MLPA and NGS approach for detecting exon deletions, we compared the normalized ratio and their standard deviations (SD) for the deleted and non-deleted exons from the present and previous studies (Ding et al., 2015; Zhang et al., 2016). For normal exons, the mean normalized ratio in NGS and MLPA method was 0.9488 (*SD*: 0.0956) and 1.04 (*SD*: 0.1023), respectively. For deleted exons, the mean normalized ratio in NGS and MLPA method was 0.5105 (*SD*: 0.0762) and 0.569 (*SD*: 0.0614), respectively. The corresponding histogram showed that the effectiveness in detecting exon deletions of the two methods is equivalent (**Figure 3A**). No overlap was seen between the diploid and haploid copy number values in both MLPA and NGS group. We found a cut-off value of 0.7 for scoring a deletion (<0.7) or normal copy number (>0.7) status for both methods.

Detailed methods and a visualized flow chart of deletion detection and precise breakpoints determination (**Supplementary Figure 2**) were available in the **Supplementary Material**.

DISCUSSION AND CONCLUSION

Here we report a novel *FLCN* intragenic deletion spanning exons 10–14 segregated in a BHDS family. In this family, the deletion was detected in two affected members (II-1 and II-2) as well as two asymptomatic carriers (**Figure 1A**), who were subsequently diagnosed with BHDS. The intragenic deletion spanning exon 10–14 was predicted to result in C-terminal truncation of FLCN (Ding et al., 2015). As previously reported, FLCN interacts with FNIP1 or FNIP2 through its C-terminus, truncating mutations result in loss of the C-terminal region of FLCN, abolish its interaction with FNIP1 and FNIP2, and consequently disrupt its normal function (Baba et al., 2006; Hasumi et al., 2008; Takagi et al., 2008; Woodford et al., 2016).

In this family we reported, for the proband, chest CT scanning reveals multiple bilaterally parenchymal lung cysts predominantly in the basal and periphery lung region (subpleural) (**Figure 1B**), which is a typical BHDS manifestation (Graham et al., 2005; Agarwal et al., 2011; Raoof et al., 2016). For II-1, despite his atypical CT manifestation, multiple, dome-shaped, whitish or skin-colored papules can be observed on his face and neck, the phenotype is highly considered as BHDS. Both patients are confirmed to be BHDS by genetic testing. However, CT scanning also reveals the pulmonary bullous changes in I-1 (**Figure 1C**), we found that the location and characteristics of CT feature is more likely an emphysematous destruction (Koo and Yoo, 2013), which is considered mainly related to cigarette smoking. He is eventually determined to be a BHDS-ruled-out case by genetic analysis.

Although chest CT scanning is essential for differential diagnosis, for suspected BHDS patients, we highlight the need for pedigree investigation. Segregation analysis combined with clinical and radiographic features may help differentiate among various cystic lung diseases, and genetic screening is imperative to make a definite diagnosis (Painter et al., 2005).

FLCN, responsible for BHDS, is a tumor suppressor gene that was firstly reported in 2002 (Nickerson et al., 2002) and has variable mutation types, most of which are small indels and nonsense mutations detected by Sanger sequencing, and there are also approximately 10% large intragenic deletions and duplications normally detected by MLPA.

Up to now, including our previous and present findings, totally 30 cases or families harboring 24 unique *FLCN* intragenic deletions/duplications have now been reported worldwide (**Figure 3B** and **Table 1**) (Kunogi et al., 2010; Sempau et al., 2010; Benhammou et al., 2011; Houweling et al., 2011; Babaei Jandaghi et al., 2013; Ding et al., 2015; Matsutani et al., 2016; Liu et al., 2017; Rossing et al., 2017; Iwabuchi et al., 2018; Schneider et al., 2018; Enomoto et al., 2020). Patients with *FLCN* deletions/duplications exhibit a high degree of interfamilial clinical variability, while no particular phenotype was seen more frequently in association with intragenic deletions ($p > 0.05$) (**Supplementary Table 4**). The main detection methods were quantitative PCR (qPCR) and MLPA, array-based comparative genomic hybridization (aCGH) was also employed in order to more finely map the deletions (Benhammou et al., 2011). Among the deletions, nine pairs of the breakpoints were determined by long range PCR, six pairs of the deletion breakpoints were flanked by Alu repeats, and one was partially Alu-mediated.

All the breakpoints were finally defined by long-range PCR followed by DNA sequencing. For deletions detected by qPCR and MLPA, multiple pairs of primers flanking the deleted exon(s) were designed for PCR reaction. While in our present study, as we optimized the NGS method by including the 5′ flank and 3′ flank sequence, the boundaries of the deletions were further minimized, only one pair of primer set was required to successfully define the breakpoints.

Although our present data shows an equal effectiveness between MLPA and NGS methods in detecting exon deletions (**Figure 3A**). We also observed that MLPA could exhibit false positive results when the probe cannot combine with DNA at a locus with small indel (Vorstman et al., 2006; Lim et al., 2010; Zhang et al., 2016), resulting in a decreased probe hybridization and ligation with the target DNA sequences, and consequently loss of amplification (**Figure 3C**). Due to the occasionally false positives, we suggest that breakpoint analysis should be conducted in patients with positive MLPA results to confirm the deletions. In addition, according to the study of other researchers, partial exonic deletions could escape MLPA detection, while NGS method has the capacity to detect them (Liu et al., 2018).

The incidence of large intragenic deletions is high, so normal screening must take this into consideration, this makes the genetic diagnostic procedure complex, as Sanger sequencing and MLPA must be both performed. Targeted next-generation sequencing method could identify the full spectrum of *FLCN* gene mutations, as well as determining deletion junctions in a single experiment, and has proven to be a highly sensitive, accurate and cost-saving tool for detecting single nucleotide variations, indels and CNVs in the previous and present studies (Zhang et al., 2016).

TABLE 1 | Previous and present reported *FLCN* intragenic deletions/duplications.

	References	cDNA change	Deleted exon(s)	Detection method	Breakpoint detected (yes/no)	Deletion size (bp)	Clinical manifestations
(1)	Kunogi et al., 2010	c.872−?_1740+?del	Exon 9–14 del	qPCR	No	–	PTX
		c.1539−?_c.1740+?del	Exon 14 del	qPCR	No	–	PTX
(2)	Sempau et al., 2010	c.1539−?_c.1740+?del	Exon 14 del	MLPA	No	–	PTX, FF, RCC
(3)	Benhammou et al., 2011	c.−227–853_c.397–295del	Exons 2–5 del	qPCR, MLPA, aCGH	Yes, flanked by AluSq and AluSx	9,189	FF, LC
		c.619−?_c.1740+?del	Exons 7–14 del	qPCR, MLPA, aCGH	No	–	FF
		c.−4174_−227–1566del	Exon 1 del	qPCR, MLPA, aCGH	Yes, flanked by AluY and AluSx	6,391	Perifollicular fibroma, RCC
		c.−5575_−228+341delins CCCCCATGG	Exon 1 del	qPCR, MLPA, aCGH	Yes, not flanked by Alu	5,688	FF, LC
		c.−6544_−228+454delins −3779_−3655inv	Exon 1 del	qPCR, MLPA, aCGH	Yes, partially Alu-mediated, AluY and AluSg	6,645	FF, LC
		c. −?_−227−?del	Exon 1 del	qPCR, MLPA, aCGH	No	–	FF, LC, PTX
		c.1063–154_1300+410dup	Exons 10–11 dup	qPCR, MLPA, aCGH	Yes, not flanked by Alu	1,341	FF, LC, RCC
(4)	Houweling et al., 2011	c.250−?_c.1740+?del	Exons 5–14 del	MLPA	No	–	PTX
		c.1301−?_c.1740+?del	Exons 12–14 del	MLPA	No	–	PTX
(5)	Babaei Jandaghi et al., 2013	c. −?_−227−?del	Exon 1 del	qPCR	No	–	PTX, RCC
(6)	Ding et al., 2015	c.−504−1303_−25+845del	Exons 1–3 del	MLPA	Yes, flanked by Alu-Sc and Alu-Sz	7,543	PTX, LC, SL
		c.872−429_1740+1763del	Exons 9–14 del	MLPA	Yes, flanked by Alu-Sx3 and Alu-Sc8	7,747	PTX, LC, SL
		c.1539–536_1740+1071del	Exon 14 del	MLPA	Yes, mediated by 2 Alu-Sq2s with 86.1% identity	1,809	PTX, LC, SL
(7)	Rossing et al., 2017	c.872−?_1062 + ?del	Exon 9 del	MLPA	No	–	LC, RCC
(8)	Matsutani et al., 2016	c.1177−?_1300+?del	Exon 11 del	Not mentioned	No	–	PTX, LC, SL
(9)	Liu et al., 2017	c.780−?_871+?del	Exon 8 del	MLPA	No	–	LC, renal cysts
(10)	Schneider et al., 2018	c.250−?_396+?del	Exon 5 del	Microarray-CGH, qPCR	No	–	RCC
(11)	Iwabuchi et al., 2018	c.1063−?_1176+?del	Exon 10 del	Not mentioned	No	–	LC, PTX, renal cysts
		c.1539−?_c.1740+?del	Exon 14 del	Not mentioned	No	–	LC, PTX, skin trichodiscoma
(12)	Enomoto et al., 2020	c. −?_−227−?del	Exon 1 del	qPCR	No	–	LC, SL, RCC
(13)	This study	c.1063–1446_1740+3410del	Exon 10–14 del	NGS, MLPA	Yes, flanked by Alu-Sg and Alu-Sx1	8,384	PTX, LC, SL

PTX, pneumothorax; FF, fibrofolliculomas; LC, lung cysts; RCC, renal cell carcinoma; SL, skin lesion (not histopathologically confirmed).
Reference genome, GRCh37/hg19; accession number: FLCN, NM_144997.7.

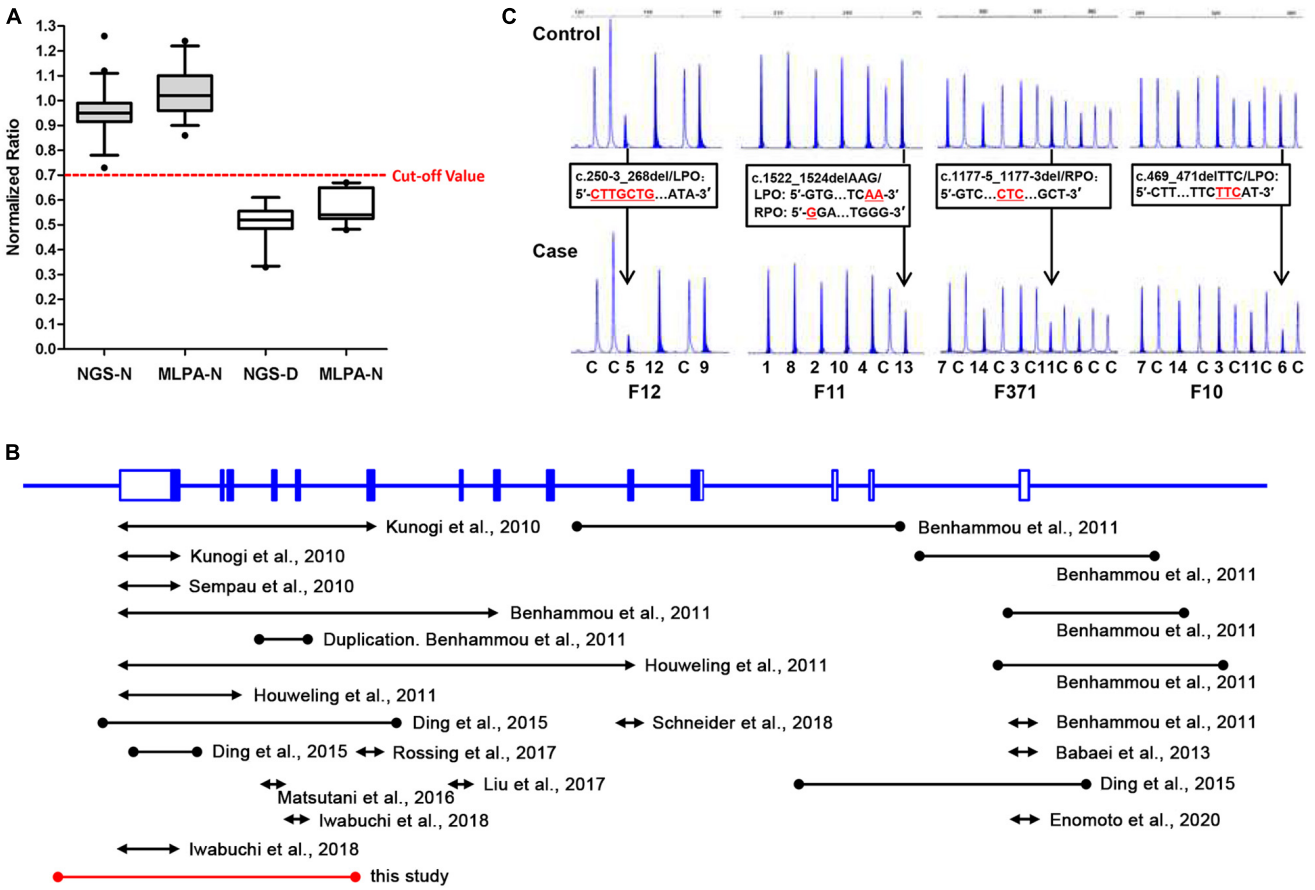


FIGURE 3 | A comparative analysis of NGS and MLPA and a review of previously reported *FLCN* intragenic deletions. **(A)** Box-plot (median, box: 5–95th percentile, whiskers) representation of haploid copy number values for *FLCN* deleted exons (NGS-D and MLPA-D) and normal controls (NGS-N and MLPA-N) by NGS and MLPA, respectively. The red dotted line represents the cut-off value; exons with values <0.7 were scored as deleted, while >0.7 as normal status. **(B)** Genomic structure of the *FLCN* gene, showing the locations and sizes of previously reported large intragenic deletions found in BHD patients; black line: deletions previously reported. Red line: deletion found in this study. The lines with solid circles or arrowheads at both ends indicate the deletions with breakpoints determined or undetermined, respectively. **(C)** Small indels can show peak heights of heterozygous exonic deletions (normalized ratio <0.7) compared with the control individuals (black arrow) in MLPA analysis. DNA changes/*FLCN* probe sequences are presented in black frames, respectively. LPO (Left Probe Oligo) is the 5′ half of the probe; RPO (Right Probe Oligo) is the 3′ half of the probe. Red underlined characters indicate deleted bases. The characters and numbers below the probes indicate exons and control probes. The number below each MLPA result indicate sample number in our private sample database.

Considering all of the above, we highly recommend targeted NGS technique, as a diagnostic application, be widely used for clinical molecular diagnosis in suspect BHDS patients.

Here we reported a Chinese BHDS family with a novel *FLCN* intragenic deletion identified by NGS. The affected family members should be followed up in the following years in case of renal cancer and other symptoms. Our report expands the mutation spectrum of the disease-causing gene.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI, FLCN: NM_144997.7 (https://www.ncbi.nlm.nih.gov/nuccore/NM_144997.7), MW600551, and MW600552.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethical Committees of Taizhou Hospital of Zhejiang Province. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MC, XZ, QG, CZ, LY, and DM conceived and designed the study, contributed to data analysis and interpretation, manuscript drafting, and critical review for intellectual content and final approval of the manuscript. MC, BC, MY, CZ, and DM co-designed experiments, contributed to collection of the clinical data, manuscript editing and discussed analyses, interpretation, and presentation. XZ, LF, SCh, AK, and SCe performed

the experiments, contributed to data analysis, and manuscript editing. All authors read and approved the final manuscript.

FUNDING

This work was supported by the grant from National Natural Science Foundation of China (Grant Nos. 81700003 and 81670003) and the Fundamental Research Funds for the Central Universities (2242017K40103).

ACKNOWLEDGMENTS

We thank the patients, their families who collaborated in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.636900/full#supplementary-material

REFERENCES

Agarwal, P. P., Gross, B. H., Holloway, B. J., Seely, J., Stark, P., and Kazerooni, E. A. (2011). Thoracic CT findings in Birt-Hogg-Dube syndrome. *AJR Am. J. Roentgenol.* 196, 349–352. doi: 10.2214/AJR.10.4757

Baba, M., Hong, S. B., Sharma, N., Warren, M. B., Nickerson, M. L., Iwamatsu, A., et al. (2006). Folliculin encoded by the BHD gene interacts with a binding protein, FNIP1, and AMPK, and is involved in AMPK and mTOR signaling. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15552–15557. doi: 10.1073/pnas.0603781103

Babaei Jandaghi, A., Daliri, S., Kikkawa, M., Khaledi, M., Soleimanifar, N., Alizadeh, A., et al. (2013). The discovery of a Persian family with a form of Birt-Hogg-Dube syndrome lacking the typical cutaneous stigmata of the syndrome. *Clin. Imaging* 37, 111–115. doi: 10.1016/j.clinimag.2012.03.003

Benhammou, J. N., Vocke, C. D., Santani, A., Schmidt, L. S., Baba, M., Seyama, K., et al. (2011). Identification of intragenic deletions and duplication in the *FLCN* gene in Birt-Hogg-Dube syndrome. *Genes Chromosomes Cancer* 50, 466–477. doi: 10.1002/gcc.20872

Birt, A. R., Hogg, G. R., and Dubé, W. J. (1977). Hereditary multiple fibrofolliculomas with trichodiscomas and acrochordons. *Arch. Dermatol.* 113, 1674–1677.

Ding, Y., Zhu, C., Zou, W., Ma, D., Min, H., Chen, B., et al. (2015). *FLCN* intragenic deletions in Chinese familial primary spontaneous pneumothorax. *Am. J. Med. Genet. A* 167A, 1125–1133. doi: 10.1002/ajmg.a.36979

Enomoto, Y., Namba, Y., Hoshika, Y., Komemushi, Y., Mitani, K., Kume, H., et al. (2020). A case of Birt-Hogg-Dube syndrome implying reduced or no wild-type folliculin without mutated protein is pathogenic. *Eur. J. Med. Genet.* 63:103820. doi: 10.1016/j.ejmg.2019.103820

Furuya, M., Yao, M., Tanaka, R., Nagashima, Y., Kuroda, N., Hasumi, H., et al. (2016). Genetic, epidemiologic and clinicopathologic studies of Japanese Asian patients with Birt-Hogg-Dube syndrome. *Clin. Genet.* 90, 403–412. doi: 10.1111/cge.12807

Graham, R. B., Nolasco, M., Peterlin, B., and Garcia, C. K. (2005). Nonsense mutations in folliculin presenting as isolated familial spontaneous pneumothorax in adults. *Am. J. Respir. Crit. Care Med.* 172, 39–44. doi: 10.1164/rccm.200501-143OC

Hasumi, H., Baba, M., Hong, S. B., Hasumi, Y., Huang, Y., Yao, M., et al. (2008). Identification and characterization of a novel folliculin-interacting protein FNIP2. *Gene* 415, 60–67. doi: 10.1016/j.gene.2008.02.022

Houweling, A. C., Gijezen, L. M., Jonker, M. A., van Doorn, M. B., Oldenburg, R. A., van Spaendonck-Zwarts, K. Y., et al. (2011). Renal cancer and pneumothorax risk in Birt-Hogg-Dube syndrome; an analysis of 115 *FLCN* mutation carriers from 35 BHD families. *Br. J. Cancer* 105, 1912–1919. doi: 10.1038/bjc.2011.463

Iwabuchi, C., Ebana, H., Ishiko, A., Negishi, A., Mizobuchi, T., Kumasaka, T., et al. (2018). Skin lesions of Birt-Hogg-Dube syndrome: Clinical and histopathological findings in 31 Japanese patients who presented with pneumothorax and/or multiple lung cysts. *J. Dermatol. Sci.* 89, 77–84. doi: 10.1016/j.jdermsci.2017.10.014

Jensen, D. K., Villumsen, A., Skytte, A. B., Madsen, M. G., Sommerlund, M., and Bendstrup, E. (2017). Birt-Hogg-Dube syndrome: a case report and a review of the literature. *Eur. Clin. Respir. J.* 4:1292378. doi: 10.1080/20018525.2017.1292378

Koo, H. K., and Yoo, C. G. (2013). Multiple cystic lung disease. *Tuberc. Respir. Dis.* 74, 97–103. doi: 10.4046/trd.2013.74.3.97

Kunogi, M., Kurihara, M., Ikegami, T. S., Kobayashi, T., Shindo, N., Kumasaka, T., et al. (2010). Clinical and genetic spectrum of Birt-Hogg-Dube syndrome patients in whom pneumothorax and/or multiple lung cysts are the presenting feature. *J. Med. Genet.* 47, 281–287. doi: 10.1136/jmg.2009.070565

Lim, D. H., Rehal, P. K., Nahorski, M. S., Macdonald, F., Claessens, T., Van Geel, M., et al. (2010). A new locus-specific database (LSDB) for mutations in the folliculin (*FLCN*) gene. *Hum. Mutat.* 31, E1043–E1051. doi: 10.1002/humu.21130

Liu, C., Deng, H., Yang, C., Li, X., Zhu, Y., Chen, X., et al. (2018). A resolved discrepancy between multiplex PCR and multiplex ligation-dependent probe amplification by targeted next-generation sequencing discloses a novel partial exonic deletion in the Duchenne muscular dystrophy gene. *J. Clin. Lab. Anal.* 32:e22575. doi: 10.1002/jcla.22575

Liu, Y., Xu, Z., Feng, R., Zhan, Y., Wang, J., Li, G., et al. (2017). Clinical and genetic characteristics of chinese patients with Birt-Hogg-Dube syndrome. *Orphanet. J. Rare Dis.* 12:104. doi: 10.1186/s13023-017-0656-7

Matsutani, N., Dejima, H., Takahashi, Y., Uehara, H., Iinuma, H., Tanaka, F., et al. (2016). Birt-Hogg-Dube syndrome accompanied by pulmonary arteriovenous malformation. *J. Thorac. Dis.* 8, E1187–E1189. doi: 10.21037/jtd.2016.09.68

Menko, F. H., van Steensel, M. A., Giraud, S., Friis-Hansen, L., Richard, S., Ungari, S. et al. (2009). Birt-Hogg-Dube syndrome: diagnosis and management. *Lancet. Oncol.* 10, 1199–1206.

Nickerson, M. L., Warren, M. B., Toro, J. R., Matrosova, V., Glenn, G., Turner, M. L., et al. (2002). Mutations in a novel gene lead to kidney tumors, lung wall defects, and benign tumors of the hair follicle in patients with the Birt-Hogg-Dubé syndrome. *Cancer Cell* 2, 157–164.

Painter, J. N., Tapanainen, H., Somer, M., Tukiainen, P., and Aittomaki, K. (2005). A 4-bp deletion in the Birt-Hogg-Dube gene (*FLCN*) causes dominantly inherited spontaneous pneumothorax. *Am. J. Hum. Genet.* 76, 522–527. doi: 10.1086/428455

Raoof, S., Bondalapati, P., Vydyula, R., Ryu, J. H., Gupta, N., Raoof, S., et al. (2016). Cystic lung diseases: algorithmic approach. *Chest* 150, 945–965. doi: 10.1016/j.chest.2016.04.026

Ren, H. Z., Zhu, C. C., Yang, C., Chen, S. L., Xie, J., Hou, Y. Y., et al. (2008). Mutation analysis of the *FLCN* gene in Chinese patients with sporadic and familial isolated primary spontaneous pneumothorax. *Clin. Genet.* 74, 178–183. doi: 10.1111/j.1399-0004.2008.01030.x

Rossing, M., Albrechtsen, A., Skytte, A. B., Jensen, U. B., Ousager, L. B., Gerdes, A. M., et al. (2017). Genetic screening of the *FLCN* gene identify six novel variants and a Danish founder mutation. *J. Hum. Genet.* 62, 151–157. doi: 10.1038/jhg.2016.118

Roth, J. S., Rabinowitz, Ad, Benson, M., and Grossman, M. E. (1993). Bilateral renal cell carcinoma in the Birt-Hogg-Dube syndrome. *J. Am. Acad. Dermatol.* 29, 1055–1056.

Schmidt, L. S., Nickerson, M. L., Warren, M. B., Glenn, G. M., Toro, J. R., Merino, M. J., et al. (2005). Germline BHD-mutation spectrum and phenotype analysis of a large cohort of families with Birt-Hogg-Dube syndrome. *Am. J. Hum. Genet.* 76, 1023–1033. doi: 10.1086/430842

Schneider, M., Dinkelborg, K., Xiao, X., Chan-Smutko, G., Hruska, K., Huang, D., et al. (2018). Early onset renal cell carcinoma in an adolescent girl with germline *FLCN* exon 5 deletion. *Fam. Cancer* 17, 135–139. doi: 10.1007/s10689-017-0008-8

Sempau, L., Ruiz, I., Gonzalez-Moran, A., Susanna, X., and Hansen, T. V. O. (2010). New mutation in the Birt Hogg Dube Gene. *Actas. Dermosifiliogr.* 101, 637–640. doi: 10.1016/j.ad.2010.03.007

Takagi, Y., Kobayashi, T., Shiono, M., Wang, L., Piao, X., Sun, G., et al. (2008). Interaction of folliculin (Birt-Hogg-Dube gene product) with a novel Fnip1-like (FnipL/Fnip2) protein. *Oncogene* 27, 5339–5347. doi: 10.1038/onc.2008.261

Toro, J. R., Pautler, S. E., Stewart, L., Glenn, G. M., Weinreich, M., Toure, O., et al. (2007). Lung cysts, spontaneous pneumothorax, and genetic associations in 89 families with Birt-Hogg-Dube syndrome. *Am. J. Respir. Crit. Care. Med.* 175, 1044–1053. doi: 10.1164/rccm.200610-1483OC

Vorstman, J. A., Jalali, G. R., Rappaport, E. F., Hacker, A. M., Scott, C., and Emanuel, B. S. (2006). MLPA: a rapid, reliable, and sensitive method for detection and analysis of abnormalities of 22q. *Hum. Mutat.* 27, 814–821. doi: 10.1002/humu.20330

Woodford, M. R., Dunn, D. M., Blanden, A. R., Capriotti, D., Loiselle, D., Prodromou, C., et al. (2016). The FNIP co-chaperones decelerate the Hsp90 chaperone cycle and enhance drug binding. *Nat. Commun.* 7:12037. doi: 10.1038/ncomms12037

Xu, K. F., Feng, R., Cui, H., Tian, X., Wang, H., Zhao, J., et al. (2016). Diffuse cystic lung diseases: diagnostic considerations. *Semin. Respir. Crit. Care. Med.* 37, 457–467. doi: 10.1055/s-0036-1580690

Zbar, B., Alvord, W. G., Glenn, G., Turner, M., Pavlovich, C. P., Schmidt, L., et al. (2002). Risk of renal and colonic neoplasms and spontaneous pneumothorax in the Birt-Hogg-Dube syndrome. *Cancer Epidemiol. Biomarkers. Prev.* 11, 393–400.

Zhang, X., Ma, D., Zou, W., Ding, Y., Zhu, C., Min, H., et al. (2016). A rapid NGS strategy for comprehensive molecular diagnosis of Birt-Hogg-Dube syndrome in patients with primary spontaneous pneumothorax. *Respir. Res* 17:64. doi: 10.1186/s12931-016-0377-9

Conflict of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

frontiers in Genetics

ORIGINAL RESEARCH
published: 14 April 2021
doi: 10.3389/fgene.2021.635458

Xp11.2 Duplication in Females: Unique Features of a Rare Copy Number Variation

*Márta Czakó[1,2], Ágnes Till[1], Judith Zima[1], Anna Zsigmond[1], András Szabó[1,2], Anita Maász[1,2], Béla Melegh[1,2] and Kinga Hadzsiev[1,2]**

[1] Department of Medical Genetics, Medical School, University of Pécs, Pécs, Hungary, [2] Szentágothai Research Centre, Pécs, Hungary

Among the diseases with X-linked inheritance and intellectual disability, duplication of the Xp11.23p11.22 region is indeed a rare phenomenon, with less than 90 cases known in the literature. Most of them have been recognized with the routine application of array techniques, as these copy number variations (CNVs) are highly variable in size, occurring in recurrent and non-recurrent forms. Its pathogenic role is not debated anymore, but the information available about the pathomechanism, especially in affected females, is still very limited. It has been observed that the phenotype in females varies from normal to severe, which does not correlate with the size of the duplication or the genes involved, and which makes it very difficult to give an individual prognosis. Among the patients studied by the authors because of intellectual disability, epilepsy, and minor anomalies, overlapping duplications affecting the Xp11.23p11.22 region were detected in three females. Based on our detailed phenotype analysis, we concluded that Xp11.23p11.22 duplication is a neurodevelopmental disorder.

Keywords: Xp11.23p11.22 duplication, array CGH, X-inactivation, speech and language delay, regression

OPEN ACCESS

Edited by:
Katalin Komlosi, Medical Center University of Freiburg, Germany

Reviewed by:
Thomas Liehr, Friedrich Schiller University Jena, Germany
Scott Hickey, Nationwide Children's Hospital, United States

***Correspondence:**
Kinga Hadzsiev
hadzsiev.kinga@pte.hu

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

Received: *30 November 2020*
Accepted: *22 March 2021*
Published: *14 April 2021*

Citation:
Czakó M, Till Á, Zima J, Zsigmond A, Szabó A, Maász A, Melegh B and Hadzsiev K (2021) Xp11.2 Duplication in Females: Unique Features of a Rare Copy Number Variation. Front. Genet. 12:635458. doi: 10.3389/fgene.2021.635458

INTRODUCTION

The technical possibilities of copy number variation (CNV) detection were significantly improved in the last 15 years, as a result the abnormality causing the pathological phenotype and the mechanism of their development have become known in many diseases (Zhang et al., 1997, 2009; Lupski, 1998; Marshall et al., 2008; Carvalho and Lupski, 2016). Chromosome microarray has become a first-line investigation tool, particularly in patients with intellectual disability (ID), multiple malformations, epilepsy and autism (Miller et al., 2010). In syndromic forms of ID, where morphological abnormalities, behavioral disorders, seizures, and abnormal growth are associated with the disease, the group showing X-linked inheritance is remarkable. This is not surprising considering that based on new data published in the last decade; we know that almost twice as many genes are associated with X-linked mental retardation as thought before (Neri et al., 2018). Among ID patients, the proportion of males is slightly higher (1, 3:1 male to female), 5–10% of the cases shows X-linked inheritance (Froyen et al., 2008). Their studies revealed the role of more than 100 genes located on chromosome X, and a number of CNVs have been described (Froyen et al., 2007; Neri et al., 2018). While among pathogenic CNVs detected on autosomes deletions occur at a higher rate, there are much more duplications occurring on chromosome X (Ropers, 2006; Whibley et al., 2010; Lubs et al., 2012; Vulto-van Silfhout et al., 2013).

The duplication affecting the Xp11.23p11.22 region is unique even among these specific CNVs of X chromosome. It is very rare occurring in both genders (Kokalj Vokac et al., 2002; Bonnet et al., 2006; Froyen et al., 2008, 2012; Monnot et al., 2008; Giorda et al., 2009; Zou and Milunsky, 2009; Holden et al., 2010; Honda et al., 2010; Broli et al., 2011; Chung et al., 2011; Edens et al., 2011; El-Hattab et al., 2011; Flynn et al., 2011; Nizon et al., 2014; Evers et al., 2015; Grams et al., 2015; Moey et al., 2016; Orivoli et al., 2016; Arican et al., 2018; Wang et al., 2020). Inherited and *de novo* forms are known, all *de novo* Xp11.23 duplications in which parent of origin has been determined have been paternally inherited (Deng et al., 1990). Generally, duplications of the X chromosome in females are often asymptomatic because X-inactivation process silences the chromosome carrying the duplication (Van den Veyver, 2001). However, in case of Xp11.23p11.22 duplications, the opposite is true: in females with skewed X-inactivation the X chromosome carrying the wild-type allele is silenced, while the abnormal one is active in the majority of the cells. Therefore, females with random X-inactivation develop a milder phenotype.

Based on the cases reported so far, the phenotypic features which develop in males and females affected by duplication of Xp11.23p11.22 are very similar (moderate to severe ID, significant delay of speech development, very specific pattern observed on electroencephalography (Broli et al., 2011) with or without seizures manifestation, and dysmorphic facial features), which is very thought provoking in terms of the pathomechanism. While in boys, the pure increase in gene dose caused by the extra copy may explain the symptoms, the gene dosage assessment in girls is much more complicated due to skewed or random X-inactivation (Bonnet et al., 2006; El-Hattab et al., 2011; Nizon et al., 2014; Orivoli et al., 2016; Wang et al., 2020).

The relationship between the genes affected by the Xp11.23p11.22 duplication and the phenotype is also very complex. Non-recurrent duplications of 0, 3–55 Mb in size have been reported in addition to the recurrent form of about 4.5 Mb, therefore the genes involved in the individual patient's duplications may be quite different. Patients with this copy number alteration, which have become known so far, show surprisingly similar symptoms despite the variability in CNV size: the resulting symptoms do not correlate with size and gene content of the affected genomic regions (Giorda et al., 2009; Zou and Milunsky, 2009; Flynn et al., 2011; Arican et al., 2018).

In particular, genotype-phenotype analysis of cases with non-recurrent duplication offered opportunity to analyze the relationship between shared symptoms and affected genomic regions. Several studies have found association between certain genes and symptoms (**Table 1**), however, the individual studies did not examine exactly the same phenotypic traits, which makes the correlation difficult. In a study of six families, Froyen et al. (2008) isolated a minimal overlapping region. Functional analysis of the genes involved identified a causal relationship between elevated gene dosage and intellectual disability only in case of the *HUWE1* gene (Froyen et al., 2012). Grams et al. (2015) based on their own analyses and the previously published cases with

TABLE 1 | Genes involved in Xp11.22p11.23 duplications.

Gene (OMIM #)	Function	Associated symptoms
TM4SF2 (300096)	Tetraspanin 7, control of neurite outgrowth	Intellectual disability (ID)
ZNF41 (314995)	Zinc finger protein 41	ID, speech delay
ZNF81	Zinc finger protein 81	ID
SYN1 (313440)	Synapsin I, regulation of neuronal development	seizures, learning difficulties, behavioral abnormalities
FTSJ1 (300499)	Homolog of Escherichia coli RNA methyltransferase FtsJ/RrmJ; takes part in the regulation of translation	ID, agressive behavior (obesity, macrocephaly)
PQBP1 (300463)	Transcriptional activator; overexpression of *PQBP1* suppressed the cell growth (stress susceptibility)	ID, microcephaly, neuronal dysfunction, short stature, spasticity
HDAC6 (300272)	Histone deacetylase 6	ID
ATP6AP2 (300423)	Renin/prorenin receptor precursor	Seizures, ID, motor and speech delay
CASK (300172)	Calcium/calmodulin-dependent serine protein kinase	ID, microcephaly, brain malform
ZNF674 (300573)	Kruppel-type zinc finger protein; transcriptional regulator	ID, learning difficulties, disturbances of adaptive behavior
MAOA (309850)	Monoamine oxidase localized in the outer mitochondrial membrane	ID and agressive behavior in males
BCOR (300485)	BCL-6 corepressor; key transcriptional regulator during early embryogenesis in eyes and central nervous system	Microphthalmia (syndromic, type 2), low weight, short stature, teeth anomalies, heart failure, seizures, scoliosis, ID, motor delay
NDP (300658)	A norrin precursor; neuroectodermal cell-cell interaction	Vitreoretinopathy, psychosis, growth failure, seizures
NYX (300278)	A nyctalopin precursor	Myopia, hyperopia, nystagmus, reduced visual acuity
RP2 (300757)	Stimulates the GTPase activity of tubulin	XL-retinitis pigmentosa 2
SYP (313475)	Synaptophysin; an integral membrane protein that regulates synaptic vesicle endocytosis	ID, epilepsy
BMP15 (300247)	Bone morphogenetic protein (BMP) 15; oocyte-specific growth and differentiation factor	Abnormal growth parameters; early puberty
KDM5C (314690)	Lysine-specific demethylase 5C; transcriptional repressor	ID, autism spectrum disorder, spastic paraplegia
HUWE1 (300697)	HECT, UBA, and WWE domains-containing protein 1; E3 ubiquitin ligase,	ID, neuronal development, proliferation, synaptogenesis
PHF8 (300560)	Zinc finger protein 422	ID
FGD1 (300546)	FYVE, RhoGEF, and PH domains-containing protein 1	ID
SLC35A2 (314375)	Solute carrier family 35 (UDP-galactose transporter), member 2	Abnormal galactosylation in neurons; seizures

(Continued)

TABLE 1 | Continued

Gene (OMIM #)	Function	Associated symptoms
SHROOM4 (300579)	SHROOM family member 4, Stocco dos Santos XLMR-syndrome; influences cytoskeletal architecture	Severe ID, delayed/no speech, seizures, hyperactivity
KCND1 (300281)	potassium voltage-gated channel, Shal-related subfamily, member 1; prominent in the repolarization phase of the action potential	Seizures
GRIPAP1 (300408)	GRIP1 associated protein 1; dendritogenesis, synaptic vesicle release, AMPA receptor exocytosis	Seizures
PRAF2 (300840)	PRA1 domain family, member 2; protein of synaptic vesicle membranes	Seizures
PLP2 (300112)	Proteolipid membrane protein, colonic epithelium-enriched differentiation-dependent protein A4	ID
CCDC22 (300859)	Coiled-coil domain-containing protein 22	ID
IQSEC2 (300522)	IQ motif- and SEC7 domain-containing protein 2; in neurons: cytoskeletal organization, dendritic spine morphology, excitatory synaptic organization	ID, autistic behavior, psychiatric problems, delayed early speech development
SMC1A (300040)	Structural maintenance of chromosomes 1A; Cornelia de Lange syndrome type 2	Seizures

duplications of Xp11.2 highlighted the importance of two smaller subregions. One of them (Region 1) contains the *SHROOM4* and *DGKK* genes, the other (Region 2) is the same as that was examined by Froyen et al., in which in addition to *HUWE1, KDM5C,* and *IQSEC2* are included as candidate genes (**Table 1** and **Figure 1**). Taken together, based only on the genes with extra copy, the expected phenotype cannot be predicted in individual patients.

Within the framework of diagnostic array CGH testing of a cohort of 448 patients presenting ID, epilepsy, and minor anomalies, we detected overlapping duplications affecting the Xp11.23p11.22 region in three female patients. In two girls, we identified the already known recurrent duplication, while in the third patient a large, non-recurrent copy number gain was detected.

MATERIALS AND METHODS

Subjects

In the Department of Medical Genetics, we collected blood samples from probands with ID, seizures and/or congenital malformations and dysmorphic features, and from their family members. Written informed consent for genetic testing was obtained in genetic counseling from all individuals examined or their guardian, as well as from their healthy relatives. Genomic DNA was isolated from peripheral blood according to standard procedures. This study was performed in accordance with the Hungarian genetic law (XXI/2008).

Patient 1

She is the first child of non-consanguineous parents. The mother has epilepsy and mood disorder; as well the father has mood disorder. Three of the first cousins of the mother are treated with hyperactivity; the maternal grandmother had hearing impairment. One of the father's first cousins was born from a consanguineous marriage, and had a muscular disorder, without a correct diagnosis. Unfortunately, only the mother was available for genetic testing, the distant relatives were not.

She was born following an uneventful pregnancy in the 39th week of gestation, per vias naturals (birth weight 3,370 g, Apgar scores 9/10). Her early psychomotor development was delayed (determined by Bayley Scales of Infant and Toddler Development); she receives neurohabilitation since her 8 month of age. As a result, she sat at 15 month of age and stood up at the age of 18 months. Delayed language development was detected, at 19 month it is limited to one word. Brain magnetic resonance imaging (MRI) at 16 month of age showed supratentorial abnormalities in the white matter, hypoplasia of corpus callosum, and plexus chorioideus cysts. At 2 years of age, compulsive behavior occurred, with bruxism, hand biting and repetitive hand movements. She does not keep eye contact, and unreasonable laughter can be observed, in addition, her speech is inarticulate. There is a stagnation and slight regression in motor development. Her first epileptic seizure developed at the age of 28 months, appropriate seizure control was achieved by valproate monotherapy (**Table 2**).

She was temporarily cared by an ophthalmologist because of divergent strabismus. She often had upper respiratory infection and inflammation of the eyes. The sequencing of *MECP2*, *FOXG1*, *CDKL5* genes gave normal results.

Patient 2

She was born from the first, uncomplicated pregnancy of her mother with cesarean section from meconium-stained amniotic fluid. The mother's sister has short stature, small feet, she is slow moving, with an IQ at the lower limit of normal and similar facial characteristics as our Patient 2 (The sister did not consent to genetic testing).

In the background of feeding difficulties and delayed development of this patient generalized muscle hypotonia was detected at the age of 8 months. As a result of neurohabilitation therapy she walked alone at 28 months of age. Her speech development was severely delayed; she used short sentences from the age of 5–6 years with articulation errors (Budapest-Binet Intelligence Scale). Now she attends a special education school. She has many friends, helps at home and loves to play with a ball. According to the parents, she would be aggressive if not handled well. Brain MRI at age of 1 year showed cerebral atrophy, subdural hygroma, and parietally on the right side a small demyelinated focus was displayed. Three years later the control MRI detected discrete supratentorial, subcortical

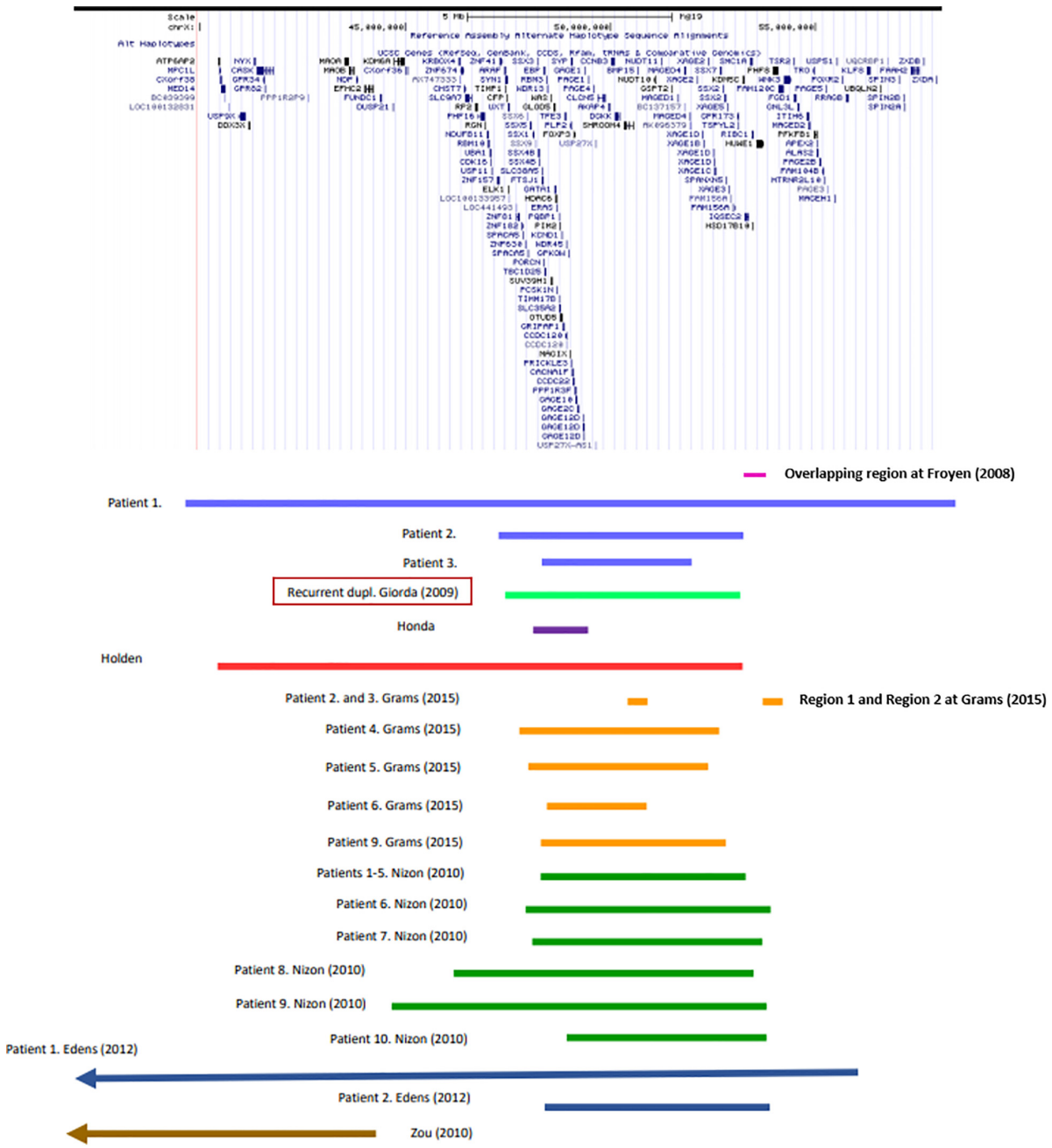


FIGURE 1 | Recurrent and non-recurrent Xp11.22p11.23 duplications of the patients listed in **Supplementary Table 1** (UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly).

white matter abnormalities. Weight gain started around the age of 3 years due to compulsive eating, she has regular endocrinological surveillance due to her obesity. Laboratory investigations excluded Prader-Willi syndrome (**Table 2**).

Patient 3

She was born at the 33rd week of gestation with a weight of 1,480 g after premature rupture of membranes. She was adopted; no family history can be obtained except that her mother has intellectual disability as well. In the perinatal period, she was treated for hypoglycemia, omphalitis, hyperbilirubinemia, and urinary tract infection. Gastroesophageal reflux was confirmed in the background of apnea. Her early psychomotor development was delayed. She was never toilet trained. Her first epileptic seizure developed at the age of 10 years, the EEG (electroencephalography) examination showed fronto-temporal epileptic discharges on the right side. She has been seizure-free with lamotrigine monotherapy for 2 years. The brain

TABLE 2 | Comparison of clinical features in our patients.

		Patient 1	Patient 2	Patient 3
At birth	Weeks of gestation	39th	42nd	33rd
	Birth weight	3,370 g	2,900 g	1,480 g
	Apgar scores	9/10	9/10	N/A
Last examination	Age	1, 5year	13 years	19 years
	Weight	8,620 g (<3 percentile)	75 kg (>97 percentile)	64 kg (50–75 percentile)
	Height	82 cm (50 percentile)	165 cm (75–90 percentile)	150 cm (< 3 percentile)
	Head circumference	43, 5 cm (<3 percentile)	N/A	54, 5 cm (25–50 percentile)
Dysmorphism	Microcephaly	+	−	−
	Biparietal diameter	−	−	decreased
	Flat occiput	+	−	−
	Low frontal hairline	+	−	−
	Low posterior hair line	−	+	−
	Flat face	+	+	−
	Round face	−	+	−
	Thick eyebrows	+	−	+
	Synophrys	+	+	−
	Eyelids	Anti-mongoloid	−	Mongoloid
	Hypertelorism	+	−	−
	Asymmetrical eyes	−	Smaller eye on the left side	−
	Short philtrum	+	+	−
	Wide nose	−	−	+
	Downward corner of the mouth	−	+	−
	Retromicrognathia	+	−	−
	Prominent mandible	−	+	−
	Low set ears	+	−	−
	Short neck	−	+	−
	Hands	Thick fingers, brittle nails	Mild syndactyly on fingers II-III-IV, tapering fingers	Clinodactyly of the 5th fingers on both sides
	Feet	−	−	Lateral deviation of the first toes on both sides
	Other	Joint laxity, hypertrichosis	Hypoplasia of the labia minora	A hemangioma capillare of 5 cm in diameter above the left elbow

MRI detected no abnormalities. Aggression, tantrums have been observed since childhood, risperidone was applied. She attends a special school because of the moderate intellectual disability (tested by Budapest-Binet Intelligence Scale). Regression has been noticed at some developmental areas. Menarche occurred at 10 years of age, the menses is irregular (**Table 2**).

GTG Banding

Karyotyping from cultured peripheral blood lymphocytes was performed by Giemsa–Trypsin (GTG) banding at 550 bands per haploid set using standard procedures (Caspersson et al., 1970).

Array CGH

Array CGH was performed using Agilent Human Genome Unrestricted G3 ISCA v2 Sureprint 8 × 60K oligo-array (Amadid 021924) (Agilent, Santa Clara, CA) (Kallioniemi et al., 1992). DNA was isolated from peripheral blood leukocytes using the NucleoSpin®Dx Blood DNA Purification Kit (Thermo Fisher Scientific, Waltham, MA) as recommended by the manufacturer. For calculation of the concentration and purity of the isolated DNA NanoDrop spectrophotometer was used. Labeling and hybridization of the samples was made according to the Agilent Oligonucleotide Array-Based CGH for Genomic DNA Analysis—Enzymatic Labeling Protocol. Washing was performed following the instructions of Agilent Protocol v7.2. The results were obtained by Agilent dual laser scanner G2565CA and processed with Agilent Feature Extraction software (v10.10.1.1.). Agilent Cytogenomics software (v4.0.1.) was used for evaluation of the CNVs. DNA sequence information refers to the public UCSC database (Assembly: Human GRCh37/hg19). The CNVs detected were compared to known aberrations available in public databases like DECIPHER (Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources), the Database of Genomic Variants, Clingen Dosage Sensitivity Map, Clinvar, and Ensembl (among others).

X-Inactivation Study

For determination of the X-chromosome inactivation pattern the human androgen receptor gene (*AR*) assay (HUMARA) was performed on peripheral leukocytes (in case of Patient 1 and

her mother, Patient 2 and her parents, and Patient 3 without family members). The assay is based on PCR analysis of the polymorphic CAG repeat containing region of the *AR* gene, comparing the pattern of DNA samples digested with *Hpa*II methylation sensitive restriction enzyme to undigested samples (Allen et al., 1992).

RESULTS

G-Banding

The karyotype of Patient 1 showed a duplication on the short arm of chromosome X. Chromosome analysis of the mother revealed the presence of the same duplication on one of the X chromosomes. In Patients 2 and 3 the laboratory investigations resulted in a normal female karyotype at 550 bhps resolution.

Array CGH

The breakpoints and sizes of the duplications of chromosome X are the following (according to McGowan-Jordan et al., 2016): Patient 1: arr[GRCh37] Xp11.4p11.21(39969653_58051765) × 3, which means a 18,082 kb duplication; Patient 2: arr[GRCh37] Xp11.23p11.22(46994270_52693966) × 3, with the size of 5,700 kb; and Patient 3.: arr[GRCh37] Xp11.23p11.22(48584351_51956858) × 3, a copy number gain of 3,373 kb, respectively. In addition, common benign variants were detected in all three of the patients. The base pair positions of the genomic imbalances refer to the February 2009 Assembly (GRCh37/hg19). In summary, duplication is of maternal origin for Patient 1 and *de novo* for Patient 2. In case of Patient 3 none of the parents were available for genetic testing.

X-Inactivation Study

The human androgen receptor assay detected random X-inactivation pattern in Patient 1 and in her mother as well. The results of the test showed similarly random X-inactivation in Patient 2. In contrast, the assay detected non-random X-inactivation pattern in Patient 3 with the same X-chromosome being preferentially inactivated in each of the cells. Unfortunately, in absence of parental samples the origin of the active X chromosome cannot be determined.

DISCUSSION

Studying the literature data on Xp11.22p11.23 duplication, several interesting observations emerge. Although a limited number of such cases have been reported so far, it can be seen that the results of each recent study are inconsistent with one of the previous observations. However, in spite of the size and gene content differing among previously reported patients with Xp11.22p11.23 duplications; it is our opinion that the clinical symptoms are similar. Attempts to link certain symptoms to one or a few genes are remarkably ineffective in this patient group.

Based on a comparison of our three patients and the cases published so far, developmental delay, intellectual disability with varying severity, seizures and different behavioral abnormalities are the most common major symptoms (**Supplementary Table 1**). This is not surprising, as this genomic region contains a number of genes associated with ID, from which *SHROOM4, DGKK, KDM5C, IQSEC2, HSD17B10,* and *HUWE1* are included in most publications (**Table 1**). Based on these features, Xp11.22p11.23 duplication could be classified as neurodevelopmental disease. These symptoms are all present in the three patients described here, although the extent of their duplication varies significantly.

To our knowledge, regression has been described rarely in similar patients so far. We observed it in Patients 1 and 3, especially in the field of motor skills. In a male patient with recurrent Xp11.22p11.23 duplication, Flynn et al. (2011) described regression in areas of speech, memory and recognition, furthermore, leading to aggressive behavior. Helm et al. (2017) reported regression of development with co-occurrence of the onset of seizures in a 20 year-old male patient. Until now, we do not know the background of this symptom either.

Behavioral abnormalities are also characteristic of our three patients: compulsive behavior with bruxism, hand biting and repetitive hand movements, no eye contact and unreasonable laughter (Patient 1), aggression (Patients 2 and 3) and tantrums (Patient 3). The authors of studies on Xp11.22p11.23 duplication in the context of autistic behavior and attention deficit and hyperactivity raise the role of *KDM5C*, *SHROOM4* and *IQSEC2* genes (**Table 1**). Our Patients 1 and 2 present such symptoms, in spite of *SHROOM4* being present in both of them (and in Patient 3 without similar symptoms), however, *KDM5C* and *IQSEC2* genes are involved only in the duplication of Patient 1. Within the area of Xp11.22p11.23 duplication, in the cases studied so far, two genes have been associated with behavioral abnormalities, *SYN1* and *ZNF674* (**Table 1**). Both genes are duplicated in case of Patient 1, only *SYN1* is affected in Patient 2. However, all three of our patients struggle with some kind of behavioral disorder.

Epilepsy is an important symptom of the disease; it has been described in about half of the cases reported so far (Giorda et al., 2009; Holden et al., 2010; Edens et al., 2011; Nizon et al., 2014; Grams et al., 2015). Seizures occurred in two of the three patients studied by us (Patient 1 and 3). **Table 1** contains 11 genes that may play a role in epilepsy.

It can be seen from the data of **Supplementary Table 1** that the somatic parameters of the patients with Xp11.22p11.23 duplication vary widely. Higher than average values for head circumference (OFC) can be seen as typical. Our most noteworthy observation on Patient 1 refers to her OFC (below the 3rd percentile). As far as we know, only two girls has been reported to date with OFC < 3rd percentile (Patient 2 at Edens et al., 2011; Arican et al., 2018). In addition, the mother of Giorda's Patient 1 should be mentioned with the same OFC value. Examining the role of the underlying genes, we can see that *PQBP1* has been associated with microcephaly. One might also speculate that for our Patient 1, it may be explained by the large duplication. However, Patient 1 by Edens et al. (2011) and that of Holden et al. (2010) have large, overlapping duplications also, nevertheless, OFC is at 75th and 90–97th percentile, respectively. The examples listed above demonstrate that studying the size

and gene content of Xp11.22p11.23 duplications alone does not provide an explanation for these differences.

Brain MRI showed aspecific abnormalities in a number of reported cases including our cases: Patient 1 and 2 share the supratentorial white matter abnormalities, in addition we observed even corpus callosum hypoplasia and plexus chorioideus cysts in Patient 1., and subdural hygroma in Patient 2 with a small demyelinated focus on the right side.

Duplication of Xp11.22p11.23 is accompanied by dysmorphism in most cases known to date, but the features observed in individual patients are not specific (**Table 2**). However, it is worth mentioning synophrys and bushy eyebrows, which are present in about half of the patients, including the three girls reported here (**Supplementary Table 1** and on the photos by Nizon et al., 2014).

In Patient 1 and 2, eye abnormalities as divergent strabismus (Patient 1), and asymmetrical eyes with smaller eye on the left side (Patient 2) have been also observed. Small eyes are characteristic also for Patient 9 at Grams (2015). In the same study myopia were described at Patient 3 (and in her mother), cataracts, in Patent 5 pseudostrabismus, hyperopia and astigmatism. Early onset myopia and bilateral hypopigmentation of the midperiphery were reported by Zou and Milunsky (2009). Hypermetropic astigmatism occurred in Patient 1 and 2, and recurrent uveitis in Patient 3 described by Giorda et al. (2009). These examples demonstrate that eye disorders are not uncommon in patients with Xp11.22p11.23 duplication, however, only four genes as *BCOR*, *NDP*, *NYX,* and *RP2* are known associated with abnormalities of the eyes and these are involved in the duplication of our Patient 1 and the case reported by Zou and Milunsky (2009). The further above mentioned patient's eye abnormalities cannot be explained by the extra copy of these four genes.

A similar conclusion can be obtained when examining the symptom of early puberty described in about 50% of the patients. As far as we know, *BMP15* is the only gene associated with early puberty. *BMP15* is affected in all three of our patients, but this symptom is present in two of them only (for similar cases see **Supplementary Table 1**). Only two patients are listed in **Supplementary Table 1** where the duplication does not affect this gene: the 3rd patient reported by Grams et al. (2015) and the case described by Honda et al. (2010). While in the latter case the author did not report data on early puberty, this symptom was described in the patient of Grams et al. (2015). Based on our three patients, age could be an explanation, as Patient 1 not affected by early puberty, is very young. This theory could be applied to 4, 5, and 8th patients of Nizon et al. (2014) (4, 5, and 6 years old girls, respectively) as well. However, **Supplementary Table 1** also contains some older patients with *BMP15* duplication without early puberty (e.g., 4th patient of Giorda and 9th patient of Grams, both of them with 14 years). Therefore, an extra copy of the *BMP15* gene alone may not be a sufficient explanation for the development of this symptom, although reduced penetrance should also be considered.

To examine the causal relationship between Xp11.22p11.23 duplication and phenotype, several factors need to be considered. Which genes and regions may play a role in the development of the symptoms, and which factors may affect the expression of these genes (regulatory elements—primarily *cis*-acting elements with respect to duplication, the copy number of affected genes, and closely related to this the active functioning copy number influenced by X inactivation).

Among the cases published so far, duplications of various size occur. Therefore, it is difficult to identify a critical region. The most experimental evidence support the role of Region 1 and 2 reported by Grams et al. (2015) (see above). The fact that the known recurrent and non-recurrent Xp11.23 duplications involve these regions is definitely an argument for it (**Figure 1**). It is an interesting observation that the patient reported by Honda et al. (2010) is the only one shown in **Figure 1** whose duplication does not affect these regions. She shows relatively fewer symptoms, mostly intellectual disability that is associated with more genes (see **Table 1**) being involved in her duplication. Particular attention should be paid to cases where duplication does not overlap with the critical region, but the phenotype does not differ from that of patients with critical region involvement, for example the girl reported by Zou and Milunsky (2009) (see **Figure 1**). This observation raises the possible role of other factors influencing gene expression, like cis-acting regulatory elements. We still have little information about these for this region now. In any case, studying the ENCODE project data for the genomic region shown in **Figure 1**, it is striking that H3K27Ac marks are located near the breakpoints of the recurrent duplication which play a role in transcription activation. Based on all this, the genomic region responsible for the phenotype features is most likely to fall into the area of recurrent duplication defined by Giorda et al. (2009).

The occurrence of the Xp11.22p11.23 duplication is well known in both genders, thus, many affected females are described in the literature (e.g., Monnot et al., 2008; Zou and Milunsky, 2009; Holden et al., 2010; Chung et al., 2011; Edens et al., 2011; Evers et al., 2015). According to the authors of Xp11.22p11.23 duplication articles, the cause of the abnormal phenotype is the functional disomy of the genes affected by duplication. This seems to be clear in the case of affected males; however, determination of actually functioning copy number in the female patients is more difficult due to X-inactivation. Not all of the published reports provide data for X-inactivation but in about half of the cases studied, a skewed X-inactivation with preferential inactivation of the normal X chromosome in the majority of the cells was found. In these females, functional disomy is an acceptable causal factor as well. Along this argument, one would expect that females with random X-inactivation show milder clinical features but this is not always the case (Evers et al., 2015; for examples of random X-inactivation see **Supplementary Table 1**). When interpreting X-inactivation data, it should be taken into account that peripheral blood cells are always examined, and the pattern of X-inactivation may vary from tissue to tissue. An additional challenge for genotype-phenotype analysis is that duplication of Xp11.22p11.23 does not occur exclusively in females with abnormal phenotype. Asymptomatic carrier mothers are reported in some familiar cases, they may have affected offspring of both genders (Giorda et al., 2009; Honda et al., 2010; Grams et al., 2015). In these cases, the normal phenotype does not always involve preferable inactivation of the

duplicated X chromosome, i.e., skewed X inactivation does not correct for the effect of duplication. This phenomenon suggests the role of additional regulatory factors also. Among our cases, Patient 1 and her mother who has the same Xp11.22p11.23 duplication have random X-inactivation pattern. Given that the phenotype of Patient 1 is much more severe compared to her mother who has only mild intellectual disability, and the random X inactivation suggests similar active functioning copy number, we must also assume the role of further factors besides the copy number of the duplicated genes. However, the most severe phenotype can be seen in Patient 3 with non-random X-inactivation which reinforces the pathogenic role of elevated copy numbers of the genes involved later on.

CONCLUSION

The duplication of the p11.22p11.23 region of the short arm of X chromosome, as well as its effect on the phenotype is known from the description of only a limited number of cases to date. The detailed description of the three patients we studied contributes with new observations to the clinical data that have become known so far related to this rare disease. The recognized cases have mostly been examined in diagnostic centers, where the examination possibilities are limited, which in addition to rare occurrence of this abnormality, may also be the reason why very little is currently known about the relationship between genotype and the resulting phenotype. As the comparison of our patients with others reported to date clearly demonstrate, that in addition to the breakpoints of the duplication and the role of the genes involved, a number of other factors influencing gene expression may affect the symptoms that appear.

Deciphering of the secret can be hoped for from systematic analysis of the genomic data, the gene expression, X-inactivation in multiple tissues, as well as other factors involved in the regulation of gene expression. This rare disease with its peculiarities offers an opportunity to gain insight into the functioning of this section of the X chromosome, primarily into the role of regulatory factors and mechanisms.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

JZ and ÁT were responsible for the patient's clinical genetic examination. AZ contributed to the clinical description. AM, MC, and AS were responsible for methodology and software analysis. KH and ÁT made the conceptualization of the study. ÁT and MC prepared the writing—original draft. KH and BM reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Research on the pathogenesis of rare diseases, developments establishing new diagnostic and therapeutic procedures–GINOP-2.3.2-15-2016-00039; Grant Manager: Ministry of Human Resources, Hungary.

ACKNOWLEDGMENTS

We would like to thank the patients and their family members who participated in the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.635458/full#supplementary-material

REFERENCES

Allen, R. C., Zoghbi, H. Y., Moseley, A. B., Rosenblatt, H. M., and Belmont, J. W. (1992). Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am. J. Hum. Genet.* 51, 1229–1239.

Arican, P., Cavusoglu, D., Gencpinar, P., Ozyilmaz, B., Ozdemir, T. R., and Dundar, N. O. (2018). A de novo Xp11.23 duplication in a girl with a severe phenotype: expanding the clinical spectrum. *J. Pediatr. Genet.* 7, 74–77. doi: 10.1055/s-0037-1612598

Bonnet, C., Grégoire, M. J., Brochet, K., Raffo, E., Leheup, B., and Jonveaux, P. (2006). Pure de-novo 5 Mb duplication at Xp11.22-p11.23 in a male: phenotypic and molecular characterization. *J. Hum. Genet.* 51:815. doi: 10.1007/s10038-006-0023-3

Broli, M., Bisulli, F., Mastrangelo, M., Fontana, E., Fiocchi, I., Zucca, C., et al. (2011). Definition of the neurological phenotype associated with dup (X)(p11.22-p11.23). *Epileptic Disord.* 13, 240–251. doi: 10.1684/epd.2011.0462

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

Caspersson, T., Zech, L., Johansson, C., and Modest, E. J. (1970). Identification of human chromosomes by DNA-binding fluorescent agents. *Chromosoma* 30, 215–227.

Chung, B. H., Drmic, I., Marshall, C. R., Grafodatskaya, D., Carter, M., Fernandez, B. A., et al. (2011). Phenotypic spectrum associated with duplication of Xp11.22-p11.23 includes autism spectrum disorder. *Eur. J. Med. Genet.* 54, e516–e520. doi: 10.1016/j.ejmg.2011.05.008

Deng, H. X., Xia, J. H., Ishikawa, M., and Niikawa, N. (1990). Parental origin and mechanism of formation of X chromosome structural abnormalities: four cases determined with RFLPs. *Jinrui Idengaku Zasshi* 35, 245–251. doi: 10.1007/BF01876853

Edens, A. C., Lyons, M. J., Duron, R. M., Dupont, B. R., and Holden, K. R. (2011). Autism in two females with duplications involving Xp11.22-p11.23. *Dev. Med. Child. Neurol.* 53, 463–466. doi: 10.1111/j.1469-8749.2010.03909.x

El-Hattab, A. W., Bournat, J., Eng, P. A., Wu, J. B., Walker, B. A., Stankiewicz, P., et al. (2011). Microduplication of Xp11.23p11.3 with effects on cognition, behavior, and craniofacial development. *Clin. Genet.* 79, 531–538. doi: 10.1111/j.1399-0004.2010.01496.x

Evers, C., Mitter, D., Strobl-Wildemann, G., Haug, U., Hackmann, K., Maas, B., et al. (2015). Duplication Xp11.22-p14 in females: does X-inactivation help in assessing their significance? *Am. J. Med. Genet. A* 167A, 553–562. doi: 10.1002/ajmg.a.36897

Flynn, M., Zou, Y. S., and Milunsky, A. (2011). Whole gene duplication of the PQBP1 gene in syndrome resembling renpenning. *Am. J. Med. Genet. A* 155A, 141–144. doi: 10.1002/ajmg.a.33756

Froyen, G., Belet, S., Martinez, F., Santos-Rebouças, C. B., Declercq, M., Verbeeck, J., et al. (2012). Copy-number gains of HUWE1 due to replication- and recombination-based rearrangements. *Am. J. Hum. Genet.* 91, 252–264. doi: 10.1016/j.ajhg.2012.06.010

Froyen, G., Corbett, M., Vandewalle, J., Jarvela, I., Lawrence, O., Meldrum, C., et al. (2008). Submicroscopic duplications of the hydroxysteroid dehydrogenase HSD17B10 and the E3 ubiquitin ligase HUWE1 are associated with mental retardation. *Am. J. Hum. Genet.* 82, 432–443. doi: 10.1016/j.ajhg.2007.11.002

Froyen, G., Van Esch, H., Bauters, M., et al. (2007). Detection of genomic copy number changes in patients with idiopathic mental retardation by high-resolution X-array-CGH: important role for increased gene dosage of XLMR genes. *Hum. Mutat.* 28, 1034–1042. doi: 10.1002/humu.20564

Giorda, R., Bonaglia, M. C., Beri, S., Fichera, M., Novara, F., Magini, P., et al. (2009). Complex segmental duplications mediate a recurrent dup(X)(p11.22-p11.23) associated with mental retardation, speech delay, and EEG anomalies in males and females. *Am. J. Hum. Genet.* 85, 394–400. doi: 10.1016/j.ajhg.2009.08.001

Grams, S. E., Argiropoulos, B., Lines, M., Chakraborty, P., Mcgowan-Jordan, J., Geraghty, M. T., et al. (2016). Genotype-phenotype characterization in 13 individuals with chromosome Xp11.22 duplications. *Am. J. Med. Genet. A* 170A, 967–977. doi: 10.1002/ajmg.a.37519

Helm, B. M., Powis, Z., Prada, C. E., Casasbuenas-Alarcon, O. L., Balmakund, T., Schaefer, G. B., et al. (2017). The role of IQSEC2 in syndromic intellectual disability: narrowing the diagnostic odyssey. *Am. J. Med. Genet. A* 173, 2814–2820. doi: 10.1002/ajmg.a.38404

Holden, S. T., Clarkson, A., Thomas, N. S., Abbott, K., James, M. R., and Willatt, L. (2010). A de novo duplication of Xp11.22-p11.4 in a girl with intellectual disability, structural brain anomalies, and preferential inactivation of the normal X chromosome. *Am. J. Med. Genet. A* 152A, 1735–1740. doi: 10.1002/ajmg.a.33457

Honda, S., Hayashi, S., Imoto, I., Toyama, J., Okazawa, H., Nakagawa, E., et al. (2010). Copy-number variations on the X chromosome in Japanese patients with mental retardation detected by array-based comparative genomic hybridization analysis. *J. Hum. Genet.* 55, 590–599. doi: 10.1038/jhg.2010.74

Kallioniemi, A., Kallioniemi, O. P., Sudar, D., Rutovitz, D., Gray, J. W., Waldman, F., et al. (1992). Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258, 818–821. doi: 10.1126/science.1359641

Kokalj Vokac, N., Seme Ciglenecki, P., Erjavec, A., Zagradisnik, B., and Zagorac, A. (2002). Partial Xp duplication in a girl with dysmorphic features: the change in replication pattern of late-replicating dupX chromosome. *Clin. Genet.* 61, 54–61. doi: 10.1034/j.1399-0004.2002.610111.x

Lubs, H. A., Stevenson, R. E., and Schwartz, C. E. (2012). Fragile X and X-linked intellectual disability: four decades of discovery. *Am. J. Hum. Genet.* 90, 579–590. doi: 10.1016/j.ajhg.2012.02.018

Lupski, J. R. (1998). Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends. Genet.* 14, 417–422. doi: 10.1016/s0168-9525(98)01555-1558

Marshall, C. R., Noor, A., Vincent, J. B., Lionel, A. C., Feuk, L., Skaug, J., et al. (2008). Structural variation of chromosomes in autism spectrum disorder. *Am. J. Hum. Genet.* 82, 477–488. doi: 10.1016/j.ajhg.2007.12.009

McGowan-Jordan, J., Simons, A., and Schmid, M. (2016). "ISCN 2016," in . *An International System for Human Cytogenomic Nomenclature*, 1st Edn, eds J. McGowan-Jordan, A. Simons, and M. Schmid (Basel: Karger).

Miller, D. T., Adam, M. P., Aradhya, S., Biesecker, L. G., Brothman, A. R., Carter, N. P., et al. (2010). Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am. J. Hum. Genet.* 86, 749–764. doi: 10.1016/j.ajhg.2010.04.006

Moey, C., Hinze, S. J., Brueton, L., Morton, J., McMullan, D. J., Kamien, B., et al. (2016). Xp11.2 microduplications including IQSEC2, TSPYL2 and KDM5C genes in patients with neurodevelopmental disorders. *Eur. J. Hum. Genet.* 24, 373–380. doi: 10.1038/ejhg.2015.123

Monnot, S., Giuliano, F., Massol, C., Fossoud, C., Cossée, M., Lambert, J. C., et al. (2008). Partial Xp11.23-p11.4 duplication with random X-inactivation: clinical report and molecular cytogenetic characterization. *Am. J. Med. Genet. A.* 146A, 1325–1329. doi: 10.1002/ajmg.a.32238

Neri, G., Schwartz, C. E., Lubs, H. A., and Stevenson, R. E. (2018). X-linked intellectual disability update 2017. *Am. J. Med. Genet. A.* 176, 1375–1388. doi: 10.1002/ajmg.a.38710

Nizon, M., Andrieux, J., Rooryck, C., de Blois, M. C., Bourel-Ponchel, E., Bourgois, B., et al. (2014). Phenotype-genotype correlations in 17 new patients with an Xp11.23p11.22 microduplication and review of the literature. *Am. J. Med. Genet. A* 167A, 111–122. doi: 10.1002/ajmg.a.36807

Orivoli, S., Pavlidis, E., Cantalupo, G., Pezzella, M., Zara, F., Garavelli, L., et al. (2016). Xp11.22 microduplications including HUWE1: case report and literature review. *Neuropediatrics.* 47, 51–56. doi: 10.1055/s-0035-1566233

Ropers, H. H. (2006). X-linked mental retardation: many genes for a complex disorder. *Curr. Opin. Genet. Dev.* 16, 260–269. doi: 10.1016/j.gde.2006.04.017

Van den Veyver, I. B. (2001). Skewed X-inactivation in X-linked disorders. *Semin. Reprod. Med.* 19, 183–191. doi: 10.1055/s-2001-15398

Vulto-van Silfhout, A. T., Hehir-Kwa, J. Y., van Bon, B. W., Schuurs-Hoeijmakers, J. H., Meader, S., Hellebrekers, C. J., et al. (2013). Clinical significance of de novo and inherited copy-number variation. *Hum. Mutat.* 34, 1679–1687. doi: 10.1002/humu.22442

Wang, Q., Chen, P., Liu, J., Lou, J., Liu, Y., and Yuan, H. (2020). Xp11.22 duplications in four unrelated Chinese families: delineating the genotype-phenotype relationship for HSD17B10 and FGD1. *BMC Med. Genom.* 13:66. doi: 10.1186/s12920-020-0728-8

Whibley, A. C., Plagnol, V., Tarpey, P. S., Abidi, F., Fullston, T., Choma, M. K., et al. (2010). Fine-scale survey of X chromosome copy number variants and indels underlying intellectual disability. *Am. J. Hum. Genet.* 87, 173–188. doi: 10.1016/j.ajhg.2010.06.017

Zhang, A., Weaver, D. D., and Palmer, C. G. (1997). Molecular cytogenetic identification of four X chromosome duplications. *Am. J. Med. Genet.* 68, 29–38.

Zhang, F., Gu, W., Hurles, M. E., and Lupski, J. R. (2009). Copy number variation in human health, disease, and evolution. *Annu. Rev. Genom. Hum. Genet.* 10, 451–481. doi: 10.1146/annurev.genom.9.081307.164217

Zou, Y. S., and Milunsky, J. M. (2009). Developmental disability and hypo-melanosis of Ito in a female with 7.3 Mb de novo duplication of Xp11.3–p11.4 and random X-inactivation. *Am. J. Med. Genet. A* 149, 2573–2577. doi: 10.1002/ajmg.a.33066

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

The handling editor KK declared a past collaboration with several of the authors MC, ÁT, AS, AM, BM, and KH.

frontiers in Genetics

REVIEW
published: 14 April 2021
doi: 10.3389/fgene.2021.634217

Germline Structural Variations in Cancer Predisposition Genes

Tímea Pócza[1†], Vince Kornél Grolmusz[1,2†], János Papp[1,2], Henriett Butz[1,2,3], Attila Patócs[1,2,3] and Anikó Bozsik[1,2]*

[1] Department of Molecular Genetics, National Institute of Oncology, Budapest, Hungary, [2] Hereditary Cancers Research Group, Hungarian Academy of Sciences, Semmelweis University, Budapest, Hungary, [3] Department of Laboratory Medicine, Semmelweis University, Budapest, Hungary

In addition to single nucleotide variations and small-scale indels, structural variations (SVs) also contribute to the genetic diversity of the genome. SVs, such as deletions, duplications, amplifications, or inversions may also affect coding regions of cancer-predisposing genes. These rearrangements may abrogate the open reading frame of these genes or adversely affect their expression and may thus act as germline mutations in hereditary cancer syndromes. With the capacity of disrupting the function of tumor suppressors, structural variations confer an increased risk of cancer and account for a remarkable fraction of heritability. The development of sequencing techniques enables the discovery of a constantly growing number of SVs of various types in cancer predisposition genes (CPGs). Here, we provide a comprehensive review of the landscape of germline SV types, detection methods, pathomechanisms, and frequency in CPGs, focusing on the two most common cancer syndromes: hereditary breast- and ovarian cancer and gastrointestinal cancers. Current knowledge about the possible molecular mechanisms driving to SVs is also summarized.

Keywords: germline mutation, structural variations, cancer–predisposing genes, copy number variation, large genomic rearrangement, structural variation

OPEN ACCESS

Edited by:
Judit Bene,
University of Pécs, Hungary

Reviewed by:
Silvia R. Rogatto,
University of Southern Denmark, Denmark
Pilar Garre,
San Carlos University Clinical Hospital, Spain
Tamás I. Orbán,
Hungarian Academy of Sciences (MTA), Hungary

***Correspondence:**
Anikó Bozsik
bozsik.aniko@oncol.hu;
genetikus@gmail.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: *27 November 2020*
Accepted: *08 March 2021*
Published: *14 April 2021*

Citation:
Pócza T, Grolmusz VK, Papp J, Butz H, Patócs A and Bozsik A (2021) Germline Structural Variations in Cancer Predisposition Genes. Front. Genet. 12:634217. doi: 10.3389/fgene.2021.634217

INTRODUCTION

The genetic diversity of the human genome is based on several types of variations from single nucleotide polymorphisms to large genomic rearrangements (LGRs). Chromosomal rearrangements comprise various types of structural variations (SVs). Some of them are copy-neutral (balanced) rearrangements, such as inversions and translocations, while others modify the dosage of chromosomal regions. These latter groups consist of copy number variations (CNVs), which are gains or losses of DNA fragments (deletions, duplications, or amplifications) constituting approximately 5% of the genome and providing the major source of genetic diversity (Zhang et al., 2009). Although CNVs may involve larger chromatin structures, the majority of them are subtle alterations of submicroscopic size generally ranging from 100 bp to 3 Mb (Zhang et al., 2009).

As opposed to recurrent genomic rearrangements, which are mainly gross recombinational events between chromosomal arms encompassing the same genomic interval in unrelated

Abbreviations: SV, structural variation; CNV, copy number variation; LGR, large genomic rearrangement; CPG, cancer predisposition gene; HBOC, hereditary breast- and ovarian cancer; FAP, familial adenomatous polyposis; HNPCC, hereditary non-polyposis colorectal cancer

individuals, SVs in cancer predisposition genes (CPGs) are mainly non-recurrent events. This means that there is no particular genomic region for the breakpoints; these can arise at any chromosomal position. Only slight hot spots or grouping of breakpoints can be discerned with especially complex chromatin architectural structures or within pseudogene regions (Puget et al., 2002).

Structural variations appear either as somatic variations, especially in tumors, or can be generated in the germline. In this case, they are heritable. When SVs fall in functionally relevant regions of the genome, especially when they alter the open reading frame of coding genes, they seriously compromise gene function. This is especially remarkable in CPGs, where these changes contribute to the hereditary mutation profile and confer an increased risk for cancer. Noteworthy, SVs can also affect non-coding genes involved in cancer susceptibility (lncRNAs, microRNAs, and other types of small RNAs), some of which also have exon/intron structures.

Cancer predisposition genes are mainly tumor suppressors, which generally act recessively: both alleles should be lost for developing a phenotype. The Knudson two-hit model sets out that the first hit is an inherited germline mutation, which is followed by a subsequent somatically acquired second hit for tumor generation (Knudson, 1971). The second hit for tumorigenesis frequently appears in one individual's lifetime, causing dominantly appearing cancer disease phenotypes.

The pathogenicity of the SVs is not directly obvious in all cases. While the majority are clear-cut mutations, there are cases, especially in certain duplications and inversions, where additional functional tests are required to assess their effect on clinical outcomes.

Continuously evolving sequencing technologies enable the detection of various types of SVs, which were formerly missed by conventional detection techniques. This allows the identification of an emerging number of rearrangements, which further broadens the spectrum of these variations.

Here, we provide a comprehensive review of germline SV types, detection methods, pathogenic mechanisms, and frequency in CPGs, focusing on three common cancer syndromes. These are hereditary breast- and ovarian cancer (HBOC) and two types of hereditary gastrointestinal cancers, i.e., familial adenomatous polyposis (FAP) and hereditary non-polyposis colorectal cancer (HNPCC or Lynch syndrome). Our work also covers rare hereditary cancer syndromes, each possessing a strong heritability factor and associated with acknowledged tumor suppressor genes.

SVs – UNDERLYING CAUSATIVE MOLECULAR MECHANISMS

The molecular mechanisms generating rearrangements may be replicational or recombinational events and are mostly related to repair processes.

Non-allelic homologous recombination (NAHR) has long been acknowledged as the principal molecular process for SV generation (Sen et al., 2006). A decade ago, novel mechanisms were also discovered as possible casual events. Below, we summarize the most relevant molecular mechanisms calling forth SVs and present them in **Figure 1** through selected examples of already proved rearrangements.

Non-allelic homologous recombination happens between false homologous alleles. This phenomenon, which is called illegitimate recombination, is the main source of recurrent rearrangements. The mechanism usually takes place during meiosis and mitosis and requires an extensive chromosomal homology region of several kilobases similar to conventional recombination (Weckselblatt and Rudd, 2015). Segmental duplications or low-copy repeats throughout the genome serve as a target for ectopic (non-allelic) alignment of the chromosomal regions (Bailey et al., 2002). Additionally, the human genome contains several thousands of transposon-related remnants and may even be located on different chromosomes, which might cause such illegitimate recombinations (Arkhipova and Yushenova, 2019). NAHR can take place between homologous chromosomes, but it can also be intrachromosomal (between sister chromatids) or take place during an intrachromatidal event. NAHR between chromosomal arms calls forth duplications and deletions. Intrachromatidal recombinations between direct repeats result in deletion, whereas inverted repeats serve as a target for inversions. Recurrent rearrangement in CPGs through NAHR is feasible in the case of extensive homology served by long pseudogene regions. Smaller-scale rearrangements are more likely to happen as a result of repair events. Homologous recombinational repair, occurring mainly during the S phase of the cell cycle, can also mispair with ectopic regions with similar sequences and can generate deletion and duplication of chromosomal fragments through a recombination-like resolution of the Holliday junction (Hastings et al., 2009b; **Figure 1A**). Non-homologous end joining (NHEJ) is a more error-prone end-joining repair, generally requiring no homology (Zhao et al., 2020). This mechanism is proposed for deletions, where no, or only a few base pairs of homology are detectable at the junction points. The insertion of additional nucleotides at the junction point (so-called scars) is characteristic of this repair (**Figure 1B**).

The breakpoints of the majority of CNVs encompassing exons of CPGs fall in Alu repetitive sequences. This indicates that Alu elements have a substantial role in the generation of exon-scale chromosomal rearrangements. Novel findings argue that NAHR events require more extensive homology than the typical 300 bp of Alu sequences (Kowalczykowski, 2015). Instead, for rearrangements between small stretches of homologies microhomology-mediated end-joining (MMEJ) mechanism was suggested (McVey and Lee, 2008; Sfeir and Symington, 2015; Sinha et al., 2016). This is a special repair mechanism at double-strand breaks, which involves 5′ strand resection and annealing of the 3′ overhangs mediated by nearby/proximal small homologies of 5–50 bases (Tournier et al., 2004; **Figure 1C**).

The combination of chromosomal segments, which are sometimes even distantly positioned, can occur as a result of replication-based molecular mechanisms: Fork Stalling and Template Switching (FoSTeS) and Microhomology-mediated Break-Induced Repair (MMBIR) (Hastings et al., 2009a). Despite

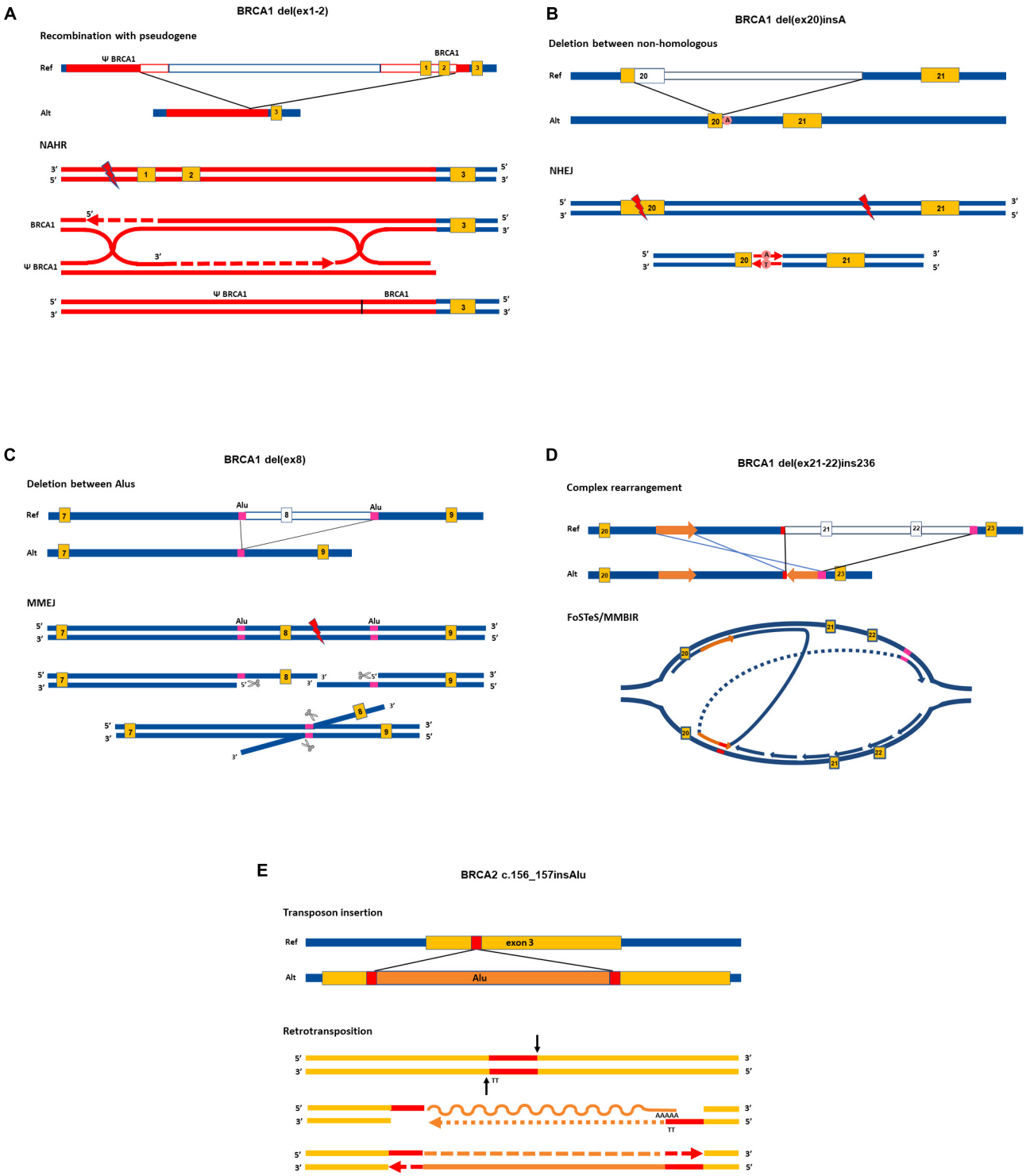


FIGURE 1 | Proposed rearrangement types in cancer susceptibility genes. Upper parts of the panels show the allelic structure in sense orientation of elected rearrangements. Lower parts of the panels depict the most probable molecular mechanisms giving rise to the respective rearrangements, focusing on the main successive steps. Enzymes and auxiliary proteins mediating the mechanisms are not indicated. Yellow boxes indicate exons, red and magenta lines denote homologies, empty lines in reference allele mark deleted regions in alternative allele. Red lightning signs stand for double-strand break. Graphical representation of exons and introns is not to scale. The running name of the rearrangements are indicated above each graph, and exact names are given in the caption with HGVS nomenclature. All SVs taken as examples are registered variations in the LOVD database. NAHR, non-allelic homologous recombination, NHEJ, non-homologous end joining; MMEJ, microhomology-mediated end-joining; MMBIR, microhomology-mediated brake-induced repair; FoSTeS, fork stalling and template switching.

(Continued)

FIGURE 1 | Continued
Ref, reference allele; Alt, alternative allele. **(A)** NG_005905.2:g.61201_98134del. Running name: BRCA1 del(ex1-2) (Puget et al., 2002). At the position of the DNA double-strand break, the 5′ ends are resected and one of the overhanging 3′ ends invades into the D-loop of the homolog *Psi-BRCA1* region annealing with its complementary strand. Synthesis proceeds further for hundreds of base pairs, harnessing the ectopic homology as a template. Extensive homology between BRCA1 and its pseudogene enables the formation of double Holliday junction. The resolution of the double cross with recombination event results in a hybrid region of *Psi-BRCA1* and *BRCA1*. Single-strand nicks are sealed by polymerase (dashed arrows) and ends are rejoined by ligase.
(B) LRG_292t1:c.5213_5278-2753delinsA. Running name: BRCA1 del(ex20)insA (Belogianni et al., 2004; Bozsik et al., 2020). DNA stretch between the two double-strand breaks is deleted and the two exposed ends are rejoined by NHEJ without homology requirement. Error-prone polymerase seals the nick by editing the sequence with an additional adenine residue at the synapsis. **(C)** LRG_292t1:c.442-1102_547+252del. Running name: BRCA1 del(ex8) (Bozsik et al., 2020). DNA ends are processed at the site of the double-strand break: 5′ strands are resected by exonucleases (marked with scissors). The overhanging 3′ strands find 26 base pairs with exact microhomology between nearby 300-bp-long and almost completely homolog AluSx and AluSp sequences (marked with purple boxes). The two strands anneal by the microhomology, and the protruding 3′ strand is eliminated by flap trimming (incision is marked with scissors). The DNA ends are rejoined by ligase. **(D)** LRG_292t1:c.5278-492_5407-128delins236. Running name: BRCA1 del(ex21-22)ins236 (Zikan et al., 2008; Bozsik et al., 2020). Replication forks stalls at structural hindrance caused by palindrome sequence. The 3′ end of the newly synthesized strand disassembles and reanneals to the complementer strand of the same replication bubble with the help of small homology of few nucleotides (indicated with red line). The synthesis proceeds in the reverse direction for 236 base pairs (marked with an orange arrow) and reanneals to the original template using another stretch of microhomology (marked with magenta line) skipping the intervening region. **(E)** NG_012772.1:g.8686_8687insAlu. Running name: BRCA2 c.156-157insAlu (Peixoto et al., 2011). RNA sequence from the AluYa5 inserts into the exon three through target primed reverse transcription method. Endonuclease incises (black arrows) at the ends of the target site (marked with red) liberating 3′ end with TT nucleotides. This serves as a complementer template for the polyA tail of the AluYa5 RNA to anneal. Reverse transcription priming is provided by the free 3′ end of the gene. After reverse transcription (indicated by dotted arrow) gaps are filled by polymerase. The RNA strand is lysed and exchanged with DNA also by 3′→5′ synthesis action of polymerase (dashed arrow). Lygase seals the ends. At the end of the retrotransposition process, the Alu sequence is inserted into the exon with the flanking duplication of the target site.

the difference in the molecular background, these processes are not distinguishable by the resulted product: both give rise to complex rearrangements. Both mechanisms are preceded by stalled replication forks: the DNA polymerase is stopped either by palindrome loops and structural hindrance (FoSTeS) or breaks in the template strand (MMBIR) (Hastings et al., 2009a). The 3′ end disengages and anneals to another replication fork through microhomology of only a few (<6 bp) bases. The new fork, though positioned adjacently, may be distant in the chromosome, or even can be located on another chromosome. The polymerase uses this new strand as a template and replicates a stretch of this region before the strand reanneals to the original fork. Moreover, the 3′ end invasion to new replication forks can be repeated several times between different chromosomal regions before reannealing, thus entailing multiple, distantly located fragments coming together in juxtapositions (Colnaghi et al., 2011). This mechanism can also give rise to deletions, duplications with misaligned homology, and even reversions when the leading strand anneals to the lagging strand. The typical FoSTeS/MMBIR-resulted rearrangement in CPGs is a characteristic pattern of some kilobase deletion combined with a short stretch of reverse oriented duplication of a neighboring intronic segment (**Figure 1D**).

Retrotransposition is also a way for copy number gain. The main mobile elements in the human genome are Long INterspersed Element-1 (L1), SINE-VNTR-Alu (SVA), and Alu, the copy number of which continuously expands with replicative copy-and-paste retrotransposition. The mechanism involves the reverse transcription of the RNA of these elements and insertion of the cDNA copy into a new genomic position with the help of a special endonuclease (Goodier, 2016; **Figure 1E**). L1 elements are the only autonomous transposons in the current human genome. Alu and SVA elements have no capacity for reverse transcription themselves but harness the enzymatic activity of L1 elements for moving (Hancks and Kazazian, 2016).

It is important to note, that the results of the different mechanisms are overlapping. Therefore, the underlying molecular event is not unequivocally identifiable by the inspection of the rearrangement pattern.

PATHOMECHANISMS OF SVs IN CANCER-PREDISPOSING GENES

There is a wide array of mechanisms by which a CNV can abrogate cancer gene function. The most typical is the deletion of one or more exons coding for indispensable domains or structural elements of the protein. Moreover, out-of-frame deletions, generated at any position of the open reading frame, result in premature termination codon (PTC) on the transcript, which either codes for a truncated protein or is eliminated by nonsense-mediated decay. When the deletion affects the promoter region, the regulation of the gene expression may be compromised. For example, a 10 kb promoter deletion in the *APC* gene affecting promoter 1B reduces the expression of *APC-1B* (Yamaguchi et al., 2016). If the rearrangements on the chromosome are more extensive, the deletion may cover the whole coding gene.

Contrary to the effect of deletions, the pathogenic effect of duplications is not straightforward. Breakpoint characterizations are needed for the detection of their exact positions and orientations for interpreting their genetic consequences. Out-of-frame tandem duplications generate PTCs and interfere with protein function similarly to deletions. The duplication of in-frame exons, resulting in two tandem copies of certain protein regions, theoretically, does not necessarily cause a severe adverse effect on the protein, if domain positions and functions are not affected. The same applies to promoter duplications, which sometimes also involve the first coding exons of the gene: in this case, there is at least one correct copy of the whole gene and optimal choice between the two promoters can help to evade

the generation of an altered transcript. For example, the tandem duplication of 357 kb upstream of the *BRCA1* gene, reaching up to *BRCA1* exons 1–19, was evaluated as a benign variation (Du et al., 2018).

A special form of gene silencing is transcriptional interference (read-through). As an example, *MSH2* silencing can occur due to the deletion of 3′ exons of the upstream *EPCAM* gene (Ligtenberg et al., 2009). The *EPCAM* deletion eliminates the transcription termination signal, thus the RNA polymerase goes further towards the neighboring *MSH2* gene, preventing the binding of transcription factors to the *MSH2* promoter, consequently hindering its transcriptional initiation. Furthermore, the long transcript usually ends up in a PTC, directing this fused RNA towards nonsense-mediated decay, resulting in a deletion effect on both *EPCAM* and *MSH2* (Kovacs et al., 2009).

Retroelement (RE) insertions are rare events of great consequence regarding the functional abrogation on CPGs. Insertion of a RE (L1, SVA, or Alu elements) into exons or intron regions of genes may cause exon skipping, exonization (Schmitz and Brosius, 2011), PTC generation, or transcriptional interference (Kaer et al., 2011). Aberrant splicing as a result of RE insertion into splice regulation regions was also described. A prominent example for this latter effect is the insertion of an Alu-like element in *MLH1* intron 7, which interferes via a canonical splice donor site, leading to complete disruption of mRNA splicing (Li et al., 2020).

SV DETECTION METHODS

Precise determination of SVs is not an easy task. Due to their heterogeneity, there is no one standard procedure that allows the correct identification of both deletions, insertion, and copy number alterations involving multiallelic loci. Generally, molecular biological methods providing quantitative differences can be used for the detection of SVs (Cantsilieris et al., 2013; Butz and Patocs, 2019). Based on the size of SVs, different assays are available. Two widely employed approaches in routine clinical practice are hybridization-based and PCR-based techniques.

Fluorescent in-situ hybridization (FISH) is typically used for the identification of large genomic alterations, such as gross chromosomal abnormalities, but current advances in the technique enable the detection of CNVs with sizes as small as 50 kb. Fluorescently labeled DNA probes complementary to the sequence of specific regions are hybridized to metaphase chromosomes or interphase nuclei (Bayani and Squire, 2004). A state-of-the-art version of FISH providing even better resolution (5–500 kb) is fiber-FISH, where probes are visualized on mechanically stretched chromosomes (Ceulemans et al., 2012). This technique is especially appropriate for determining complex CNVs. Applying different fluorescence dyes, multiple DNA targets can be tested simultaneously, allowing for whole-genome analysis (multi-color FISH) (Ceulemans et al., 2012).

Another hybridization-based method is Southern blotting. Recently, due to its highly labor-intensive workflow, radioisotope labeling, and the requirement of high quantity and quality DNA, it has been mostly replaced by other techniques.

Of these approaches, microarrays, which belong to high-throughput techniques, are used to analyze the expression, genotype, or copy number of multiple genes simultaneously. In germline testing, array-based genotyping platforms (i.e., single nucleotide polymorphism-SNP arrays) are applied (Zhang et al., 2009). SNP arrays covering the entire genome or selected genetic regions using disease-specific SNP panels are also employed. The principle is based on the hybridization of fluorescently labeled probes detecting each genotype. By virtue of the intensity of the fluorescence signal, hetero-, hemi-, and homozygous variants may be distinguished, so the presence of either deletions or insertions of SVs can be determined. Loss of heterozygosity can be demonstrated by a parallel evaluation of normal and somatic DNAs of the same patient.

Array comparative genome hybridization (array CGH) uses a small glass slide (chip) that contains thousands of probes specific for certain regions of the genome. Fragmented sample and reference DNA are labeled with different fluorescent dyes, combined, and hybridized to the DNA probes on the array slide. After detection of the two fluorescent signals, the results are given as the ratio of test DNA to reference DNA at each probe. Depending on the chosen platform's design and probe density, the resolution of CGH can vary from whole chromosomes to a few kilobases in size (Davies et al., 2005). In clinical practice, this is the most frequently employed cytogenetic assay. It is applied for the analysis of LGRs as well as submicroscopic structural alterations with unclear clinical importance. There are several databases (Database of Chromosomal Imbalance and Phenotype in Humans using Ensemble Resources; DECIPHER (Firth et al., 2009); International Standards for Cytogenomic Arrays Consortium; ISCA Consortium (Riggs et al., 2013), which aid in the interpretation of results. In tumor genetics, CGH is useful for the detection of somatic changes, including tumor heterogeneity and somatic mosaicism.

Optical genome mapping, an accurate high-throughput assay originally designed for aiding contiguous genome assemblies, is also applicable for the identification of all classes of SVs in the human genome (involving balanced events). Ultra-long, linearized DNA molecules are fluorescently labeled and digested with a combination of restriction enzymes. The nicks at the cleavages are optically detected as fluorescent signal discontinuities, which give a characteristic high-resolution restriction pattern for the DNA sequence. Dedicated software can assemble DNA stretches according to pattern similarities and comparative analysis of the strands enables the detection of divergent regions > 500 bp caused by SVs. Optical mapping offers a significantly higher resolution than karyotyping on a similar scale to fiber-FISH.

Copy number changes can also be detected based on relative quantitations by qPCR or QMPSF (Quantitative Multiplex PCR of Short Fluorescent fragments). In both cases, the quantity of the examined region is compared to that of control regions with surely two copies. With the qPCR analysis, deletions are readily detectable by the difference of Ct values, each unit of Ct corresponding to two copy differences (Schmittgen and Livak, 2008). Limitations for duplications, however, do exist since the

detection of a 2:3 ratio is not feasible. In contrast, QMPSF, where the area under the curve of the sample and control peaks yielded by multiplex PCR are compared, is amenable also for the detection of duplications (Ceulemans et al., 2012).

Inverse PCR is a suitable method for the verification of single inversions with known breakpoints. The principle of the detection is, that a PCR product is generated only when the primers hybridizing to the same strand in a reference template get into opposing orientations as a result of inversion (Wagner et al., 2002). New inversions may be discovered by allelic dropout test following long-range PCR. It is based on the phenomenon, that amplicons covering an inversion breakpoint appear as spurious deletions of one allele, presenting as stretches of homozygosity spanning the position of the inversion variant (Rhees et al., 2014). The detection of insertions is similarly demanding since their insertion point reside mainly in introns, which are not genotyped routinely. cDNA-level analysis of the genes may shed light on a part of these rearrangements since some of them generate new exons (exonization) (Schmitz and Brosius, 2011).

Multiplex ligation-dependent probe amplification (MLPA) is a semi-high-throughput technique developed to detect copy number alteration of up to 50 genomic DNA sequences in a single multiplex PCR-based mode (Kozlowski et al., 2008). Both internal control probes and positive-negative control samples have to be used during the analysis. First normalizing to internal controls (positions that are typically not affected by copy number alterations) in each sample, and then normalizing to control samples yield the relatively quantitative determination of the dosage in each probed locus. MLPA is an efficient way for detecting large deletions even in the hemizygote state. In a molecular genetic analysis of hereditary cancer syndromes, assays and complex reagents are available, and some of them have been already approved for in vitro diagnostic applications[1]. It has to be noted, that sequence polymorphisms within the ligation site can disturb the ligation sufficiently to cause a false positive deletion call (Serizawa et al., 2010).

Next-generation sequencing (NGS) is a high-throughput technology allowing simultaneous sequencing of multiple DNA samples. Currently, NGS-based procedures are the most widely used techniques in the routine molecular genetic diagnosis of hereditary cancer syndromes (Sarkadi et al., 2018). These approaches allow simultaneous determination of germline mutations and somatic alterations in sporadic tumors but have several requirements both from the sample and investigator sides (Robson et al., 2015). Complex laboratory workflow followed by bioinformatic analysis is needed for data mining (Krumm et al., 2012). Genotyping based on read depth analysis allows absolute copy number determination. Basically, the number of sequencing reads that map to a specific region is proportional to the number of copies of this region in the genome. The original hypothesis applies the Poisson distribution of sequencing reads, which means that a region assumed to be deleted or duplicated has fewer or more mapped reads than expected, respectively. Regarding instrumentation and sequencing chemistry, a wide selection of NGS analyses can be performed. In everyday practice, targeted gene panel sequencing (i.e., cancer panels, metabolic panels, pharmacogenetic panels, etc.) and whole exome sequencing are the most widely employed. Copy number determination from exome sequencing data is challenging because the coverage of coding exomes by sequencing reads is not uniform and can be biased by sequence capture design (Robson et al., 2015; Deans et al., 2017). Recent advances in computational approaches allow increasingly accurate determination of SVs and by unraveling the whole sequence, the correct breakpoints can also be determined (Oliver et al., 2015).

In summary, there are numerous methods available for the determination of SVs, but there remains no gold standard approach. Based on clinical practice, a combination of these techniques (i.e., array CGH and MLPA, NGS and MLPA, or qPCR) would allow the best diagnostic accuracy. The introduction of NGS technology and the development of computational data analysis will significantly increase the throughput and improve the accuracy of determining SVs.

EXAMPLES OF SVs IN CANCER-PREDISPOSING GENES

The size of germline deletions affecting cancer-predisposing genes ranges from few hundreds of base pairs to several kilobases (Bozsik et al., 2020). Chromosome regions characterized by abundant directly oriented repeats, especially Alu sequences, are markedly prone to deletions primarily through the MMEJ mechanism (Smith et al., 2016). Frequently occurring deletion types are single exon deletions and multi-exon deletions. A typical example for the former is *BRCA1* del(ex8) (Sluiter and van Rensburg, 2011; Bozsik et al., 2020; **Figure 1B**) and for the latter *CHEK2* del(ex9-10) (Cybulski et al., 2007), and both variants cause frameshifts at the transcript level. The deletion of the full *BRCA1* gene, del(ex1-24) has been previously detected in various populations (de la Hoya et al., 2006; Engert et al., 2008; Fachal et al., 2014). Pseudogene regions of a gene can serve as long sequence stretches with considerable homology for NAHR events, thus providing hot-spots for illegitimate recombinations that often result in deletions. There are several rearrangements with different breakpoints between *BRCA1* and its pseudogene (*Psi-BRCA1*), generating a ∼37 kb deletion involving the *BRCA1* promoter and exons 1–2 (Puget et al., 2002). Similarly, the *PMS2* locus also has multiple pseudogenes, especially *PMS2CL*, which has an almost 100% sequence identity with *PMS2* exons 12–15. This exact sequence homology enables dynamic gene conversions and recombinations between the two regions (Kohlmann and Gruber, 1993). Concerning genotype-phenotype correlations, there is no evidence for HBOC genes, *BRCA1* and *BRCA2*; whole exon deletions manifest in a more severe phenotype of the disease than smaller-scale indels (Gad et al., 2003; Walsh et al., 2006; Smith et al., 2016; Bozsik et al., 2020). Whole exon deletions in Lynch syndrome genes, *MLH1* and *MSH2*, however, are associated with a slightly earlier age of onset for colorectal cancer than small truncating variants, but this difference does not reach the nominal significance of 0.05 (Smith et al., 2016). In contrast, whole gene deletions may have altered phenotypic

[1] www.mlpa.com

consequences in some syndromes: whole *NF1* deletions tend to cause a more severe phenotype, whereas whole *NF2* deletions generally result in a milder phenotype than truncating point mutations (Smith et al., 2016). Nevertheless, genetic alterations affecting additional causative genes may correlate with disease phenotype: in the case of a 7.4 Mb deletion encompassing *NF2* and neighboring genes corresponds to a more severe phenotype (Smith et al., 2016). In another example of contiguous gene deletion, germline 10q chromosomal deletion resulted in the loss of both *PTEN* and *BMPR1A*, and this corresponds to distinct pathological features of polyposis syndromes, underlining the complex interactions of these genes in tumorigenesis (Delnatte et al., 2006). Similarly, contiguous gene deletion within the 2p16-p21 chromosomal region, encompassing *MSH2, MSH6, EPCAM*, and 24 additional genes, causes Lynch syndrome with distinct phenotypic features (Salo-Mullen et al., 2018).

The majority of genomic duplications are directly oriented tandem repeats in CPGs (Sluiter and van Rensburg, 2011) and genome-wide (Newman et al., 2015) as well. The most prevalent tandem duplication in the *BRCA1* locus is *BRCA1* dup(ex13), which was detected in high frequency in nearly all European populations (The BEDSG, 2000). The bulk of single-exon or multi-exon duplications in CPGs reported so far are unambiguously pathogenic, although duplications encompassing the whole promoter together with a various number of downstream exons are evaluated as variants with unknown significance. For example, the examination of the *BRCA1* dup(ex1-2) variant by Fachal et al. (2014) failed to identify any aberrant transcripts (Fachal et al., 2014). Pathogenicity of other exon duplications detected by dosage-sensitive genotyping tests must also be confirmed by precise breakpoint assessment, as it was done for *BRCA2* dup(ex22-24) by van Luttikhuizen et al. (2020). They revealed, that the duplicated region was arranged in tandem and direct orientation, generating a PTC (van Luttikhuizen et al., 2020).

Particular types of copy gains include the insertion of mobile REs. Qian et al. (2017), conducted a large pan-cancer study on a panel of 26 genes and found that RE insertions were identifiable in 10 of the 26 genes tested (Qian et al., 2017). Indeed, RE insertions were detected in several genes (*BRCA1/2*, *APC*, *ATM*, *PMS2*, *MLH1*, and *MSH2*) by other groups studying hereditary breast and gastrointestinal cancers (Kaer et al., 2011; Li et al., 2020). Insertions of Alu repetitive motifs into exonic or intronic regions are the most prevalent transposition events in cancer-predisposing genes (Kaer et al., 2011). An Alu insertion in exon 3 of the *BRCA2* gene caused exon 3 skipping, and this is a founder mutation in the Portuguese population (Machado et al., 2007; Peixoto et al., 2011). In the Lynch syndrome-associated gene *PMS2*, insertion of an SVA nonautonomous retrotransposon element in intron 7 causes partial exonization of SVA using cryptic splice sites (van der Klift et al., 2012). *APC*, the germline susceptibility gene for FAP is disrupted by the insertion of an L1 sequence into exon 16 (Miki et al., 1992).

The detection of inversions can be challenging, therefore, their contribution to the SV pool is underestimated. In HNPCC, a 10 Mb paracentric inversion involving exons 1–7 of the *MSH2* gene was described first (Chen, 2008). This inversion was found to be a frequent cause of Lynch syndrome in a US population, accounting for an appreciable percentage of the mutational burden of this gene (Rhees et al., 2014). Later, another cryptic paracentric inversion of exons 2–6 of the same gene was detected (Liu et al., 2016). Germline inversion has also been shown for the *MLH1* locus, another major susceptibility gene in HNPCC. In this latter case, the inversion breakpoints are in intron 15 of *MLH1* and intron 3 of the neighboring *LRRFIP2* genes, generating two fusion transcripts between *MLH1* and *LRRFIP2* (Morak et al., 2011).

Translocations of whole chromosome arms are not typical events for disrupting tumor suppressor genes. However, two isolated cases with different chromosomal arm interchanges were described so far, each affecting the *APC* gene—a constitutional reciprocal translocation *t*(5;10) (van der Luijt et al., 1995) and a *t*(5;7) translocation (Sahnane et al., 2016).

The combination of rearrangement types manifests in complex genomic rearrangements. Despite these rearrangements possessing more than one junction point, they often arise from one molecular event, typically FoSTeS/MMBIR in cancer susceptibility genes. The characteristic pattern of deletion together with reverse duplication of some hundred base pairs occurs in various independent CNVs. *BRCA1* del(ex21-22) with reverse-oriented insertion of 236 bp of an intronic repeat is a founder complex CNV in the Czech population (Zikan et al., 2008; Ticha et al., 2010; **Figure 1D** and **Table 1**) and has also been reported as a recurrent variant in other European countries (Sluiter and van Rensburg, 2011; Bozsik et al., 2020). A complex recombination event characterized by the deletion of exons 5–10 and the insertion of a 35-bp nucleotide stretch in inverted orientation derived from the intron 3 sequence of the *BRCA1* gene is also a Czech founder mutation (Ticha et al., 2010). A deletion of exons 6–8 of *MLH1*, with the retention of 349-bp of intron 6 is also a complex rearrangement reported in one patient with colorectal cancer (McVety et al., 2005).

FREQUENCY OF GERMLINE SVs IN CANCER-PREDISPOSING GENES

The type and frequency of germline SVs in cancer susceptibility genes show wide differences across various populations. This is conceivably due in part to the different detection methods applied, the various and sometimes limited number of patients tested, as well as the founder alterations specific to certain populations. Founder variants are genetic alterations with a common origin, which are generated in an ancestor and spread through generations in an isolated ethnic group, thus these are recurrent and characteristic of a population. For example, haplotype analysis revealed, that the recurrent *BRCA1* deletion of exons 23 and 24 is a Greek founder mutation (Apostolou et al., 2017). Similarly, a recurrent exon 22 deletion in the *BRCA1* gene was found in the Netherlands (Sluiter and van Rensburg, 2011). Duplication of *BRCA1* exon 13 of Northern British origin (The BEDSG, 2000), as well as the above-mentioned deletion of exon 21-22 in *BRCA1* of Czech origin, are also frequently occurring CNVs in all populations in Europe (Ticha et al., 2010; Sluiter

TABLE 1 | Examples of founder SVs and frequencies in various populations.

Gene	SV (Running name)	Variant (HGVS)*	Population	Cancer syndrome	Frequency relative to gene mutations	Frequency in families of the syndrome	References
BRCA1	del(ex22)	NG_005905.2:g.168752_169261del	Dutch (Holland)	HBOC	36% of BRCA1(+)	NA	Petrij-Bosch et al., 1997
BRCA1	del(ex13)	NG_005905.2:g.133766_137600del					
BRCA1	del(ex3–16)	NC_000017:g.8655_55240del46586 NM_007294.3:c.81-1018_4986 +716del46586	Danish	HBOC	9/642 BRCA1/2(–)	NA	Hansen et al., 2009
BRCA1	del(ex17)	L78833:g.58530_61209delNG_005905. 2:g.147782_150460del	German	HBOC	NA	NA	Engert et al., 2008
BRCA1	del(ex5–14)	NG_005905.2:g.110966_142550del NM_007294.3:c.135-485_4485- 913del31583	Czech	HBOC	NA	4/239	Ticha et al., 2010
BRCA1	del(ex1–17)	NM_007294.3:c.1-21434_5075- 1084del80496					
BRCA1	del(ex21–22)	NG_005905.2:g.166375_170153delins:g. 162086_162321			NA	1/96, 2/172	Vasickova et al., 2007; Zikan et al., 2008
BRCA2	c.156_157insAlu	NG_012772.1:g.8686_8687insAlu	Portuguese	HBOC	NA	NA	Teugels et al., 2005; Machado et al., 2007
BRCA1	del(ex23–24)	NM_007294.3:c.5406+664_*8273del 11052L78833:g.80280_91331del NG_005905.2:g.169527_180579del	Greek	HBOC	22/181 BRCA1(+)	35/2092	Konstantopoulou et al., 2014; Apostolou et al., 2017
BRCA1	del(ex20)	NM_007294.3:c.5256_5277+3179del 3200L78833:g.71660_74860del3200			7/181 BRCA1(+)	7/760	
BRCA1	del(ex24)	NM_007294.3:c.5468- 285_5592+4019del4429_insCACAGL 78833:g.82651_87079del4429_ins5			13/181 BRCA1(+)	13/720	Konstantopoulou et al., 2014
BRCA1	dup(ex13)	L78833:g.44369_50449dupNG_005905. 1:g.133622_139702dup	Northern British	HBOC	NA	NA	The BEDSG, 2000
BRCA1	del(ex9–12)	NG_005905.1:g.118955_133611del	Hispanic	HBOC	4/106 BRCA1/2(–)	NA	Weitzel et al., 2007
BRCA1	del(ex3–5)	L78833:g.8097_22733delNG_005905. 2:g.97346_111983del	Eastern Spanish	HBOC	10,97% of BRCA1(+)	NA	Palanca et al., 2013
CHEK2	del(ex9–10)	NM_007194.3:c.909- 2028_1095+330del5395	Czech	HBOC	NA	NA	Cybulski et al., 2007
MLH1	del(ex17–19)	NM_000249.3:c.1896+280_oLRRFIP 2:c.1750-678del	Portuguese	HNPCC	17% of MMR(+)	NA	Pinhero et al., 2011
MSH2	del(ex7)	NM_000251.2:c.1077-3513_1276 +5655	Spanish	HNPCC	47% of MSH2(+)	7/160	Perez-Cabornero et al., 2011
MSH2	del(ex4–8)	NM_000251.2:c.646- 1019_1386+2420del					
MSH2	del(ex1–6)	chr2:g.47,618,487_47,650,860delins (155); hg19	United States	HNPCC	NA	NA	van der Klift et al., 2005
APC	del(promB)	chr5:g.112,703,831-112,710,688; GRCh38/hg38	Italian	FAP	NA	NA	Marabelli et al., 2017

**Reference sequences and nomenclature for variants are taken from the articles cited.*
Running names are the short forms of the respective variants.
SV, structural variation; HBOC, hereditary breast and ovarian syndrome; HNPCC, hereditary non-polyposis cancer syndrome; FAP, familial adenomatous polyposis; BRCA1(+), BRCA1 mutation carriers; BRCA1/2(–), BRCA1/2 mutation non-carriers; MMR(+), MLH1, MSH2, and MSH6 mutation positives; MSH2(+), MSH2 mutation carriers.

and van Rensburg, 2011; Bozsik et al., 2020). A large screening study in the US revealed, that 70.8% of all *BRCA1* rearrangements of Western and Northern European origin are made up of five founder CNVs (Judkins et al., 2012). **Table 1** summarizes some examples of founder pathogenic SVs and their frequencies in the source populations.

Deletions are the most prevalent CNV types in cancer susceptibility genes (Sluiter and van Rensburg, 2011; Mancini-DiNardo et al., 2019; Bozsik et al., 2020), contributing to approximately 80–85% of all rearrangements (Mancini-DiNardo et al., 2019). In contrast, duplications account for only 10–15% of rearrangements (Mancini-DiNardo et al., 2019;

Bozsik et al., 2020). The predominance of molecular processes resulting in genomic deletions compared to duplications might explain the observed difference between the frequencies of these alterations (Hastings et al., 2009b). RE insertions also represent a significant SV type as they accounted for one in every 325 unique pathogenic variants detected in a large pan-cancer study (Qian et al., 2017). In this cohort, 92% of all RE events were retrotransposition of Alu elements, while the most frequently affected genes with unique RE insertions were *BRCA2* (45.9%) and *ATM* (16.2%) (Qian et al., 2017). Mechanistically, there is no reason for the observed predominance of *BRCA2* in RE events.

A hereditary pan-cancer gene panel survey of 376,159 individuals in the US revealed 3,461 LGRs in 27 genes (Mancini-DiNardo et al., 2019). In general, SVs accounted for 7.2% of all pathogenic variants detected. The largest proportion of pathogenic LGRs were identified in *BRCA1* (27.4%), followed by *PMS2* (11.7%), *CHEK2* (11.1%), and *MSH2* (8.9%) (Mancini-DiNardo et al., 2019). In a separate study focusing on point-mutation-negative HNPCC patients, 11% of cases harbored large rearrangements in four predisposition genes (*MLH1*, *MSH2*, *MSH6*, and *PMS2*) among which 29.6% affected the *MSH2* gene (van der Klift et al., 2005). Similarly, 15% of the point mutation-negative patients with classical FAP had a genomic deletion in *APC* (Michils et al., 2005). On the contrary, various studies in different populations focusing on *BRCA1*/2 mutation-negative HBOC patients revealed that large rearrangements of *BRCA1* and *BRCA2* genes contributed to only 2–3% of the cases (Agata et al., 2006; Preisler-Adams et al., 2006; Thomassen et al., 2006).

The proportion of pathogenic SVs in a given locus as a fraction of the clear-cut mutations of the gene shows a different ratio pattern. **Table 2** summarizes the reported SV ratios in the most relevant susceptibility genes of various cancer predisposition syndromes. Typically, SVs represent 10% (ranging from 0.1 to 60.7%) of the acknowledged mutations of these genes. The high extreme was detected in the *STK11* gene, where 30–60% of all mutations are CNVs (Aretz et al., 2005; Mancini-DiNardo et al., 2019). The ratios are also high in the case of *MSH2* and *PMS2*, where large deletions account for ∼20 and ∼25% of mutations, respectively (Kohlmann and Gruber, 1993; Mancini-DiNardo et al., 2019). *MUTYH* is the less abundant in SVs with its ratio of 0.1% relative to all mutations of the genes (Mancini-DiNardo et al., 2019). The differences with regard to SV frequencies in different genes are conceivably a function of their genetic surroundings and chromosomal complexity. For example, the higher proportion of Alu repeats may contribute to the higher rate of genomic rearrangements in *MSH2* compared to that of *MLH1* (van der Klift et al., 2005). The prevalence of *BRCA1* rearrangements over *BRCA2* is explained also by the differences in the ratio of intronic Alu repeats between the two genes (Judkins et al., 2012; Sluiter and van Rensburg, 2011). Similarly, *PMS2*, due to its extensive pseudogene regions is an especially good subject for rearrangements through recombinations (Smith et al., 2016).

CNVs in *BRCA1* and *BRCA2* genes, being the most penetrant HBOC predisposition loci, are extensively studied in various populations worldwide. To date, more than 100 LGRs have been characterized in *BRCA1*, whereas much fewer have been

TABLE 2 | Relative ratios of germline SVs compared to all mutations of the susceptibility gene in various cancer syndromes.

Gene	Syndrome	Ratio of SVs in all mutations of the gene	References
STK11	Juvenile polyposis syndrome	60.7%	Mancini-DiNardo et al., 2019
		30%	Borun et al., 2015
SMAD4		10%	Calva-Cerqueira et al., 2009
BMPR1A		10%	Calva-Cerqueira et al., 2009
SMAD4 & BMPR1A		30%	Aretz et al., 2005
APC	Familial adenomatous polyposis	6%	Kerr et al., 2013
		8.3%	Mancini-DiNardo et al., 2019
MUTYH		0.1%	Mancini-DiNardo et al., 2019
MLH1	Hereditary non-polyposis colorectal cancer syndrome	10%	Smith et al., 2016
MSH2		24%	Smith et al., 2016
MSH6		2.7%	Mancini-DiNardo et al., 2019
PMS2		25%	Mancini-DiNardo et al., 2019
		21%	Senter et al., 2008
		37%	Vaughn et al., 2010
PTCH	Gorlin syndrome	15%	Smith et al., 2016
VHL	Von Hippel-Lindau disease	16.6%	Smith et al., 2016
		25%	Maher et al., 1996
NF1	Neurofibromatosis type 1	12%	Smith et al., 2016
NF2	Neurofibromatosis type 2	20%	Smith et al., 2016
MEN1	Multiple endocrine neoplasia type 1 syndrome	12%	Pardi et al., 2017
CHEK2	Hereditary breast and ovarian cancer syndrome	14%	Nizic-Kos et al., 2020
		15.26%	Kleiblova et al., 2019
PALB2		9.6%	Mancini-DiNardo et al., 2019
		18%	Janatova et al., 2013
RAD51C		21%	Mancini-DiNardo et al., 2019
BARD1		10.2%	Mancini-DiNardo et al., 2019
BRIP1		4.7%	Mancini-DiNardo et al., 2019
ATM	Ataxia telangectasia	5.8%	Mancini-DiNardo et al., 2019
CDH1	Hereditary diffuse castric cancer	14.4%	Mancini-DiNardo et al., 2019
		16.7%	Molinaro et al., 2014
TP53	Li-Fraumeni syndrome	10%	Smith et al., 2016

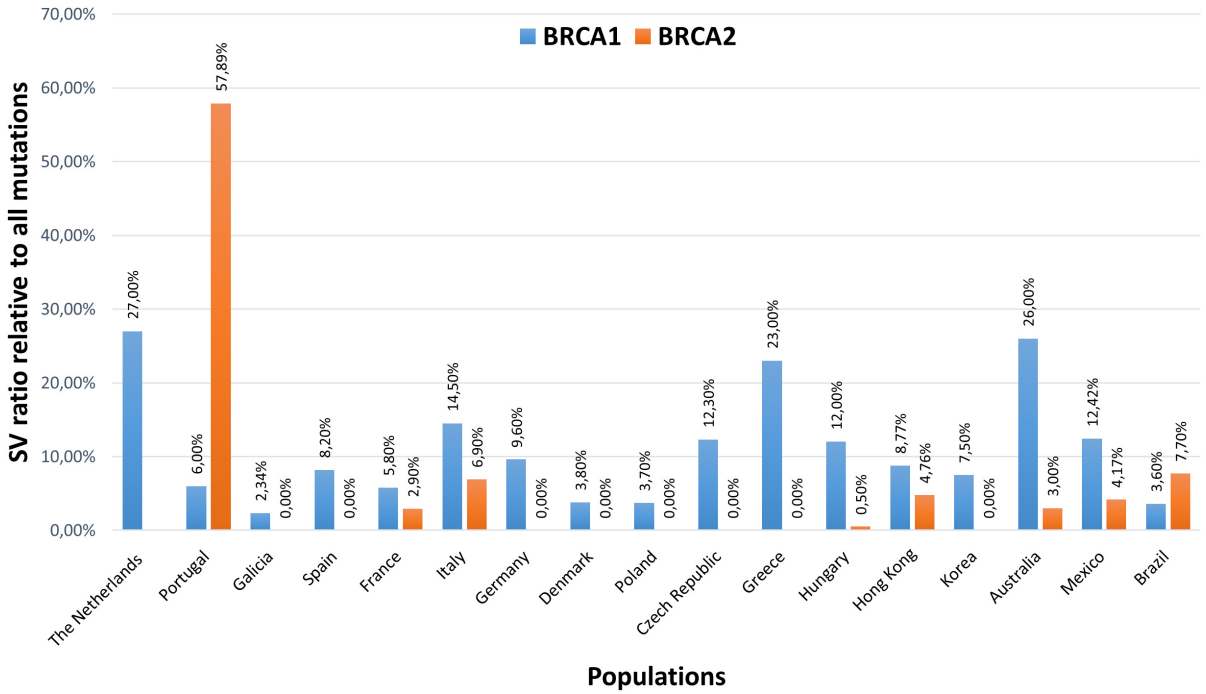


FIGURE 2 | Structural variation ratios of *BRCA1* and *BRCA2* genes relative to all pathogenic mutations detected in HBOC probands of various ethnicities. The Netherlands (Hogervorst et al., 2003), Portugal (Peixoto et al., 2011), Galicia (Fachal et al., 2014), Spain (de la Hoya et al., 2006), France (Caux-Moncoutier et al., 2011), Italy (Concolino et al., 2018), Germany (Engert et al., 2008), Denmark (Thomassen et al., 2006), Poland (Rudnicka et al., 2013), Czech Republic (Ticha et al., 2010), Greece (Armaou et al., 2009), Hungary (Bozsik et al., 2020), Hong Kong (Kwong et al., 2015), Korea (Seong et al., 2014), Australia (James et al., 2015), Mexico (Lopez-Urrutia et al., 2019), Brazil (Palmero et al., 2018).

characterized in *BRCA2* (Sluiter and van Rensburg, 2011). The CNV ratios of *BRCA1* and *BRCA2* genes relative to all pathogenic mutations detected in HBOC probands of various ethnicities are visualized on the histogram in **Figure 2**. On average, CNVs account for 10% of all pathogenic mutations of the *BRCA1* gene; the differences in ratios in various ethnicities are mainly attributed to founder mutations. *BRCA2* locus has only low contribution (<0.5%) to CNVs (Hogervorst et al., 2003; de la Hoya et al., 2006; Thomassen et al., 2006; Engert et al., 2008; Ticha et al., 2010; Fachal et al., 2014). A remarkable percentage of *BRCA2* large rearrangements have only been detected in Portugal, where the c.156_157insAlu founder mutation constitutes the bulk of the cases (Machado et al., 2007). An elevated ratio of *BRCA2* CNVs is also observed in male breast cancer patients (Tournier et al., 2004).

Additional association studies are seeking to identify further pathogenic CNVs in breast cancer contributing to the disease phenotype. Kumaran et al. (2017), identified 200 common germline CNVs associated with breast cancer in a whole-genome sequencing study of 422 breast cancer cases and 348 controls (Kumaran et al., 2017). Moreover, they also confirmed, that germline CNVs conferred dosage effects on gene expression in breast tissue (Kumaran et al., 2017). Similarly, another study of genome-wide germline CNVs identified 275 unique rearrangements that potentially contribute to breast cancer initiation and/or progression (Masson et al., 2014).

SUMMARY

Within germline SVs in CPGs for hereditary cancer syndromes, copy number changes are the prevailing alterations. The predominant CNVs are deletions, affecting various portions of the genes. The main structural source of these deletions is intronic Alu sequences. Double-strand break repairs and additional molecular mechanisms harness these sequence homologies and may result in copy changes through ectopic alignments. Studies conducted in different populations confirmed that SVs generally account for 7–10% of all mutations in CPGs, thus their contribution of mutational burden is significant. Precise detection of these types of alterations is essential to provide an optimal genetic diagnosis. Differences in neighboring genetic architecture, as well as various applied detection techniques, may contribute to the wide range of variations in the exact ratios of pathogenic CNVs compared to point mutations (Mancini-DiNardo et al., 2019) [etc.].

The association with the clinico-pathological phenotype is straightforward in the majority of germline SVs, however, in a few cases, and especially within some duplications, pathogenic effects cannot be addressed unambiguously. Moreover, several studies proposed that copy number changes, although larger in size, do not necessarily associate with a more severe pathological phenotype than smaller-scale indels (Walsh et al., 2006; Rhees et al., 2014; Smith et al., 2016; Bozsik et al., 2020). On the other

hand, extensive rearrangements affecting the whole gene together with several neighboring genes may elicit a complex phenotype due to the putative interfering effects of the respective proto-oncogenes and tumor suppressors (Delnatte et al., 2006; Smith et al., 2016). However, since the number of cases harboring such rearrangements is limited, further studies are needed to ascertain these correlations.

Due to the continuous development of dosage-sensitive detection modes and validations, an increasing number of germline structural rearrangements are being discovered in several CPGs. Note that synthesis of the data highlighted differences regarding structural variation types and frequencies between the studied CPGs. This has raised the possibility, that some rearrangement types, mainly inversions and insertions, may be underrepresented as a consequence of genotyping insufficiency, and a significant portion of heritability may remain unexplained with current genotyping assays. For example, in NGS sequencing results spurious deletions, not validated as real copy number losses may be a consequence of allelic dropout or failed alignment of the reads due to possible breakpoints of other types of rearrangements. Genotyping techniques that also enable sequencing of introns are preferred since the majority of rearrangement breakpoints reside in these regions. Equally important, several deletions affecting more genes may manifest in a multilocus phenotype, modulating the typical symptoms of diseases. Therefore, careful evaluation of the syndromic spectrum is warranted for determining the genotyping eligibility criteria.

DATABASES

The following curated databases register the detected SVs: InSight LOVD for gastrointestinal hereditary tumors (http://insight-database.org). Breast-and ovarian cancer: ENIGMA, BRCA Exchange (https://brcaexchange.org/variants) and LOVD Fanconi anemia mutation database (https://databases.lovd.nl/shared/genes/BRCA1). General: Database of Genomic Variants (http://dgv.tcag.ca/dgv/app/home). Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php).

AUTHOR CONTRIBUTIONS

AB, TP, VG, and AP wrote and carried out the original draft preparation. TP, JP, HB, and AP reviewed and edited the manuscript. AP carried out the funding acquisition. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was funded by the Hungarian Research Grants TUDFO/51757/2019-ITM and TKP2020-NKA-26.

REFERENCES

Agata, S., Viel, A., Della Puppa, L., Cortesi, L., Fersini, G., Callegaro, M., et al. (2006). Prevalence of BRCA1 genomic rearrangements in a large cohort of Italian breast and breast/ovarian cancer families without detectable BRCA1 and BRCA2 point mutations. *Genes Chromosomes Cancer* 45, 791–797. doi: 10.1002/gcc.20342

Apostolou, P., Pertesi, M., Aleporou-Marinou, V., Dimitrakakis, C., Papadimitriou, C., Razis, E., et al. (2017). Haplotype analysis reveals that the recurrent BRCA1 deletion of exons 23 and 24 is a Greek founder mutation. *Clin. Genet.* 91, 482–487. doi: 10.1111/cge.12824

Aretz, S., Stienen, D., Uhlhaas, S., Loff, S., Back, W., Pagenstecher, C., et al. (2005). High proportion of large genomic STK11 deletions in Peutz-Jeghers syndrome. *Hum. Mutat.* 26, 513–519. doi: 10.1002/humu.20253

Arkhipova, I. R., and Yushenova, I. A. (2019). Giant transposons in eukaryotes: is bigger better? *Genome Biol. Evol.* 11, 906–918. doi: 10.1093/gbe/evz041

Armaou, S., Pertesi, M., Fostira, F., Thodi, G., Athanasopoulos, P. S., Kamakari, S., et al. (2009). Contribution of BRCA1 germ-line mutations to breast cancer in Greece: a hospital-based study of 987 unselected breast cancer cases. *Br. J. Cancer* 101, 32–37. doi: 10.1038/sj.bjc.6605115

Bailey, J. A., Gu, Z., Clark, R. A., Reinert, K., Samonte, R. V., Schwartz, S., et al. (2002). Recent segmental duplications in the human genome. *Science* 297, 1003–1007. doi: 10.1126/science.1072047

Bayani, J., and Squire, J. A. (2004). Fluorescence in situ Hybridization (FISH). *Curr. Protoc. Cell Biol.* Chapter 22:Unit 22.4. doi: 10.1002/0471143030.cb2204s23

Belogianni, I., Apessos, A., Mihalatos, M., Razi, E., Labropoulos, S., Petounis, A., et al. (2004). Characterization of a novel large deletion and single point mutations in the BRCA1 gene in a Greek cohort of families with suspected hereditary breast cancer. *BMC Cancer* 4:61. doi: 10.1186/1471-2407-4-61

Borun, P., De Rosa, M., Nedoszytko, B., Walkowiak, J., and Plawski, A. (2015). Specific Alu elements involved in a significant percentage of copy number variations of the STK11 gene in patients with Peutz-Jeghers syndrome. *Fam. Cancer* 14, 455–461. doi: 10.1007/s10689-015-9800-5

Bozsik, A., Pocza, T., Papp, J., Vaszko, T., Butz, H., Patocs, A., et al. (2020). Complex characterization of germline large genomic rearrangements of the BRCA1 and BRCA2 genes in high-risk breast cancer patients-novel variants from a Large National Center. *Int. J. Mol. Sci.* 21:4650. doi: 10.3390/ijms21134650

Butz, H., and Patocs, A. (2019). "Brief summary of the most important molecular genetic methods (PCR, qPCR, microarray, next-generation sequencing, etc.)," in *Genetics of Endocrine Diseases and Syndromes*, eds P. Igaz and A. Patócs (Cham: Springer), 33–52.

Calva-Cerqueira, D., Chinnathambi, S., Pechman, B., Bair, J., Larsen-Haidle, J., and Howe, J. R. (2009). The rate of germline mutations and large deletions of SMAD4 and BMPR1A in juvenile polyposis. *Clin. Genet.* 75, 79–85. doi: 10.1111/j.1399-0004.2008.01091.x

Cantsilieris, S., Baird, P. N., and White, S. J. (2013). Molecular methods for genotyping complex copy number polymorphisms. *Genomics* 101, 86–93. doi: 10.1016/j.ygeno.2012.10.004

Caux-Moncoutier, V., Castera, L., Tirapo, C., Michaux, D., Remon, M. A., Lauge, A., et al. (2011). EMMA, a cost- and time-effective diagnostic method for simultaneous detection of point mutations and large-scale genomic rearrangements: application to BRCA1 and BRCA2 in 1,525 patients. *Hum. Mutat.* 32, 325–334. doi: 10.1002/humu.21414

Ceulemans, S., van der Ven, K., and Del-Favero, J. (2012). Targeted screening and validation of copy number variations. *Methods Mol. Biol.* 838, 311–328. doi: 10.1007/978-1-61779-507-7_15

Chen, J. M. (2008). The 10-Mb paracentric inversion of chromosome arm 2p in activating MSH2 and causing hereditary nonpolyposis colorectal cancer: re-annotation and mutational mechanisms. *Genes Chromosomes Cancer* 47, 543–545. doi: 10.1002/gcc.20556

Colnaghi, R., Carpenter, G., Volker, M., and O'Driscoll, M. (2011). The consequences of structural genomic alterations in humans: genomic disorders, genomic instability and cancer. *Semin. Cell Dev. Biol* . 22, 875–885. doi: 10.1016/j.semcdb.2011.07.010

Concolino, P., Rizza, R., Mignone, F., Costella, A., Guarino, D., Carboni, I., et al. (2018). A comprehensive BRCA1/2 NGS pipeline for an immediate Copy Number Variation (CNV) detection in breast and ovarian cancer molecular diagnosis. *Clin. Chim. Acta* 480, 173–179. doi: 10.1016/j.cca.2018.02.012

Cybulski, C., Wokolorczyk, D., Huzarski, T., Byrski, T., Gronwald, J., Gorski, B., et al. (2007). A deletion in CHEK2 of 5,395 bp predisposes to breast cancer in Poland. *Breast Cancer Res. Treat.* 102, 119–122. doi: 10.1007/s10549-006-9320-y

Davies, J. J., Wilson, I. M., and Lam, W. L. (2005). Array CGH technologies and their applications to cancer genomes. *Chromosome Res.* 13, 237–248. doi: 10.1007/s10577-005-2168-x

de la Hoya, M., Gutierrez-Enriquez, S., Velasco, E., Osorio, A., Sanchez de Abajo, A., Vega, A., et al. (2006). Genomic rearrangements at the BRCA1 locus in Spanish families with breast/ovarian cancer. *Clin. Chem.* 52, 1480–1485. doi: 10.1373/clinchem.2006.070110

Deans, Z. C., Costa, J. L., Cree, I., Dequeker, E., Edsjo, A., Henderson, S., et al. (2017). Integration of next-generation sequencing in clinical diagnostic molecular pathology laboratories for analysis of solid tumours; an expert opinion on behalf of IQN Path ASBL. *Virchows Arch.* 470, 5–20. doi: 10.1007/s00428-016-2025-7

Delnatte, C., Sanlaville, D., Mougenot, J. F., Vermeesch, J. R., Houdayer, C., Blois, M. C., et al. (2006). Contiguous gene deletion within chromosome arm 10q is associated with juvenile polyposis of infancy, reflecting cooperation between the BMPR1A and PTEN tumor-suppressor genes. *Am. J. Hum. Genet.* 78, 1066–1074. doi: 10.1086/504301

Du, C., Mark, D., Wappenschmidt, B., Bockmann, B., Pabst, B., Chan, S., et al. (2018). A tandem duplication of BRCA1 exons 1-19 through DHX8 exon 2 in four families with hereditary breast and ovarian cancer syndrome. *Breast Cancer Res. Treat.* 172, 561–569. doi: 10.1007/s10549-018-4957-x

Engert, S., Wappenschmidt, B., Betz, B., Kast, K., Kutsche, M., Hellebrand, H., et al. (2008). MLPA screening in the BRCA1 gene from 1,506 German hereditary breast cancer cases: novel deletions, frequent involvement of exon 17, and occurrence in single early-onset cases. *Hum. Mutat.* 29, 948–958. doi: 10.1002/humu.20723

Fachal, L., Blanco, A., Santamarina, M., Carracedo, A., and Vega, A. (2014). Large genomic rearrangements of BRCA1 and BRCA2 among patients referred for genetic analysis in Galicia (NW Spain): delimitation and mechanism of three novel BRCA1 rearrangements. *PLoS One* 9:e93306. doi: 10.1371/journal.pone.0093306

Firth, H. V., Richards, S. M., Bevan, A. P., Clayton, S., Corpas, M., Rajan, D., et al. (2009). DECIPHER: database of chromosomal imbalance and phenotype in humans using ensembl resources. *Am. J. Hum. Genet.* 84, 524–533. doi: 10.1016/j.ajhg.2009.03.010

Gad, S., Bieche, I., Barrois, M., Casilli, F., Pages-Berhouet, S., Dehainault, C., et al. (2003). Characterisation of a 161 kb deletion extending from the NBR1 to the BRCA1 genes in a French breast-ovarian cancer family. *Hum. Mutat.* 21:654. doi: 10.1002/humu.9148

Goodier, J. L. (2016). Restricting retrotransposons: a review. *Mob. DNA* 7:16. doi: 10.1186/s13100-016-0070-z

Hancks, D. C. Jr., and Kazazian, H. H. Jr. (2016). Roles for retrotransposon insertions in human disease. *Mob. DNA* 7:9. doi: 10.1186/s13100-016-0065-9

Hansen, T., Jonson, L., Albrechtsen, A., Andersen, M. K., Ejlertsen, B., and Nielsen, F. C. (2009). Large BRCA1 and BRCA2 genomic rearrangements in Danish high risk breast-ovarian cancer families. *Breast Cancer Res. Treat.* 115, 315–323. doi: 10.1007/s10549-008-0088-0

Hastings, P. J., Ira, G., and Lupski, J. R. (2009a). A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet.* 5:e1000327. doi: 10.1371/journal.pgen.1000327

Hastings, P. J., Lupski, J. R., Rosenberg, S. M., and Ira, G. (2009b). Mechanisms of change in gene copy number. *Nat. Rev. Genet.* 10, 551–564. doi: 10.1038/nrg2593

Hogervorst, F. B., Nederlof, P. M., Gille, J. J., McElgunn, C. J., Grippeling, M., Pruntel, R., et al. (2003). Large genomic deletions and duplications in the BRCA1 gene identified by a novel quantitative method. *Cancer Res.* 63, 1449–1453.

James, P. A., Sawyer, S., Boyle, S., Young, M. A., Kovalenko, S., Doherty, R., et al. (2015). Large genomic rearrangements in the familial breast and ovarian cancer gene BRCA1 are associated with an increased frequency of high risk features. *Fam. Cancer* 14, 287–295. doi: 10.1007/s10689-015-9785-0

Janatova, M., Kleibl, Z., Stribrna, J., Panczak, A., Vesela, K., Zimovjanova, M., et al. (2013). The PALB2 gene is a strong candidate for clinical testing in BRCA1- and BRCA2-negative hereditary breast cancer. *Cancer Epidemiol. Biomark. Prev.* 22, 2323–2332. doi: 10.1158/1055-9965.EPI-13-0745-T

Judkins, T., Rosenthal, E., Arnell, C., Burbidge, L. A., Geary, W., Barrus, T., et al. (2012). Clinical significance of large rearrangements in BRCA1 and BRCA2. *Cancer* 118, 5210–5216. doi: 10.1002/cncr.27556

Kaer, K., Branovets, J., Hallikma, A., Nigumann, P., and Speek, M. (2011). Intronic L1 retrotransposons and nested genes cause transcriptional interference by inducing intron retention, exonization and cryptic polyadenylation. *PLoS One* 6:e26099. doi: 10.1371/journal.pone.0026099

Kerr, S. E., Thomas, C. B., Thibodeau, S. N., Ferber, M. J., and Halling, K. C. (2013). APC germline mutations in individuals being evaluated for familial adenomatous polyposis: a review of the Mayo Clinic experience with 1591 consecutive tests. *J. Mol. Diagn.* 15, 31–43. doi: 10.1016/j.jmoldx.2012.07.005

Kleiblova, P., Stolarova, L., Krizova, K., Lhota, F., Hojny, J., Zemankova, P., et al. (2019). Identification of deleterious germline CHEK2 mutations and their association with breast and ovarian cancer. *Int. J. Cancer* 145, 1782–1797. doi: 10.1002/ijc.32385

Knudson, A. G. Jr. (1971). Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl. Acad. Sci. U.S.A.* 68, 820–823. doi: 10.1073/pnas.68.4.820

Kohlmann, W., and Gruber, S. B. (1993). "Lynch syndrome," in *GeneReviews((R))*, eds M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. Stephens, et al. (Seattle, WA: University of Washington).

Konstantopoulou, I., Tsitlaidou, M., Fostira, F., Pertesi, M., Stavropoulou, A. V., Triantafyllidou, O., et al. (2014). High prevalence of BRCA1 founder mutations in Greek breast/ovarian families. *Clin. Genet.* 85, 36–42. doi: 10.1111/cge.12274

Kovacs, M. E., Papp, J., Szentirmay, Z., Otto, S., and Olah, E. (2009). Deletions removing the last exon of TACSTD1 constitute a distinct class of mutations predisposing to Lynch syndrome. *Hum. Mutat.* 30, 197–203. doi: 10.1002/humu.20942

Kowalczykowski, S. C. (2015). An overview of the molecular mechanisms of recombinational DNA repair. *Cold Spring Harb. Perspect. Biol.* 7:a016410. doi: 10.1101/cshperspect.a016410

Kozlowski, P., Jasinska, A. J., and Kwiatkowski, D. J. (2008). New applications and developments in the use of multiplex ligation-dependent probe amplification. *Electrophoresis* 29, 4627–4636. doi: 10.1002/elps.200800126

Krumm, N., Sudmant, P. H., Ko, A., O'Roak, B. J., Malig, M., Coe, B. P., et al. (2012). Copy number variation detection and genotyping from exome sequence data. *Genome Res.* 22, 1525–1532. doi: 10.1101/gr.138115.112

Kumaran, M., Cass, C. E., Graham, K., Mackey, J. R., Hubaux, R., Lam, W., et al. (2017). Germline copy number variations are associated with breast cancer risk and prognosis. *Sci. Rep.* 7:14621. doi: 10.1038/s41598-017-14799-7

Kwong, A., Chen, J., Shin, V. Y., Ho, J. C., Law, F. B., Au, C. H., et al. (2015). The importance of analysis of long-range rearrangement of BRCA1 and BRCA2 in genetic diagnosis of familial breast cancer. *Cancer Genet.* 208, 448–454. doi: 10.1016/j.cancergen.2015.05.031

Li, Y., Salo-Mullen, E., Varghese, A., Trottier, M., Stadler, Z. K., and Zhang, L. (2020). Insertion of an Alu-like element in MLH1 intron 7 as a novel cause of Lynch syndrome. *Mol. Genet. Genomic Med.* 8:e1523. doi: 10.1002/mgg3.1523

Ligtenberg, M. J., Kuiper, R. P., Chan, T. L., Goossens, M., Hebeda, K. M., Voorendt, M., et al. (2009). Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. *Nat. Genet.* 41, 112–117. doi: 10.1038/ng.283

Liu, Q., Hesson, L. B., Nunez, A. C., Packham, D., Williams, R., Ward, R. L., et al. (2016). A cryptic paracentric inversion of MSH2 exons 2-6 causes Lynch syndrome. *Carcinogenesis* 37, 10–17. doi: 10.1093/carcin/bgv154

Lopez-Urrutia, E., Salazar-Rojas, V., Brito-Elias, L., Coca-Gonzalez, M., Silva-Garcia, J., Sanchez-Marin, D., et al. (2019). BRCA mutations: is everything said? *Breast Cancer Res. Treat.* 173, 49–54. doi: 10.1007/s10549-018-4986-5

Machado, P. M., Brandao, R. D., Cavaco, B. M., Eugenio, J., Bento, S., Nave, M., et al. (2007). Screening for a BRCA2 rearrangement in high-risk breast/ovarian cancer families: evidence for a founder effect and analysis of the associated phenotypes. *J. Clin. Oncol.* 25, 2027–2034. doi: 10.1200/JCO.2006.06.9443

Maher, E. R., Webster, A. R., Richards, F. M., Green, J. S., Crossey, P. A., Payne, S. J., et al. (1996). Phenotypic expression in von Hippel-Lindau disease: correlations

with germline VHL gene mutations. *J. Med. Genet.* 33, 328–332. doi: 10.1136/jmg.33.4.328

Mancini-DiNardo, D., Judkins, T., Kidd, J., Bernhisel, R., Daniels, C., Brown, K., et al. (2019). Detection of large rearrangements in a hereditary pan-cancer panel using next-generation sequencing. *BMC Med. Genomics* 12:138. doi: 10.1186/s12920-019-0587-3

Marabelli, M., Gismondi, V., Ricci, M. T., Vetro, A., Abou Khouzam, R., Rea, V., et al. (2017). A novel APC promoter 1B deletion shows a founder effect in Italian patients with classical familial adenomatous polyposis phenotype. *Genes Chromosomes Cancer* 56, 846–854. doi: 10.1002/gcc.22488

Masson, A. L., Talseth-Palmer, B. A., Evans, T. J., Grice, D. M., Hannan, G. N., and Scott, R. J. (2014). Expanding the genetic basis of copy number variation in familial breast cancer. *Hered. Cancer Clin. Pract.* 12:15. doi: 10.1186/1897-4287-12-15

McVety, S., Younan, R., Li, L., Gordon, P. H., Wong, N., Foulkes, W. D., et al. (2005). Novel genomic insertion– deletion in MLH1: possible mechanistic role for non-homologous end-joining DNA repair. *Clin. Genet.* 68, 234–238. doi: 10.1111/j.1399-0004.2005.00486.x

McVey, M., and Lee, S. E. (2008). MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet.* 24, 529–538. doi: 10.1016/j.tig.2008.08.007

Michils, G., Tejpar, S., Thoelen, R., van Cutsem, E., Vermeesch, J. R., Fryns, J. P., et al. (2005). Large deletions of the APC gene in 15% of mutation-negative patients with classical polyposis (FAP): a Belgian study. *Hum. Mutat.* 25, 125–134. doi: 10.1002/humu.20122

Miki, Y., Nishisho, I., Horii, A., Miyoshi, Y., Utsunomiya, J., Kinzler, K. W., et al. (1992). Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res.* 52, 643–645.

Molinaro, V., Pensotti, V., Marabelli, M., Feroce, I., Barile, M., Pozzi, S., et al. (2014). Complementary molecular approaches reveal heterogeneous CDH1 germline defects in Italian patients with hereditary diffuse gastric cancer (HDGC) syndrome. *Genes Chromosomes Cancer* 53, 432–445. doi: 10.1002/gcc.22155

Morak, M., Koehler, U., Schackert, H. K., Steinke, V., Royer-Pokora, B., Schulmann, K., et al. (2011). Biallelic MLH1 SNP cDNA expression or constitutional promoter methylation can hide genomic rearrangements causing Lynch syndrome. *J. Med. Genet.* 48, 513–519. doi: 10.1136/jmedgenet-2011-100050

Newman, S., Hermetz, K. E., Weckselblatt, B., and Rudd, M. K. (2015). Next-generation sequencing of duplication CNVs reveals that most are tandem and some create fusion genes at breakpoints. *Am. J. Hum. Genet.* 96, 208–220. doi: 10.1016/j.ajhg.2014.12.017

Nizic-Kos, T., Krajc, M., Blatnik, A., Stegel, V., Skerl, P., Novakovic, S., et al. (2020). Bilateral disease common among slovenian CHEK2-positive breast cancer patients. *Ann. Surg. Oncol.* [Epub ahead of print]. doi: 10.1245/s10434-020-09178-y

Oliver, G. R., Hart, S. N., and Klee, E. W. (2015). Bioinformatics for clinical next generation sequencing. *Clin. Chem.* 61, 124–135. doi: 10.1373/clinchem.2014.224360

Palanca, S., de Juan, I., Perez-Simo, G., Barragan, E., Chirivella, I., Martinez, E., et al. (2013). The deletion of exons 3-5 of BRCA1 is the first founder rearrangement identified in breast and/or ovarian cancer Spanish families. *Fam. Cancer* 12, 119–123. doi: 10.1007/s10689-012-9579-6

Palmero, E. I., Carraro, D. M., Alemar, B., Moreira, M. A. M., Ribeiro-Dos-Santos, A., Abe-Sandes, K., et al. (2018). The germline mutational landscape of BRCA1 and BRCA2 in Brazil. *Sci. Rep.* 8:9188. doi: 10.1038/s41598-018-27315-2

Pardi, E., Borsari, S., Saponaro, F., Bogazzi, F., Urbani, C., Mariotti, S., et al. (2017). Mutational and large deletion study of genes implicated in hereditary forms of primary hyperparathyroidism and correlation with clinical features. *PLoS One* 12:e0186485. doi: 10.1371/journal.pone.0186485

Peixoto, A., Santos, C., Pinheiro, M., Pinto, P., Soares, M. J., Rocha, P., et al. (2011). International distribution and age estimation of the Portuguese BRCA2 c.*156*_157insAlu founder mutation. *Breast Cancer Res. Treat.* 127, 671–679. doi: 10.1007/s10549-010-1036-3

Perez-Cabornero, L., Borras Flores, E., Infante Sanz, M., Velasco Sampedro, E., Acedo Becares, A., Lastra Aras, E., et al. (2011). Characterization of new founder Alu-mediated rearrangements in MSH2 gene associated with a Lynch syndrome phenotype. *Cancer Prev. Res.* 4, 1546–1555. doi: 10.1158/1940-6207.CAPR-11-0227

Petrij-Bosch, A., Peelen, T., van Vliet, M., van Eijk, R., Olmer, R., Drusedau, M., et al. (1997). BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients. *Nat. Genet.* 17, 341–345. doi: 10.1038/ng1197-341

Pinhero, R., Pazhekattu, R., Marangoni, A. G., Liu, Q., and Yada, R. Y. (2011). Alleviation of low temperature sweetening in potato by expressing Arabidopsis pyruvate decarboxylase gene and stress-inducible rd29A : a preliminary study. *Physiol. Mol. Biol. Plants* 17, 105–114. doi: 10.1007/s12298-011-0056-8

Preisler-Adams, S., Schonbuchner, I., Fiebig, B., Welling, B., Dworniczak, B., and Weber, B. H. (2006). Gross rearrangements in BRCA1 but not BRCA2 play a notable role in predisposition to breast and ovarian cancer in high-risk families of German origin. *Cancer Genet. Cytogenet.* 168, 44–49. doi: 10.1016/j.cancergencyto.2005.07.005

Puget, N., Gad, S., Perrin-Vidoz, L., Sinilnikova, O. M., Stoppa-Lyonnet, D., Lenoir, G. M., et al. (2002). Distinct BRCA1 rearrangements involving the BRCA1 pseudogene suggest the existence of a recombination hot spot. *Am. J. Hum. Genet.* 70, 858–865. doi: 10.1086/339434

Qian, Y., Mancini-DiNardo, D., Judkins, T., Cox, H. C., Brown, K., Elias, M., et al. (2017). Identification of pathogenic retrotransposon insertions in cancer predisposition genes. *Cancer Genet.* 21, 159–169. doi: 10.1016/j.cancergen.2017.08.002

Rhees, J., Arnold, M., and Boland, C. R. (2014). Inversion of exons 1-7 of the MSH2 gene is a frequent cause of unexplained Lynch syndrome in one local population. *Fam. Cancer* 13, 219–225. doi: 10.1007/s10689-013-9688-x

Riggs, E. R., Wain, K. E., Riethmaier, D., Savage, M., Smith-Packard, B., Kaminsky, E. B., et al. (2013). Towards a universal clinical genomics database: the 2012 international standards for cytogenomic arrays consortium meeting. *Hum. Mutat.* 34, 915–919. doi: 10.1002/humu.22306

Robson, M. E., Bradbury, A. R., Arun, B., Domchek, S. M., Ford, J. M., Hampel, H. L., et al. (2015). American society of clinical oncology policy statement update: genetic and genomic testing for cancer susceptibility. *J. Clin. Oncol.* 33, 3660–3667. doi: 10.1200/JCO.2015.63.0996

Rudnicka, H., Debniak, T., Cybulski, C., Huzarski, T., Gronwald, J., Lubinski, J., et al. (2013). Large BRCA1 and BRCA2 genomic rearrangements in Polish high-risk breast and ovarian cancer families. *Mol. Biol. Rep.* 40, 6619–6623. doi: 10.1007/s11033-013-2775-0

Sahnane, N., Bernasconi, B., Carnevali, I., Furlan, D., Viel, A., Sessa, F., et al. (2016). Disruption of the APC gene by t(5;7) translocation in a Turcot family. *Cancer Genet.* 209, 107–111. doi: 10.1016/j.cancergen.2015.12.003

Salo-Mullen, E. E., Lynn, P. B., Wang, L., Walsh, M., Gopalan, A., Shia, J., et al. (2018). Contiguous gene deletion of chromosome 2p16.3-*p*21 as a cause of Lynch syndrome. *Fam. Cancer* 17, 71–77. doi: 10.1007/s10689-017-0006-x

Sarkadi, B., Grolmusz, V. K., Butz, H., Kovesdi, A., Liko, I., Nyiro, G., et al. (2018). [Evolution of molecular genetic methods in the clinical diagnosis of hereditary endocrine tumour syndromes]. *Orv. Hetil.* 159, 285–292. doi: 10.1556/650.2018.31036

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* 3, 1101–1108. doi: 10.1038/nprot.2008.73

Schmitz, J., and Brosius, J. (2011). Exonization of transposed elements: a challenge and opportunity for evolution. *Biochimie* 93, 1928–1934. doi: 10.1016/j.biochi.2011.07.014

Sen, S. K., Han, K., Wang, J., Lee, J., Wang, H., Callinan, P. A., et al. (2006). Human genomic deletions mediated by recombination between Alu elements. *Am. J. Hum. Genet.* 79, 41–53. doi: 10.1086/504600

Senter, L., Clendenning, M., Sotamaa, K., Hampel, H., Green, J., Potter, J. D., et al. (2008). The clinical phenotype of Lynch syndrome due to germ-line PMS2 mutations. *Gastroenterology* 135, 419–428. doi: 10.1053/j.gastro.2008.04.026

Seong, M. W., Cho, S. I., Kim, K. H., Chung, I. Y., Kang, E., Lee, J. W., et al. (2014). A multi-institutional study of the prevalence of BRCA1 and BRCA2 large genomic rearrangements in familial breast cancer patients. *BMC Cancer* 14:645. doi: 10.1186/1471-2407-14-645

Serizawa, R. R., Ralfkiaer, U., Dahl, C., Lam, G. W., Hansen, A. B., Steven, K., et al. (2010). Custom-designed MLPA using multiple short synthetic probes: application to methylation analysis of five promoter CpG islands in tumor and

urine specimens from patients with bladder cancer. *J. Mol. Diagn.* 12, 402–408. doi: 10.2353/jmoldx.2010.090152

Sfeir, A., and Symington, L. S. (2015). Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway? *Trends Biochem. Sci.* 40, 701–714. doi: 10.1016/j.tibs.2015.08.006

Sinha, S., Villarreal, D., Shim, E. Y., and Lee, S. E. (2016). Risky business: microhomology-mediated end joining. *Mutat. Res.* 788, 17–24. doi: 10.1016/j.mrfmmm.2015.12.005

Sluiter, M. D., and van Rensburg, E. J. (2011). Large genomic rearrangements of the BRCA1 and BRCA2 genes: review of the literature and report of a novel BRCA1 mutation. *Breast Cancer Res. Treat.* 125, 325–349. doi: 10.1007/s10549-010-0817-z

Smith, M. J., Urquhart, J. E., Harkness, E. F., Miles, E. K., Bowers, N. L., Byers, H. J., et al. (2016). The contribution of whole gene deletions and large rearrangements to the mutation spectrum in inherited tumor predisposing syndromes. *Hum. Mutat.* 37, 250–256. doi: 10.1002/humu.22938

Teugels, E., De Brakeleer, S., Goelen, G., Lissens, W., Sermijn, E., and De Greve, J. (2005). De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes. *Hum. Mutat.* 26:284. doi: 10.1002/humu.9366

The BEDSG (2000). The exon 13 duplication in the BRCA1 gene is a founder mutation present in geographically diverse populations. The BRCA1 Exon 13 duplication screening group. *Am. J. Hum. Genet.* 67, 207–212.

Thomassen, M., Gerdes, A. M., Cruger, D., Jensen, P. K., and Kruse, T. A. (2006). Low frequency of large genomic rearrangements of BRCA1 and BRCA2 in western Denmark. *Cancer Genet. Cytogenet.* 168, 168–171. doi: 10.1016/j.cancergencyto.2005.12.016

Ticha, I., Kleibl, Z., Stribrna, J., Kotlas, J., Zimovjanova, M., Mateju, M., et al. (2010). Screening for genomic rearrangements in BRCA1 and BRCA2 genes in Czech high-risk breast/ovarian cancer patients: high proportion of population specific alterations in BRCA1 gene. *Breast Cancer Res. Treat.* 124, 337–347. doi: 10.1007/s10549-010-0745-y

Tournier, I., Paillerets, B. B., Sobol, H., Stoppa-Lyonnet, D., Lidereau, R., Barrois, M., et al. (2004). Significant contribution of germline BRCA2 rearrangements in male breast cancer families. *Cancer Res.* 64, 8143–8147. doi: 10.1158/0008-5472.CAN-04-2467

van der Klift, H., Wijnen, J., Wagner, A., Verkuilen, P., Tops, C., Otway, R., et al. (2005). Molecular characterization of the spectrum of genomic deletions in the mismatch repair genes MSH2, MLH1, MSH6, and PMS2 responsible for hereditary nonpolyposis colorectal cancer (HNPCC). *Genes Chromosomes Cancer* 44, 123–138. doi: 10.1002/gcc.20219

van der Klift, H. M., Tops, C. M., Hes, F. J., Devilee, P., and Wijnen, J. T. (2012). Insertion of an SVA element, a nonautonomous retrotransposon, in PMS2 intron 7 as a novel cause of Lynch syndrome. *Hum. Mutat.* 33, 1051–1055. doi: 10.1002/humu.22092

van der Luijt, R. B., Tops, C. M., Khan, P. M., van der Klift, H. M., Breukel, C., van Leeuwen-Cornelisse, I. S., et al. (1995). Molecular, cytogenetic, and phenotypic studies of a constitutional reciprocal translocation t(5;10)(q22;q25) responsible for familial adenomatous polyposis in a Dutch pedigree. *Genes Chromosomes Cancer* 13, 192–202. doi: 10.1002/gcc.2870130309

van Luttikhuizen, J. L., Bublitz, J., Schubert, S., Schmidt, G., Hofmann, W., Morlot, S., et al. (2020). From a variant of unknown significance to pathogenic: reclassification of a large novel duplication in BRCA2 by high-throughput sequencing. *Mol. Genet. Genomic Med.* 8:e1045. doi: 10.1002/mgg3.1045

Vasickova, P., Machackova, E., Lukesova, M., Damborsky, J., Horky, O., Pavlu, H., et al. (2007). High occurrence of BRCA1 intragenic rearrangements in hereditary breast and ovarian cancer syndrome in the Czech Republic. *BMC Med. Genet.* 8:32. doi: 10.1186/1471-2350-8-32

Vaughn, C. P., Robles, J., Swensen, J. J., Miller, C. E., Lyon, E., Mao, R., et al. (2010). Clinical analysis of PMS2: mutation detection and avoidance of pseudogenes. *Hum. Mutat.* 31, 588–593. doi: 10.1002/humu.21230

Wagner, A., van der Klift, H., Franken, P., Wijnen, J., Breukel, C., Bezrookove, V., et al. (2002). A 10-Mb paracentric inversion of chromosome arm 2p inactivates MSH2 and is responsible for hereditary nonpolyposis colorectal cancer in a North-American kindred. *Genes Chromosomes Cancer* 35, 49–57. doi: 10.1002/gcc.10094

Walsh, T., Casadei, S., Coats, K. H., Swisher, E., Stray, S. M., Higgins, J., et al. (2006). Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. *JAMA* 295, 1379–1388. doi: 10.1001/jama.295.12.1379

Weckselblatt, B., and Rudd, M. K. (2015). Human structural variation: mechanisms of chromosome rearrangements. *Trends Genet.* 31, 587–599. doi: 10.1016/j.tig.2015.05.010

Weitzel, J. N., Lagos, V. I., Herzog, J. S., Judkins, T., Hendrickson, B., Ho, J. S., et al. (2007). Evidence for common ancestral origin of a recurring BRCA1 genomic rearrangement identified in high-risk Hispanic families. *Cancer Epidemiol. Biomark. Prev.* 16, 1615–1620. doi: 10.1158/1055-9965.EPI-07-0198

Yamaguchi, K., Nagayama, S., Shimizu, E., Komura, M., Yamaguchi, R., Shibuya, T., et al. (2016). Reduced expression of APC-1B but not APC-1A by the deletion of promoter 1B is responsible for familial adenomatous polyposis. *Sci. Rep.* 6:26011. doi: 10.1038/srep26011

Zhang, F., Gu, W., Hurles, M. E., and Lupski, J. R. (2009). Copy number variation in human health, disease, and evolution. *Annu. Rev. Genomics Hum. Genet.* 10, 451–481. doi: 10.1146/annurev.genom.9.081307.164217

Zhao, B., Rothenberg, E., Ramsden, D. A., and Lieber, M. R. (2020). The molecular basis and disease relevance of non-homologous DNA end joining. *Nat. Rev. Mol. Cell Biol.* 21, 765–781. doi: 10.1038/s41580-020-00297-8

Zikan, M., Pohlreich, P., Stribrna, J., Kleibl, Z., and Cibula, D. (2008). Novel complex genomic rearrangement of the BRCA1 gene. *Mutat. Res.* 637, 205–208. doi: 10.1016/j.mrfmmm.2007.08.002

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 Genetics

ORIGINAL RESEARCH
published: 29 April 2021
doi: 10.3389/fgene.2021.635480

Systemic Screening for 22q11.2 Copy Number Variations in Hungarian Pediatric and Adult Patients With Congenital Heart Diseases Identified Rare Pathogenic Patterns in the Region

*Gloria Kafui Esi Zodanu[1], Mónika Oszlánczi[2], Kálmán Havasi[2], Anita Kalapos[2], Gergely Rácz[2], Márta Katona[3], Anikó Ujfalusi[4], Orsolya Nagy[4], Márta Széll[1] and Dóra Nagy[1]**

[1] Department of Medical Genetics, Faculty of Medicine, University of Szeged, Szeged, Hungary, [2] Second Department of Internal Medicine and Cardiology Centre, Faculty of Medicine, University of Szeged, Szeged, Hungary, [3] Department of Pediatrics, Faculty of Medicine, University of Szeged, Szeged, Hungary, [4] Division of Clinical Genetics, Department of Laboratory Medicine, Faculty of Medicine, University of Debrecen, Debrecen, Hungary

OPEN ACCESS

Edited by:
Katalin Komlosi, Medical Center University of Freiburg, Germany

Reviewed by:
Rincic Martina, University of Zagreb, Croatia
Társis Paiva Vieira, State University of Campinas, Brazil
Rafael Rosa, Federal University of Health Sciences of Porto Alegre, Brazil

***Correspondence:**
Dóra Nagy
nagydor@gmail.com

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *30 November 2020*
Accepted: *07 April 2021*
Published: *29 April 2021*

Citation:
Zodanu GKE, Oszlánczi M, Havasi K, Kalapos A, Rácz G, Katona M, Ujfalusi A, Nagy O, Széll M and Nagy D (2021) Systemic Screening for 22q11.2 Copy Number Variations in Hungarian Pediatric and Adult Patients With Congenital Heart Diseases Identified Rare Pathogenic Patterns in the Region. Front. Genet. 12:635480. doi: 10.3389/fgene.2021.635480

Congenital heart defects (CHD) are the most common developmental abnormalities, affecting approximately 0.9% of livebirths. Genetic factors, including copy number variations (CNVs), play an important role in their development. The most common CNVs are found on chromosome 22q11.2. The genomic instability of this region, caused by the eight low copy repeats (LCR A-H), may result in several recurrent and/or rare microdeletions and duplications, including the most common, ∼3 Mb large LCR A-D deletion (classical 22q.11.2 deletion syndrome). We aimed to screen 22q11.2 CNVs in a large Hungarian pediatric and adult CHD cohort, regardless of the type of their CHDs. All the enrolled participants were cardiologically diagnosed with non-syndromic CHDs. A combination of multiplex ligation-dependent probe amplification (MLPA), chromosomal microarray analysis and droplet digital PCR methods were used to comprehensively assess the detected 22q11.2 CNVs in 212 CHD-patients. Additionally, capillary sequencing was performed to detect variants in the *TBX1* gene, a cardinal gene located in 22q11.2. Pathogenic CNVs were detected in 5.2% (11/212), VUS in 0.9% and benign CNVs in 1.8% of the overall CHD cohort. In patients with tetralogy of Fallot the rate of pathogenic CNVs was 17% (5/30). Fifty-four percent of all CNVs were typical proximal deletions (LCR A-D). However, nested (LCR A-B) and central deletions (LCR C-D), proximal (LCR A-D) and distal duplications (LCR D-E, LCR D-H, LCR E-H, LCR F-H) and rare combinations of deletions and duplications were also identified. Segregation analysis detected familial occurrence in 18% (2/11) of the pathogenic variants. Based on in-depth clinical information, a detailed phenotype–genotype comparison was performed. No pathogenic variant was identified in the *TBX1* gene. Our findings confirmed the previously described large

phenotypic diversity in the 22q11.2 CNVs. MLPA proved to be a highly efficient genetic screening method for our CHD-cohort. Our results highlight the necessity for large-scale genetic screening of CHD-patients and the importance of early genetic diagnosis in their clinical management.

Keywords: 22q11.2 deletion syndrome, *TBX1* gene, multiplex ligation-dependent probe amplification, copy number variations, droplet digital PCR, syndromic and non-syndromic congenital heart defects, chromosomal microarray analysis

INTRODUCTION

Congenital heart defects (CHDs) are the most common congenital developmental defects and affect approximately 0.9% of livebirths (van der Linde et al., 2011). Thirty to forty percent of CHDs are syndrome-associated and are caused by copy number variants (CNVs) or a mutation in a single gene. The most common human CNVs affect chromosomal region 22q11.2 (Fahed et al., 2013; Digilio and Marino, 2016). Proximal microdeletions of 1.5–3 Mb in this chromosomal region typically include the sequence between low copy repeat regions A and D (LCR A-D, LCR A-B) and may lead to the classical phenotype of 22q11.2 deletion syndrome, also known as DiGeorge syndrome. Central (LCR B-D, LCR C-D) or distal deletions (LCR C-H) may cause other, variable phenotypes (Burnside, 2015; Kruszka et al., 2017). Duplications have also been identified in this chromosomal region and are associated with even more significant phenotypic variability than deletions (Wentzel et al., 2008).

DiGeorge syndrome (also known as velocardiofacial syndrome and conotruncal anomaly face syndrome) is mostly characterized by CHD, thymus hypoplasia, immunodeficiency and skeletal, gastrointestinal and urogenital defects as well as by developmental delay, learning difficulties, susceptibility to neuropsychiatric disorders and, in some cases, by mild to moderate intellectual disability. 22q11.2 CNVs have reduced penetrance and incomplete expression and may be detected in asymptomatic or mildly affected individuals; approximately 7–10% of the cases are familial (McDonald-McGinn et al., 1993-2020; Campbell et al., 2018).

The *T-box transcription factor 1 (TBX1)* gene is located within the proximal 22q11.2 region, encodes a transcription factor that plays an important role in early embryonic development and is hypothesized to contribute to 22q11.2 deletion phenotype as well as to non-syndromic CHDs. (Griffin et al., 2010; Heike et al., 2010; Guo et al., 2011).

Clinical diagnosis may be challenging and significantly delayed due to the large phenotypic spectrum resulting from 22q11.2 CNVs (from asymptomatic appearance to multiple defects) (van Engelen et al., 2010). Previous studies have drawn attention to the importance of routine screening for 22q11.2 CNVs in patients with congenital heart defects, especially with conotruncal anomalies (Wozniak et al., 2010; Huber et al., 2014; Goldmuntz, 2020).

Based on the low number of patients referred for 22q11.2 CNV analysis at our genetic department over the last decade, we hypothesized that some patients with 22q11.2 CNVs—especially in the adult population—may have remained undiagnosed. The aim of our study was therefore to test for 22q11.2 CNVs and *TBX1* gene variants for the pediatric and adult patients of the Southern-Hungarian CHD Registry, cardiologically diagnosed with non-syndromic CHDs, and to carry out genotype–phenotype comparison in positive cases based on in-depth clinical data.

MATERIALS AND METHODS

Overall, 212 unrelated patients (110 females, 112 males; mean age: 26.9 years; age range: 2 weeks to 74 years) previously cardiologically diagnosed with non-syndromic congenital heart defects were enrolled in the study at the University of Szeged between 2016 and 2019. The distribution of the patients with the different CHD types are presented in **Table 1**.

The DNA of 211 Hungarian individuals with no CHDs (confirmed with cardiological examination), and with no family history of CHD (144 females, 67 males, mean age: 37 years, age range: 8–73 years) was used as controls for the comparative analyses.

In positive cases, genetic testing was offered to all first-degree family members.

All investigations were performed according to the Helsinki Declaration 2008 and approved by the National Medical Research Council (No CHD-01/2016—IF-6299-8/2016) and the Local Ethical Committee of the University of Szeged (No 105/2016-SZTE). Participants/legal guardians/parents gave their informed consent to the study.

Sample Preparation and Multiplex Ligation-Dependent Probe Amplification (MLPA)

DNA was extracted from peripheral blood with the QIAamp DNA Blood Mini Kit (QIAGEN, Gödöllõ, Hungary).

To detect CNVs in the 22q11.2 locus, all patient samples were processed using the P250-B2 DiGeorge SALSA MLPA Probemix (IVD, MRC-Holland, Amsterdam) according to the manufacturer's instructions. The MLPA probe mix contained 48 probes, 29 of which are located in the 22q11.2 region (24 in the LCR A to H region and 5 in the Cat-Eye syndrome region) and 19 in regions 4q35, 8p23, and 9q34 (Kleefstra syndrome), 10p14 (DiGeorge syndrome 2) and 17p13 and 22q13 (Phelan-McDermid syndrome), deletions in the latter may result in phenotypical similarity to DGS. Amplicon fragment length analysis was performed on an ABI 3500 Genetic Analyzer

TABLE 1 | The distribution of different types of congenital heart defects among patients (*N* = 212).

Type of congenital heart defect	Number of patients (%)
Ventricular septal defect	**36 (16.9%)**
VSD alone	25
VSD + ASD + PDA	5
VSD + ASD	2
VSD + PDA	2
VSD + PS	2
Atrial septal defect	**35 (16.5%)**
ASD alone	27
ASD + PDA	3
ASD + VSD	3
ASD + PS	2
Congenital aorta stenosis	**31 (14.6%)**
AoS alone	17
AoS + bicuspid aortic valve	13
AoS + ASD	1
Fallot IV	**30 (14.1%)**
TGA	**21 (9.9%)**
Bicuspid aortic valve	**19 (8.9%)**
Coarctation of the aorta	**17 (8%)**
CoA alone	10
CoA + bicuspid aortic valve	3
CoA + VSD + PDA	4
Atrioventricular septal defect	5 (2.4%)
Anomalous pulmonary venous drainage	4 (2%)
TAPVD alone	2
TAPVD + PA + VSD	1
PPAVR	1
Pulmonary stenosis (congenital)	4 (2%)
Univentricular heart	3 (1.4%)
Hypoplastic left heart syndrome	2 (0.9%)
Pulmonary atresia	2 (0.9%)
Truncus arteriosus communis	1 (0.5%)
Double outlet right ventricle	1 (0.5%)
Ebstein anomaly	1 (0.5%)

AoS, congenital aorta stenosis; ASD, atrial septal defect; CoA, coarctation of the aorta; Fallot IV, Tetralogy of Fallot; PA, pulmonary atresia; PDA, patent ductus arteriosus; PPAVR, partial pulmonary anomalous venous return; PS, pulmonary stenosis; TGA, transposition of the great arteries; TAPVD, total anomalous pulmonary venous drainage; VSD, ventricular septal defect. The most frequent CHD groups in the cohort are indicated in bold.

(Thermo Fisher Scientific, Waltham, MA) and analyzed by Coffalyser.net software (MRC-Holland, Amsterdam).

Validation of Positive Cases: FISH, Chromosomal Microarray Analysis, ddPCR

MLPA was repeated for all samples in which CNVs were found. Deletions and duplications were confirmed with an independent method, including FISH (Vysis DiGeorge Region LSI N25 SO/ARSA SGn Probes, Abbott Molecular Inc., Des Plaines, IL, United States, and SureFISH 22q11.21 CRKL, Agilent Technologies, Cedar Creek, TX, United States), a supplementary MLPA kit (P372-SALSA MLPA Microdeletions 6, MRC-Holland, Amsterdam, Netherlands) or chromosomal microarray analysis (CMA, Affymetrix, CytoScan 750 K, Thermo Fisher Scientific, Waltham, MA, United States). CMA was performed as described by Nagy et al. (2019). In cases where one probe was deleted, the probe region was sequenced with bidirectional capillary sequencing to exclude MLPA-interfering SNPs in the sample DNA. These validation methods confirmed all positive MLPA results (i.e., no false positives).

A droplet digital PCR (ddPCR) method was designed for the confirmation of recurrent single-probe CNVs in the *TOP3B* gene from CHD patient samples. This method was also used to determine the frequency of *TOP3B* CNVs in the control cohort as well. The analysis was performed on the QX100 Droplet Digital PCR system (Bio-Rad Laboratories, Hercules, CA, United States), according to the manufacturer's instruction. Primers and probes were designed for *TOP3B* exon 7 and for the *PRDM15* gene as reference region on chromosome 21 (**Supplementary Material**). *TOP3B* CNVs found in the controls with ddPCR were confirmed with MLPA.

Sequencing of the *TBX1* Gene

Bidirectional capillary sequencing of *TBX1* coding regions was performed for all patient samples with an ABI 3500 Genetic Analyzer. The primers used are listed in **Supplementary Material**. The non-synonymous variants (all located in exon 9 of *TBX1* gene) were tested in the control cohort as well.

CNV and Variant Interpretation

Identified CNVs and single nucleotide variants (SNVs) were classified according to the standards and guidelines of the American College of Medical Genetics (Richards et al., 2015; Riggs et al., 2020). The following websites and databases were used for CNV interpretation: Database of Chromosomal Imbalance, Phenotype of Humans using Ensemble Resources (DECIPHER, Firth et al., 2009), Database of Genomic Variation (DGV, MacDonald et al., 2013), PubMed and GeneReviews (McDonald-McGinn et al., 1993-2020). For SNV interpretation, VarSome (Kopanos et al., 2011), ClinVar (Landrum et al., 2020), and Genome Aggregation Database (GnomAD, Karczewski et al., 2020) databases were used.

Statistical Analysis

GraphPad Prism (GraphPad Software, San Diego, California, United States), version 4.00 for Windows, was used for statistical analysis. The frequency of *TOP3B* CNVs and *TBX1* variants in the patient cohort was compared with the frequency in the control cohort and also with the frequency in the global dataset of GnomAD using the Fisher exact test and χ^2-test. $P < 0.05$ was considered to be statistically significant.

RESULTS

Distribution of CHD Types in Patients

In the CHD cohort, the four most common CHD types were ventricular septal defect (VSD), atrial septal defect (ASD), congenital aorta stenosis (AoS) and tetralogy of Fallot

(TOF) (**Table 1**). In 81% of the patients, only one cardiac entity was diagnosed; whereas, in 19% of the cases, two or more CHDs occurred together. The distribution of the different CHDs among the South-Hungarian Registry patients corresponded well with the frequency described in the literature (van der Linde et al., 2011).

Distribution of Positive MLPA Results and Classification of the Detected CNVs

Overall, 17 cases of 212 patients (8%) diagnosed prior with non-syndromic CHD were yielded positive MLPA result, and after evaluation 11 of these copy number changes (5.2%) were interpreted as pathogenic variant, two as variant of unknown significance (VUS, 0.9%) and four as benign (1.8%) (**Figure 1** and **Supplementary Material**). The most frequent CNVs of the positive MLPA results were microdeletions (8/17); however, microduplications (7/17) and a combination of deletions and duplications (2/17) were also observed.

Among pathogenic CNVs 7 microdeletions, 2 duplications and 1 combination of a deletion and a duplication was detected, while among the VUS one duplication and one combined CNV and among the benign variants one deletion and three duplications.

Pathogenic results were observed most frequently in the TOF group: in 17% of all TOF patients, followed by the group of bicuspid aortic valve with 10% (**Figure 1** and **Table 2**).

Based on the interpretation guidelines (Riggs et al., 2020), 11 CNVs were interpreted as pathogenic:

6 typical deletions of LCR A-D, one proximal nested deletion of LCR A-B, two duplications of LCR A-D, one combination of the proximal deletion of LCR A-D with a duplication of LCR E-H and one duplication of LCR D-H. Two further CNVs (one combination of a central deletion of LCR C-D with duplication of LCR D-E and one duplication of LCR F-H) were classified as VUS (**Figure 1** and **Table 3**). Four CNVs (three 268 kbp duplications and one 278 kbp deletion) were detected in the *TOP3B* gene and resulted from one probe alteration in the MLPA reaction. These CNVs were confirmed with chromosomal microarray analysis (**Supplementary Material**). Considering the relatively high proportion of *TOP3B* CNVs in our patient cohort (overall 4/212, 1.9%: deletion in 1/212, 0.5% and duplication in 3/212, 1.4%), we decided to perform an independent analysis with ddPCR to determine the frequency of *TOP3B* CNVs in the healthy controls. The *TOP3B* deletion was detected in one control sample (0.5%) and a duplication in four control samples (1.9%); i.e., CNVs were identified in 2.4% of the controls (5/211). The difference in the CNV

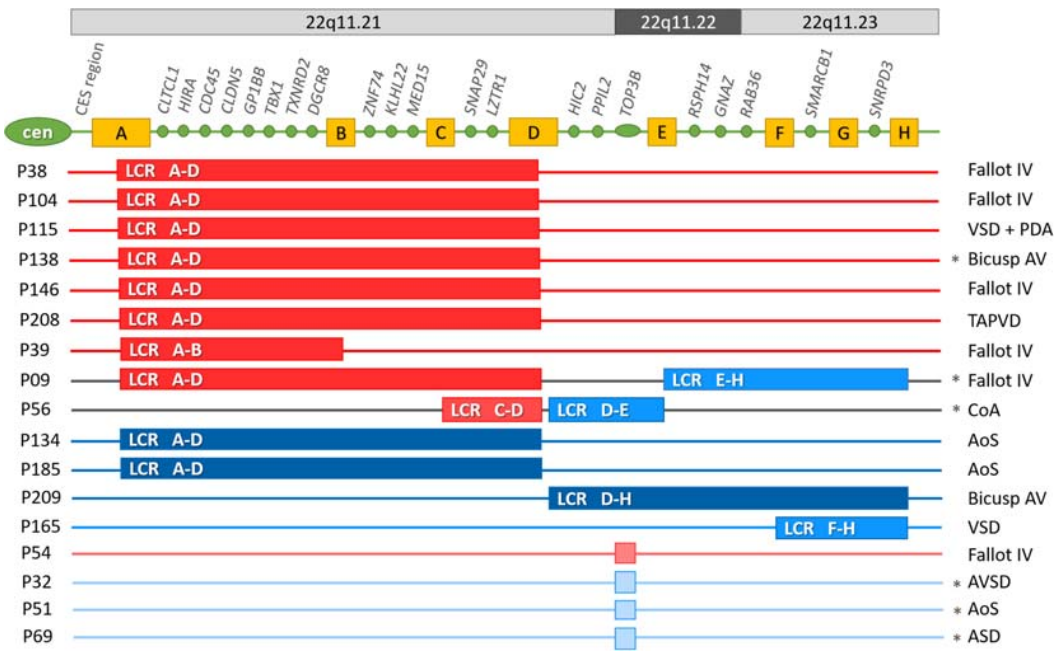


FIGURE 1 | Pathogenic variants, variants of uncertain significance and benign 22q11.2 copy number variations in the CHD cohort. Cen, centromere; CES, Cat eye syndrome region; A-H yellow boxes: low copy repeat regions in locus 22q11.2. Genes indicated between the LCR regions, are the ones that have corresponding probes in the P250-B2 DiGeorge SALSA MLPA Probemix; LCR: low copy repeat region. Dark red: pathogenic microdeletion; Dark blue: pathogenic microduplication; Middle red: microdeletion of uncertain significance; Middle blue: microduplication of uncertain significance; Light red: benign microdeletion; Light blue: benign microduplication. LCR A-D: the typical ∼2.5–3 Mb microdeletion in region 22q11.21; LCR A-B: ∼1.5 Mb proximal microdeletion in region 22q11.21; LCR C-D: ∼0.5 Mb central microdeletion in region 22q11.21; LCR D-E: ∼1.2 Mb central microduplication in region 22q11.21q11.22; LCR D-H: ∼3.1–3.5 Mb central-distal microduplication in region 22q11.21q23; LCR E-H: ∼1.55–2 Mb distal microduplication in region 22q11.22q11.23; LCR F-H: ∼1–1.2 Mb distal microduplication in region 22q11.23; Asterix denotes familial CNVs. AoS, congenital aorta stenosis; AVSD, atrioventricular septal defect; bicuspid AV: bicuspid aortic valve; CoA, coarctation of the aorta; Fallot IV, tetralogy of Fallot; PDA, patent ductus arteriosus; TAPVD, total anomalous pulmonary venous drainage; VSD, ventricular septal defect.

TABLE 2 | The distribution of the pathogenic and VUS 22q11 CNVs in the different CHD groups.

Type of CHD	Number of pathogenic CNVs or VUS/Number of patients with CHD (%)	Del	Dupl	Del + Dupl
Fallot IV	**5 path/30 (17%)**	**4**	0	**1**
Bicuspid aortic valve	2 path/19 (10%)	**1**	**1**	0
AoS	2 path/31 (6.5%)	0	**2**	0
CoA	1 VUS/17 (6%)	0	0	**1**
VSD	1 path/36 (2.7%)	**1**		0
	1 VUS/36 (2.7%)		**1**	
TAPVD + VSD + PA	1 path/4 (25%)*	**1**	0	0
Total	**11 path/212 (5.2%)**	**7**	**3**	**1**
	2 VUS/212 (0.9%)	**0**	**1**	**1**

*VUS, variant of uncertain significance; CHD, congenital heart defect; CoA, coarctation of the aorta; AoS, congenital aorta stenosis Fallot IV, tetralogy of Fallot; TAPVD, total anomalous pulmonary venous drainage; VSD, ventricular septal defect; PA, pulmonary atresia; DEL, 22q11 microdeletion; DUPL, 22q11 microduplication. *This proportion is biased, since this patient could be also classified in VSD or PA group.*

frequency between patients and controls was not significant (p = 0.751). Thus, we ultimately classified *TOP3B* CNVs as rare benign variants, which are more frequent in the Hungarian population than in the global database (frequency in DECIPHER: 0.36%).

Familial Segregation

It was possible to perform segregation analysis for 14 of the 17 positive cases. Six cases proved to be familial (**Figure 1**), two of these were for patients with pathogenic CNVs, one for a patient with VUS and three for patients with benign *TOP3B* variants. In addition to the *TOP3B* microduplication, proband P32 also had a 201-bp microduplication on chromosome 17p13.3, which included the *YWHAE* gene (**Supplementary Material**). Both chromosome imbalances were inherited from an asymptomatic parent. Two asymptomatic siblings also carried the *YWHAE* microduplication but without the *TOP3B* CNV. Therefore, the *YWHAE* CNV was interpreted as a rare benign variant. In case of the other *TOP3B* microduplications (P51 and P69) one healthy parent carried also the variant. In the patient with the *TOP3B* deletion, segregation analysis was not performed, the proband had two healthy children. The individuals with *TOP3B* CNVs were excluded from the genotype–phenotype comparison based on these results, the high frequency of *TOP3B* CNV in controls and the fact that these patients displayed no other malformations or comorbidities in addition to CHD.

The segregation analysis detected familial occurrence for 18% (2/11) of the pathogenic CNVs and 50% (1/2) of VUS. In these three familial cases (P138, P09 and P56), the proband's mother carried the same chromosome imbalance. The phenotypes of the mothers were the same severity (P138) or milder (P09, P56) than the probands'. The suspicion of an underlying 22q11.2 CNV prior to the genetic testing was not raised at any of the affected family members.

For proband P09 and for the proband's mother, the typical ∼2.5–3 Mb 22q11.2 microdeletion was combined with a distal ∼1.5–2 Mb 22q11.2 microduplication. The segregation analysis of the family showed that the maternal grandmother and one sibling of the mother carried only the duplication with no cardiological symptom or developmental malformation, although they had been diagnosed with mild anxiety disorder and depression. The deletion occurred most probably *de novo* in the mother and was transferred to the child.

The mother of proband P56 had only bicuspid aortic valve (clinically diagnosed only after the genetic diagnosis) and similar facial features as the proband without torticollis or severe scoliosis.

The clinical features of patients and family members are shown in **Table 3**.

Genotype-Phenotype Comparison

The probands' age at the genetic diagnosis with pathogenic or VUS 22q11.2 CNVs ranged from 2 months to 52 years (median age: 21 years). Three patients out of 13 were diagnosed in childhood, one child in the first year of life. The two oldest patients and the affected family members were born before the molecular diagnostic era. No correlation could be observed between the severity of the phenotype and the age at the diagnosis.

The prevalence of common clinical features for different CNVs is comparable to previously reported prevalence data in the literature (**Table 4**). Patients presented with more marked phenotypic features for 22q11.2 microdeletions than with microduplications in the same region. In addition to the CHDs, the typical microdeletions of LCR A-D—with or without accompanying CNVs—resulted in the classical phenotype of 22q11.2 deletion syndrome. The co-occurring duplication in proband P09 and mother has not modified their phenotype significantly compared to other LCR A-D microdeletion phenotypes. In proband P56 and his mother with the combination of central deletion and distal duplication, the phenotype differed completely from that of 22q11.2 deletion syndrome (except the CHD) and the impact of the duplication could not be determined precisely (**Table 3**).

Among patients with deletions, Fallot tetralogy was the most common CHD. Among patients with duplications, congenital aorta stenosis, coarctation of the aorta and bicuspid aortic valve were the most common CHD types (**Table 2**). These three entities may be considered on the spectrum for one disease.

CHDs were overrepresented in our 22q11.2 CNV patients and their affected family members compared to data in the literature (94% vs. 74%), which may be the result of the patient enrollment criteria. Neuropsychiatric disorders were underrepresented among our patients (19% vs. 60%). Other characteristics (facial features, velopharyngeal insufficiency, immunodeficiency, hypocalcemia, skeletal anomalies, developmental delay, and learning difficulties) had a distribution in our cohort similar to that described in the literature (**Table 4**). The presence of immunodeficiency was deduced from the recurrence of respiratory and ear infections occurring mostly in childhood. Based on regular laboratory check-ups, the average absolute

TABLE 3 | Clinical features of probands and parents with pathogenic and VUS MLPA results.

Proband no./Age*	CNV Extension (size)	Classification**	Type of CHD	Extracardiac manifestations	DD/ID	Facial features	Classical phenotype***	Hypocalcemia	Miscellaneous	Familial inheritance or *de novo* mutation
P38/ Adulthood	**DEL LCR A-D** in 22q11.21 (~2.5-3 Mb)	Path	TOF	Recurrent bronchitis and otitis media, tonsillectomy, vesicoureteral reflux, renal cyst, velopharyngeal insufficiency, nasal speech, early teeth lost, small stature, kyphoscoliosis, block vertebrae, lower leg cramps, dyslexia, anxiety disorder, microcephaly, juvenile cataract, autoimmune hypothyroidism	DD/mild ID	Narrow face, micrognathia, low-set ears, narrow, small palpebral fissures, hypertelorism, hypoplastic alae nasi, pointed ear tips, thin lips	Yes	Yes	Obstipation, GOR, feeding difficulties in childhood, hypomagnesemia	*De novo*
P104/ Adulthood	**DEL LCR A-D** in 22q11.21 (~2.5–3 Mb)	Path	TOF	Arrythmia (radio frequent ablation), recurrent otitis media in childhood (Grommet tubes, adenotonsillectomy), inguinal hernia, scoliosis, mild kyphosis, thorax asymmetry, narrow shoulders, recurrent urinary infections, nephrolithiasis, nasal speech, learning difficulties.	DD/low normal IQ	Narrow, long face, narrow palpebral fissures, deep-set eyes, marked hypertelorism, large ear lobes, thin small lip, hypoplastic alae nasi, malar flattening	Yes	Yes	Lumbago, pulmonary embolism. Deceased postoperative in nosocomial infection before the genetic diagnosis	*De novo*
P115/ Adulthood	**DEL LCR A-D** in 22q11.21 (~2.5–3 Mb)	Path	VSD, PDA	Right main bronchus stenosis, congenital lacrimal duct stenosis, many recurrent upper and lower airway infections until puberty, severe scoliosis, hernia diaphragm + severe GOR (fundoplication), inguinal hernia (operated), palatoschisis, velopharyngeal insufficiency, thorax and hip deformity, learning difficulties, nasal speech, episodic hand tremor and foot paresthesia, small stature, microcephaly	DD/low normal IQ	Narrow, long face low-set ears, narrow, small palpebral fissures, hypertelorism, hypoplastic alae nasi, pointed ear tips, thin small lips, malar flattening, mild facial asymmetry	Yes	–	Nasogastric tube feeding in infancy	*De novo*
P138/ Childhood	**DEL LCR A-D** in 22q11.21 (~2.5–3 Mb)	Path	Bicuspid aortic valve	Recurrent lower and upper airway infections and otitis media, nasal speech, hypermetropy, astigmia, no developmental delay, normal kindergarten, and preschool	No/No	Long face, small mouth, straight nose, narrow eyelids, hypertelorism, pointed ear tip, fleshy ear lobes	Yes	Yes	Severe obesity (due to diet failure), secondary hypertonia	Maternally inherited: mother has similar outer appearance, umbilical hernia, vestibular neuronitis, impaired hearing, but no CHD
P146/ Adulthood	**DEL LCR A-D** in 22q11.21 (~2.5–3 Mb)	Path	TOF	Recurrent respiratory infection and otitis media, mastoiditis, cryptorchism (orchidopexy), severe scoliosis, thorax asymmetry (small right scapula), nasal speech, neuropsychiatric problems, learning difficulties, special school	Speech delay/mild ID	Narrow, long face, straight nose, low-set ears, narrow palpebral fissures	Yes	Yes	Brain MRI and abdominal ultrasound: normal	*De novo*
P208/ Childhood	**DEL LCR A-D** in 22q11.21 (~2.5–3 Mb)	Path	TPAVD, PA, VSD	Thymus aplasia, many respiratory infections in small childhood, prolonged Candidiasis, small stature	DD/mild ID	Narrow face, micrognathia, low-set ears, narrow, small palpebral fissures	Yes	Yes	Epidermal skin problems	*De novo*

(Continued)

TABLE 3 | Continued

Proband no./Age*	CNV Extension (size)	Classification**	Type of CHD	Extracardiac manifestations	DD/ID	Facial features	Classical phenotype***	Hypocalcemia	Miscellaneous	Familial inheritance or *de novo* mutation
P39/ Adulthood	**DEL LCR A-B** in 22q11.21 (~2.5–3 Mb)	Path	TOF	Recurrent respiratory infections, learning difficulties, special school, neuropsychiatric problems	DD/low normal IQ	Narrow face, low-set ears, narrow, small palpebral fissures	Yes	–	–	ND
P09/ Childhood	**DEL LCR A-D** in 22q11.21 (~2.5–3 Mb) + **DUPL LCR E-H** 22q11.22q11.23 (~1.55–2 Mb)	Path + VUS	TOF	Congenital laryngeal stenosis, thymus aplasia, respiratory infections (postoperative as well), pes calcaneovalgus	?/?	Small mouth, pointed ear tips, hypoplastic alae nasi, low-set ears	Yes	No	Small for gestational age, transient nasogastric tube feeding	Maternally inherited. mother: TOF, severe scoliosis, nasal speech, classical DGS phenotype, low normal IQ, anxiety disorder
P56/ Adulthood	**DEL LCR C-D** in 22q11.21 (~0.5 Mb) + **DUPL LCR D-E** 22q11.21q11.22 (~1.2 Mb)	VUS + VUS	CoA	Torticollis, scoliosis, nasal speech, socially withdrawn, studies in higher education	No/No	Facial asymmetry, pointed ear lobes, small philtrum, low-set ears, triangular chin, retrognathia	No	No	–	Maternally inherited: mother has bicuspid aortic valve, similar facial features without asymmetry and torticollis
P134/ Adulthood	**DUPL LCR A-D** in 22q11.21 (~2.5–3 Mb)	Path	AoS	None	No/No	No	No	No	Atopy, asthma	*De novo*
P185/ Adulthood	**DUPL LCR A-D** in 22q11.21 (~2.5–3 Mb)	Path	AoS	None	No/No	No	No	No	–	ND
P209/ Adulthood	**DUPL LCR D-H** 22q11.21q11.23 (~3.1–3.5 Mb)	Path	Bicuspid aortic valve	Horseshoe kidney, pyeloureteral stenosis, Ewing-sarcoma in childhood, frequent tonsillitis (tonsillectomy), primary amenorrhea, special school	Speech delay/mild ID	Hypertelorism, divergent strabismus, prominent long mandible, uvula elongata	No	No	Obesity. Twin sibling died of pulmonary atresia after birth	No CNV in father, mother not tested
P165/ Adulthood	**DUPL LCR F-H** in 22q11.23 (~1–1.2 Mb)	VUS	VSD	Myopia, bilateral inguinal hernia, truncal obesity, learning difficulties	No/No	Micrognathia	No	No	Preterm birth, normal catch-up development, bronchial asthma	*De novo*

CHD, congenital heart disease; CNV, copy number variation; LCR, low copy repeat region; TOF, Tetralogy of Fallot; AoS, congenital stenosis of the aorta; PA, pulmonal atresia; VSD, ventricular septal defect; TPAVD, Total anomalous pulmonary venous return; CHD, congenital heart disease; CoA, coarctation of the aorta; AVSD, atrioventricular septal defect. ND, not done; DD, developmental delay; ID, intellectual disability, GOR, gastro-esophageal reflux; Path, pathological; VUS, variant of uncertain significance.

**Denotes the life period when the genetic diagnosis was set up. **Denotes the classification of CNVs based on the joint consensus of ACMG and ClinGen Guidelines (Riggs et al., 2020). ***Classical phenotype of 22q11.2 deletion syndrome, previously described as DiGeorge syndrome.*

TABLE 4 | The prevalence of the common clinical features in all probands and family members carrying pathogenic CNVs and VUS in the 22q11.2 region compared to the prevalence in the literature.

Symptoms	Classical LCR A-D deletion	All other deletions*	All duplications alone	All CNVs of this study	Prevalence in the literature**
CHD	8/9 (89%)	3/3 (100%)	4/4 (100%)	**15/16 (94%)**	64–74%
Facial dysmorphia	9/9 (100%)	3/3 (100%)	2/4 (50%)	**14/16 (87.5%)**	46–88%
Classical facial features in 22q11.2 deletion	9/9 (100%)	1/3 (33%)	0/4	**10/16 (62.5%)**	ND
Velopharyngeal insufficiency	7/9 (78%)	1/3 (33%)	0/4	**8/16 (50%)**	55–69%
Immunodeficiency, recurrent infections	8/9 (89%)	1/3 (33%)	1/4 (25%)	**10/16 (62.5%)**	50–77%
Skeletal anomalies	6/9 (67%)	1/3 (33%)	0/4	**7/16 (44%)**	15–50%
Other anomalies***	7/9 (78%)	0/3	2/4 (50%)	**9/16 (56%)**	67–81%
Developmental delay (motor ± speech) in childhood	6/9 (67%)	1/3 (33%)	1/4 (25%)	**8/16 (50%)**	70–90%
Intellectual disability	4/9 (44%)	0/3	1/4 (25%)	**5/16 (31%)**	28–31%
Learning difficulties	6/9 (67%)	1/3 (33%)	1/4 (25%)	**8/16 (50%)**	66–93%
Neuropsychiatric problems	3/9 (33%)	0/3	0/4	**3/16 (19%)**	60–73%
Hypocalcemia	5/9 (55.5%)	0/3	0/4	**5/16 (31%)**	17–60%

CHD, congenital heart defect; ID, intellectual disability.
In this category LCR A-B nested deletion and LCR C-D deletion combined with LCR D-E duplication are included; ** Data for 22q11.2 deletions described by McDonald-McGinn et al. (1993-2020), Burnside (2015); Campbell et al. (2018), and Niarchou et al. (2019) are presented; * Including clinically diagnosed thymus aplasia, respiratory, gastrointestinal and/or urogenital abnormalities; ND: not determined. Velopharyngeal insufficiency, nasal speech could not be assessed in the two youngest probands. Intellectual disability, learning difficulties could be only assessed in individuals in school age or older. Psychiatric disorders are taken into account only if they were clinically diagnosed.*

lymphocyte count was in the lower normal range (2.13 G/L, normal range: 1.5–3.2 G/L); whereas the average relative lymphocyte count was below normal (23.8%, normal range: 27–34%). Flow-cytometry and serum immunoglobulin levels were not measured regularly. The immune status and infections of 22q11.2 CNV patients were not strictly controlled before the genetic diagnosis. Before genetic diagnosis, proband P104 died of a fulminant postoperative infection (**Table 3**).

Hypocalcemia (average serum calcium level: 1.84 mmol/l, normal range: 2.2–2.55 mmol/l) was often present in patients with the typical 22q11.2 microdeletions—with or without clinical symptoms. However, hypocalcemia was not considered relevant for therapy before genetic diagnosis. Severe hypomagnesemia was also detected in one 22q11.2 microdeletion patient. The thrombocyte count was in the low normal range with an average of 156×10^9/l. Thyroid and parathyroid hormone levels and vitamin D levels were not measured in these patients before genetic diagnosis.

These laboratory abnormalities could not be consistently identified for patients with CNVs other than the typical microdeletion.

Results of the *TBX1* Gene Sequencing

No apparently pathogenic variant was detected in the *TBX1* gene. For CHD patients, three missense variants were found in exon 9: c.1189A>A; p.Asn397His with a 21% minor allele frequency (MAF), c.1049G>A; p.Gly350Asp with 0.48% MAF and c.1341_1342insCCGCACGCGCAT; p.Ala450_His453dup with 0.24% MAF (**Table 5**). The frequency of the p.Asn397His variant was also 21% for the controls. The two less frequent variants were also detected in one of the proband's healthy parents and are listed in the Hungarian or the global database with very low frequencies (**Table 5**). Of the 10 probands and mothers with proximal 22q11.2 microdeletions encompassing the *TBX1* gene, two (20%) carried the common p.Asn397His variant in hemizygous form (P09 and P138). Proband P09 exhibited a severe phenotype; whereas proband P138 presented only milder symptoms. The two rare variants were not detected in any of the microdeletion patients. Based on the allele frequencies, ACMG criteria and segregation analyses, all three variants were ultimately classified as benign.

DISCUSSION

This was the first systemic, large-scale genetic screening study of Hungarian CHD patients. All patients with cardiologically verified CHDs were enrolled in the study without further selection. Although the enrolled patients were cardiologically diagnosed with non-syndromic CHDs prior to this study, 13 were found to be syndromic after the genetic screening.

We observed a higher median age (21 years) and a similar or wider age range (0.17–52 years) at the genetic diagnosis in our cohort as compared to previously described cohorts (median age: 17.3 years, range: 0.1–59.4 years in Canadian patients; median age: 2.9 years, range: 0–17.6 years in American patients) (Palmer et al., 2018). This difference may partly be explained by the fact, that the 22q11.2 duplication patients with more variable phenotypes were also included in the present study, whereas only 22q11.2 deletions were analyzed by Palmer and colleagues.

The frequency of CHDs was representative and corresponded to the frequency described in large epidemiological studies (van der Linde et al., 2011).

All types of CNVs in the 22q11.2 chromosomal region were present in 8% of the CHD cohort, while pathogenic CNVs in 5.2%, VUS in 0.9% and benign CNVs in 1.8%. Our patients

TABLE 5 | Minor allele frequencies of *TBX1* variants in the CHD cohort and control cohort compared to the allele frequency in the global database.

Variant	MAF in CHD	MAF in controls	*p*	MAF in GnomAD	*p*	*In silico* prediction*
c.1189A>A; p.Asn397His (rs72646967)	21%	21%	*0.809*	23.19%	*0.3583*	Benign
c.1049G>A; p.Gly350Asp (rs781731042)	0.48%	0.95%	*0.686*	0.0402%	*0.0138*	Benign
c.1341_1342insCCGCACGCGCAT; p.Ala450_His453dup (rs1341195668)	0.24%	0%	*0.498*	0.00325%	*0.0267*	VUS

In TBX1 locus CHD cohort contained overall 418 alleles, control cohort 422 alleles.
MAF, minor allele frequency; CHD, congenital heart defect.
*GnomAD, Genome Aggregation Database (Karczewski et al., 2020). TBX1 transcript number: NM_080647. *In silico variant prediction was performed with VarSome, based on the ACMG criteria (Kopanos et al., 2011; Richards et al., 2015).*

presented pathogenic 22q11.2 CNVs more often compared to other CHD cohorts, such as 1.27% in Brazilian, 2.8% in Cameroonian and 2.9% in Chinese population (Huber et al., 2014; Wonkam et al., 2017; Li et al., 2019). However, this difference may be also explained by the fact that most of these studies focused on the detection of 22q11.2 deletion but not on duplications.

For tetralogy of Fallot, the proportion of pathogenic CNVs was significantly higher, 17% in our cohort, which is in agreement with the fact that 22q11.2 CNVs are common in conotruncal heart defects (McDonald-McGinn et al., 1993-2020). 22q11.2 deletion can be detected in approximately 20% of all conotruncal heart defects (Wozniak et al., 2010), within this category its prevalence can be as high as ∼50% in interrupted aortic arch type B, ∼35% in truncus arteriosus or 10–25% in tetralogy of Fallot (Goldmuntz, 2020). All these suggest an absolute indication for 22q11.2 CNV analysis in these CHD groups, especially when co-occurring with at least one extracardiac manifestation or dysmorphic traits (Wozniak et al., 2010).

Congenital bicuspid aortic valve is common (0.5–2%) and, without complication of stenosis, regurgitation or dissection, considered a largely benign congenital heart defect (Li et al., 2017). However, based on our results, it should not be ignored in genetic testing.

The most common (64%) pathogenic CNV among our patients was the typical microdeletion of the LCR A-D region on chromosome 22q11.2, which is in agreement with the literature (Du et al., 2020). The frequency of deletions decreased toward the LCR F-H region, which was reflected in our results as well, since nested and central deletions were rare, and distal deletions were not detected. Proximal and distal duplication as well as two combined CNVs were also identified. Although most patients with the typical LCR A-D deletion showed the majority of the characteristic features of 22q11.2 deletion syndrome (velopharyngeal insufficiency, skeletal malformation, gastrointestinal and nephrological anomalies, hypocalcemia, frequent infections due to immunodeficiency and common facial features), these symptoms were present less frequently (**Table 4**) with the non-typical deletions and the duplications, as expected (Burnside, 2015; Du et al., 2020). In addition to the presence of CHDs, no typical common characteristics could be found for these patients. This may be due to the low number of patients with single CNVs in our cohort or to the even wider phenotypic spectrum of these CNVs, e.g., in the case of 22q11.2 duplications (Yu et al., 2019). Therefore, several individuals with only mild symptoms or no detectable malformation or dysmorphia may remain undetected.

Hypocalcemia, hypomagnesemia, lymphocytopenia, thrombocytopenia, and abnormalities of the thyroid, parathyroid hormone or vitamin D levels may remain undiscovered in 22q11.2 deletion patients, especially without genetic diagnosis. However, these conditions may significantly contribute to co-morbidities, such as increased susceptibility to infections, bleeding diathesis and heightened prevalence of autoimmune disorders, and, thus, should be considered for treatment (Lambert et al., 2018; Goldmuntz, 2020; Legitimo et al., 2020).

Neuropsychiatric disorders (attention deficit hyperactivity disorder, autism spectrum disorder, schizophrenia, anxiety symptoms and sleep disturbances) are frequent in patients with 22q11.2 CNVs (Brzustowicz and Bassett, 2012; Moulding et al., 2020). Complex presentation of three or more psychiatric traits may occur in 73% of patients with 22q11.2 CNVs (Niarchou et al., 2014; Chawner et al., 2019; Niarchou et al., 2019). However, these were markedly underrepresented (19%) in our patient cohort (**Table 4**), and this is most probably due to the lack of awareness and screening rather than to their absence. This result further emphasizes the importance of multidisciplinary management of patients with 22q11.2 CNVs.

The large phenotypic variability of 22q11.2 microdeletions has recently been the focus of much research but is still not yet fully understood. The haploinsufficiency of the coding genes, including *TBX1, DGCR8, CRKL* among others, alone does not seem to account for the highly variable phenotypes and incomplete penetrance of affected individuals (Du et al., 2020). Some recent studies have investigated the role of possible genetic and epigenetic factors which contribute to the diversity of phenotypes associated with 22q11.2 deletions (Brzustowicz and Bassett, 2012; Bertini et al., 2017; Du et al., 2020). Breakpoint analysis of the LCR A-D region showed that small variations in the deletion size within this region have no significant role on phenotypic variability (Bertini et al., 2017). Pathogenic sequential variations in the remaining single copy of the genes (with an emphasis on *TBX1* gene) encompassed in the deleted region, were not yet revealed by previous investigation (Brzustowicz and Bassett, 2012). And this was further supported by our *TBX1* sequencing results, since no pathogenic *TBX1* variant

was detected in our patients with 22q11.2 CNVs or in the overall CHD-cohort. Thus, pathogenic *TBX1* mutations may be causal most probably only in a small fraction of CHD patients and 22q11.2 CNV patients, if at all. It was also hypothesized that sequential variations elsewhere in the genome (for example, *de novo* mutations in histone modifying genes) may collectively contribute to this diversity (Zaidi et al., 2013). Bertini and colleagues have investigated additional rare and common CNVs in typical 22q11.2 patients (2017). According to their results, these additional CNVs often contain miRNA genes or mitochondrial genes, which may interact with 22q11.2 deletion and lead to metabolic and energetic problems rather than a decreased dosage of morphogenetic genes. The *DGCR8* gene is located within the typical 22q11.2 region and plays a crucial role in miRNA biosynthesis and, in combination with other CNV-miRNAs, may orchestrate highly variable phenotypic outcomes (Bertini et al., 2017). Previously, these additional CNVs, miRNAs have been investigated exclusively for 22q11.2 microdeletion patients, but not for duplication patients. Studying these duplication patients may further refine the diversity.

The family segregation study proved to be beneficial in cases with pathogenic CNVs, since in 18% further affected family members were identified. The number of familial cases was higher in our cohort than the previously described 6–10% (McDonald-McGinn et al., 1993-2020; Goldmuntz, 2020). The affected family members in our cohort exhibited similar or milder symptoms than the probands. This phenomenon has already been observed (Wozniak et al., 2010; Goldmuntz, 2020).

22q11.2 is considered one of the most unstable regions of the human genome, due to the low-copy repeat regions on chromosome 22. This instability predisposes the region to deletions and duplications through non-allelic homologous recombination events. Hence, the presence of a parental CNV may trigger the development of another CNV in the same or nearby chromosomal region in the offspring, as seen in the family P09 and as described by Capra et al. (2013).

In conclusion, based on the present results and on those described in the literature (Wozniak et al., 2010; Li et al., 2019; Goldmuntz, 2020), we suggest the implementation of the genetic screening of CNVs in the postnatal management of CHD patients, regardless of the type of CHDs. For this purpose, MLPA is a cost-effective, fast and specific method suitable for the screening of a large number of samples. Patients and families benefit greatly from early diagnosis, through the regular cardiological, orthopedic, endocrinological, immunological, neurodevelopmental, and psychiatric follow-ups, the more aggressive infection control and the possibility of positive family planning.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the National Medical Research Council (No. CHD-01/2016—IF-6299-8/2016) and by the Local Ethical Committee of the University of Szeged (No. 105/2016-SZTE). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

DN and MS: conceptualization, review, editing, and supervision. DN, GZ, MO, AK, KH, GR, MK, AU and ON: methodology, investigation, and validation. DN and MO: data curation. GZ and DN: writing and original draft preparation. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the Hungarian Scientific Research Fund (Grant No. 5S441-A202) and GINOP-2.3.2-15-2016-00039 grant.

ACKNOWLEDGMENTS

We thank Zsuzsanna Horváth-Gárgyán, Blanka Godza, Dóra Isaszegi, and Anikó Gárgyán for their skilled technical assistance and Dr. Shannon Frances for providing language help.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.635480/full#supplementary-material

REFERENCES

Bertini, V., Azzarà, A., Legitimo, A., Milone, R., Battini, R., Consolini, R., et al. (2017). Deletion extents are not the cause of clinical variability in 22q11.2 deletion syndrome: does the interaction between DGCR8 and miRNA-CNVs play a major role? *Front. Genet.* 8:47. doi: 10.3389/fgene.2017.00047

Brzustowicz, L. M., and Bassett, A. S. (2012). MiRNA-mediated risk for schizophrenia in 22q11.2 deletion syndrome. *Front. Genet.* 3:291. doi: 10.3389/fgene.2012.00291

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

Campbell, I. M., Sheppard, S. E., Crowley, T. B., McGinn, D. E., Bailey, A., McGinn, M. J., et al. (2018). What is new with 22q? An update from the 22q and you center at the children's hospital of Philadelphia. *Am. J. Med. Genet. A* 176, 2058–2069. doi: 10.1002/ajmg.a.40637

Capra, V., Mascelli, S., Garrè, M. L., Nozza, P., Vaccari, C., Bricco, L., et al. (2013). Parental imbalances involving chromosomes 15q and 22q may predispose to

the formation of de novo pathogenic microdeletions and microduplications in the offspring. *PLoS One* 8:e57910. doi: 10.1371/journal.pone.0057910

Chawner, S. J. R. A., Owen, M. J., Holmans, P., Raymond, F. L., Skuse, D., Hall, J., et al. (2019). Genotype-phenotype associations in children with copy number variants associated with high neuropsychiatric risk in the UK (IMAGINE-ID): a case-control cohort study. *Lancet Psychiatry* 6, 493–505. doi: 10.1016/S2215-0366(19)30123-3

Digilio, M. C., and Marino, B. (2016). What is new in genetics of congenital heart defects? *Front. Pediatr*. 4:120. doi: 10.3389/fped.2016.00120

Du, Q., de la Morena, M. T., and van Oers, N. S. C. (2020). The genetics and epigenetics of 22q11.2 deletion syndrome. *Front. Genet*. 10:1365. doi: 10.3389/fgene.2019.01365

Fahed, A. C., Gelb, B. D., Seidman, J. G., and Seidman, C. E. (2013). Genetics of congenital heart disease: the glass half empty. *Circ. Res.* 112, 707–720. doi: 10.1161/CIRCRESAHA.112.300853

Firth, H. V., Richards, S. M., Bevan, A. P., Clayton, S., Corpas, M., Rajan, D., et al. (2009). DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. *Am. J. Hum. Genet*. 84, 524–533. doi: 10.1016/j.ajhg.2009.03.010

Goldmuntz, E. (2020). 22q11.2 deletion syndrome and congenital heart disease. *Am. J. Med. Genet*. 184C, 64–72. doi: 10.1002/ajmg.c.31774

Griffin, H. R., Töpf, A., Glen, E., Zweier, C., Stuart, A. G., Parsons, J., et al. (2010). Systematic survey of variants in TBX1 in non-syndromic tetralogy of Fallot identifies a novel 57 base pair deletion that reduces transcriptional activity but finds no evidence for association with common variants. *Heart* 96, 1651–1655. doi: 10.1136/hrt.2010.200121

Guo, T., McDonald-McGinn, D., Blonska, A., Shanske, A., Bassett, A. S., Chow, E., et al. (2011). Genotype and cardiovascular phenotype correlations with TBX1 in 1,022 velo-cardio-facial/DiGeorge/22q11.2 deletion syndrome patients. *Hum. Mutat*. 32, 1278–1289. doi: 10.1002/humu.21568

Heike, C. L., Starr, J. R., Rieder, M. J., Cunningham, M. L., Edwards, K. L., Stanaway, I. B., et al. (2010). Single nucleotide polymorphism discovery in TBX1 in individuals with and without 22q11.2 deletion syndrome. *Birth Defects Res. A Clin. Mol. Teratol*. 88, 54–63. doi: 10.1002/bdra.20604

Huber, J., Peres, V. C., de Castro, A. L., dos Santos, T. J., da Fontoura Beltrão, L., de Baumont, A. C., et al. (2014). Molecular screening for 22Q11.2 deletion syndrome in patients with congenital heart disease. *Pediatr. Cardiol*. 35, 1356–1362. doi: 10.1007/s00246-014-0936-0

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581, 434–443. doi: 10.1038/s41586-020-2308-7

Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Albarca Aguilera, M., Meyer, R., et al. (2011). VarSome: the human genomic variant search engine. *Bioinformatics* 35, 1978–1980. doi: 10.1093/bioinformatics/bty897

Kruszka, P., Addissie, Y. A., McGinn, D. E., Porras, A. R., Biggs, E., Share, M., et al. (2017). 22q11.2 deletion syndrome in diverse populations. *Am. J. Med. Genet. A* 173, 879–888. doi: 10.1002/ajmg.a.38199

Lambert, M. P., Arulselvan, A., Schott, A., Markham, S. J., Crowley, T. B., Zackai, E. H., et al. (2018). The 22q11.2 deletion syndrome: cancer predisposition, platelet abnormalities and cytopenias. *Am. J. Med. Genet. A* 176, 2121–2127. doi: 10.1002/ajmg.a.38474

Landrum, M. J., Chitipiralla, S., Brown, G. R., Chen, C., Gu, B., Hart, J., et al. (2020). ClinVar: improvements to accessing data. *Nucleic Acids Res*. 48, D835–D844. doi: 10.1093/nar/gkz972

Legitimo, A., Bertini, V., Costagliola, G., Baroncelli, G. I., Morganti, R., Valetto, A., et al. (2020). Vitamin D status and the immune assessment in 22q11.2 deletion syndrome. *Clin. Exp. Immunol.* 200, 272–286. doi: 10.1111/cei.13429

Li, Y., Wei, X., Zhao, Z., Liao, Y., He, J., Xiong, T., et al. (2017). Prevalence and complications of bicuspid aortic valve in chinese according to echocardiographic database. *Am. J. Cardiol.* 120, 287–291. doi: 10.1016/j.amjcard.2017.04.025

Li, Z., Huang, J., Liang, B., Zeng, D., Luo, S., Yan, T., et al. (2019). Copy number variations in the GATA4, NKX2-5, TBX5, BMP4 CRELD1, and 22q11.2 gene regions in Chinese children with sporadic congenital heart disease. *J. Clin. Lab. Anal*. 33:e22660. doi: 10.1002/jcla.22660

MacDonald, J. R., Ziman, R., Yuen, R. K., Feuk, L., and Scherer, S. W. (2013). The database of genomic variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res.* 42, D986–D992. doi: 10.1093/nar/gkt958

McDonald-McGinn, D. M., Hain, H. S., Emanuel, B. S., and Zackai, E. H. (1993-2020). "22q11.2 Deletion Syndrome," in *GeneReviews® [Internet]*, eds M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. Stephens, et al. (Seattle, WA: University of Washington).

Moulding, H. A., Bartsch, U., Hall, J., Jones, M. W., Linden, D. E., Owen, M. J., et al. (2020). Sleep problems and associations with psychopathology and cognition in young people with 22q11.2 deletion syndrome (22q11.2DS). *Psychol. Med.* 50, 1191–1202. doi: 10.1017/S0033291719001119

Nagy, O., Szakszon, K., Biró, B. O., Mogyorósy, G., Nagy, D., Nagy, B., et al. (2019). Copy number variants detection by microarray and multiplex ligation-dependent probe amplification in congenital heart diseases. *J. Biotechnol*. 299, 86–95. doi: 10.1016/j.jbiotec.2019.04.025

Niarchou, M., Chawner, S. J. R. A., Fiksinski, A., Vorstman, J. A. S., Maeder, J., Schneider, M., et al. (2019). Attention deficit hyperactivity disorder symptoms as antecedents of later psychotic outcomes in 22q11.2 deletion syndrome. *Schizophr. Res*. 204, 320–325. doi: 10.1016/j.schres.2018.07.044

Niarchou, M., Zammit, S., van Goozen, S. H., Thapar, A., Tierling, H. M., Owen, M. J., et al. (2014). Psychopathology and cognition in children with 22q11.2 deletion syndrome. *Br. J. Psychiatry* 204, 46–54. doi: 10.1192/bjp.bp.113.132324

Palmer, L. D., Butcher, N. J., Boot, E., Hodgkinson, K. A., Heung, T., Chow, E. W. C., et al. (2018). Elucidating the diagnostic odyssey of 22q11.2 deletion syndrome. *Am. J. Med. Genet*. A 176, 936–944. doi: 10.1002/ajmg.a.38645

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

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

van der Linde, D., Konings, E. E., Slager, M. A., Witsenburg, M., Helbing, W. A., Takkenberg, J. J., et al. (2011). Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis. *J. Am. Coll. Cardiol*. 58, 2241–2247. doi: 10.1016/j.jacc.2011.08.025

van Engelen, K., Topf, A., Keavney, B. D., Goodship, J. A., van der Velde, E. T., Baars, M. J., et al. (2010). 22q11.2 Deletion Syndrome is under-recognised in adult patients with tetralogy of Fallot and pulmonary atresia. *Heart* 96, 621–624. doi: 10.1136/hrt.2009.182642

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

Wonkam, A., Toko, R., Chelo, D., Tekendo-Ngongang, C., Kingue, S., Dahoun, S., et al. (2017). The 22q11.2 deletion syndrome in congenital heart defects: prevalence of microdeletion syndrome in cameroon. *Glob. Heart* 12, 115–120. doi: 10.1016/j.gheart.2017.01.003

Wozniak, A., Wolnik-Brzozowska, D., Wisniewska, M., Glazar, R., Materna-Kiryluk, A., Moszura, T., et al. (2010). Frequency of 22q11.2 microdeletion in children with congenital heart defects in western poland. *BMC Pediatr.* 10:88. doi: 10.1186/1471-2431-10-88

Yu, A., Turbiville, D., Xu, F., Ray, J. W., Britt, A. D., Lupo, P. J., et al. (2019). Genotypic and phenotypic variability of 22q11.2 microduplications: an institutional experience. *Am. J. Med. Genet. A* 179, 2178–2189. doi: 10.1002/ajmg.a.61345

Zaidi, S., Choi, M., Wakimoto, H., Ma, L., Jiang, J., Overton, J. D., et al. (2013). De novo mutations in histone-modifying genes in congenital heart disease. *Nature* 498, 220–223. doi: 10.1038/nature12141

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: 07 May 2021
doi: 10.3389/fgene.2021.615072

Overdosage of *HNF1B* Gene Associated With Annular Pancreas Detected in Neonate Patients With 17q12 Duplication

Feifan Xiao[1,2†], Xiuyun Liu[1,2†], Yulan Lu[2], Bingbing Wu[2], Renchao Liu[2], Bo Liu[2], Kai Yan[3], Huiyao Chen[1,2], Guoqiang Cheng[3], Laishuan Wang[3], Qi Ni[1,2], Gang Li[2], Ping Zhang[2], Xiaomin Peng[2], Yun Cao[3], Chun Shen[4], Huijun Wang[2] and Wenhao Zhou[1,2,3]**

[1] Center for Molecular Medicine, Children's Hospital of Fudan University, National Children's Medical Center, Institutes of Biomedical Sciences, Fudan University, Shanghai, China, [2] Center for Molecular Medicine, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai, China, [3] Division of Neonatology, Key Laboratory of Neonatal Diseases, Ministry of Health, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai, China, [4] Department of Pediatric Surgery, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai, China

OPEN ACCESS

Edited by:
Judit Bene,
University of Pécs, Hungary

Reviewed by:
Rincic Martina,
University of Zagreb, Croatia
Anikó Ujfalusi,
University of Debrecen, Hungary

***Correspondence:**
Huijun Wang
huijunwang@fudan.edu.cn
Wenhao Zhou
zhouwenhao@fudan.edu.cn

[†]These authors share first authorship

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: 08 October 2020
Accepted: 01 April 2021
Published: 07 May 2021

Citation:
Xiao F, Liu X, Lu Y, Wu B, Liu R, Liu B, Yan K, Chen H, Cheng G, Wang L, Ni Q, Li G, Zhang P, Peng X, Cao Y, Shen C, Wang H and Zhou W (2021) Overdosage of HNF1B Gene Associated With Annular Pancreas Detected in Neonate Patients With 17q12 Duplication. Front. Genet. 12:615072. doi: 10.3389/fgene.2021.615072

The annular pancreas (AP) is a congenital anomaly of the pancreas that can cause acute abdominal pain and vomiting after birth. However, the genetic cause of AP is still unknown, and no study has reported AP in patients with 17q12 duplication. This study retrospectively analyzed the next-generation sequencing (NGS) data of individuals from January 2016 to June 2020 for 17q12 duplication. To identify the function of the key gene of *HNF1B* in the 17q12 duplication region, human *HNF1B* mRNA was microinjected into LiPan zebrafish transgenic embryos. A total of 19 cases of 17q12 duplication were confirmed. AP was diagnosed during exploratory laparotomy in four patients (21.1%). The other common features of 17q12 duplication included intellectual disability (50%), gross motor delay (50%), and seizures/epilepsy (31.58%). The ratio of the abnormal pancreas in zebrafish was significantly higher in the *HNF1B* overexpression models. In conclusion, we first reported AP in patients with duplication of the 17q12 region, resulting in the phenotype of 17q12 duplication syndrome. Furthermore, our zebrafish studies verified the role of the *HNF1B* gene in pancreatic development.

Keywords: annular pancreas, 17q12 duplication, CNV, *HNF1B*, zebrafish

INTRODUCTION

Annular pancreas (AP) is a morphological anomaly that results in the pancreatic tissue completely or incompletely surrounding the duodenum (Kiernan et al., 1980). The incidence of AP is estimated to range from 0.015 to 0.05% (Maker et al., 2003; Rondelli et al., 2016). The common symptoms of AP are abdominal pain, vomiting, acute or chronic pancreatitis, and swollen belly (Zyromski et al., 2008). The degree of manifestation depends on the severity of the intestinal blockage.

AP has been reported to be associated with some congenital anomalies (Jimenez et al., 2004; Wang et al., 2018). Some chromosome disorders, such as Down's syndrome, have been reported in 8–25% of patients with AP (Jimenez et al., 2004; Wang et al., 2018). Jacobsen syndrome has also been reported to be associated with AP (Fernández González et al., 2002). Chromosome

17q12 recurrent deletion syndrome (MIM: 614527), caused by the presence of a 1.4-Mb deletion at the approximate position of 36,458,167–37,854,616 (GRCh37/hg19), presents various clinical phenotypes including kidney anomalies, maturity-onset diabetes of the young type 5, and agenesis of the dorsal pancreas (Mitchel et al., 1993; Andersen and Schaffalitzky de Muckadell, 2019). Furthermore, chromosome 17q12 recurrent duplication syndrome (MIM: 614526), caused by a 1.4-Mb duplication at the approximate position of chr17:34,815,072–36,192,492 (GRCh37/hg19), presents various clinical phenotypes, including behavioral abnormalities, neurological symptoms, and brain abnormalities (Mefford et al., 1993; Mefford et al., 2007). However, AP has not been reported in either 17q12 deletion syndrome or 17q12 duplication syndrome.

There are 15 genes in the 17q12 recurrent duplication/deletion regions, namely, *ATF*, *ACACA*, *C17orf78*, *DDX52*, *DHRS11*, *DUSP14*, *GGNBP2*, *HNF1B*, *LHX1*, *MRM1*, *MYO19*, *PIGW*, *SYNRG*, *TADA2A*, and *ZNHIT3* (Mefford et al., 1993). Not all genes are haploinsufficient. Among these genes, variants in three genes (*ACACA*, *PIGW*, and *ZNHIT3*) can cause autosomal recessive inheritance diseases, whereas variants in *HNF1B* cause two autosomal dominant inheritance diseases including non-insulin-dependent diabetes mellitus (MIM: 125853) and renal cysts and diabetes syndrome (MIM: 137920). The *HNF1B* gene encodes a member of the homeodomain-containing superfamily of transcription factors and is regarded as an important transcription factor that controls the development of the pancreas (Coffinier et al., 1999).

The 17q12 recurrent deletion/duplication can be detected by array-based comparative genomic hybridization (aCGH), exome sequencing (ES) with copy number variant (CNV) calling, genome sequencing, or targeted deletion analysis. Our recent studies (Dong et al., 2020; Wang et al., 2020) have reported that next-generation sequencing (NGS) has good performance in detecting CNVs. In this study, we describe 19 patients with duplication of 17q12, 4 of whom present with AP as identified by NGS. Furthermore, functional studies were conducted in zebrafish. This study aimed to describe additional clinical characteristics and provide experimental data to understand 17q12 duplication.

MATERIALS AND METHODS

NGS Data Collection and CNV Calling

This study retrospectively collected the NGS data of individuals referred to the Center for Molecular Medicine of the Children's Hospital of Fudan University (CHFU) for genetic testing from January 1, 2016 to June 1, 2020. This study was approved by the CHFU Ethics Committee (2020-440).

Clinical exome sequencing (CES) data of 2,720 genes and ES data were included. Sequence data were aligned to the reference human genome (GRCh37/hg19). The detailed procedure was described in our previous study (Yang et al., 2019). This study developed an in-house CNV detection pipeline based on CANOES and HMZDelFinder and combined it with PhenoPro to prioritize phenotype-related genetic analysis (Li et al., 2019). The clinical significance of the CNVs was determined based on the following literature and genetic databases: UCSC Genome Browser[1], DECIPHER[2], and ClinGen[3]. Agilent SurePrint G3 aCGH and SNP 4 × 180 K microarray (Agilent Technologies, United States) were used to confirm the CNVs detected by NGS following the manufacturer's instructions. Data were processed using the DNA analytics software (Agilent Cytogenomics 4.0).

Clinical Information Collection and Patient Follow-Up

The inclusion criteria of individuals were as follows: (1) 17q12 duplication detected using NGS data analysis and (2) 17q12 duplication confirmed using aCGH. Clinical information was collected from medical records and *via* phone-call follow-up.

In vitro Study of the *HNF1B* Gene in Zebrafish

In vitro Synthesis of Wild-Type Human *HNF1B* and GFP mRNAs

The coding region of human *HNF1B* (NM_000458), which was synthesized by a biotechnology company (TsingKe, Beijing, China), and green fluorescent protein (GFP) was inserted into recombinant plasmids (pCS2+). Furthermore, 3 μg of each plasmid was digested with Not I. The insert containing the *HNF1B* or GFP cDNA was gel-purified and transcribed with SP6 RNA polymerase using the mMESSAGE mMACHINE^{TM} SP6 *in vitro* transcription kit (AM1340; Invitrogen) according to the manufacturer's instructions. The mRNA was diluted with diethylpyrocarbonate-treated water at a final concentration of 100 ng/μl and stored at -80°C until use.

Microinjection of Zebrafish Eggs

Zebrafish eggs were obtained by random mating of LiPan zebrafish with wild-type zebrafish (TU strain). *In vitro* synthesized *HNF1B* and GFP mRNAs were injected into embryos at the 1–2-cell stage together with the dye tracer phenol red solution (P0290; Sigma). After microinjection, embryos were maintained in egg water with ∼0.0005% methylene blue in a standard laboratory environment (28.5°C) and a 14-h light/10-h dark cycle according to a standard protocol (Kalueff et al., 2014) until analysis. Egg water was refreshed every day, and embryos with obvious deformities were discarded.

Microscopy and Image Analysis

At 7 days post-fertilization, larvae with exocrine pancreas-specific GFP expression were selected for analysis. The larvae were anesthetized with 0.08% tricaine (E10521; Sigma) and immobilized in 3% methylcellulose (M0521; Sigma). Observations of live embryos were performed using a Nikon stereoscope (SMZ800N), and the number of zebrafish with abnormal pancreas was recorded. Photographs were obtained using a Leica confocal microscope (TCS-SP8), and the length of the pancreas was measured.

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

[2]https://decipher.sanger.ac.uk/

[3]https://clinicalgenome.org/

Statistical Analysis

Data were presented as the mean ± standard error of the mean. Statistical analyses were performed, and graphs were plotted using the GraphPad Prism software (version 8.0). Student's *t*-test (two-tailed) was used to analyze the changes between different larval groups. The minimal criterion of significance was set at $P < 0.05$.

RESULTS

17q12 Duplication Identified From NGS Data

As shown in **Figure 1A**, a total of 41,171 individuals with suspected genetic diseases underwent genetic testing (CES or ES) in our laboratory from January 1, 2016 to June 1, 2020. All CES/ES data from these patients were analyzed for CNV calling. A total of 19 cases with 17q12 duplications were identified. Eight of these were tested using ES. Four cases (patients 7, 8, 11, and 15) had a 1,262.765-kb duplication (**Figure 1B**); however, in the remaining four cases, ES called four different sizes of duplication in the region of 17q12 (1,581.241 kb in patient 10, 1,532.705 kb in patient 6, 1,308.066 kb in patient 18, and 40.314 kb in patient 13). In the cases tested by CES, a 663.315-kb duplication (**Figure 1C**) was detected in all 11 cases (patients 1, 2, 3, 4, 5, 9, 12, 14, 16, 17, and 19). To confirm the size of the CNVs called by CES, aCGH was performed, and the length of the CNV in patient 5 was verified to be 1,516.456 kb (**Figure 1D**). All duplication regions included the *HNF1B* gene (**Supplementary Figure 1**). The detailed position of the 17q12 duplication in each patient is shown in **Supplementary Table 1**. No other pathogenic or likely pathogenic variants/CNVs were found in these 19 patients.

Six of the patients (31.58%) were neonates (<28 days), and 13 (68.42%) were pediatric patients (median age, 24 months). Among the 19 patients, AP was identified in 4 neonates (21.1%) who underwent exploratory laparotomy. The detailed clinical symptoms of the 19 patients are shown in **Supplementary Table 2**.

Clinical Characteristics of Four Patients With AP

Patient 1 was the first child of a non-consanguineous couple. The breathing of the patient stopped twice at the age of 3 days. He was then presented to our hospital for further treatment. Physical examination revealed a yellowish skin color. Serum biochemistry tests showed a high level of direct bilirubin (34.5 μmol/L; normal range: 0–6.8 μmol/L) and total bilirubin (231.4 μmol/L; normal range: 0–17.1 μmol/L). Plain abdominal radiography revealed a duodenal ileus (**Figure 2A**). Electroencephalography showed that the patient experienced neonatal seizures. Moreover, he was diagnosed with renal abnormalities. Facial features were normal. In addition, there were no ophthalmologic, endocrine, and cardiac abnormalities in this patient. Subsequently, the patient was diagnosed with AP during exploratory laparotomy. The patient did not present with intellectual disability, gross motor delay, and behavioral abnormalities during follow-up.

Patient 2 was a 28-day-old male infant. The patient was presented to a local hospital because of vomiting at the age of 3 days. He had been diagnosed with duodenal ileus by upper gastrointestinal contrast study and underwent surgery. However, the patient did not recover and was transferred to our hospital for further treatment. Physical examination revealed normal facial features. The upper gastrointestinal series (UGI) showed duodenal ileus (**Figure 2B**). No other abnormalities

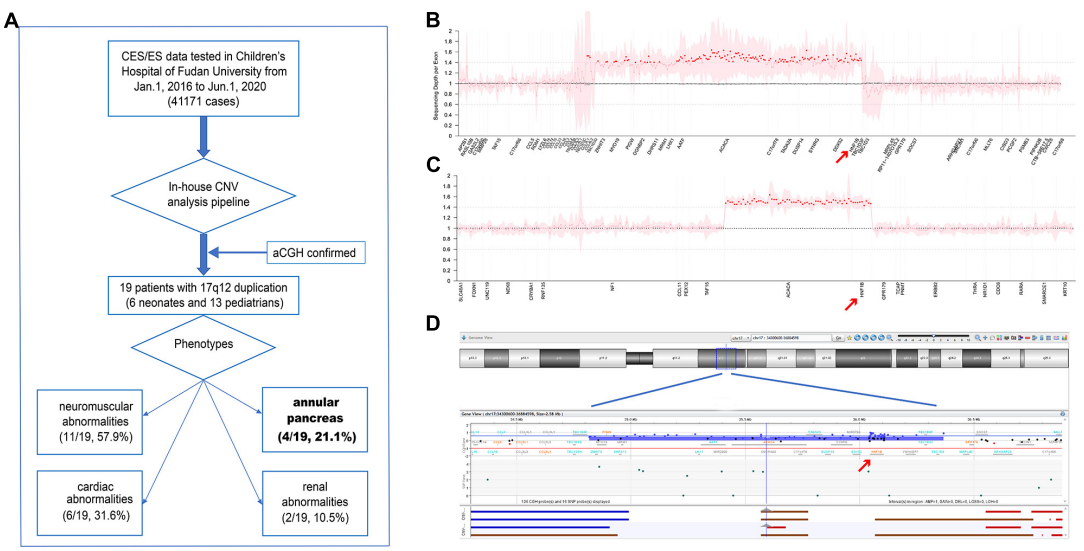


FIGURE 1 | (A) Flow diagram of this study. **(B,C)** Normalized sequencing depth of exons in samples with 17q12 duplication, for exome sequencing (ES) **(B)** and clinical exome sequencing (CES) **(C)** samples, separately. Sequencing reads depth on each exon were normalized and summarized. The mean values were marked as dots (pink dots for normal regions and red dots for duplication regions), and 95% confidential interval was shown as pink shades. X-axis was re-scaled by exon numbers other than gene length. **(D)** 17q12 duplication verified in patient 5 by aCGH.

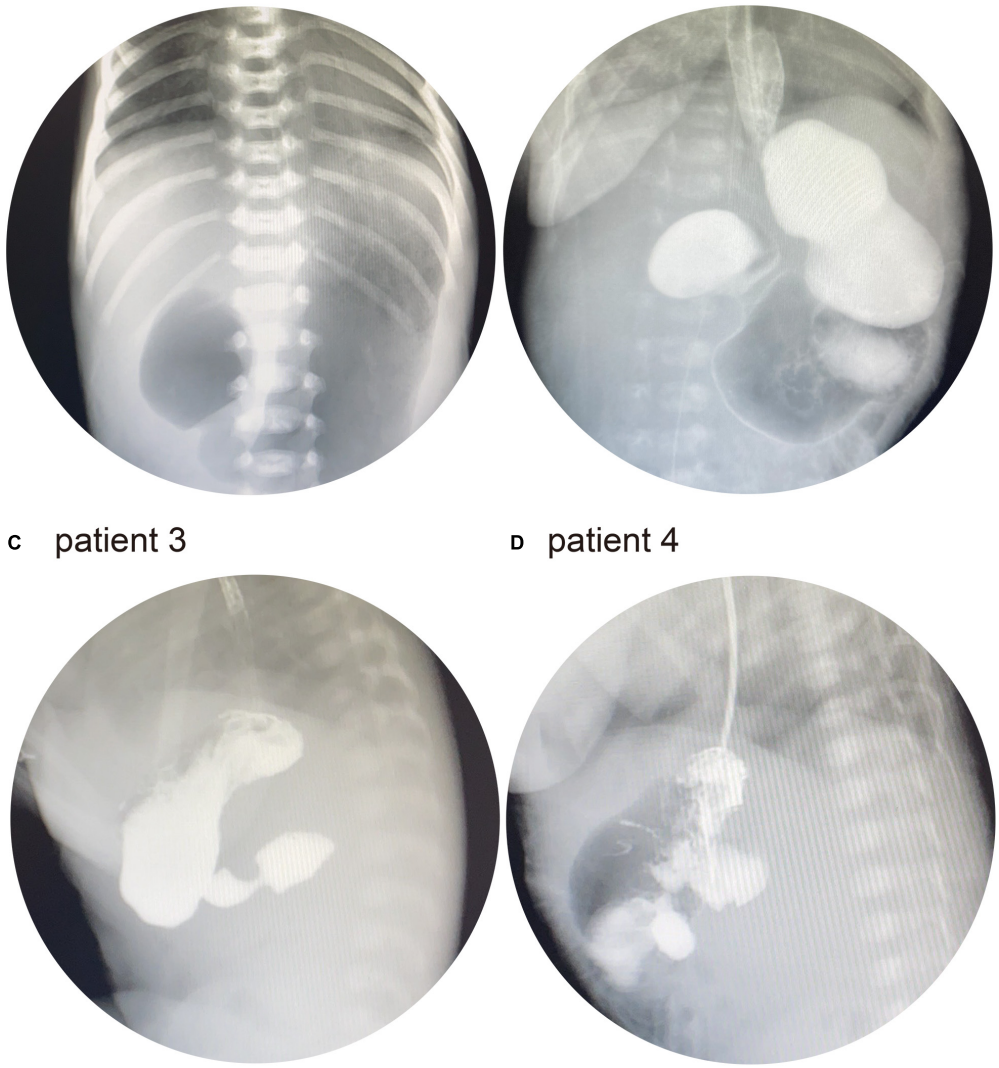


FIGURE 2 | (A) Patient 1: abdominal X-ray showed the "double bubble" sign. **(B)** Patient 2: upper gastrointestinal series (UGI) showed distended stomach and duodenal bulb. **(C)** Patient 3: UGI showed a distended duodenal bulb. **(D)** Patient 4: UGI showed a distended duodenal bulb.

were observed in patient 2. At the age of 32 days, he underwent exploratory laparotomy in our hospital and was diagnosed with AP.

Patient 3 was a 16-day-old male infant who was born *via* cesarean section. The patient was presented to our hospital owing to vomiting since birth. He had normal facial, ophthalmologic, and endocrine features, and the UGI showed a duodenal ileus (**Figure 2C**). Echocardiography revealed a patent arterial duct and a patent foramen ovale. At the age of 20 days, he underwent exploratory laparotomy and was diagnosed with AP.

Patient 4 was a 3-h-old male infant who was born *via* cesarean section. Prenatal color ultrasound indicated that the patient had duodenal ileus. Thus, the patient was transferred to our hospital for further treatment after birth. He had normal facial features. Ophthalmologic and endocrine tests were normal. Moreover, there was no abnormal cardiac malformation. His UGI revealed duodenal ileus (**Figure 2D**). The patient had been diagnosed with AP during surgery at the age of 5 days.

Clinical Characteristics of Patients With 17q12 Duplication Detected in This Study and in the Literature

Among the 19 patients, different neurodevelopmental abnormalities were observed. Among the 10 patients available for intelligence quotient and language assessment, only 5 patients (5/10, 50%) were presented with intellectual disability, and 1 patient (1/10, 10%) had speech delay. Moreover, information on gross development could be obtained from 12 patients, and half of them (6/12, 50%) showed gross motor

TABLE 1 | Summarized characteristics of patients with 17q12 duplication in the present study and published studies.

	19 Patients in this study (%)	108 Published patients (%)
Sex		
Male	12/19 (63.16)	52/94 (55.32)
Female	7/19 (36.84)	42/94 (44.68)
Age		
Neonate (0–28 days)	6/19 (31.58)	1/80 (1.25)
Children (>28 days–18 years)	13/19 (68.42)	60/80 (75)
Adult (>18 years)	0 (0)	19/80 (23.75)
Clinical characteristics		
Neurodevelopmental abnormalities		
Intellectual disability	5/10 (50.00)	51/69 (73.91)
Speech delay	1/10 (10.00)	37/49 (75.51)
Gross motor delay	6/12 (50.00)	30/53 (56.60)
Behavioral abnormalities	2/19 (10.53)	42/66 (63.63)
Seizures/epilepsy	6/19 (31.58)	30/59 (50.85)
Hypotonia	2/19 (10.53)	11/19 (57.89)
Structure abnormalities		
Dysmorphic facial features	1/19 (5.26)	34/57 (59.65)
Skeletal abnormalities	0/19 (0)	10/17 (58.82)
Cardiac abnormalities	6/19 (31.58)	10/26 (38.46)
Renal abnormalities	2/19 (10.53)	17/44 (38.64)
Annular pancreas	4/19 (21.1)	0 (0)
Endocrine abnormalities	0/19 (0)	12/40 (30.00)
Ophthalmologic abnormalities	3/19 (15.79)	15/46 (32.61)

delay. Other neurodevelopmental abnormalities included seizures/epilepsy (6/19, 31.58%), hypotonia (2/19, 10.53%), and behavioral abnormalities (2/19, 10.53%). Some patients had congenital abnormalities. Six patients (6/19, 31.58%) had cardiac abnormalities, four patients (4/19, 21.1%) had AP, two patients had renal abnormalities (2/19, 10.53%), and one patient (1/19, 5.26%) had dysmorphic facial features. Moreover, three patients (3/19, 15.79%) had ophthalmologic abnormalities, and none of them had endocrine abnormalities. The detailed clinical characteristics of the patients are shown in **Supplementary Table 1**.

In addition, the clinical symptoms of patients with 17q12 duplication were reviewed and summarized. A total of 108 patients were included in the analysis. Among them, 75% (60/80) were children (>28 days–18 years), and 23.75% were adults (19/80) (>18 years). As shown in **Table 1**, intellectual disability (73.9%, 51/69) and speech delay (75.5%, 37/49) were the most common neurodevelopmental abnormalities. Other common findings included behavioral abnormalities (63.63%, 42/66) and hypotonia (57.9%, 11/19). The most common congenital abnormalities included dysmorphic facial features (59.6%, 34/57) and skeletal abnormalities (58.82%, 10/17). To the best of our knowledge, AP has not been reported in patients with 17q12 duplication.

Overexpression of *HNF1B* mRNA in Zebrafish Embryos

Human *HNF1B* mRNA was microinjected into LiPan transgenic embryos with GFP fluorescence in the exocrine pancreas (Korzh et al., 2008) at the 1-cell stage (50 pg/embryo). In addition, GFP mRNA was synthesized and injected into embryos to confirm that the *in vitro* synthesis system worked well and to exclude the possibility that the phenomenon observed in *HNF1B* overexpression zebrafish results from random microinjection. As expected, GFP expression was observed in zebrafish at 24 h post-fertilization, and pancreatic morphology was similar to that in groups without injection (data not shown). Thus, embryos without injection and embryos injected with GFP mRNA were used as controls for the experiment.

Compared with the control, the morphology of the exocrine pancreas was obviously affected in the *HNF1B* overexpression groups. In the control group, the exocrine pancreas in zebrafish had a large anterior head region and a posterior tail elongated to the end of the yolk sac, whereas the exocrine pancreas in the *HNF1B* overexpression groups exhibited various abnormal morphologies, such as short pancreas and irregular or blurred shape of the pancreas (**Figure 3A**). In this study, we analyzed the ratio of zebrafish with an abnormal exocrine pancreas and the length of the pancreas. The ratio of zebrafish with abnormal pancreas was significantly higher (**Figure 3B**), and the length of the pancreas was significantly lower (**Figure 3C**) in the *HNF1B* overexpression groups than in the control group.

DISCUSSION

Individuals with AP may remain asymptomatic and may be diagnosed incidentally on imaging, surgery, or autopsy. However, a fraction of patients can present with intestinal

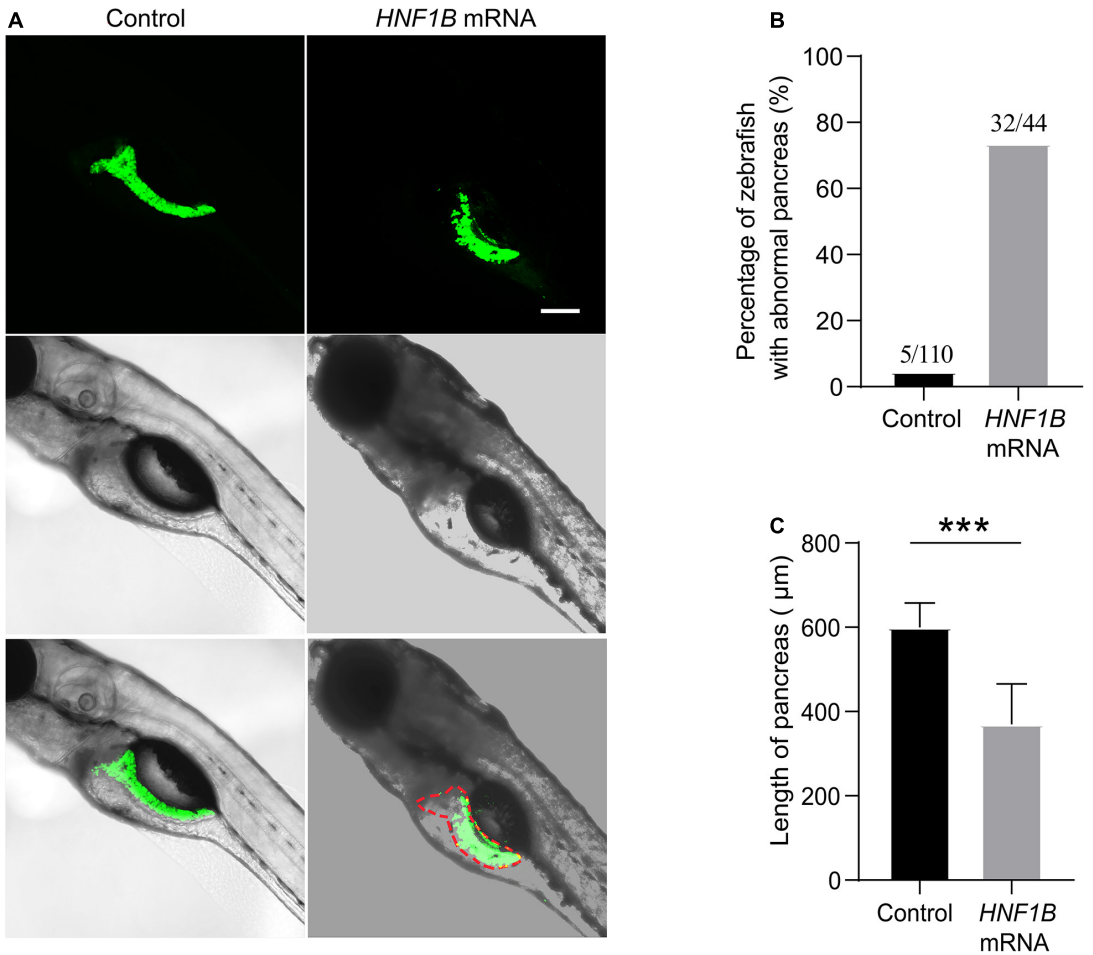


FIGURE 3 | (A) Different features of zebrafish in the *HNF1B* overexpression group and the control group. **(B)** The ratio of zebrafish with abnormal pancreas in the *HNF1B* overexpression group and the control group. **(C)** The length of the pancreas in the *HNF1B* overexpression group and the control group.

obstruction in infancy or abdominal pain, peptic ulcer disease, and pancreatitis during adulthood (Sandrasegaran et al., 2009). In the present study, the symptoms of AP included vomiting and hyperbilirubinemia, which are common manifestations of AP (Wang et al., 2019). Imaging examination revealed duodenal ileus in four patients. Furthermore, all four patients were diagnosed with AP during surgery.

In 2006, Sharp et al. (2006) first reported chromosome 17q12 duplication in a patient with intellectual disability. The prevalence of 17q12 duplication syndrome is estimated to range from 0.037 to 0.25% (Stefansson et al., 2014; Mitchell et al., 2015). Early in 2001, Sun and Hopkins (2001) identified that the loss of *HNF1B* can lead to underdeveloped pancreas in zebrafish embryos. A subsequent study (Haumaitre et al., 2005) showed that a lack of *HNF1B* causes pancreatic agenesis in mice. Further functional research found that deletion of *HNF1B* can decrease proliferation and increase apoptosis in pancreatic multipotent progenitor cells, which leads to severe pancreatic hypoplasia and perinatal lethality (De Vas et al., 2015). In the *Xenopus laevis* model, overexpression of *Hnf1b* can lead to expansion of the pancreatic progenitor domain, but it has a limited influence on other genes in the adjoining region (Gere-Becker et al., 2018). This study used LiPan transgenic zebrafish embryos with GFP fluorescence in the exocrine pancreas to help visualize the development of the pancreas. Human *HNF1B* mRNA was microinjected into LiPan transgenic embryos to analyze the function of the *HNF1B* gene in pancreatic development. Compared with blank injection, *HNF1B* mRNA zebrafish showed various abnormal morphologies and a short length of the pancreas. Our data confirmed that overexpression of *HNF1B* plays a vital role in the development of the pancreas. Detailed mechanisms should be explored in the future.

To date, 108 patients with 17q12 duplication syndrome have been reported. The clinical characteristics of patients with 17q12 duplication are summarized in **Table 1**, and the details are reviewed in **Supplementary Table 1**. This study reported 19 patients with 17q12 duplications. The prevalence in our study was 0.046% (19/41,171), which is similar to that reported in the literature (Stefansson et al., 2014; Mitchell et al., 2015). The diagnostic age of patients in this study was less than

that of the 108 patients reported in previous studies. All 19 patients were children, 6 of them were neonates, the median age was 24 months, and the average age was 50 months in the other 13 pediatric patients. Therefore, the evaluation of neurodevelopmental abnormalities related to intellectual disability, speech delay, and gross motor delay could only be performed in some patients. The behavioral abnormalities of some patients may occur later, and further examinations or follow-up is needed for the younger patients, particularly for the six neonates.

The diagnosed congenital abnormalities of dysmorphic facial features (59.65 vs. 5.26%), skeletal abnormalities (58.82 vs. 0%), and renal abnormalities (38.64 vs. 10.53%) of 19 patients were lower than those of 108 reported patients. AP was identified in 4 of 19 neonates with 17q12 duplication in this study. However, AP was not present in 108 reported patients. As AP may remain asymptomatic, we think that AP cannot be excluded in the reported 108 patients with 17q12 duplication. Approximately 40% of patients with AP diagnoses are diagnosed at surgery (Urayama et al., 1995), similar to our patients. Barr et al. (2020) reported an 11-year-old girl with DiGeorge syndrome who had been diagnosed with AP, even though the girl had a history of intermittent vomiting since birth. In this study, in 15 patients with 17q12 duplication, AP was not excluded. Thus, follow-up is necessary.

The duplication is recurrent and mediated by segmental duplications, and the reported size may be larger if adjacent segmental duplications are based on the design of the test method. However, although the capture regions of CES were different from those of ES, all the CNVs presented the key gene of 17q12 of *HNF1B*.

This study had two limitations. First, we only studied AP in patients with 17q12 duplication, but not with 17q12 deletion. The AP phenotype has not yet been identified. Second, the number of patients presented with the AP phenotype was only 4 of 19 with 17q12 duplication. In the other 15 patients, AP was not excluded, and follow-up was necessary for these patients to confirm if they had AP. Our future study will focus on whether the AP phenotype is present in patients with 17q12 deletion.

In conclusion, we first reported AP in patients with duplication of the 17q12 region that expanded the phenotype of 17q12 duplication syndrome. Further zebrafish studies have shown that the *HNF1B* gene plays an important role in the development of the pancreas.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

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

HW and WZ conceived and supervised the project. FX and XL designed and implemented the methods. YL, BW, RL, BL, KY, HC, GC, LW, QN, GL, PZ, XP, YC, and CS contributed to the data acquisition and analysis. HW, WZ, FX, and XL wrote the manuscript. All authors approved the manuscript.

FUNDING

This study was funded by the Shen Kang Hospital Development Center Clinical Science and Technology Innovation Project of Shanghai (SHDC12017110 to WZ) and the Shanghai Municipal Science and Technology Major Project (2017SHZDZX01 to WZ). The funding source did not have any role in the study design, data collection, data analysis, interpretation of the results, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

ACKNOWLEDGMENTS

We thank Dr. Li Qiang from the zebrafish center in Children's Hospital of Fudan University who provided suggestive opinion and technical support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.615072/full#supplementary-material

Supplementary Figure 1 | Locations of 17q12 duplication in our 19 patients.

REFERENCES

Andersen, B. B., and Schaffalitzky de Muckadell, O. B. (2019). [17q12 deletion as a possible cause of agenesis of the dorsal pancreas and polycystic kidney disease]. *Ugeskr. Laeger* 181:V081 90452.

Barr, M. M., Gilbert, J., and Murrell, Z. (2020). Delayed diagnosis of annular pancreas in 11-year-old girl with DiGeorge syndrome. *J. Pediatr. Surg. Case Rep.* 60, 101528. doi: 10.1016/j.epsc.2020.101528

Coffinier, C., Thépot, D., Babinet, C., Yaniv, M., and Barra, J. (1999). Essential role for the homeoprotein vHNF1/HNF1beta in visceral endoderm differentiation. *Development* 126, 4785–4794.

De Vas, M. G., Kopp, J. L., Heliot, C., Sander, M., Cereghini, S., and Haumaitre, C. (2015). Hnf1b controls pancreas morphogenesis and the generation of Ngn3+ endocrine progenitors. *Development* 142, 871–882. doi: 10.1242/dev.110759

Dong, X., Liu, B., Yang, L., Wang, H., Wu, B., Liu, R., et al. (2020). Clinical exome sequencing as the first-tier test for diagnosing developmental disorders covering both CNV and SNV: a Chinese cohort. *J. Med. Genet.* 57, 558–566. doi: 10.1136/jmedgenet-2019-106377

Fernández González, N., Prieto Espuñes, S., Ibáñez Fernández, A., Fernández Colomer, B., López Sastre, J., and Fernández Toral, J. (2002). [Deletion 11q23 – > qter (Jacobsen Syndrome) associated with duodenal atresia and annular pancreas]. *An. Esp. Pediatr.* 57, 249–252.

Gere-Becker, M. B., Pommerenke, C., Lingner, T., and Pieler, T. (2018). Retinoic acid-induced expression of Hnf1b and Fzd4 is required for pancreas development in Xenopus laevis. *Development* 145, dev161372. doi: 10.1242/dev.161372

Haumaitre, C., Barbacci, E., Jenny, M., Ott, M. O., Gradwohl, G., and Cereghini, S. (2005). Lack of TCF2/vHNF1 in mice leads to pancreas agenesis. *Proc. Natl. Acad. Sci. U.S.A.* 102, 1490–1495. doi: 10.1073/pnas.0405776102

Jimenez, J. C., Emil, S., Podnos, Y., and Nguyen, N. (2004). Annular pancreas in children: a recent decade's experience. *J. Pediatr. Surg.* 39, 1654–1657. doi: 10.1016/j.jpedsurg.2004.07.003

Kalueff, A. V., Stewart, A. M., and Gerlai, R. (2014). Zebrafish as an emerging model for studying complex brain disorders. *Trends Pharmacol. Sci.* 35, 63–75. doi: 10.1016/j.tips.2013.12.002

Kiernan, P. D., ReMine, S. G., Kiernan, P. C., and ReMine, W. H. (1980). Annular pancreas: may clinic experience from 1957 to 1976 with review of the literature. *Arch. Surg.* 115, 46–50. doi: 10.1001/archsurg.1980.01380010038007

Korzh, S., Pan, X., Garcia-Lecea, M., Winata, C. L., Pan, X., Wohland, T., et al. (2008). Requirement of vasculogenesis and blood circulation in late stages of liver growth in zebrafish. *BMC Dev. Biol.* 8:84. doi: 10.1186/1471-213x-8-84

Li, Z., Zhang, F., Wang, Y., Qiu, Y., Wu, Y., Lu, Y., et al. (2019). PhenoPro: a novel toolkit for assisting in the diagnosis of Mendelian disease. *Bioinformatics* 35, 3559–3566. doi: 10.1093/bioinformatics/btz100

Maker, V., Gerzenshtein, J., and Lerner, T. (2003). Annular pancreas in the adult: two case reports and review of more than a century of literature. *Am. Surg.* 69, 404–410.

Mefford, H. C., Clauin, S., Sharp, A. J., Moller, R. S., Ullmann, R., Kapur, R., et al. (2007). Recurrent reciprocal genomic rearrangements of 17q12 are associated with renal disease, diabetes, and epilepsy. *Am. J. Hum. Genet.* 81, 1057–1069. doi: 10.1086/522591

Mefford, H. C., Mitchell, E., and Hodge, J. (1993). "17q12 recurrent duplication," in *GeneReviews(®)*, eds M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, G. Mirzaa, et al. (Seattle (WA): University of Washington, Seattle). Copyright © 1993-2021, University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved.).

Mitchel, M. W., Moreno-De-Luca, D., Myers, S. M., Finucane, B., Ledbetter, D. H., and Martin, C. L. (1993). "17q12 recurrent deletion syndrome," in *GeneReviews(®)*, eds M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. Stephens, et al. (Seattle (WA): University of Washington, Seattle). Copyright © 1993-2020, University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved.).

Mitchell, E., Douglas, A., Kjaegaard, S., Callewaert, B., Vanlander, A., Janssens, S., et al. (2015). Recurrent duplications of 17q12 associated with variable phenotypes. *Am. J. Med. Genet. A* 167a, 3038–3045. doi: 10.1002/ajmg.a.37351

Rondelli, F., Bugiantella, W., Stella, P., Boni, M., Mariani, E., Crusco, F., et al. (2016). Symptomatic annular pancreas in adult: Report of two different presentations and treatments and review of the literature. *Int. J. Surg. Case Rep.* 20(Suppl.), 21–24. doi: 10.1016/j.ijscr.2016.02.001

Sandrasegaran, K., Patel, A., Fogel, E. L., Zyromski, N. J., and Pitt, H. A. (2009). Annular pancreas in adults. *AJR Am. J. Roentgenol.* 193, 455–460. doi: 10.2214/ajr.08.1596

Sharp, A. J., Hansen, S., Selzer, R. R., Cheng, Z., Regan, R., Hurst, J. A., et al. (2006). Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat. Genet.* 38, 1038–1042. doi: 10.1038/ng1862

Stefansson, H., Meyer-Lindenberg, A., Steinberg, S., Magnusdottir, B., Morgen, K., Arnarsdottir, S., et al. (2014). CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* 505, 361–366. doi: 10.1038/nature12818

Sun, Z., and Hopkins, N. (2001). vhnf1, the MODY5 and familial GCKD-associated gene, regulates regional specification of the zebrafish gut, pronephros, and hindbrain. *Genes Dev.* 15, 3217–3229. doi: 10.1101/gad946701

Urayama, S., Kozarek, R., Ball, T., Brandabur, J., Traverso, L., Ryan, J., et al. (1995). Presentation and treatment of annular pancreas in an adult population. *Am. J. Gastroenterol.* 90, 995–999.

Wang, D., Kang, Q., Shi, S., and Hu, W. (2018). Annular pancreas in China: 9 years' experience from a single center. *Pediatr. Surg. Int.* 34, 823–827. doi: 10.1007/s00383-018-4299-0

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

Wang, L., Xue, J., Chen, Y., Lyu, C., Huang, S., Tou, J., et al. (2019). [Clinical analysis of annular pancreas in neonates]. *Zhejiang Da Xue Xue Bao Yi Xue Ban* 48, 481–486.

Yang, L., Kong, Y., Dong, X., Hu, L., Lin, Y., Chen, X., et al. (2019). Clinical and genetic spectrum of a large cohort of children with epilepsy in China. *Genet. Med.* 21, 564–571. doi: 10.1038/s41436-018-0091-8

Zyromski, N. J., Sandoval, J. A., Pitt, H. A., Ladd, A. P., Fogel, E. L., Mattar, W. E., et al. (2008). Annular pancreas: dramatic differences between children and adults. *J Am Coll Surg* 206, 1019–102; discussion 1025–10175. doi: 10.1016/j.jamcollsurg.2007.12.009

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: 08 June 2021
doi: 10.3389/fgene.2021.673025

Genotype-Phenotype Associations in Patients With Type-1, Type-2, and Atypical *NF1* Microdeletions

*Gergely Büki[1,2], Anna Zsigmond[1], Márta Czakó[1,2], Renáta Szalai[1,2], Gréta Antal[1], Viktor Farkas[3], György Fekete[4], Dóra Nagy[5], Márta Széll[5], Marianna Tihanyi[6], Béla Melegh[1,2,7], Kinga Hadzsiev[1,2] and Judit Bene[1,2,7]**

[1] Department of Medical Genetics, Clinical Center, Medical School, University of Pécs, Pécs, Hungary, [2] Szentágothai Research Centre, University of Pécs, Pécs, Hungary, [3] 1st Department of Pediatrics, Semmelweis University, Budapest, Hungary, [4] 2nd Department of Pediatrics, Semmelweis University, Budapest, Hungary, [5] Department of Medical Genetics, Faculty of Medicine, University of Szeged, Szeged, Hungary, [6] Genetic Laboratory, Szent Rafael Hospital of Zala County, Zalaegerszeg, Hungary, [7] Full member of the European Reference Network on Genetic Tumour Risk Syndromes (ERN GENTURIS) – Project ID No. 739547, Pécs, Hungary

OPEN ACCESS

Edited by:
Ramcés Falfán-Valencia,
Instituto Nacional de Enfermedades Respiratorias-México, Mexico

Reviewed by:
Feng Zhang,
Fudan University, China
Hildegard Kehrer-Sawatzki,
Ulm University Medical Center, Germany
Eric Pasmant,
Université de Paris, France

***Correspondence:**
Judit Bene
bene.judit@pte.hu

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: 26 February 2021
Accepted: 12 May 2021
Published: 08 June 2021

Citation:
Büki G, Zsigmond A, Czakó M, Szalai R, Antal G, Farkas V, Fekete G, Nagy D, Széll M, Tihanyi M, Melegh B, Hadzsiev K and Bene J (2021) Genotype-Phenotype Associations in Patients With Type-1, Type-2, and Atypical NF1 Microdeletions. Front. Genet. 12:673025. doi: 10.3389/fgene.2021.673025

Neurofibromatosis type 1 is a tumor predisposition syndrome inherited in autosomal dominant manner. Besides the intragenic loss-of-function mutations in *NF1* gene, large deletions encompassing the *NF1* gene and its flanking regions are responsible for the development of the variable clinical phenotype. These large deletions titled as *NF1* microdeletions lead to a more severe clinical phenotype than those observed in patients with intragenic *NF1* mutations. Around 5-10% of the cases harbor large deletion and four major types of *NF1* microdeletions (type 1, 2, 3 and atypical) have been identified so far. They are distinguishable in term of their size and the location of the breakpoints, by the frequency of somatic mosaicism with normal cells not harboring the deletion and by the number of the affected genes within the deleted region. In our study genotype-phenotype analyses have been performed in 17 mostly pediatric patients with *NF1* microdeletion syndrome identified by multiplex ligation-dependent probe amplification after systematic sequencing of the *NF1* gene. Confirmation and classification of the *NF1* large deletions were performed using array comparative genomic hybridization, where it was feasible. In our patient cohort 70% of the patients possess type-1 deletion, one patient harbors type-2 deletion and 23% of our cases have atypical *NF1* deletion. All the atypical deletions identified in this study proved to be novel. One patient with atypical deletion displayed mosaicism. In our study *NF1* microdeletion patients presented dysmorphic facial features, macrocephaly, large hands and feet, delayed cognitive development and/or learning difficulties, speech difficulties, overgrowth more often than patients with intragenic *NF1* mutations. Moreover, neurobehavior problems, macrocephaly and overgrowth were less frequent in atypical cases compared to type-1 deletion. Proper diagnosis is challenging in certain patients since several clinical manifestations show age-dependency. Large tumor load exhibited more frequently in this type of disorder, therefore better understanding of genotype-phenotype correlations

and progress of the disease is essential for individuals suffering from neurofibromatosis to improve the quality of their life. Our study presented additional clinical data related to *NF1* microdeletion patients especially for pediatric cases and it contributes to the better understanding of this type of disorder.

Keywords: copy number variation, type-1 *NF1* microdeletion, type-2 *NF1* microdeletion, atypical *NF1* microdeletion, 17q11.2 deletion syndrome, array-CGH, multiplex ligation-probe dependent amplification, *NF1* gene

INTRODUCTION

Neurofibromatosis type 1 (NF1; MIM#162200), also known as von Recklinghausen disease, is an autosomal dominant disorder caused by loss-of-function mutations in the neurofibromin 1 (*NF1)* gene. The incidence of NF1 at birth is approximately 1 in 2500-3000 and the disease frequency shows no gender or racial predilection (Lammert et al., 2005; Uusitalo et al., 2015). The typical clinical features of NF1 are the hyperpigmented skin macules, called as café-au-lait spots (CALs), freckling of the axillary and inguinal regions, the pathognomonic neurofibromas and Lisch nodules. The neurofibromas are mostly benign tumors, localized on or under the skin (Huson and Hughes, 1994). They consist of a mixed cell types including Schwann cells, perineural cells, mast cells and fibroblasts. However, neurofibromatosis has a tremendous spectrum of clinical variability, including skeletal abnormalities, vascular disease, central nervous system tumors and cognitive dysfunction (attention deficit, learning disabilities) as well. Skeletal abnormalities such as dysplasia of the long bones are also characteristic for NF1 patients. Many features increase in frequency with aging and shows age-dependent manifestations. Moreover, strong intra- and interfamilial phenotypic variability can be observed among individuals carrying the same pathogenic mutations (Jett and Friedman, 2010).

Neurofibromin 1 gene is located on the long arm of the chromosome 17 (17q11.2) and codes for neurofibromin, a tumor suppressor that functions in the RAS/MAPK and mTOR pathways and controls the cell growth and proliferation (Jett and Friedman, 2010). The penetrance is complete and the mutation rate is high. Most of the intragenic *NF1* mutations are of paternal origin. Half of the known patients inherit the mutation, and the other half have a spontaneous mutation. Novel mutations occur primarily in paternally derived chromosomes, and the probability of these mutations increases with the paternal age (Stephens et al., 1992). A great number of germline mutations are intragenic and their effect causes a truncated neurofibromin (Park and Pivnick, 1998). Currently approximately 2000 mutations (nonsense, frameshift, point mutations etc.) are dispersed through the gene (Abramowicz and Gos, 2014).

The general NF1 population is mostly affected by point mutations or small indels, although a number of cases reported large deletions encompassing the *NF1* gene and its flanking regions. These large deletions titled as *NF1* microdeletions lead to a more severe clinical phenotype than those observed in patients with intragenic *NF1* gene mutations. These severe clinical features include large numbers of early-onset neurofibromas, cognitive deficits, dysmorphic features and an increased risk for the development of malignant peripheral nerve sheath tumors (MPNSTs) (Kehrer-Sawatzki et al., 2017).

Approximately 5-10% of NF1 patients have large deletions and the numbers are continuously increasing as a result of technological innovations (Cnossen et al., 1997; Kluwe et al., 2004; Zhang et al., 2015). Four major types of *NF1* microdeletions (type 1, 2, 3 and atypical) have been identified so far. The main difference among them are the breakpoint location, the size of the deletion, and the number of the affected genes within the deleted region (Kehrer-Sawatzki et al., 2017). The most frequent form is the type-1 *NF1* microdeletion, which is 1.4 Mb long and includes 14 protein-coding genes and four microRNA genes as well (Dorschner et al., 2000; Lopez-Correa et al., 2001). Type-1 deletions account for 70-80% of all large *NF1* deletions (Pasmant et al., 2010; Messiaen et al., 2011). Type-2 *NF1* deletions are less common than type-1 and they represent ca. 10-20% of all large *NF1* deletions (Mautner et al., 2010; Pasmant et al., 2010; Messiaen et al., 2011). Type-2 deletions are 1.2 Mb in size and result in the deletion of 13 genes. In contrast to type-1 and type-2 *NF1* deletions, type-3 *NF1* deletions are very rare, their occurrence is around 1-4% of patients with *NF1* microdeletions (Bengesser et al., 2010; Pasmant et al., 2010; Messiaen et al., 2011). This type of deletion spans 1 Mb and leads to the loss of 9 protein coding genes.

Type-1, 2, and 3 NF1 microdeletions are generated by non-allelic homologous recombination (NAHR) between low-copy repeats (LCRs) during either meiosis (type-1, type-3), or mitosis (type-2) (Dorschner et al., 2000; Jenne et al., 2001; Lopez-Correa et al., 2001; Bengesser et al., 2010; Pasmant et al., 2010; Roehl et al., 2010; Zickler et al., 2012; Hillmer et al., 2016). Type-1 cases are usually maternally inherited germline deletions (Neuhausler et al., 2018), while type-2 ones are predominantly of postzygotic origin (Kehrer-Sawatzki et al., 2004; Steinmann et al., 2008; Vogt et al., 2012). Besides these three types of recurrent microdeletions, atypical *NF1* deletions have been identified in a number of patients. In atypical deletions non-recurrent breakpoints have been discovered, thereby the size of the deletion and the number of the affected genes also vary (Kehrer-Sawatzki et al., 2003, 2005, 2008; Mantripragada et al., 2006; Pasmant et al., 2010; Messiaen et al., 2011). Non-homologous end joining mechanism has been associated mostly with atypical deletions (Venturin et al., 2004a). However, either aberrant DNA double strand break repair and/or replication, and retrotransposon-mediated mechanisms have also been supposed to be involved in the background of their formation (Vogt et al., 2014). Atypical microdeletions may occur

approximately in 8-10% of all patients with NF1 microdeletions (Pasmant et al., 2010).

Somatic mosaicism with normal cells not harboring large *NF1* deletion can be observed with different frequencies in different types of *NF1* deletions. This phenomenon is rare among type-1 deletions, vast majority (more than 95%) of the patients with type-1 deletion is non-mosaic (Messiaen et al., 2011; Summerer et al., 2019). Contrast to type-1 deletion, somatic mosaicism is quite common in type-2 *NF1* deletions, it occurs in at least 63% of all type-2 deletions (Vogt et al., 2012). Atypical *NF1* deletions also display mosaicism frequently. In a study reported by Vogt et al. (2014), approximately 60% of the cases were associated with somatic mosaicism (Vogt et al., 2014). It is worth to note that somatic mosaicism with normal cells without the deletion has a considerable effect on the disease phenotype, however it is difficult to assess its presence.

In addition to the extent of somatic mosaicism, the age of the patients is also an important confounding factor in phenotypic comparisons of NF1 patient cohort, since many symptoms are progressive in onset and some of them appears later in life (Cnossen et al., 1998).

Several research groups have investigated different aspects of *NF1* microdeletions, however only a few studies presented profound clinical examinations. Here we report clinical and genotype data from 17 patients, mainly (82%) children and adolescents, carrying different types of microdeletion. One of the patients with atypical deletion showed somatic mosaicism. The aim of our study was to characterize the detected deletions in our patient cohort and elucidate genotype-phenotype correlations through clinical data collection.

MATERIALS AND METHODS

Participants

Between 2009 and 2019, our laboratory tested 640 unrelated patients with suspected neurofibromatosis. After Sanger sequencing of the *NF1* gene or NGS analyses of *NF1, NF2, KIT, PTPN11, RAF1, SMARCB1, SPRED1* genes no disease-causing mutations have been identified in 252 patients. Of these, 17 patients (7 females, 10 males; mean age at time of examination:12.9 years, age range:2-36 years) with large *NF1* deletion were identified by MLPA and were enrolled into this study. Our patient cohort mostly (14 out of 17) consisted of children between the ages of 2 and 17. Two patients inherited the deletion from their mothers (patients 85 and 260), while in the remaining 15 patients the deletions had *de novo* origin based on the negative MLPA results of the parents or the absence of a clinically affected parent. However, in the latter case low grade or tissue specific mosaicism cannot be ruled out. The mother of patient 260 (patient 134) was clinically affected as well, therefore she was also included in the analysis. The mother of patient 85 was sine morbo. As a control, age and sex matched 33 patients (14 females, 19 males; mean age at the time of examination: 15.2 years, age range:6 months-47 years) with intragenic *NF1* mutations were enrolled into the study as well.

The study was approved by the ethics committee of the University of Pecs (Protocol 8581-7/2017/EUIG). Written informed consent was obtained from all patients or their legal guardians and peripheral blood samples were collected. All experiments were performed in accordance with the Helsinki Declaration of 1975 and with the Hungarian legal requirements of genetic examination, research and biobanking.

All of the patients fulfilled the diagnostic NIH criteria for NF1. Main clinical characteristics of our patient cohort are summarized in **Table 1**. Phenotypic data was obtained from our Genetic counseling unit and from our collaborator clinicians.

Sample Preparation and MLPA Analysis

DNA was isolated from peripheral blood leukocytes with E.Z.N.A.® Blood DNA Maxi kit (Omega BIO-TEK, Norcross, United States). The concentration and purity of extracted DNAs were measured with the NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, United States).

Multiplex ligation-dependent probe amplification (MLPA) assays were performed for screening large deletions or duplications in *NF1* gene using the commercially available SALSA MLPA kits P081-D1 and P082-C2 (MRC-Holland, Amsterdam, The Netherlands). The two probemixes contained together one probe for each exon, three probes for exon 1, one probe for intron 1, and two probes for the exons 15, 21, 23, 51, and 58 of the *NF1* gene. Additionally, one upstream and one downstream probe of *NF1* gene and two probes for the *OMG* gene (located within intron 36 of *NF1* gene) were applied. Moreover, SALSA MLPA kit P122-D1 *NF1* area mix was used for the examination of the contiguous genes in the flanking regions. The probemix contained 20 probes for 16 genes (*MYO1D*, *PSMD11*, *ZNF207*, *LRRC37B*, *SUZ12*, *UTP6*, *RNF135*, *ADAP2*, *ATAD5*, *CRLF3*, *SUZ12P*, *CPD*, *BLMH*, *TRAF4*, *PMP22*, *ASPA*), which were localized upstream and downstream as well. Besides, it also contained probes for five distinct *NF1* exons (1, 17, 30, 49, 57). According to the manufacturer's instructions, a total of 100–200 ng of genomic DNA of each patient and the same amount of three control genomic DNA was used for hybridization. Amplification products were separated by capillary electrophoresis on an ABI 3130 Genetic Analyzer (Life Technologies, United States) and the results were analyzed using Coffalyser software (MRC-Holland, Amsterdam, Netherlands). Each MLPA signal was normalized and compared to the corresponding peak area obtained from the three control samples. Deletions and duplications of the targeted regions were suspected when the signal ratio exceeded 30% deviation. Positive results were confirmed by repeated MLPA experiments and further investigated with array CGH.

Whole Genome Array Comparative Genomic Hybridization Analysis

Array comparative genomic hybridization (aCGH) was performed using the Affymetrix CytoScan 750 K Array. Genomic DNA samples were digested, ligated, amplified, fragmented, labeled, and hybridized to the CytoScan 750 K Array platform according to the manufacturer's instructions. The raw data were

TABLE 1 | Clinical features of our patients with different type of *NF1* microdeletions.

	Deletion type	**Type 1**								**Type 1**				**Type 2**	**Atypical**			
	Applied method	**aCGH**								**MLPA**				**aCGH**	**aCGH**	**MLPA**		
	Patients	**68/NF**	**115/NF**	**255NF**	**428NF**	**467/2016**	**532/NF**	**629/NF**	**761/NF**	**9/NF**	**271/NF**	**387/NF**	**483/NF**	**85/NF**	**556/NF**	**125/NF**	**134/NF**	**260/NF**
	Gender	**M**	**F**	**M**	**M**	**F**	**M**	**F**	**M**	**M**	**M**	**F**	**M**	**F**	**M**	**F**	**F**	**M**
	Age of onset	**26 y**	**5 mo**	**at birth**	**at birth**	**N/A**	**12 y**	**at birth**	**at birth**	**at birth**	**at birth**	**at birth**	**5 y**	**1 mo**	**6.5 y**	**at birth**	**3 y**	**at birth**
	Age at examination	**36 y**	**9 y**	**14 y**	**5 y**	**9 y**	**14 y**	**4.5y**	**9 y**	**21 y**	**4 y**	**17 y**	**7.5 y**	**13 y**	**10 y**	**2 y**	**40 y**	**8 y**
Dysmorphic features	Facial dysmorphism	X	X	X	X	-	X	X	-	X	-	-	X	-	-	-	-	X
	Hypertelorism	X	X	X	X	-	X	X	-	-	-	-	X	-	-	-	X	X
	Facial asymmetry	-	-	-	-	-	X	-	X	X	-	-	-	-	-	-	-	-
	Coarse face	X	-	X	X	-	X	X	X	X	-	-	X	X	-	-	-	-
	Broad neck	-	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Large hands, feet	-	X	X	X	-	X	X	X	X	-	-	X	X	-	-	-	-
Skin manifestations	CALs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Freckling	-	X	X	X	X	-	X	X	X	X	X	X	-	X	X	-	X
	Excess soft tissue	-	-	X	X	-	-	X	-	X	-	-	-	X	-	-	-	-
	SBC neurofibromas	X	X	X	X	-	-	-	X	-	X	X	-	-	-	-	X	-
	CT neurofibromas	-	-	-	-	-	-	-	-	X	-	-	-	-	-	-	-	-
	PL neurofibromas*	-	-	-	-	-	-	-	-	X	-	X	-	-	-	-	-	-
Education and behavior problems	SDiCD	X	-	X	X	X	X	X	X	X	-	-	X	-	-	-	-	X
	Learning difficulties	X	-	X	-	X	X	X	X	X	-	X	X	X	-	-	-	-
	Speech difficulties	-	-	X	X	X	X	X	X	-	-	X	X	-	-	-	-	-
	IQ < 70	-	-	-	–	-	-	-	-	X	-	-	-	-	-	-	-	-
	ADHD	-	-	-	X	-	-	-	-	X	-	-	-	-	-	-	-	-
Skeletal manifestations	Skeletal anomalies	X	X	X	X	X	X	X	X	X	X	X	-	X	X	X	X	X
	Scoliosis	X	-	X	-	-	X	-	-	X	X	-	-	X	-	-	X	-
	Pectus excavatum	-	X	-	X	-	X	-	-	X	X	-	-	-	-	X	-	X
	Bone cysts	X	n.d.	-	n.d.	n.d.	-	-	-	-	-	-	-	-	-	-	-	-
	Joint hyperflexibility	-	-	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-
	Macrocephaly	-	X	X	X	X	-	X	X	-	-	X	-	X	-	-	-	X
Neurological manifestations	Muscular hypotonia	X	-	X	-	-	-	-	X	-	-	-	-	-	-	-	-	-
	Headache	-	-	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-
	Coordination problem	-	-	X	X	X	-	-	X	-	-	-	-	-	-	-	-	-
	MPNST	X	-	-	-	-	-	-	-	-	-	X	-	-	-	-	-	-

(Continued)

TABLE 1 | Continued

	Deletion type	Type 1														Type 1				Type 2		Atypical			
	Applied method	aCGH														MLPA				aCGH	aCGH	MLPA			
	Patients	68/NF	115/NF	255NF	428NF	467/2016	532/NF	629/NF	761/NF	9/NF	271/NF	387/NF	483/NF	85/NF	556/NF	125/NF	134/NF	260/NF							
	Gender	M	F	M	M	F	M	F	M	M	M	F	M	F	M	F	M	M							
	Age of onset	26 y	5 mo	at birth	at birth	N/A	12 y	at birth	at birth	at birth	at birth	at birth	5 y	1 mo	6.5 y	at birth	3 y	at birth							
	Age at examination	36 y	9 y	14 y	5 y	9 y	14 y	4.5y	9 y	21 y	4 y	17 y	7.5 y	13 y	10 y	2 y	40 y	8 y							
	Spinal neurofibromas	-	n.d.	n.d.	n.d.	-	-	n.d.	X	-	-	X	-	n.d.	-	-	n.d.	n.d.							
	T2 hyperintensities	X	X	X	X	-	-	X	X	X	X	X	X	X	-	X	n.d.	X							
Ocular manifestations	Visual disturbance	-	-	-	-	-	X	-	-	-	-	X	-	X	-	-	-	-							
	Lisch nodules	-	-	X	-	-	-	-	-	X	-	X	-	X	-	-	-	-							
	Strabismus	-	-	-	-	-	X	-	-	-	-	-	X	-	-	-	-	-							
	OPG	-	-	-	-	-	-	X	-	-	-	X	-	-	X	-	-	X							
Development. problems	Tall stature	-	X	X	-	-	X	X	X	X	-	X	-	-	-	-	-	-							

*CALs, café-au-lait spots; CT/SBC/PL, cutaneous/subcutaneous/plexiform neurofibroma; SDiCD, significant delay in cognitive development; ADHD, Attention deficit hyperactivity disorder; MPNST, malignant peripheral nerve sheath tumors; OPG, Optic Pathway Gliomas. * means externally observable plexiform neurofibroma. -, absent; X, present, n.d., not determined.*

analyzed by ChAS v2.0 Software (Affymetrix, Thermo Fisher Scientific, Waltham, MA).

CNV Interpretation

DNA sequence information of the identified CNVs refer to the public UCSC database (GRCh37/hg19). CNV interpretation was performed with the help of the following databases and websites: DECIPHER (Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources) (Firth et al., 2009), DGV (Database of Genomic Variants), Ensembl and ECARUCA (European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations) (Vulto-van Silfhout et al., 2013). The estimated size of the deletions and the estimated breakpoints were assessed using the known locations of the last proximal and first distal deleted probes.

Somatic Mosaicism Determination

In patients examined by aCGH assay, allele difference plot and B allele frequency (BAF) plot were evaluated together with Log2 ratios and weighted Log2 ratios with the help of ChAS software to assess the presence and extent or absence of somatic mosaicism. In those samples investigated by MLPA, the ratio values for each MLPA probe were used to assess mosaicism. Values between 0.4-0.6 were considered as non-mosaic deletion, values around 0.7 or up to 0.8 were considered as mosaic deletion.

Clinical Investigation

Phenotypic features of the 17 microdeletion and the 33 control patients were collected using the same standardized questionnaire collection protocol in four HCPs (health care provider). The same patient was always examined and followed up by the same clinician. Most features were identified by physical examination. Dysmorphic features were assessed by expert clinical syndromologist following international guidelines[1] (Allanson et al., 2009; Hall et al., 2009). Lisch nodules and other ocular manifestations were diagnosed by an ophthalmologist. To evaluate childhood overgrowth age and race-related percentile curve was applied. All the patients were investigated by cranial MRI. To evaluate intellectual functions, developmental delay and learning disabilities, patients were assessed by various psychological tests appropriate to their age (Walter Strassmeier's developmental scale: ages between 0 and 5 years (Strassmeier, 1980), Bayley Scales test (BSID-III): ages between 1 and 42 months (Bayley, 2006), Budapest Binet test: ages between 3 and 14 years (Bass et al., 1989)). When IQ was not measured, it was estimated to be > 70 based on the fact that the patient attended a regular kindergarten or school (with special educational needs). ADHD was diagnosed following international guidelines[2]. The term "speech difficulties" was used in those cases when the patient did not speak or he or she had a problem with the language content, language structure and expressive vocabulary and grammar. We assigned it to delayed language development and not neurological symptoms (dysarthria or orofacial dyskinesis).

[1]http://elementsofmorphology.nih.gov/

[2]https://www.nhs.uk/conditions/attention-deficit-hyperactivity-disorder-adhd/symptoms/

Statistical Analysis

All statistical analyses were performed with SPSS version 27 (SPSS Inc,. Chicago, IL, United States). Two-tailed Fisher's exact test was used to assess whether there is a difference in the frequency of clinical features between patients with type-1 *NF1* microdeletion and patients with intragenic *NF1* mutations. A difference with $p < 0.05$ was considered as significant.

RESULTS

Characterization of the *NF1* Microdeletions

A total of 252 patients in whom mutation analysis did not find any pathogenic *NF1* point mutations or intragenic insertions/deletions were screened for large *NF1* rearrangements by MLPA. Heterozygous deletions of the entire *NF1* gene and its flanking regions were identified in 17 patients using SALSA P081/082 assay. To determine the contiguous genes involved in the deletion, the SALSA P122 assay was applied. As a result, majority of our cases (12/17) had type-1 deletion. Moreover, the MLPA analysis revealed atypical deletions in 5 patients. The estimated proximal and distal breakpoints, preceding and following marker locations and the estimated size of the deletions identified by MLPA are summarized in **Supplementary Table 1**. To confirm the MLPA results, array comparative genomic hybridization (aCGH) analyses were performed in 10 patients (8 patients with type-1 and 2 patients with atypical deletions). The estimated location of proximal and distal breakpoints, preceding and following markers and the estimated size of the deletions determined by aCGH are summarized in **Supplementary Table 2**. The classification by MLPA and by aCGH were found to be the same in eight cases (7 type-1 deletion and 1 atypical). In patient 85/NF the aCGH finally revealed the existence of type-2 deletion although the MLPA showed atypical deletion. In patient 4672016 the aCGH test showed atypical deletion whereas MLPA detected a type-1 deletion, finally we considered this patient has type-1 deletion. The discrepancy between the MLPA and aCGH results in these cases may originate from the different localization of the probes. Type-2 deletions are characterized by breakpoints located within *SUZ12* gene and its pseudogene *SUZ12P*. SALSA P122 probe set contains only one probe for *SUZ12* gene (SUZ12-10: localized within exon 10) and 2 probes for *SUZ12P* pseudogene (SUZ12P-3, SUZ12P-1: probe localization within exon 3 and exon 1, respectively). The breakpoints of the deletion detected in our patient (85/NF) were localized within the region covered by *SUZ12* and *SUZ12P* probes of P-122 set. The applied CytoScan 750K chip contains more probes, at least 50 and 7 for *SUZ12* and *SUZ12P*, respectively. Therefore, aCGH was capable to identify this type-2 deletion. Breakpoints of type-1 deletions are located within the low-copy repeats NF1-REPa and NF1-REPc. In patient 4672016 the estimated proximal breakpoint detected by aCGH can be found within NF1-REPa and the estimated distal breakpoint detected by MLPA can be found within NF1-REPc, therefore we considered 4672016 patient as having type-1 deletion. In the remaining 7 cases (4 patients with type-1 and three with atypical deletions), aCGH tests were not feasible due to the quality of the available samples. After all, 8 type-1 deletions, 4 potential type-1 deletions (altogether 12 type-1 deletions), one type-2 deletion and 3 atypical deletions in four patients were identified in our patient cohort. No type-3 microdeletion was found in our patients. Among the type-1 deletions aCGH analyses revealed identical estimated breakpoints in four cases with an approximately 1.37 Mb deletion size. Among atypical cases three distinct novel deletions were detected. Patient 134/NF and 260/NF are close relatives (mother and child), so they possess the same deletion. The results of our MLPA and aCGH analyses with the localization of the MLPA probes are visualized in **Figure 1**. Novel atypical deletions identified in this study, together with the already known atypical NF1 cases, are demonstrated in **Figure 2** and **Tables 4**, **5**. Two out of three novel atypical deletions were identified by MLPA. SALSA P122 probe set contains 23 probes within the 17q region and the distance between the adjacent probes are quite variable from 11 kb up to 1500 kb. The preceding markers of the estimated proximal breakpoint and the following markers of the estimated distal breakpoint are localized far from the breakpoint boundaries. The distance between the preceding markers and the estimated proximal breakpoints are ca. 270 kb and 27 kb in case 125/NF and 260/NF (134/NF), respectively. The distance between the following markers and the estimated distal breakpoints are ca. 80 kb and 500 kb in case 125/NF and 260/NF (134/NF), respectively. MLPA is able to identify only estimated location of breakpoints, the exact localization of the breakpoints can be determined precisely by breakpoint-spanning PCR (Summerer et al., 2018). In our cases the actual breakpoints are presumably located somewhere between two MLPA probes. Therefore, the regions in proximal direction from the first probe or in distal direction from the last probe affected by the deletion until the adjacent probe are suggested as potential deleted region and represented in **Figure 2** with dotted lines.

Assessment of Somatic Mosaicism

Among 10 patients investigated by aCGH, only one subject (556/NF) with atypical *NF1* microdeletion displayed somatic mosaicism with an extent of ca. 30%. In 7 patients examined by MLPA, the ratio values do not imply the presence of any mosaicism. However, neither aCGH, nor MLPA measurements are capable to detect low-grade mosaicism below 20% due to the nature of these techniques. In this study we investigated only blood samples, so to completely rule out mosaicism, examination of additional tissues such fibroblast, buccal or urine cells are necessary. In type-1 *NF1* microdeletion the occurrence of somatic mosaicism is known to be very rare (Summerer et al., 2019), so based on our results our type-1 patients can be considered as non-mosaic cases. The only one patient with type-2 deletion inherited the deletion from her mother, consequently she does not possess somatic mosaicism. Anyway, this is compatible with the aCGH result as well. Among our four patients with atypical *NF1* deletion, the results indicated ca. 30% mosaicism in only one case (556/NF). Patient 260/NF inherited the deletion from his mother, therefore this patient is considered as non-mosaic. His mother (134/NF) is supposed to be a non-mosaic case as

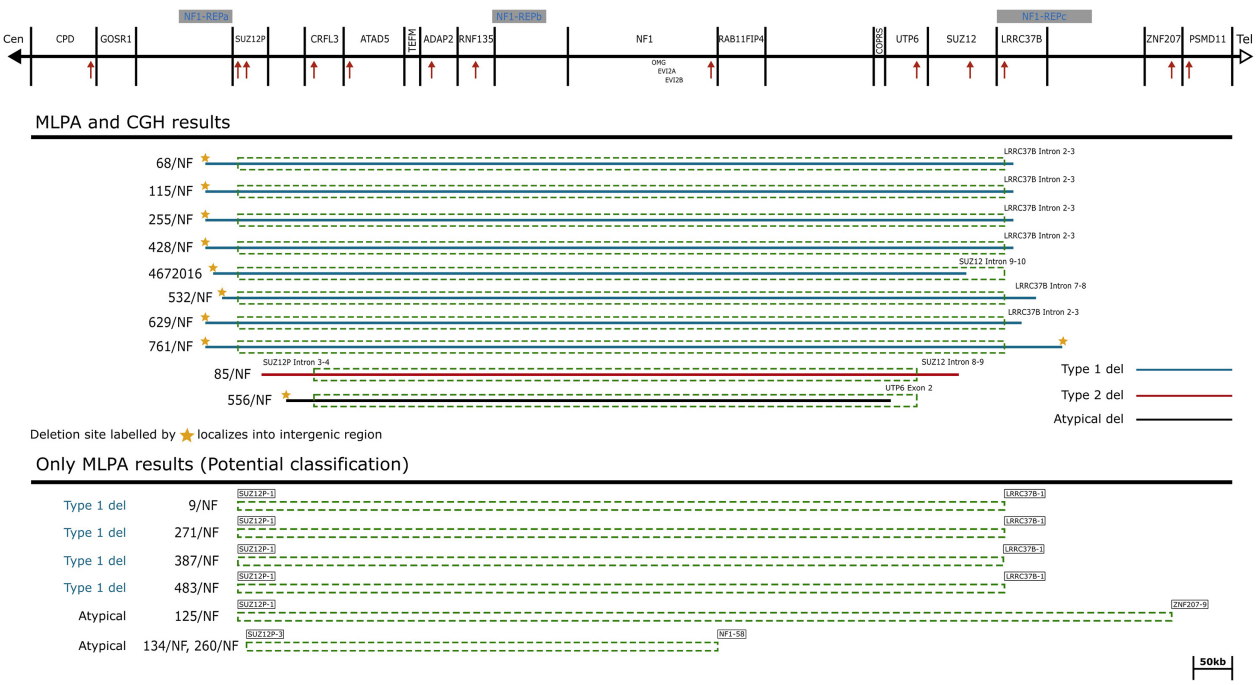


FIGURE 1 | Schematic representation of the *NF1* gene and flanking regions. The affected genes and NF1-REP regions are schematically displayed at the top of the figure. Localization of MLPA probes are demonstrated by red arrows. Solid lines symbolize the deletion range with known breakpoints determined by aCGH probes. Dotted rectangles correspond to the deleted range determined by MLPA probes. Deletion types are marked by colored solid lines, blue: type-1 deletions, red: type-2 deletion and black: the suggested atypical deletion. The last probes contained by the deletion are explicitly displayed at the ends of the deletion ranges.

well, since she has a positive family history (her mother and her grandmother were also affected, however, without laboratory diagnosis) and the MLPA results (peak ratios were between 0.49-0.55) also supported this assumption. MLPA peak ratios were between 0.49 and 0.55 also for patient 125/NF, therefore we supposed this patient to be a non-mosaic as well.

Clinical Characterization of Our Patients With Different Type of *NF1* Microdeletion

Several clinical features and neuropsychological manifestations belonging to eight major categories were selected for consideration for genotype-phenotype association analysis (**Table 1**). The frequency of each clinical feature that appeared in patients with type-1 *NF1* microdeletion is compared with frequencies observed in our control group, i.e., patients with intragenic *NF1* mutation (**Supplementary Tables 4**, **5**).

Dysmorphic Features

Facial dysmorphism was described in 9 of the 17 patients investigated (53%). It was present in 8 out of 12 patients with type-1 *NF1* deletion (67%) and in 1 out of 4 atypical *NF1* deletion (25%) patient cohort. The prevalence of hypertelorism was similar to that of facial dysmorphism, however the distribution among the deletion types was different. This clinical feature was found to roughly the same extent in type-1 deletion and atypical deletion cases (58% vs 50%, respectively). Facial asymmetry was noted only in 3 out of 12 patients with type-1 deletion. Coarse facial appearance was frequent in type-1 deletion patients (8 out of 12 patients, 67%), it was present also in the type-2 deletion patient, though it was absent in our atypical cases. Large hand and feet seem to be a characteristic dysmorphic feature of *NF1* microdeletion patients as well, since the majority of our patients with type-1 deletion (67%, 8 out of 12) showed this trait and it was also noted in the type-2 patient. Dysmorphic features were rare events in our intragenic *NF1* patient population. Of the examined dysmorphic traits only hypertelorism and facial asymmetry were found with the frequency of 18% (6 out of 33 controls) or 6% (2 out of 33 controls), respectively.

Skin Manifestations

Café-au-lait spots (CALs) were observed in each patient in our study regardless of the type of the deletion they have. Axillary and inguinal freckling occurred also in high frequency in our patient cohort. It was more common within the type-1 deletion group, 10 out of 12 patients (83%) presented this skin manifestation. In atypical deletion group 3 out of 4 patients (75%) displayed this feature, however, it was absent in the type-2 deletion patient. Moreover, another skin manifestation, i.e. excess soft tissue in hands and feet was observed among our patients, though at a lower frequency. In type-1 deletion group it was noted in 4 out of 12 patients (33%), it developed in a patient with type-2 deletion also, in contrast, it was not found in the atypical deletion patients. Skin manifestations are characteristic for intragenic *NF1* patients as well. CALs were presented in 91% (30 out of 33) of our patients

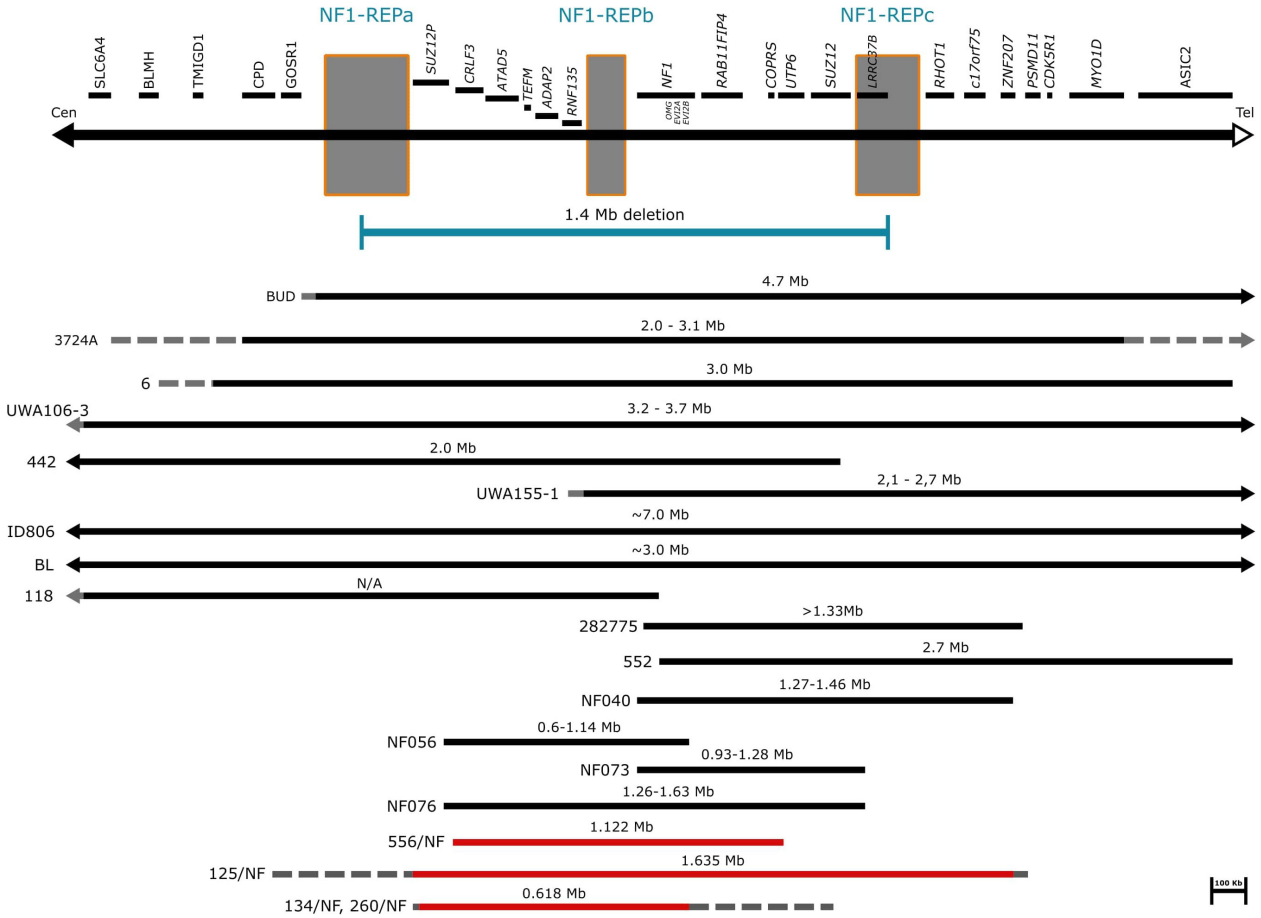


FIGURE 2 | Schematic representation of atypical NF1 deletions. The affected genes and NF1-REP regions schematically are displayed at the top of the figure. Horizontal black bars represent the already known atypical *NF1* cases. Solid lines indicate the deleted regions, dotted lines indicate the possibly deleted regions. Horizontal red bars refer to our cases. Solid lines represent the deleted regions, while dotted lines suggest the potential deletion range.

and the frequency of axillary and inguinal freckling was 52% (17 out of 33 controls).

Neurofibromas and Other Tumors

Subcutaneous neurofibromas were found more common in type-1 deletion patient cohort compared to type-2 and atypical groups. They were observed in 7 out of 12 patients (58%) with type-1 deletion, in 1 out of 4 patients (25%) with atypical microdeletion, though none occurred in the patient with type-2 deletion. The prevalence of cutaneous neurofibromas appears to be less frequent in our patient cohort, it was observed in only one patient with type-1 deletion. However, it is important to mention that 14 out of 17 patients were children and furthermore 10 out of 14 were under 10 years old at the age of examination.

Externally observable plexiform neurofibromas were seen in only 2 patients with type-1 deletion, in a 21-year-old boy and a 17-year-old girl. None of the patients with type-2 or atypical microdeletions presented this type of neurofibromas. However, this is worth to mention that whole-body MRI was not performed routinely in our patients, therefore we have no information about the internally occurring plexiform neurofibromas.

Spinal neurofibromas were found in the type-1 microdeletion group only, however, within this group, the prevalence was low, it developed in 2 out of 12 patients (17%). However, the observed low occurrence is probably the result of the fact, that spinal MRI is not part of the routine procedure in our patient management.

Optic pathway glioma (OPG) was detected by MRI in 4 patients and it was not symptomatic in any of these cases. It was more common in the atypical group with 50% prevalence. Moreover, it developed in 2 out of 12 patients (17%) with type-1 deletion but it was absent in the patient with type-2 deletion. Among the control patients 2 symptomatic and 2 asymptomatic OPG were observed.

Malignant peripheral nerve sheath tumors (MPNST) were observed in 2 of our patients, both belonging to type-1 deletion group. None of the patients with type-2 or atypical microdeletions displayed this type of tumor. MPNSTs show age-related penetrance and our patient cohort consisted of mainly children under 17 years, therefore it is not surprising to detect low occurrence among our patients. However, both patients presenting MPNSTs were adult or nearly adult (36 years and

17 years old, respectively), consequently the frequency of this type of tumor was high (50%, 2 out of 4) among adult patients.

Among our intragenic *NF1* patients, subcutaneous fibromas were found with 30% (10 out of 33) frequency, the occurrence of cutaneous and plexiform neurofibromas were 18% (6 out of 33) or 6% (2 out of 33), respectively. Spinal neurofibromas were observed in 3% (1 out of 33) of our patients. Moreover, 12% (4 out of 33) of this patient cohort developed optic pathway glioma, however, no malignant peripheral nerve sheath tumors occurred.

Skeletal Anomalies

Anomalies of the skeletal system were detected in almost all of our patients (94%, 16 out of 17). The most frequent skeletal anomaly was macrocephaly, which was observed in 9 out of 17 patients (53%). This clinical feature was common in type-1 microdeletion cohort with 58% prevalence, whereas in atypical cohort only one patient (25%) presented this symptom.

Scoliosis was noted in 7 out of 17 patients studied here (41%). It was more frequent in patients with type-1 *NF1* microdeletion than in patients with other type of *NF1* microdeletions. Interestingly, there were only 2 patients who presented scoliosis together with macrocephaly.

Pectus excavatum was observed in 35% of our patient cohort. In contrast to scoliosis, this skeletal anomaly was more frequently observed in patients with atypical microdeletion (50%) as compared to type-1 deletion group (33%).

Bone cysts were found in only one patient with type-1 microdeletion.

None of our patient displayed pes cavus, however, other foot deformities such as pes planus was observed in 3 patients.

Interestingly, skeletal anomalies were the leading manifestations in our patient with type-2 deletion. She had macrocephaly, scoliosis, bilateral dislocation of the elbow and wrist joint. Moreover, absorption of the tibial malleolus was observed and she developed osseous malignancy as well.

Skeletal anomalies were less frequently observed in the intragenic *NF1* patient group (33%). Of these, scoliosis occurred most frequently with 21% prevalence. Macrocephaly and pectus excavatum were noted in 9% of the patients and 3% of them presented pes cavus.

Ocular Manifestations

Ocular manifestations were observed in 7 of 17 our patients (41%). Lisch nodule, one of the characteristic hallmarks of type 1 neurofibromatosis, was noted only in 3 out of 12 patients with type-1 deletion and in the patient with type-2 deletion, however, it was not observed in the atypical patient cohort. Moreover, other ocular manifestations, such as visual disturbance, strabismus and proptosis were noticed in 2 patients with type-1 deletion and in the type-2 deletion patient. One of the patients had hypermetropia, while the others had myopia. The frequency of ocular manifestations was similar in the intragenic *NF1* patient cohort. Lisch nodule was noted in 21% (7 out of 33) of the patients and 15% (5 out of 33) presented visual disturbances as well. One patient had myopia, two patients had hypermetropia, and two other patients had anisometropia. However, strabismus was not observed.

Neuropsychological Manifestations

Significant delay in cognitive development and general learning difficulties were observed with high frequency (75%, 9 out of 12) in type-1 patients. Furthermore, along with the previous features, speech difficulties occurred in 67% (8 out of 12) of this patient group. One patient had an IQ below 70 and 2 patients showed attention deficit hyperactivity disorder (ADHD). IQ measurement was performed in only two among our type-1 patients (761/NF IQ:77, 9/NF IQ:47), however, all of our pediatric patients attended regular kindergarten or school, except the one with IQ = 47, and five of them have special educational needs. Therefore, we supposed these patients are not intellectually disabled, so we marked them as negative for IQ < 70 criteria in **Table 1**. Majority of these neuropsychological features were not found in atypical patient cohort (patient 556/NF IQ:89) and in the type-2 patient. Only a significant delay in cognitive development was noted in 25% (1 out of 4) of atypical patients and the type-2 patient suffered from general learning difficulties.

Structural brain abnormalities were not observed in our patients, however, T2 hyperintensities were found in the majority of our patients. It was present with 75% (9 out of 12) prevalence in type-1 deletion patient cohort, with 25% (1 out of 4) prevalence in atypical group and also in the patient with type-2 deletion. Nevertheless, we did not find any correlation between the age of our patients and the T2 signal intensities.

Muscular hypotonia and coordination problems (25% and 33%, respectively) were documented in patients with type-1 deletion. None of these neurological symptoms were found in our type-2 and atypical deletion groups.

Epilepsy and nerve pain were not noted in our patients. One patient with type-1 deletion complained of headache.

Neuropsychological manifestations were not common among the patients with *NF1* intragenic mutation. 3% (1 out of 33) of our patients presented significant delay in cognitive development, speech difficulties and epilepsy. Moreover, general learning difficulties were noted with a bit higher frequency (15%, 5 out of 33). Muscular hypotonia was observed in 12% (4 out of 33) of our patients and T2 hyperintensities were found in 39% (13 out of 33) of them.

Connective Tissue Anomalies and Cardiac Abnormalities

Connective tissue anomalies and heart abnormalities were a very rare event in our patient cohort. Hyperflexibility of joints was observed in 2 out of 12 type-1 deletion patients (17%). Such manifestation was not present in our patients with type-2 or atypical deletions. Among the cardiac abnormalities atrial septal defect was observed in one patient with atypical microdeletion. Moreover, hypertrophic cardiomyopathy was observed in one patient (8%) and patent ductus arteriosus (PDA) occurred in another patient (8%) with type-1 microdeletion. No congenital heart defect, pulmonary stenosis, ventricular septal defect, aortic stenosis, aortic dissection, mitral valve prolapses, mitral valve insufficiency, aortic valve insufficiency was found in any of the deletion groups. It should mention that two of our patients were not investigated by cardiac ultrasound.

These manifestations were rare in our patients with *NF1* intragenic mutation as well. Among the cardiac abnormalities only ventricular septal defect was observed at birth in one patient and 6% (2 out of 33) of our patients developed joint laxity.

Other Features

Some rare clinical manifestations were observed in our patient group. Obesity, hearing impairment, immune deficiency and milk protein allergy, however it is hard to tell whether these symptoms are associated with the observed large deletion or the results of an independent event.

DISCUSSION

The *NF1* gene was discovered in Viskochil et al. (1990), somewhat later the first case with large *NF1* microdeletion was published in Kayes et al. (1992). Several attempts were made to establish genotype-phenotype correlations which finally suggested a more severe clinical phenotype among patients with *NF1* microdeletion than patients with intragenic *NF1* mutations. However, certain variability of clinical symptoms has been observed among individuals with *NF1* microdeletions.

In this study, we have identified 17 patients with large *NF1* microdeletion. Among them 8 proved to be a type-1 microdeletion carrier by aCGH, 4 more patients are supposed to belong to type-1 group based on MLPA results, 1 patient has type-2 deletion and 4 patients possess atypical deletions. Somatic mosaicism with an extent of ca. 30% was detected in one patient with atypical *NF1* microdeletion. Comparison of clinical characterization of our patients with the published data on intragenic and microdeletion NF1 patients was performed to reveal distinct phenotype-genotype correlations. Moreover, the frequencies of phenotypic features in our patients with *NF1* microdeletion and with type-1 deletion were compared to frequencies observed in our patients with intragenic *NF1* mutation as well (**Supplementary Tables 3–5**).

A similar difference was found between our patients with intragenic *NF1* mutation and *NF1* microdeletion in several clinical features when comparing to those previously published by others (**Table 2**). Mainly the occurrence of dysmorphic features, subcutaneous neurofibromas, skeletal anomalies and neurobehavior problems showed significant difference. Moreover, remarkable differences in certain clinical features were observed between our patients with *NF1* microdeletion and the previously published cases with large *NF1* deletions. However, it is important to emphasize that the majority of our patients (13 out of 17) were less than 15 years old at the time of the examination. There are only few studies (Kehrer-Sawatzki et al., 2020) that demonstrated pediatric clinical data, the majority of phenotypic data published previously originated mainly from adult patient populations.

Type-1 deletion represents the largest group of *NF1* microdeletion cohort with an estimated 70-80% prevalence (Pasmant et al., 2010; Messiaen et al., 2011). The occurrence of this type of deletion among our patients was somewhat similar (70%). Significant number of articles were published on this type of deletion, however, these reports indicate that the clinical phenotype associated with *NF1* microdeletions show a certain degree of variability in the frequency of some clinical features (**Table 2**) (Mensink et al., 2006; Mautner et al., 2010; Pasmant et al., 2010; Bianchessi et al., 2015; Zhang et al., 2015). Dysmorphic features are common in individuals with large *NF1* deletions, whereas they occur rarely among intragenic NF1 patient population. Among these features facial dysmorphism is one of the most characteristic hallmarks of patients with *NF1* microdeletion. In our type-1 patient cohort 67% of the affected individuals possess this manifestation. At the same time in a large study performed by Mautner et al. involving 29 patients (Mautner et al., 2010), the majority of the cases (ca 90%) had facial dysmorphism. However, Pasmant and Zhang observed this feature with lower frequency (Pasmant et al., 2010; Zhang et al., 2015). Nevertheless, all of these data indicate that facial dysmorphic features are very frequent in type-1 deletions. Another dysmorphic feature which can be seen more often in microdeletion patients is the observed large hands and feet. It occurred with 67% prevalence in our patient cohort, it was observed in 46% of patients by Mautner (Mautner et al., 2010), however, it was not stated by others. Another observable difference can be seen in the number of the detected neurofibromas. Previous studies established an early-onset of neurofibromas among *NF1* microdeletion patients. While the frequency of the detected subcutaneous neurofibromas in our patients was close to that observed by others (58 vs 76%), the occurrence of cutaneous or plexiform neurofibromas was remarkably lower in our patients compared to other patient groups (8 vs. 86% and 17 vs. 76%, respectively). However, it is worth to highlight, that our patient cohort mainly consisted of children and adolescents, and 9 out of 17 were less than 10 years old at the time of examination. Cutaneous neurofibromas show age-related penetrance and they usually appear in adulthood, therefore this may contribute to the difference in frequency observed by us and by others. Nevertheless, a high frequency (60%) of cutaneous neurofibromas was observed among children by Kehrer-Sawatzki in a recent study (Kehrer-Sawatzki et al., 2020). The high prevalence of subcutaneous neurofibromas in type-1 *NF1* patients is important to consider, since they are associated with mortality in NF1 disease (Tucker et al., 2005). Patients with subcutaneous neurofibromas possess a higher risk for the development of MPNSTs. In addition, the presence of plexiform neurofibromas possess a risk for development of malignant tumor as well (Waggoner et al., 2000). More pronounced alteration can be seen in the cognitive ability. Although, significant delay in cognitive development was found more frequently in our type-1 patients, the prevalence of intellectual disability was less pronounced. Moreover, overgrowth, which is characteristic for type-1 *NF1* microdeletion, was observed as much as by others, however, connective tissue anomalies were fairly less frequent among our patients. It was common among Mautner's patients (72%), but it was rare (8%) in our patient cohort.

Type-1 deletion harbors 14 protein coding genes and 4 microRNA genes. Some of the genes co-deleted with *NF1* may have an influence on the clinical manifestation observed in

patients with *NF1* microdeletion, thus affecting the severity of the disease (Kehrer-Sawatzki et al., 2017). Haploinsufficiency of certain genes may contribute to dysmorphic facial features, overgrowth and reduced cognitive capability (*RNF135*) (Tastet et al., 2015) or heart defects (*ADAP2*) (Venturin et al., 2014), whereas others might have tumor suppressive function, thus their deletion promote tumor development (*SUZ12, ATAD5*) (Bell et al., 2011; Zhang et al., 2014). Although the size of the deletion and the gene content is almost the same in all patients with type-1 deletion, they demonstrate a notable clinical variability. This observation may suggest that differences in the unique genomic architecture of the patients may also contribute to the observed variability of the clinical phenotypes.

Type-2 deletions account for 10-20% of *NF1* large deletion cases according to previous studies. In our patient cohort one patient and her asymptomatic mother carries this type of large *NF1* deletion. Because of the missing phenotypic signs, we suppose that the mother should be a mosaic patient. In type-2 deletions existence of somatic mosaicism is a frequently observed phenomenon, these deletions arise during post-zygotic cell division and are associated with a milder clinical phenotype. Vogt et al. reported 18 patients with type-2 deletion, 16 of whom proved to be mosaic cases (Vogt et al., 2011). In another study the same research group identified 27 of 40 patients with mosaicism determined by FISH. That paper did not contain clinical information, because it was focused on the possible molecular mechanism behind type-2 deletion formation (Vogt et al., 2012). Only a few non-mosaic type-2 cases with detailed phenotype have been published so far (**Table 3**; Vogt et al., 2011; Zhang et al., 2015). These patients share common features, half of which can be found in our patient as well. However, some characteristic hallmarks of *NF1* microdeletion symptoms are missing from our patient's phenotype or they are presented in a mild form. This may originate from her young age (13 years). She does not have any type of externally observable neurofibromas, cardiac manifestations, those that may manifest as early as childhood, and neurobehavioral problems, whereas these features were noted in the majority of the published cases. Moreover, frequent skin manifestation such as freckling was not observed in our patient. These traits occurred in other known type-2 patients. The unique feature of our patient is that the whole clinical picture is dominated by skeletal anomalies. She underwent a number of operations affecting the skeletal system. Moreover, absorption of the tibial malleolus was observed and she developed osseous malignancy as well. After all her clinical picture possesses many features frequently observed in patients with large *NF1* deletion. Although type-2 deletions are typically 1.2 Mb in size, the exact localization of the breakpoints are presumably different in our patient and in the published cases. This may result in the removal of certain regulatory factors which may finally lead to the observed variability in the phenotype.

Atypical deletions form a heterogeneous group of *NF1* microdeletions regarding the clinical manifestations they cause as well as the size and location of the deletion. Moreover, somatic mosaicism can be frequently observed among these patients which may lead to a milder phenotype. The occurrence of atypical cases is around 8-10% among patients with *NF1* microdeletion, however, in our patient cohort we observed a higher frequency (23%) and only one patient displayed mosaicism. Around 20 patients with atypical deletion were published so far without recurrent breakpoints (Kayes et al., 1992; Upadhyaya et al., 1996; Cnossen et al., 1997; Dorschner et al., 2000; Riva et al., 2000; Kehrer-Sawatzki et al., 2003, 2005, 2008; Venturin et al., 2004a,b; Mantripragada et al., 2006; Zhang et al., 2015). In our study three distinct, novel deletions were identified. The deletions in the published cases show remarkable overlaps with those observed in our patients, though in our cases the deletions are typically smaller (**Figure 2**). However, the clinical pictures of the known cases show hardly any overlapping symptoms apart from the major diagnostic criteria for NF1 (**Table 4**). Remarkable difference can be seen in dysmorphic features, neuropsychological manifestations and the presence of various neurofibromas. Dysmorphic features such as facial dysmorphia, coarse face, facial asymmetry and large hands and feet are characteristic hallmarks of *NF1* microdeletions. They were observed in the majority of patients with type-1 *NF1* microdeletion (**Table 2**) and it was noted at least in half of the atypical cases identified so far, however, in our patient cohort only one patient displayed facial dysmorphia and another had hypertelorism. Moreover, these features were not observed in patients described by Zhang et al. (2015). In addition, notable divergence can be observed in the occurrence of various neurofibromas among the atypical *NF1* microdeletion patients. All the patients in Zhang's study manifested cutaneous or plexiform neurofibromas, 6 out of 11 other published cases had various type of neurofibromas, whereas in our study only one patient has developed subcutaneous neurofibromas. This discrepancy may be related to the age of the patients. It is a known phenomenon that the number of the neurofibromas may increase with the age of the patient. Among atypical cases the majority of the patients who presented any type of neurofibromas were teenagers or young adults. In our patient cohort, which consisted of mainly children under 10 years, the only one who had subcutaneous neurofibroma was 40 years old. In addition, observable difference can be found among the neuropsychological manifestation. These features were almost absent in our patients, only one showed significant delay in cognitive development, however, moderate to severe intellectual disability or severe learning disability were noted in almost all patients carrying larger deletion than our patients. In an atypical deletion the gene content of the deleted region has an effect on the phenotypic manifestations, particularly the genes with intolerance of haploinsufficiency are likely to have pathological consequences. **Table 5** summarized the haploinsufficiency intolerant genes in all cases published so far including this study. Although in 3 out of 4 patients of ours only MLPA measurements were feasible, the deletion of one more haploinsufficiency intolerant gene, namely *RAB11FIP4,* may be expected beyond those demonstrated in **Table 5**. The exact role of this gene in the disease pathogenesis is not clear. Previous studies (Descheemaeker et al., 2004; Ottenhoff et al., 2020) revealed that *NF1* microdeletion genotype is associated with a lower cognitive ability compared with intragenic *NF1* genotype. Co-deletion of

TABLE 2 | Clinical features of patients with type-1 *NF1* microdeletion.

System involvement/ manifestations	Clinical features	Frequency in patients with type-1 *NF1* microdeletions (%)					Frequency in *NF1* non-deleted patients (%)	
		This study (*n* = 12)	Kehrer-Sawatzki et al., 2017 (*n* = 29)	Pasmant et al., 2010 (*n* = 44)	Zhang et al., 2015 (*n* = 7)	Bianchessi et al., 2015 (*n* = 11)	This study (*n* = 33)	Kehrer-Sawatzki et al., 2017 (*n* = 29)
Dysmorphic features	Facial dysmorphism	67	90	54.8	43	n.d.	0	n.d.
	Hypertelorism	58	86	n.d.	n.d.	n.d.	18	n.d.
	Facial asymmetry	25	28	n.d.	n.d.	n.d.	6	8
	Coarse face	67	59	n.d.	n.d.	n.d.	0	n.d.
	Broad neck	8	31	n.d.	n.d.	n.d.	0	n.d.
	Large hands and feet	67	46	n.d.	n.d.	n.d.	0	n.d.
Skin manifestations	Café-au-lait spots	100	93	20.8	100	100	91	86-99
	Axillary and inguinal freckling	83	86	86.4	57	72.7	52	86-89
	Excess soft tissue in hands and feet	33	50	n.d.	n.d.	n.d.	0	n.d.
	Subcutaneous neurofibromas	58	76	37.2-41.8	29	45.5[#]	30	48
	Cutaneous neurofibromas	8	86	15.4-48.7	57	45.5[#]	18	38-84
	Plexiform neurofibromas	17	76	0.6	29	27.3	6	15-54
Education and behavior problems	SDiCD	75	48	n.d.	14	36.4	3	17
	General learning difficulties	75	45	85.7	n.d.	18.2	15	31-47
	Speech difficulties	67	48	n.d.	29	0	3	20-55
	IQ < 70	8	38	n.d.	14	36.4	0	7-8
	ADHD	17	33	n.d.	n.d.	0	6	38-49
Skeletal manifestations	Skeletal anomalies	92	76	31+	14	45.5+	33	31
	Scoliosis	42	43	31	0	9.1	21	10-28
	Pectus excavatum	33	31	n.d.	n.d.	n.d.	9	12-50
	Bone cysts	8	50	n.d.	n.d.	0	0	1
	Hyperflexibility of joints	8	72	n.d.	n.d.	n.d.	6	n.d.
	Pes cavus	n.d.	17	n.d.	n.d.	n.d.	3	n.d.
	Macrocephaly	58	39	11.5	14	45.5	9	24-45
Neurological manifestations	Muscular hypotonia	25	45	n.d.	n.d.	n.d.	12	27
	Epilepsy	0	7	n.d.	n.d.	0	3	4-13
	MPNST	17	21	7.1	0	*	0	2-7
	Spinal neurofibromas	17	64	n.d.	n.d.	n.d.	3	24-30
	T2 hyperintensities	75	45	n.d.	29	n.d.	39	34-79
Ocular manifestations	Visual disturbance	17	n.d.	n.d.	14	n.d.	15	n.d.
	Lisch nodules	25	93	40	14	45.5	21	63-93
	Strabismus	17	NA	n.d.	14	n.d.	0	NA
	Optic pathway gliomas	17	19	15	n.d.	0	12	11-19
Developmental problem	Tall-for-age stature	58	46	22.2	n.d.	n.d.	0	n.d.
Heart problems	Congenital heart defects	0	29	n.d.	n.d.	n.d.	0	2

*n.d., not determined; NA, not assessed or no data available; [#]no straightforward information (only referenced as neurofibroma); *it is not clear from the manuscript (it was mentioned that 18.2% of patient had tumors); + it may be higher (there were data for scoliosis and macrocephaly only); SDiCD, significant delay in cognitive development; MPNST, malignant peripheral nerve sheath tumors; ADHD, attention deficit hyperactivity disorder.*

TABLE 3 | Clinical features of patients with type-2 *NF1* microdeletions.

Clinical features of patients with type-1 *NF1* microdeletions (frequency observed,%)			**Presence or absence of the features in patients with "non-mosaic" type-2 *NF1* deletions**			
Patients	*n* = 29	*n* = 12	078	P. 2429	P. 2358	85/NF
Reference	Kehrer-Sawatzki et al., 2017	This study	Zhang et al., 2015	Roehl et al., 2010; Vogt et al., 2012	Roehl et al., 2010; Vogt et al., 2012	This study
CALs	93%	100%	+	+	+	+
Freckling	86%	83%	−	+	+	−
Lisch nodule	93%	25%	?	+	+	+
Cutaneous neurofibromas	86%	8%	+	+ (multiple)	−	−
Subcutaneous neurofibromas	76%	58%	+	+ (multiple)	+	−
Plexiform neurofibromas	76%	17%	−	+ (multiple)	+	−
Facial dysmorphism	90%	67%	−	+	+	−
Large hands and feet	46%	67%	N/A	+	+	+
Macrocephaly	39%	58%	−	+	+	+
Tall stature	46%	58%	N/A	−	−	−
Learning disabilities	48%	75%	?	+	+ (mild)	+
Attention deficits	33%	17%	?	+	+	−
Scoliosis	43%	42%	+	−	N/A	+
Hyperflexibility of the joints	72%	8%	N/A	+	+	−
MPNST	21%	17%	−	+	−	−
T2 hyperintensities	45%	75%	N/A	−	+	+
Muscular hypotonia	45%	25%	N/A	N/A	+	−
Congenital heart defects	21%	0%	N/A	+	+	−

−, absent; +, present; N/A, not assessed or no data available; ? unclear result from the original article. CALs, café-au-lait spots; MPNST, malignant peripheral nerve sheath tumors.

genes adjacent to *NF1*, such as *OMG* and *RNF135* are supposed to contribute to the observed decreased cognitive ability (Kehrer-Sawatzki et al., 2017). *OMG* gene encodes the oligodendrocyte myelin glycoprotein which plays an important role in early brain development (Martin et al., 2009). Moreover, *OMG* is associated with intellectual disability and neuropsychiatric disorders (Bernardinelli et al., 2014). In addition, a rare allele of *RNF135* gene has been found with higher frequency in patients with autism (Tastet et al., 2015). Although the deletion identified in our patients encompass *OMG* and *RNF135* genes as well, our patients hardly displayed neuropsychiatric symptoms. This observation implies that beyond the *OMG* and *RNF135* deletion further factors are also necessary for the development of intellectual disability or neuropsychiatric manifestations in patients with *NF1* microdeletions. Contrary to our cases, high load of internal tumors were observed in a number of patients with larger atypical deletion. Several genes (*ATAD5, COPRS, UTP6* and *SUZ12*) in the 17q11.2 region were supposed to be involved in tumorigenesis (Kehrer-Sawatzki et al., 2017). *ATAD5* was affected in our two patients, co-deletion of *ATAD5, COPRS* and *UTP6* was observed in another one. However, none of these patients of ours developed internal tumors. Co-deletion of *ATAD5, COPRS, UTP6* and *SUZ12* genes with *NF1* may possess an increased risk for high tumor load which might lead to the observed high number of tumors in patients with larger atypical deletion. In one of our patients the atypical deletion harbors all of these four genes, however, perhaps due to her young age (i.e., 2 years) no tumors were found at the age of her examination.

Genotype-phenotype analyses among our patients revealed that ones with *NF1* microdeletion more often presented dysmorphic facial features, macrocephaly, large hands and feet, delayed cognitive development and/or learning difficulties, speech difficulties, overgrowth and subcutaneous neurofibromas compared to those with intragenic *NF1* mutations. These features seemed to be characteristic for the patient group with type-1 *NF1* microdeletion, however, some of the above-mentioned traits were absent from the type-2 and atypical *NF1* microdeletion

TABLE 4 | Clinical features of patients with atypical *NF1* microdeletions.

Patient	Age (y)	Gender	Skin manifestations	Neurofibromas	Dysmorphic features	Skeletal manifestations	Ocular Manifestations	Neuropsychological manifestations	Other	References
BUD	14; 18	N/A	CALs, F	Many CNF, SNF	Coarse face	SCS, genu valgum, joint laxity	N/A	SDiCD, ID, T2 hyperintensities	Many ST	Kehrer-Sawatzki et al., 2003
3724A	13	Female	CALs, F	Few CNF	Coarse face, FA, hypertelorism, ptosis, broad lips and nose	PE	LiN	Moderate ID	-	Cnossen et al., 1997
6	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Venturin et al., 2004a,b
UWA106-3	18	Male	CALs, F	Many CNF, PNF, spinal NF	Coarse face, large hands	MA	N/A	SDiCD, IQ 46	Many ST	Dorschner et al., 2000; Kayes et al., 1992
442	18; 26	Male	CALs, F	Multiple SCNF, and many CNF, PNF	Coarse face	SCS	LiN	IQ 76, severe LD	Many ST	Kehrer-Sawatzki et al., 2005
BL	13,5	Male	CALs, F	-	FD, hypertelorism	Skeletal anomalies	-	Severe ID	-	Riva et al., 2000
ID806	3 mo; 3; 4	Male	CALs, F	-	Narrow palpebral fissures, ptosis, low set, rotated ears, prominent maxilla	-	-	Marked developmental delay, SP, seizure	-	Upadhyaya et al., 1996
UWA155-1	27	N/A	-	Multiple CNF, spinal NF	Coarse face, ptosis, large hands and feet	MA	-	Moderate ID	MPNST	Dorschner et al., 2000
118	5	Male	CALs, F	N/A	-	-	OPG	Seizure, no LD	-	Venturin et al., 2004b
282775	n.d.	N/A	CALs	-	Noonan-like FD	-	-	PD, SP	-	Mantripragada et al., 2006
552	20	Female	CALs, F	2 PNF, 4 SIN NF	Large hands and feet	PE, lumbar lordosis, pedes valgoplanus	LiN, visual disturbance	Mild ID, severe LD, SP, hypotonia	-	Kehrer-Sawatzki et al., 2008
NF040	1	Female	CALs	PNF	-	-	*	*	-	Zhang et al., 2015
NF056	60	Female	CALs, F	CNF	-	-	*	*	-	
NF073	25	Female	CALs, F	CNF	-	-	*	*	-	
NF076	36	Female	CALs	CNF	-	-	*	*	-	
556/NF	10	Male	CALs, F	-	-	Bilateral PP	OPG	-	-	this study
125/NF	2	Female	CALs, F	-	-	PE	-	-	-	
134/NF	40	Female	CALs	SCNF	Hypertelorism	SCS	-	-	-	
260/NF	8	Male	CALs, F	-	FD, hypertelorism	PE, MA	OPG	SDiCD, T2 hyperintensities	ASD	

*CALs, café-au-lait spots; F, freckling; FA, facial asymmetry; FD, facial dysmorphy; CNF, cutaneous neurofibroma; SCNF, subcutaneous neurofibroma; PNF, plexiform neurofibroma; SIN NF, small intramuscular nodular neurofibroma; ST, spinal tumors; MPNST, malignant peripheral nerve sheath tumors; SDiCD, significant delay in cognitive development; ID, intellectual disability; LD, learning difficulties; SP, speech delay; PD, psychomotor delay; SCS, scoliosis; PE, pectus excavatum; MA, macrocephaly; PP, pes planus; LiN, Lisch nodule; ASD, atrial septal defect. * unclear results in the original article. NA, no data available.*

TABLE 5 | Size of the deletions and haploinsufficient genes located within the atypical *NF1* deletions.

Patient	Deletion size (Mb)	Haploinsufficient genes (by gnomAD pLI)	References
BUD	4.7	*ATAD5, NF1, OMG, RAB11FIP4, SUZ12, PSMD11, CDK5R1, ASIC2*	Kehrer-Sawatzki et al., 2003
3724A	2.0-3.1	*ATAD5, NF1, OMG, RAB11FIP4, SUZ12, PSMD11, CDK5R1, ASIC2*	Cnossen et al., 1997
6	3	*ATAD5, NF1, OMG, RAB11FIP4, SUZ12, PSMD11, CDK5R1, ASIC2*	Venturin et al., 2004a,b
UWA106-3	3.2-3.7	*ATAD5, NF1, OMG, RAB11FIP4, SUZ12, PSMD11, CDK5R1, ASIC2*	Dorschner et al., 2000; Kayes et al., 1992; Kayes et al., 1994
442	2	*ATAD5, NF1, OMG, RAB11FIP4, SUZ12*	Kehrer-Sawatzki et al., 2005
BL	∼3	*ATAD5, NF1, OMG, RAB11FIP4, SUZ12, PSMD11, CDK5R1, ASIC2*	Riva et al., 2000
ID806	∼7	*ATAD5, NF1, OMG, RAB11FIP4, SUZ12, PSMD11, CDK5R1, ASIC2*	Upadhyaya et al., 1996
UWA155-1	2.1-2.7	*NF1, OMG, RAB11FIP4, SUZ12, PSMD11, CDK5R1, ASIC2*	Upadhyaya et al., 1996
118	N/A	*ATAD5, NF1*	Venturin et al., 2004b
282775	> 1.33	*NF1, OMG, RAB11FIP4, SUZ12*	Mantripragada et al., 2006
552	2.7	*NF1, OMG, RAB11FIP4, SUZ12, PSMD11, CDK5R1, ASIC2*	Kehrer-Sawatzki et al., 2008
40	1.27-1.46*	*NF1, OMG, RAB11FIP4, SUZ12,*	Zhang et al., 2015
56	0.60-1.14*	*ATAD5, NF1, OMG*	
73	0.93-1.28*	*NF1, OMG, RAB11FIP4, SUZ12*	
76	1.26-1.63*	*ATAD5, NF1, OMG, RAB11FIP4, SUZ12*	
556/NF	1.122	*ATAD5, NF1, OMG, RAB11FIP4, SUZ12*	This study
125/NF	1.635*	*ATAD5, NF1, OMG, RAB11FIP4, SUZ12*	
134/NF	0.618*	*ATAD5, NF1, OMG*	
260/NF	0.618*	*ATAD5, NF1, OMG*	

**Results originated from MLPA probes location. The probability of loss of function (pLI) metric were provided by the gnomAD browser (https://gnomad.broadinstitute.org/). According to official description, a transcript's intolerance to variation is measured by predicting the number of variants expected to be seen in the gnomAD dataset and comparing those expectations to the observed amount of variation. The range scales from 0 to 1, where the closer the pLI value is to 1, the more intolerant the gene appears to be to loss of function (LoF) variants. We determined as haploinsufficient a gene if the pLI value was above 0.9, which indicates extreme intolerance to LoF variants (Karczewski et al., 2020).*

patient cohort. Our patient with non-mosaic type-2 *NF1* large deletion had only a few of the typical clinical signs: macrocephaly, large hands and feet as well as learning difficulties. On the other hand, she has a strong skeletal involvement. In our atypical *NF1* microdeletion patient cohort only the facial dysmorphism, delayed cognitive development, macrocephaly and the presence of subcutaneous neurofibromas were noted. Certain clinical symptoms such as congenital heart defects, joint laxity, muscular hypotonia and bone cysts were reported by others in type-1 *NF1* microdeletion patients (Mautner et al., 2010; Kehrer-Sawatzki et al., 2017), but these were not pronounced in our patients. It is worth to mention that manifestations of several symptoms are age dependent, therefore a comprehensive study on the clinical course of patients with different type of *NF1* microdeletion could help to establish diagnostic milestones in these patients' group.

CONCLUSION

In conclusion, in our patient cohort three different types of *NF1* microdeletion have been identified. Although these deletions were associated with different clinical manifestations, possibly due to the deleted gene contents or the deletion of other regulatory DNA elements, patients with *NF1* large deletion showed more severe clinical phenotype compared to individuals with intragenic *NF1* mutations. The identification and in some cases the classification of the *NF1* microdeletions have been feasible using MLPA, a simple, cost-effective technique. This method enabled us to recognize *NF1* microdeletion patients easily among the general NF1 patients. Our study presented additional clinical data related to *NF1* microdeletion patients especially for pediatric patients and it contributes to the better understanding of this type of disorder.

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 Ethics Committee of the University of Pecs (Protocol 8581-7/2017/EUIG). 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

JB conceived and designed the research. GB, RSz, GA, and MCz performed the genetic investigations. KH, AZs, MSz, GyF, VF,

MT, and DN performed the patient examinations. GB and JB completed data analysis and drafted the manuscript. GB prepared the figures and tables. GB, BM, KH, and JB edited and revised the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from the Medical School, University of Pécs (KA 2020-27), GINOP-2.3.3-15-2016-00025 and the National Scientific Research Program (NKFI) K-119540.

ACKNOWLEDGMENTS

We would like to thank all the patients and their family members who participated in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.673025/full#supplementary-material

REFERENCES

Abramowicz, A., and Gos, M. (2014). Neurofibromin in neurofibromatosis type 1 – mutations in NF1gene as a cause of disease. *Dev. Period Med.* 18, 297–306.

Allanson, J. E., Cunniff, C., Hoyme, H. E., McGaughran, J., Muenke, M., and Neri, G. (2009). Elements of morphology: standard terminology for the head and face. *Am. J. Med. Genet. A* 149A, 6–28. doi: 10.1002/ajmg.a.32612

Bass, L., Borbély, S., Jászberényi, M., Lányiné, E. Á, Sarkady, K., Gerebenné, V. K., et al. (1989). "A budapesti binet-teszttel végzett vizsgálatokról," in *A Differenciált Beiskolázás Néhány Mérőeszköze*, eds K. G. Várbíró and T. Vidákovich (Cambridge, MA: Akadémiai Kiadó), 29–49. (in Hungarian).

Bayley, N. (2006). *Bayley Scales of Infant and Toddler Development: Bayley-III*. San Antonio, TX: Harcourt Assessment Psychological Corporation, 7.

Bell, D. W., Sikdar, N., Lee, K. Y., Price, J. C., Chatterjee, R., Park, H. D., et al. (2011). Predisposition to cancer caused by genetic and functional defects of mammalian Atad5. *PLoS Genet.* 7:e1002245. doi: 10.1371/journal.pgen.1002245

Bengesser, K., Cooper, D. N., Steinmann, K., Kluwe, L., Chuzhanova, N. A., Wimmer, K., et al. (2010). A novel third type of recurrent NF1 microdeletion mediated by nonallelic homologous recombination between LRRC37B-containing low-copy repeats in 17q11.2. *Hum. Mutat.* 31, 742–751. doi: 10.1002/humu.21254

Bernardinelli, Y., Nikonenko, I., and Muller, D. (2014). Structural plasticity: mechanisms and contribution to developmental psychiatric disorders. *Front. Neuroanat.* 8:123. doi: 10.3389/fnana.2014.00123

Bianchessi, D., Morosini, S., Saletti, V., Ibba, M. C., Natacci, F., Esposito, S., et al. (2015). 126 novel mutations in Italian patients with neurofibromatosis type 1. *Mol. Genet. Genomic Med.* 3, 513–525. doi: 10.1002/mgg3.161

Cnossen, M. H., de Goede-Bolder, A., van den Broek, K. M., Waasdorp, C. M., Oranje, A. P., Stroink, H., et al. (1998). A prospective 10 year follow up study of patients with neurofibromatosis type 1. *Arch. Dis. Child* 78, 408–412. doi: 10.1136/adc.78.5.408

Cnossen, M. H., van der Est, M. N., Breuning, M. H., van Asperen, C. J., Breslau-Siderius, E. J., van der Ploeg, A. T., et al. (1997). Deletions spanning the neurofibromatosis type 1 gene: implications for genotype-phenotype correlations in neurofibromatosis type 1? *Hum. Mutat.* 9, 458–464. doi: 10.1002/(SICI)1098-100419979:5<458::AID-HUMU13<3.0.CO;2-1

Descheemaeker, M. J., Roelandts, K., De Raedt, T., Brems, H., Fryns, J. P., and Legius, E. (2004). Intelligence in individuals with a neurofibromatosis type 1 microdeletion. *Am. J. Med. Genet. A* 131, 325–326. doi: 10.1002/ajmg.a.30346

Dorschner, M. O., Sybert, V. P., Weaver, M., Pletcher, B. A., and Stephens, K. (2000). NF1 microdeletion breakpoints are clustered at flanking repetitive sequences. *Hum. Mol. Genet.* 9, 35–46. doi: 10.1093/hmg/9.1.35

Firth, H. V., Richards, S. M., Bevan, A. P., Clayton, S., Corpas, M., Rajan, D., et al. (2009). DECIPHER: database of chromosomal imbalance and phenotype in humans using ensembl resources. *Am. J. Hum. Genet.* 84, 524–533. doi: 10.1016/j.ajhg.2009.03.010

Hall, B. D., Graham, J. M. Jr., Cassidy, S. B., and Opitz, J. M. (2009). Elements of morphology: standard terminology for the periorbital region. *Am. J. Med. Genet. A* 149A, 29–39. doi: 10.1002/ajmg.a.32597

Hillmer, M., Wagner, D., Summerer, A., Daiber, M., Mautner, V. F., Messiaen, L., et al. (2016). Fine mapping of meiotic NAHR-associated crossovers causing large NF1 deletions. *Hum. Mol. Genet.* 25, 484–496. doi: 10.1093/hmg/ddv487

Huson, S. M., and Hughes, R. A. C. (1994). *The Neurofibromatoses: A Pathogenetic and Clinical Overview*. London, NY: Chapman & Hall.

Jenne, D. E., Tinschert, S., Reimann, H., Lasinger, W., Thiel, G., Hameister, H., et al. (2001). Molecular characterization and gene content of breakpoint boundaries in patients with neurofibromatosis type 1 with 17q11.2 microdeletions. *Am. J. Hum. Genet.* 69, 516–527. doi: 10.1086/323043

Jett, K., and Friedman, J. M. (2010). Clinical and genetic aspects of neurofibromatosis 1. *Genet. Med.* 12, 1–11. doi: 10.1097/GIM.0b013e3181bf15e3

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alfoldi, J., Wang, Q., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581, 434–443. doi: 10.1038/s41586-020-2308-7

Kayes, L. M., Riccardi, V. M., Burke, W., Bennett, R. L., and Stephens, K. (1992). Large de novo DNA deletion in a patient with sporadic neurofibromatosis 1, mental retardation, and dysmorphism. *J. Med. Genet.* 29, 686–690. doi: 10.1136/jmg.29.10.686

Kayes, L. M., Burke, W., Riccardi, V. M., Bennett, R., Ehrlich, P., Rubenstein, A., et al. (1994). Deletions spanning the neurofibromatosis 1 gene: identification and phenotype of five patients. *Am. J. Hum. Genet.* 54, 424–436.

Kehrer-Sawatzki, H., Kluwe, L., Funsterer, C., and Mautner, V. F. (2005). Extensively high load of internal tumors determined by whole body MRI scanning in a patient with neurofibromatosis type 1 and a non-LCR-mediated 2-Mb deletion in 17q11.2. *Hum. Genet.* 116, 466–475. doi: 10.1007/s00439-005-1265-4

Kehrer-Sawatzki, H., Kluwe, L., Salamon, J., Well, L., Farschtschi, S., Rosenbaum, T., et al. (2020). Clinical characterization of children and adolescents with NF1 microdeletions. *Childs Nerv. Syst.* 36, 2297–2310. doi: 10.1007/s00381-020-04717-0

Kehrer-Sawatzki, H., Kluwe, L., Sandig, C., Kohn, M., Wimmer, K., Krammer, U., et al. (2004). High frequency of mosaicism among patients with neurofibromatosis type 1 (NF1) with microdeletions caused by somatic recombination of the JJAZ1 gene. *Am. J. Hum. Genet.* 75, 410–423. doi: 10.1086/423624

Kehrer-Sawatzki, H., Mautner, V. F., and Cooper, D. N. (2017). Emerging genotype-phenotype relationships in patients with large NF1 deletions. *Hum. Genet.* 136, 349–376. doi: 10.1007/s00439-017-1766-y

Kehrer-Sawatzki, H., Schmid, E., Funsterer, C., Kluwe, L., and Mautner, V. F. (2008). Absence of cutaneous neurofibromas in an NF1 patient with an atypical deletion partially overlapping the common 1.4 Mb microdeleted region. *Am. J. Med. Genet. A* 146A, 691–699. doi: 10.1002/ajmg.a.32045

Kehrer-Sawatzki, H., Tinschert, S., and Jenne, D. E. (2003). Heterogeneity of breakpoints in non-LCR-mediated large constitutional deletions of the 17q11.2 NF1 tumour suppressor region. *J. Med. Genet.* 40:e116. doi: 10.1136/jmg.40.10.e116

Kluwe, L., Siebert, R., Gesk, S., Friedrich, R. E., Tinschert, S., Kehrer-Sawatzki, H., et al. (2004). Screening 500 unselected neurofibromatosis 1 patients for deletions of the NF1 gene. *Hum. Mutat.* 23, 111–116. doi: 10.1002/humu.10299

Lammert, M., Friedman, J. M., Kluwe, L., and Mautner, V. F. (2005). Prevalence of neurofibromatosis 1 in German children at elementary school enrollment. *Arch. Dermatol.* 141, 71–74. doi: 10.1001/archderm.141.1.71

Lopez-Correa, C., Dorschner, M., Brems, H., Lazaro, C., Clementi, M., Upadhyaya, M., et al. (2001). Recombination hotspot in NF1 microdeletion patients. *Hum. Mol. Genet.* 10, 1387–1392. doi: 10.1093/hmg/10.13.1387

Mantripragada, K. K., Thuresson, A. C., Piotrowski, A., Diaz de Stahl, T., Menzel, U., Grigelionis, G., et al. (2006). Identification of novel deletion breakpoints bordered by segmental duplications in the NF1 locus using high resolution array-CGH. *J. Med. Genet.* 43, 28–38. doi: 10.1136/jmg.2005.033795

Martin, I., Andres, C. R., Vedrine, S., Tabagh, R., Michelle, C., Jourdan, M. L., et al. (2009). Effect of the oligodendrocyte myelin glycoprotein (OMgp) on the expansion and neuronal differentiation of rat neural stem cells. *Brain Res.* 1284, 22–30. doi: 10.1016/j.brainres.2009.05.070

Mautner, V. F., Kluwe, L., Friedrich, R. E., Roehl, A. C., Bammert, S., Hogel, J., et al. (2010). Clinical characterisation of 29 neurofibromatosis type-1 patients with molecularly ascertained 1.4 Mb type-1 NF1 deletions. *J. Med. Genet.* 47, 623–630. doi: 10.1136/jmg.2009.075937

Mensink, K. A., Ketterling, R. P., Flynn, H. C., Knudson, R. A., Lindor, N. M., Heese, B. A., et al. (2006). Connective tissue dysplasia in five new patients with NF1 microdeletions: further expansion of phenotype and review of the literature. *J. Med. Genet.* 43:e8. doi: 10.1136/jmg.2005.034256

Messiaen, L., Vogt, J., Bengesser, K., Fu, C., Mikhail, F., Serra, E., et al. (2011). Mosaic type-1 NF1 microdeletions as a cause of both generalized and segmental neurofibromatosis type-1 (NF1). *Hum. Mutat.* 32, 213–219. doi: 10.1002/humu.21418

Neuhausler, L., Summerer, A., Cooper, D. N., Mautner, V. F., and Kehrer-Sawatzki, H. (2018). Pronounced maternal parent-of-origin bias for type-1 NF1 microdeletions. *Hum. Genet.* 137, 365–373. doi: 10.1007/s00439-018-1888-x

Ottenhoff, M. J., Rietman, A. B., Mous, S. E., Plasschaert, E., Gawehns, D., Brems, H., et al. (2020). Examination of the genetic factors underlying the cognitive variability associated with neurofibromatosis type 1. *Genet. Med.* 22, 889–897. doi: 10.1038/s41436-020-0752-2

Park, V. M., and Pivnick, E. K. (1998). Neurofibromatosis type 1 (NF1): a protein truncation assay yielding identification of mutations in 73% of patients. *J. Med. Genet.* 35, 813–820. doi: 10.1136/jmg.35.10.813

Pasmant, E., Sabbagh, A., Spurlock, G., Laurendeau, I., Grillo, E., Hamel, M. J., et al. (2010). NF1 microdeletions in neurofibromatosis type 1: from genotype to phenotype. *Hum. Mutat.* 31, E1506–E1518. doi: 10.1002/humu.21271

Riva, P., Corrado, L., Natacci, F., Castorina, P., Wu, B. L., Schneider, G. H., et al. (2000). NF1 microdeletion syndrome: refined FISH characterization of sporadic and familial deletions with locus-specific probes. *Am. J. Hum. Genet.* 66, 100–109. doi: 10.1086/302709

Roehl, A. C., Vogt, J., Mussotter, T., Zickler, A. N., Spoti, H., Hogel, J., et al. (2010). Intrachromosomal mitotic nonallelic homologous recombination is the major molecular mechanism underlying type-2 NF1 deletions. *Hum. Mutat.* 31, 1163–1173. doi: 10.1002/humu.21340

Steinmann, K., Kluwe, L., Cooper, D. N., Brems, H., De Raedt, T., Legius, E., et al. (2008). Copy number variations in the NF1 gene region are infrequent and do not predispose to recurrent type-1 deletions. *Eur. J. Hum. Genet.* 16, 572–580. doi: 10.1038/sj.ejhg.5202002

Stephens, K., Kayes, L., Riccardi, V. M., Rising, M., Sybert, V. P., and Pagon, R. A. (1992). Preferential mutation of the neurofibromatosis type 1 gene in paternally derived chromosomes. *Hum. Genet.* 88, 279–282. doi: 10.1007/BF00197259

Strassmeier, W. (1980). Early intervention programs for handicapped and retarded children from age 0 to 5. *Int. J. Rehabil. Res.* 3, 533–535.

Summerer, A., Mautner, V. F., Upadhyaya, M., Claes, K. B. M., Hogel, J., Cooper, D. N., et al. (2018). Extreme clustering of type-1 NF1 deletion breakpoints co-locating with G-quadruplex forming sequences. *Hum. Genet.* 137, 511–520. doi: 10.1007/s00439-018-1904-1

Summerer, A., Schafer, E., Mautner, V. F., Messiaen, L., Cooper, D. N., and Kehrer-Sawatzki, H. (2019). Ultra-deep amplicon sequencing indicates absence of low-grade mosaicism with normal cells in patients with type-1 NF1 deletions. *Hum. Genet.* 138, 73–81. doi: 10.1007/s00439-018-1961-5

Tastet, J., Decalonne, L., Marouillat, S., Malvy, J., Thepault, R. A., Toutain, A., et al. (2015). Mutation screening of the ubiquitin ligase gene RNF135 in French patients with autism. *Psychiatr. Genet.* 25, 263–267. doi: 10.1097/YPG.0000000000000100

Tucker, T., Wolkenstein, P., Revuz, J., Zeller, J., and Friedman, J. M. (2005). Association between benign and malignant peripheral nerve sheath tumors in NF1. *Neurology* 65, 205–211. doi: 10.1212/01.wnl.0000168830.79997.13

Upadhyaya, M., Roberts, S. H., Maynard, J., Sorour, E., Thompson, P. W., Vaughan, M., et al. (1996). A cytogenetic deletion, del(17)(q11.22q21.1), in a patient with sporadic neurofibromatosis type 1 (NF1) associated with dysmorphism and developmental delay. *J. Med. Genet.* 33, 148–152. doi: 10.1136/jmg.33.2.148

Uusitalo, E., Leppavirta, J., Koffert, A., Suominen, S., Vahtera, J., Vahlberg, T., et al. (2015). Incidence and mortality of neurofibromatosis: a total population study in Finland. *J. Invest. Dermatol.* 135, 904–906. doi: 10.1038/jid.2014.465

Venturin, M., Carra, S., Gaudenzi, G., Brunelli, S., Gallo, G. R., Moncini, S., et al. (2014). ADAP2 in heart development: a candidate gene for the occurrence of cardiovascular malformations in NF1 microdeletion syndrome. *J. Med. Genet.* 51, 436–443. doi: 10.1136/jmedgenet-2013-102240

Venturin, M., Gervasini, C., Orzan, F., Bentivegna, A., Corrado, L., Colapietro, P., et al. (2004a). Evidence for non-homologous end joining and non-allelic homologous recombination in atypical NF1 microdeletions. *Hum. Genet.* 115, 69–80. doi: 10.1007/s00439-004-1101-2

Venturin, M., Guarnieri, P., Natacci, F., Stabile, M., Tenconi, R., Clementi, M., et al. (2004b). Mental retardation and cardiovascular malformations in NF1 microdeleted patients point to candidate genes in 17q11.2. *J. Med. Genet.* 41, 35–41. doi: 10.1136/jmg.2003.014761

Viskochil, D., Buchberg, A. M., Xu, G., Cawthon, R. M., Stevens, J., Wolff, R. K., et al. (1990). Deletions and a translocation interrupt a cloned gene at the neurofibromatosis type 1 locus. *Cell* 62, 187–192. doi: 10.1016/0092-8674(90)90252-a

Vogt, J., Bengesser, K., Claes, K. B., Wimmer, K., Mautner, V. F., van Minkelen, R., et al. (2014). SVA retrotransposon insertion-associated deletion represents a novel mutational mechanism underlying large genomic copy number changes with non-recurrent breakpoints. *Genome Biol.* 15:R80. doi: 10.1186/gb-2014-15-6-r80

Vogt, J., Mussotter, T., Bengesser, K., Claes, K., Hogel, J., Chuzhanova, N., et al. (2012). Identification of recurrent type-2 NF1 microdeletions reveals a mitotic nonallelic homologous recombination hotspot underlying a human genomic disorder. *Hum. Mutat.* 33, 1599–1609. doi: 10.1002/humu.22171

Vogt, J., Nguyen, R., Kluwe, L., Schuhmann, M., Roehl, A. C., Mussotter, T., et al. (2011). Delineation of the clinical phenotype associated with non-mosaic type-2 NF1 deletions: two case reports. *J. Med. Case Rep.* 5:577. doi: 10.1186/1752-1947-5-577

Vulto-van Silfhout, A. T., van Ravenswaaij, C. M., Hehir-Kwa, J. Y., Verwiel, E. T., Dirks, R., van Vooren, S., et al. (2013). An update on ECARUCA, the European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations. *Eur. J. Med. Genet.* 56, 471–474. doi: 10.1016/j.ejmg.2013.06.010

Waggoner, D. J., Towbin, J., Gottesman, G., and Gutmann, D. H. (2000). Clinic-based study of plexiform neurofibromas in neurofibromatosis 1. *Am. J. Med. Genet.* 92, 132–135.

Zhang, J., Tong, H., Fu, X., Zhang, Y., Liu, J., Cheng, R., et al. (2015). Molecular characterization of NF1 and neurofibromatosis type 1 genotype-phenotype correlations in a Chinese population. *Sci. Rep.* 5:11291. doi: 10.1038/srep11291

Zhang, M., Wang, Y., Jones, S., Sausen, M., McMahon, K., Sharma, R., et al. (2014). Somatic mutations of SUZ12 in malignant peripheral nerve sheath tumors. *Nat. Genet.* 46, 1170–1172. doi: 10.1038/ng.3116

Zickler, A. M., Hampp, S., Messiaen, L., Bengesser, K., Mussotter, T., Roehl, A. C., et al. (2012). Characterization of the nonallelic homologous recombination hotspot PRS3 associated with type-3 NF1 deletions. *Hum. Mutat.* 33, 372–383. doi: 10.1002/humu.21644

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: 11 June 2021
doi: 10.3389/fgene.2021.670727

CFH and *CFHR* Copy Number Variations in C3 Glomerulopathy and Immune Complex-Mediated Membranoproliferative Glomerulonephritis

*Rossella Piras[1], Matteo Breno[1], Elisabetta Valoti[1], Marta Alberti[1], Paraskevas Iatropoulos[1], Caterina Mele[1], Elena Bresin[1], Roberta Donadelli[1], Paola Cuccarolo[1], Richard J. H. Smith[2], Ariela Benigni[1], Giuseppe Remuzzi[1] and Marina Noris[1]**

[1] *Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Bergamo, Italy,* [2] *Molecular Otolaryngology and Renal Research Laboratories, Carver College of Medicine, University of Iowa, Iowa City, IA, United States*

OPEN ACCESS

Edited by:
Attila Gyenesei, University of Pécs, Hungary

Reviewed by:
Prasenjit Mitra, All India Institute of Medical Sciences, Jodhpur, India
Saumel Ahmadi, Washington University in St. Louis, United States

***Correspondence:**
Marina Noris
marina.noris@marionegri.it

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

Received: *22 February 2021*
Accepted: *14 May 2021*
Published: *11 June 2021*

Citation:
Piras R, Breno M, Valoti E, Alberti M, Iatropoulos P, Mele C, Bresin E, Donadelli R, Cuccarolo P, Smith RJH, Benigni A, Remuzzi G and Noris M (2021) CFH and CFHR Copy Number Variations in C3 Glomerulopathy and Immune Complex-Mediated Membranoproliferative Glomerulonephritis. Front. Genet. 12:670727. doi: 10.3389/fgene.2021.670727

C3 Glomerulopathy (C3G) and Immune Complex-Mediated Membranoproliferative glomerulonephritis (IC-MPGN) are rare diseases characterized by glomerular deposition of C3 caused by dysregulation of the alternative pathway (AP) of complement. In approximately 20% of affected patients, dysregulation is driven by pathogenic variants in the two components of the AP C3 convertase, complement C3 (*C3*) and Factor B (*CFB*), or in complement Factor H (*CFH*) and Factor I (*CFI*), two genes that encode complement regulators. Copy number variations (CNVs) involving the *CFH*-related genes (*CFHRs*) that give rise to hybrid FHR proteins also have been described in a few C3G patients but not in IC-MPGN patients. In this study, we used multiplex ligation-dependent probe amplification (MLPA) to study the genomic architecture of the *CFH-CFHR* region and characterize CNVs in a large cohort of patients with C3G (n = 103) and IC-MPGN (n = 96) compared to healthy controls (n = 100). We identified new/rare CNVs resulting in structural variants (SVs) in 5 C3G and 2 IC-MPGN patients. Using long-read single molecule real-time sequencing (SMRT), we detected the breakpoints of three SVs. The identified SVs included: 1) a deletion of the entire *CFH* in one patient with IC-MPGN; 2) an increased number of *CFHR4* copies in one IC-MPGN and three C3G patients; 3) a deletion from *CFHR3*-intron 3 to *CFHR3-3′UTR* ($CFHR3_{4-6}\Delta$) that results in a FHR3-FHR1 hybrid protein in a C3G patient; and 4) a $CFHR3_{1-5}$-$CFHR4_{10}$ hybrid gene in a C3G patient. This work highlights the contribution of *CFH-CFHR* CNVs to the pathogenesis of both C3G and IC-MPGN.

Keywords: C3 glomerulopathy (C3G), immune complex-mediated membranoproliferative glomerulonephritis (IC-MPGN), factor H (FH), factor H-related proteins (FHRs), complement, copy number variations (CNVs), structural variants (SVs), single molecule real-time (SMRT)

INTRODUCTION

Membranoproliferative glomerulonephritis (MPGN) is a heterogeneous group of rare glomerular diseases associated with complement dysregulation, which leads to the deposition of C3 and its cleavage products in glomeruli. Diagnosis requires a kidney biopsy, as the clinical presentation and course are variable, with patients manifesting asymptomatic haematuria and proteinuria, hypertension, nephritic or nephrotic syndrome, and/or acute kidney injury. Approximately 50% of patients develop chronic kidney disease (CKD) and progress to end-stage renal failure (ESRF) over a 10-year period (Sethi and Fervenza, 2011; Noris and Remuzzi, 2015). Current classification is based on glomerular deposits detected by immunofluorescence (IF) microscopy (Pickering et al., 2013). Cases with glomerular C3 staining in combination with significant immunoglobulin (IgGs) deposition are defined as immune-complex-mediated MPGN (IC-MPGN). C3 Glomerulopathy (C3G) is diagnosed in cases with dominant C3 staining at least two orders of magnitude greater than any other immunoreactant. Electron microscopy (EM) allows further differentiation of C3G into either dense deposit disease type (DDD), which is characterized by intramembranous highly electron-dense deposits, or C3 glomerulonephritis (C3GN), in which the deposits are less dense and have mesangial and/or subendothelial and subepithelial localization (Pickering et al., 2013; Hou et al., 2014).

Both C3G and IC-MPGN are complement-mediated diseases. The complement cascade is the cornerstone of innate immunity and can be initiated by three different pathways – the alternative (AP), classical (CP), or mannose-binding lectin (LP) pathways – that generate proteolytic complexes known as C3 convertases (**Figure 1**). The C3 convertase of the AP is C3bBb, while that of the CP and LP is C4bC2a. Both C3 convertases are so named because they cleave C3 into C3a, an anaphylatoxin, and C3b, which associates with factor B to generate additional C3bBb thereby amplifying the complement response. Binding of C3b to C3 convertases generates C5 convertases, which cleave C5 to produce C5a, another anaphylatoxin, and C5b, which initiates the terminal complement cascade by associating with other complement components (C6–C9) to form the terminal complement complex C5b-9 (Muller-Eberhard, 1986; Bhakdi and Tranum-Jensen, 1988; Morgan, 1999).

IC-MPGN has typically been linked to the activation of the complement CP following infections, autoimmune diseases or malignancies, while C3G has primarily been linked to activation of the complement AP (Sethi and Fervenza, 2011). In both C3G and IC-MPGN, genetic defects in complement AP genes like *CFH*, *C3*, *CFI*, and *CFB* (Servais et al., 2012; Iatropoulos et al., 2016, 2018), and acquired factors, such autoantibodies that stabilize the C3 convertase complex C3bBb (called C3-nephritic factors, C3NeFs) or against FH, FB and C3b have been identified (Zhang et al., 2012, 2020; Blanc et al., 2015; Marinozzi et al., 2017; Donadelli et al., 2018). These findings indicate that the dysregulation of the complement AP may underlie the pathogenesis of both diseases (**Figure 2**).

To gain further insights into the pathophysiology of these diseases, we have used unsupervised hierarchical cluster analysis based on histological, biochemical, genetic and clinical data at disease onset to divide our patient population into four clusters, each of which is defined by specific underlying pathophysiologic mechanisms (**Figure 3**) (Iatropoulos et al., 2018). In clusters 1, 2, and 3, serum C3 is low and the frequency of complement genetic variants and C3NeFs is high. Clusters 1 and 2 differentiated themselves from cluster 3 by very high sC5b-9 levels, which are indicative of dysregulated terminal pathway activity. Cluster 2 uniquely exhibits strong C1q, IgG and IgM glomerular deposition, suggesting that CP activity plays an important role in initiating disease in this cluster. Cluster 4 is characterized by normal C3 and sC5b-9 levels, and rare C3NeFs and complement genetic variants, despite intense C3 glomerular staining, indicating local glomerular complement activity (Iatropoulos et al., 2018).

Interestingly, genetic variants in *CFH*, which encodes factor H, the main regulatory protein of the AP complement pathway, are found in all 4 clusters indicating a complex pattern of functional consequences resulting in variable phenotypes (Iatropoulos et al., 2018).

The *CFH* gene family includes six genes - *CFH*, *CFHR3*, *CFHR1*, *CFHR4*, *CFHR2*, and *CFHR5*- on chromosome 1q31.3 that arose from *CFH* as a consequence of tandem genomic duplication events (Diaz-Guillen et al., 1999). The translated proteins, FH and FHR1-5s, are circulating proteins, organized in short consensus repeats (SCRs). The C-terminal region of the five *CFHRs* exhibits a high degree of sequence identity with the C-terminal domains of *CFH*, suggesting that FHR proteins can bind similar surface ligands as FH. However, FHRs do not contain the regulatory domains of FH (N-terminal region), suggesting they do not possess direct complement regulatory activity (Skerka et al., 2013).

The genomic region of the *CFH* gene family is characterized by large segmental duplications (SDs) and interspersed repetitive sequences that predispose to genomic rearrangements such as duplications, deletions and inversions (Lupski and Stankiewicz, 2005) that, when larger than 1kb, are called structural variants (SVs) (Feuk et al., 2006). The most common SV described in the *CFH* gene family is the ∼84 kb deletion of *CFHR3* and *CFHR1* (*CFHR3-CFHR1* del) with an allele frequency ranging from 2 to 51%, depending on ethnicity (Holmes et al., 2013). The absence of both copies of *CFHR3* and *CFHR1* is also associated with a lower risk of age-related macular degeneration (AMD) (Hughes et al., 2006) and IgA nephropathy (Gharavi et al., 2011), and a higher risk of atypical haemolytic uremic syndrome (aHUS) (Moore et al., 2010) and systemic lupus erythematosus (SLE) (Zhao et al., 2011).

Rare SVs involving *CFHRs* have been described in DDD and C3GN, most of which generate abnormal fusion proteins (Gale et al., 2010; Malik et al., 2012; Tortajada et al., 2013; Chen et al., 2014; Medjeral-Thomas et al., 2014; Togarsimalemath et al., 2017; Xiao et al., 2016). The classic example was identified in Greek Cypriot patients with C3GN (often called *CFHR5* nephropathy) that results from a mutant FHR5 protein encoded by a *CFHR5* gene with an internal duplication of exons 2 and 3 ($FHR5_{1,2}$-FHR5) (Gale et al., 2010). Other FHR fusion proteins linked to C3G include $FHR2_{1,2}$-FHR5 (Chen et al., 2014), $FHR5_{1,2}$-FHR2

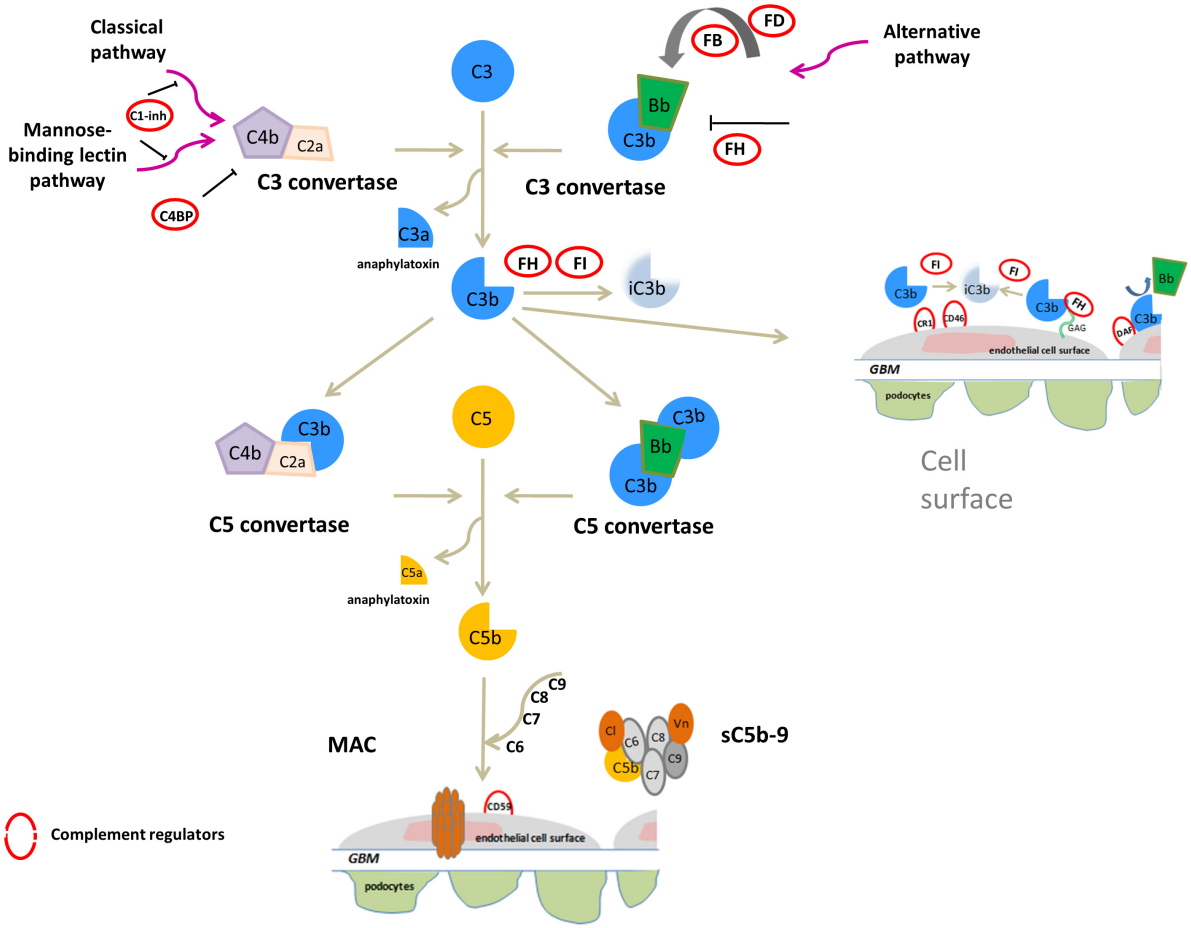


FIGURE 1 | Overview of activation and regulation of complement system. The complement system is activated by three pathways: the classical (CP), the mannose-binding lectin (LP) and alternative (AP) pathways. All three activated cascades generate the C3 convertases (C4bC2a and C3bBb), proteolytic complexes that cleave C3 into C3a and C3b. C3a acts as an anaphylatoxin. C3b can covalently bind to surface membranes (e.g., intact host cells, microbial membranes, and modified host surfaces). Binding of an additional C3b molecule to C3 convertase generates C5 convertases (C4b2aC3b and C3bBbC3b) that cleave C5 into the potent anaphylatoxin C5a, and C5b. C5b recruits other complement components (C6, C7, C8, and C9) to assemble the soluble terminal C5b-9 complex (sC5b-9), which causes inflammation, or the membrane attack complex (MAC), leading to pore formation and target cell lysis. On healthy host cells the complement system is controlled at various steps by soluble or membrane regulators (indicated by red circles in the figure). FH binds to C3b and glycosaminoglycans (GAGs) on the cell surface and inactives C3b to iC3b in the presence of FI, and also accelerates the decay of the AP C3 convertase. C3b is inactivated to iC3b by FI also in the presence of membrane cofactor protein (CD46/MCP) or complement receptor 1 (CR1/CD35). In addition, among membrane complement regulators, DAF destabilizes and dissociates the C3/C5 convertases of the classical and alternative pathways while the CD59 (or protectin) binds C5b-8 complexes, inhibiting the recruitment of C9, thus preventing MAC generation. C1 inhibitor (C1-INH) and C4b-binding protein (C4BP) regulate the CP and LP. Vitronectin (Vn) and Clusterin (Cl) bind soluble C5b-7-8-9 complexes, blocking their incorporation into cell membranes.

(Xiao et al., 2016), FHR1-FHR5 (Togarsimalemath et al., 2017), $FHR1_{1-4}$-FHR1 (Tortajada et al., 2013), and $FHR3_{1,2}$-FHR1 (Malik et al., 2012). These reports highlight the importance of the *CFH-CFHR* region in C3G and yet, with the exception of the fusion protein endemic to Cyprus, all fusion proteins thus far described have been identified in small families.

Comprehensive studies of the *CFH-CFHR* region in large cohorts of patients with C3G and IC-MPGN have not been reported. We sought to address this knowledge gap by identifying common and rare SVs and their distribution among the 4 C3G/IC-MPGN clusters we have described (Iatropoulos et al., 2018). SVs were detected using Multiplex Ligation-dependent Probe Amplification (MLPA) followed by PacBio long-read sequencing (SMRT, Single- Molecule Real-Time, sequencing) to provide base-pair resolution of selected genomic rearrangements.

MATERIALS AND METHODS

Patients

Patients (n = 199) were recruited by the Italian Registry of MPGN, coordinated by the *Aldo e Cele Daccò* Clinical Research Center for Rare Diseases at the Mario Negri Institute. Clinical, demographic and laboratory data from patients were collected in a case report form. Blood, plasma and serum were also

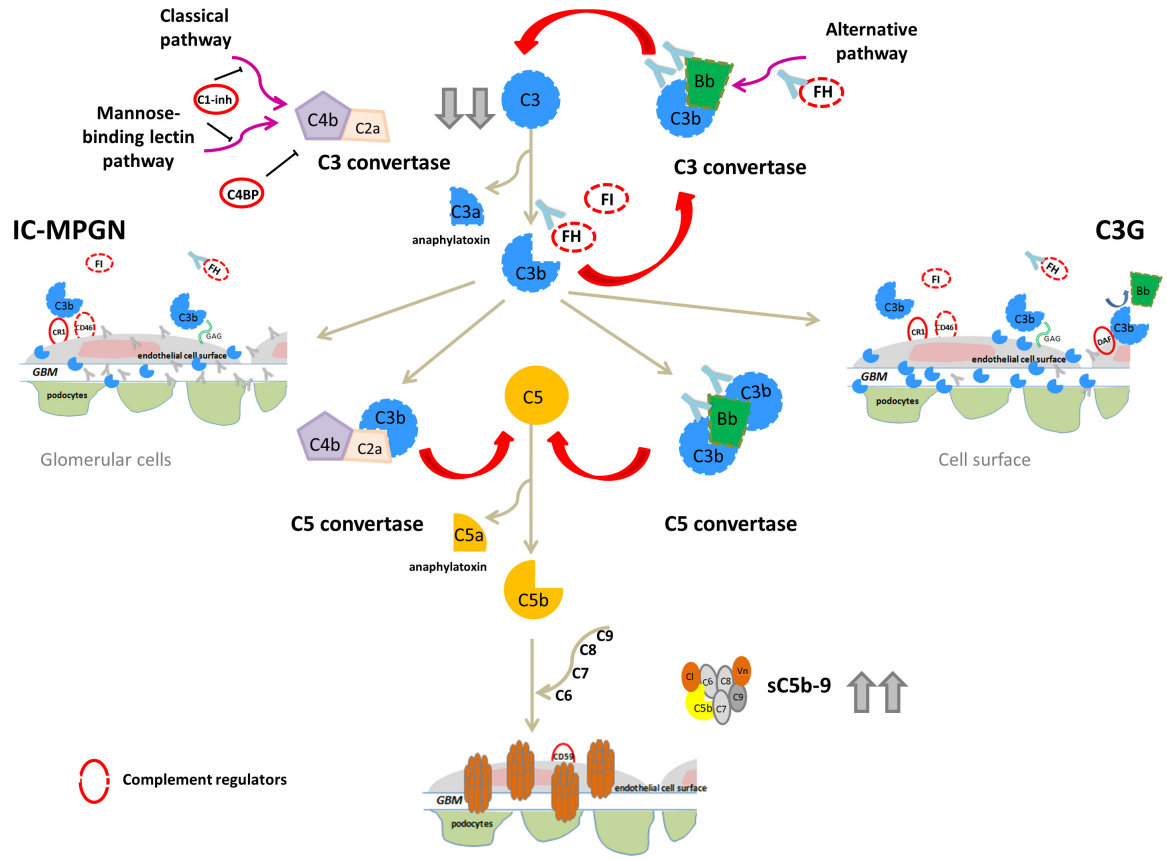


FIGURE 2 | Complement dysregulation in C3G and IC-MPGN. Variants in complement alternative pathway (AP) genes (*CFH*, *C3*, *CFI*, and *CFB;* dashed lines) and/or autoantibodies (indicated by light blue Y-shaped forms) that bind FH, FB, C3b or that stabilize the C3/C5 convertase (nephritic factors, C3NeFs/C5NeFs) are the main drivers of complement AP dysregulation. This results in complement hyperactivation and glomerular deposition of C3 compounds (C3G). In some patients there is the concomitant activation of the classical pathway by infections or immune-complexes (IC) resulting in both C3 and IC deposits (IC-MPGN). In these patients AP dysregulation provides an activation loop exacerbating C3 glomerular deposition. In both cases, abnormal C3 convertase activation causes a consumption of circulating C3 that explains low serum levels in patients (indicated by gray arrows). Complement activation can proceed until the terminal pathway, causing high sC5b-9 plasma levels and glomerular C5b-9 deposits.

collected for biochemical and genetic tests. Controls included biological samples from blood donors (n = 214), which were analyzed for copy number abnormalities identified in C3G/IC-MPGN patients. The samples used for the research were stored at the Centro Risorse Biologiche (CRB) "Mario Negri", biobank Malattie Rare e Malattie Renali.

The study was approved by the Ethics Committee of Bergamo (Italy). All participants received detailed information on the purpose and design of the study, according to the guidelines of the Declaration of Helsinki.

Diagnosis

All kidney biopsy reports were independently reviewed by two pathologists at the Mario Negri Institute and discordances were resolved through face-to-face discussion (Iatropoulos et al., 2018). The diagnosis of MPGN was based on light microscopy findings, according to the Cook HT and Pickering MC (Cook and Pickering, 2015). MPGN patients were further classified by immunofluorescence (IF) as (Sethi and Fervenza, 2011; Pickering et al., 2013; Iatropoulos et al., 2016; Marinozzi et al., 2017): (1) Immune-complex-mediated MPGN (IC-MPGN) – C3 and IgG IF similar or differing by less than two orders of magnitude; or, (2) C3 Glomerulopathy (C3G) – C3 IF at least two orders of magnitude greater than any other immune reactant (scale of 0 to 3) (**Figure 4**).

Based on electron microscopy (EM) findings, C3G was further classified as either DDD or C3GN. Patients with secondary MPGN, a previous diagnosis of aHUS, MPGN on allograft but without biopsy of native kidney, and without IF or EM studies, were excluded from this study.

All patients from the Registry who fulfilled the above inclusion criteria were included in this study.

Cluster Analysis

We used a three-step algorithm to assign patients to different clusters, as reported in Iatropoulos et al. (2018). The algorithm is based on four features available at disease onset: genetic findings (presence of rare variants), C3NeF, serum C3 levels,

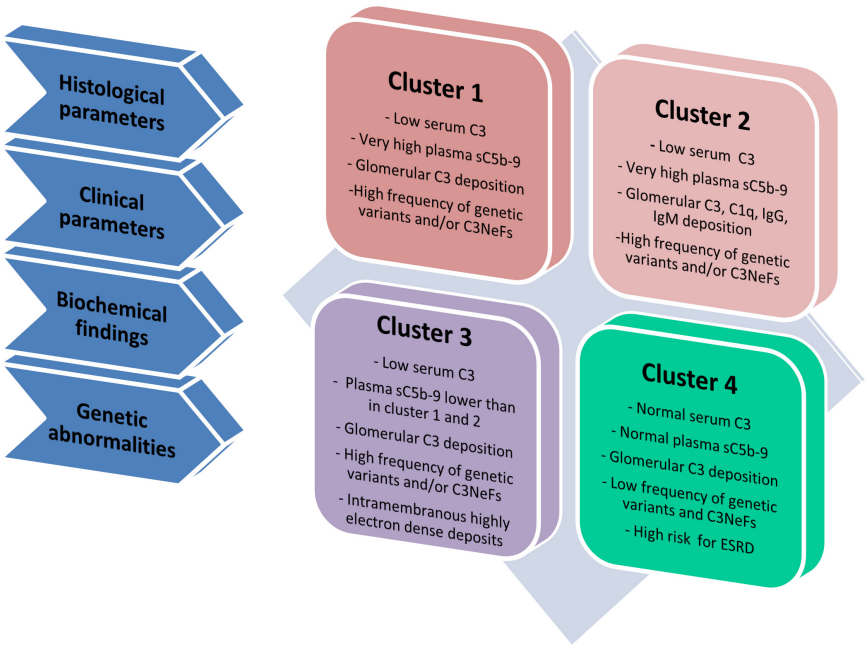


FIGURE 3 | Schematic representation of four clusters. Cluster analysis was based on 34 variables, including histological, clinical, biochemical and genetic data and divided patients in four groups called clusters (Iatropoulos et al., 2018). Cluster 1, 2, and 3 have low C3 levels and high frequency of genetic variants and/or C3NeFs. Cluster 1 and 2 differentiate themselves from cluster 3 because of highly increased plasma levels of sC5b-9, indicative of high terminal pathway activity. Compared with cluster 1, cluster 2 includes patients with strong IgG, IgA and C1q glomerular deposition, indicating the concomitant activation of the classical pathway. At variance with cluster 1–3, cluster 4 is separated from the others, since it is characterized by normal C3 and sC5b-9 levels in face of intense glomerular C3 deposits, low frequency of genetic variants and/or C3NeFs and a high risk of developing end-stage renal disease (ESRD).

biopsy findings (presence of intramembranous highly electron-dense deposits).

DNA Samples

Genomic DNA (gDNA) was extracted from peripheral blood using either the Nucleon^{TM} BACC2 Genomic DNA extraction kit (GE Healthcare, Little Chalfont, United Kingdom) or NucleoSpin Blood columns (Macherey-Nagel). DNA integrity and quality were verified by 0.8% agarose gel electrophoresis and NanoDrop Spectometer (ND-1000; Thermo Fisher), respectively. Before genetic analyses, DNA was quantified using a Qubit fluorometer (dsDNA HS Assay kit; Invitrogen).

Complement Component Assays

Serum C3 and C4 concentrations were measured by kinetic nephelometry (Noris et al., 2010). sC5b-9 levels were assessed using the MicroVue SC5b-9 Plus EIA commercial kit (SC5b-9 Plus, Quidel). IgGs purified from plasma were used to test C3NeF activity. The assay consisted in measuring IgG ability to stabilize the AP C3 convertase (C3bBb), as previously described (Fremeaux-Bacchi et al., 1994; Donadelli et al., 2018). The presence of anti-FH autoantibodies was evaluated by an Enzyme-Linked ImmunoSorbent Assay (ELISA), as reported (Valoti et al., 2019).

A in-house sandwich ELISA was developed to measure plasma or serum FH. In brief, Nunc MaxiSorp ELISA plates (Nunc, Roskilde, Denmark) were coated with 100 μL of diluted sheep polyclonal anti-factor H antibody (dilution 1:6333; Abcam) and were incubated overnight at 4°C. The next day, plates were washed with PBS and 0.05% Tween20, and blocked with PBS and 1% BSA for 1 h at RT. After washing, 100 μL of each diluted sample (1:10000 in PBS-BSA 1%) was added. After incubation for 2 h at RT, plates were washed with PBS and 0.05% Tween20. 100 μL mouse monoclonal anti-human Factor H (diluted 1:10000; OX-23, LS-C58560, LSBio), which specifically detects FH and FH-like (FHL1), was added to each well. After 2 h of incubation at RT, wells were washed and 100 μL of diluted goat anti-mouse IgG HRP conjugated (dilution 1:2000; Thermo Fischer Scientific) was added (1 h of incubation at RT). After washing, TMB was used as substrate to detect enzymatic activity. Enzymatic reactions were terminated using 100 μL of sulphuric acid and absorbance was read at 450 nm. All samples were tested in duplicate. Sample concentrations were extrapolated from sigmoidal curve. Serum/plasma samples of 102 healthy subjects were tested to establish normal FH levels (≥193 mg/L).

Genetic Screening

Genetic analyses were performed by a next generation sequencing (NGS) diagnostic minipanel for simultaneous sequencing of 6 complement genes (complement factor H, *CFH*, NG_007259.1; complement factor I, *CFI*, NG_007569.1; membrane cofactor protein, *CD46/MCP*, NG_007569.1; complement factor B, *CFB*, NG_008191.1; complement C3, *C3*, NG_009557.1; and thrombomodulin, *THBD*, NG_012027.1). Amplicons were

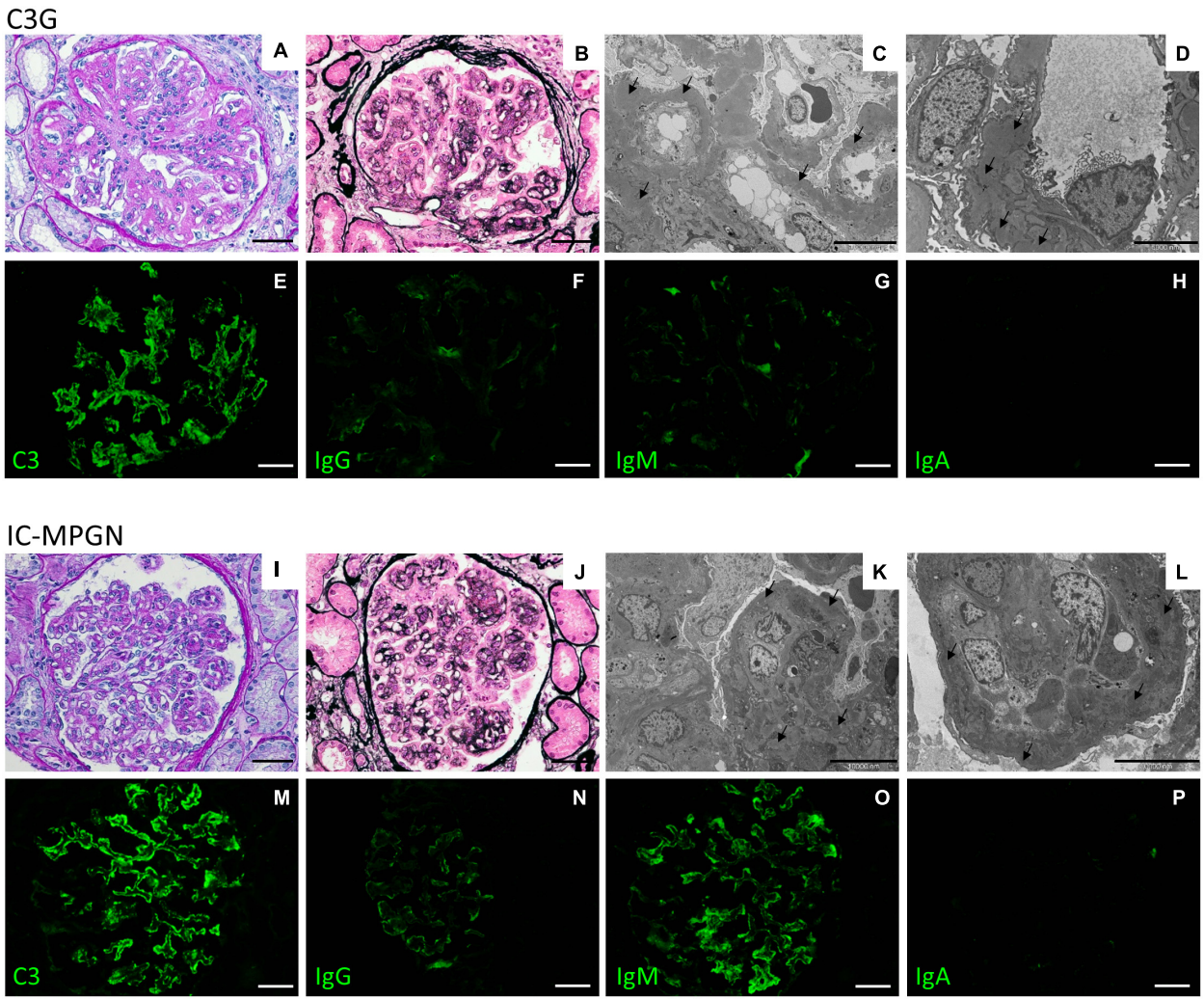


FIGURE 4 | Representative biopsy findings from patients diagnosed with C3GN (upper panel) and IC-MPGN (lower panel). Upper panel: **(A,B)** C3GN patient. Light microscopy revealed endocapillary proliferation, leukocyte infiltration accompanied by lobulation of the glomerular tuft (**A**, periodic acid-Schiff staining; **B**, Jone's silver staining). **(C,D)** Electron micrographs show intramembranous, subendothelial and mesangial deposits (arrows). **(E–H)** Representative immunofluorescence images reveal strong positivity for C3 staining along the capillary tuft **(E)**, while IgG **(F)**, IgM **(G)**, and IgA **(H)** staining is absent or present only in traces. (Scale bars: 50 μm in **A,B,E–H**; 10,000 nm in **C** and 5,000 in **D**). Lower panel: **(I–L)** IC-MPGN patient. Light microscopy and electron microscopy findings are similar to those observed in the C3GN patient (**I**, periodic acid-Schiff staining; **J**, Jone's silver staining; **K,L**, transmission electron micrographs). **(M–P)** At variance with C3GN patients, immunofluorescence analysis shows the typical pattern of IC-MPGN, with abundant deposition of C3 in the glomerular tuft **(M)**, accompanied by moderate-to-strong positivity for IgG and IgM immunoglobulin staining (**N**, IgG; **O**, IgM); IgA staining is negative (**P**, IgA). (Scale bars: 50 μm in **I,J,M–P**; 10,000 nm in **K,L**).

obtained by highly multiplex PCR using the Ion AmpliSeq^{TM} Library Kit 2.0 (Life Technologies, LT). Targets were then subjected to clonal amplification on Ion PGM^{TM} Template OT2 200 Kit and finally sequenced on Ion Torrent Personal Genome Machine Sequencer (PGM, LT), as previously described (Iatropoulos et al., 2016). In the patients with abnormal CNVs, we evaluated the presence of genetic variants in *CFHR1-5* by NGS studies, using either a panel called CasCADE, developed at the University of Iowa, or an updated version of the diagnostic minipanel (Bu et al., 2014).

Genetic variants in coding and splicing regions of complement genes with minor allele frequency (MAF) in the gnomAD database <0.001 and with a Combined Annotation Dependent Depletion (CADD) phred score ≥ 10 were considered rare variants (RVs). RVs were further classified into "pathogenic, (P)", "likely pathogenic, (LPV)," and "variants of uncertain significance, (VUS)" using guidelines from the American College of Medical Genetics and Genomic (ACMG) and from the KDIGO conference on aHUS and C3G (Kircher et al., 2014; Richards et al., 2015; Goodship et al., 2017).

Copy Number Variations (CNVs)

MLPA using the SALSA MLPA kit P236-A3 (MRC Holland) and in-house probes for *CFHR4* and *CFHR5* (**Supplementary Table 1**) were used to screen for rearrangements/deletions/duplications in the *CFH-CFHR5*

genomic region in 199 patients (195 unrelated and 4 relatives) and in 100 healthy subjects.

Two hundred fourteen healthy subjects were also screened for the novel *CFHR4* CNVs using multiplex polymerase chain reaction (mPCR) that amplified intron 1 and exon 2 of *CFHR4* and intron 3 of *CFHR1* (Moore et al., 2010).

Single Molecule Real-Time (SMRT) Sequencing

Probes targeting *CFH-CFHRs* on the human genome reference hg19 (from chr1:196619000 to chr1:196979303) were designed using online Nimble Design Software (Roche Sequencing, Pleasanton, CA, United States). Samples from 10 patients (new or rare SVs, n = 6; heterozygous *CFHR3-CFHR1* del, n = 1; homozygous *CFHR3-CFHR1* del, n = 1; heterozygous *CFHR1-CFHR4* del, n = 1; *CFHR3-CFHR1* del and *CFHR1-CFHR4* del compound heterozygote, n = 1) and 7 healthy controls (normal copy number, n = 4; heterozygous *CFHR3-CFHR1* del, n = 3) were sequenced at the Norwegian Sequencing Centre[1]. Patient #1678, in whom the boundaries of the $CFHR3_{1-5}$-$CFHR4_{10}$ fusion gene had been previously characterized by Sanger sequencing was included as a positive control. Libraries were prepared using the Pacific Biosciences (PacBio) protocol for Target Sequence Capture using SeqCap® EZ Libraries with PacBio® Barcoded Adapters. Briefly, 2 μg of DNA were sheared to 7 kb. Amplified and barcoded DNA were size selected using BluePippin. After pooling, the template was hybridized using *CFH-CFHR* probes. Following amplification, libraries were size selected by BluePippin with a 5 kb cut-off and then sequenced using Pacbio Sequel system.

Data were obtained as multiplexed subreads and were demultiplexed with the PacBio read demultiplexer *lima*, retaining only those subreads with a barcode quality greater than 45. To ensure high quality sequencing data, we used PacBio Circular Consensus Sequencing (CCS, also known as HiFi) reads, produced by obtaining a consensus sequence from subreads. The CCS reads were obtained with a PacBio tool called *ccs* with the following parameters: *–minLength=1000*, *–min-rq=0.99* and *–maxLength=10,000*. The length of the resulting CCS reads ranged from 1,351 to 10,108 bp, and the number of sequencing passes ranged from 3 to 114. CCS reads were mapped to hg19 with two long-read mappers: NGMLR (with-*min-identity=0.95*) and *minimap2* (using pre-set CCS). SV calling was carried out with *Sniffles* for NGMLR-aligned reads and with *pbsv* for both aligners. While the results for patient #1678 matched those previously obtained by MLPA and Sanger sequencing (positive control), some SVs involved large repeated regions and were difficult to resolve.

As shown in **Supplementary Figure 1**, the target region is characterized by two intralocus large SDs and a number of shorter repeats. The first duplicated region (b1 and b2, blue in **Supplementary Figure 1**) is 28,650 bp long (b1) and has an identity of around 98% with its counterpart (b2), which is 28,726 bp long. The second duplicated region (r1 and r2, red in **Supplementary Figure 1**) is 40,218 bp long (r1) and has an identity of ∼97% to its 39,726 bp long counterpart (r2).The *CFHR3-CFHR1* del CNV occurs across the b1/b2 duplications, while the *CFHR1-CFHR4* del occurs across the r1/r2 duplications. These regions are much longer than our average CCS read length (∼6,000 bp) and therefore while the "signature" of SVs involving these repeated regions typically could be detected by inspecting alignments (for example, as split-read alignments) or by reviewing the SV caller output, similar "signatures" were also observed in non-carriers (false positives). **Supplementary Figure 2** shows 3 individuals, *CFHR1-CFHR4* del, *CFHR3-CFHR1* del, normal control, who all show split-read alignments across the duplicated regions in spite of different genotypes. This example of a false positive likely reflects mapping errors caused by fragments originating in one region but mapping to the paralogous region, thereby generating a pattern similar to that associated with true SVs. We were, however, able to identify and locate SV breakpoints outside the repeated regions (see "Results" section) either by inspecting the aligned reads with Integrative Genomics Viewer (IGV) or based on the SV callers.

Western Blot

The molecular pattern of FH-FHRs was studied by Western Blot (WB) using serum/plasma (diluted 1:40 for FHRs and 1:80 for FH). Proteins were separated by 10–12% SDS-PAGE (Mini-Protean TGX Precast Gels, Bio-Rad) under non-reducing conditions and transferred by electroblotting to polyvinylidene Difluoride (PVDF) membrane (*Trans*-Blot® Turbo™ Midi PVDF Transfer; Bio-Rad). Membranes were blocked in 5% fat free (skim) milk and developed using specific FH/FHR antibodies: the FHR3 polyclonal antiserum and the monoclonal anti-FHR1 antibody (JHD) were a kind gift from Prof. Zipfel (Skerka et al., 2013) while the anti-FHR1-2-5 monoclonal antibody was kindly provided by Prof. de Cordoba (Goicoechea de Jorge et al., 2013). Factor H was detected using the commercial monoclonal anti-human Factor H (OX-23, LSBio). Incubation with primary antibodies was followed by horseradish peroxidase (HRP) conjugated secondary antibodies and ECL chemiluminescence detection system (Amersham).

Statistical Analysis

Chi-square or Fisher's exact tests were used to analyze categorical variables, while ANOVA was used to test continuous variables. Correction for multiple tests was applied.

RESULTS

Patients

One hundred ninety-nine patients with primary C3G or IC-MPGN were recruited from the Italian Registry of MPGN (IC-MPGN: n = 96, 48.2%; C3G: n = 103, 51.8%, including C3GN: n = 74; DDD: n = 29), 159 of whom have been described in a previous study (Iatropoulos et al., 2018). The mean age at diagnosis was 18.6 ± 14.9 years (range: 0.3–72 years; IC-MPGN: 19.9 ± 15.1 years; C3GN: 18.3 ± 16 years; DDD: 15.2 ± 10.5 years).

[1] www.sequencing.uio.no

TABLE 1 | Histologic diagnosis, complement assessment and genetic screening of patients recruited selected from the Italian Registry of MPGN and classified into clusters using the three-step algorithm.

	Cluster 1 (*n* = 66)	Cluster 2 (*n* = 49)	Cluster 3 (*n* = 33)	Cluster 4 (*n* = 51)	Overall *P*-value
IC-MPGN (*n* = 96)	21.2%	89.8%	12.1%	66.7%	<0.0001
C3GN (*n* = 74)	78.8%	10.2%	0%	33.3%	<0.0001
DDD (*n* = 29)	0%	0%	87.9%	0%	<0.0001
Sex,% men	57.6%	49%	63.6%	60.8%	0.54
Age (yr)-Mean (*SD*)	14 (±10.7)	18.6 (±13.7)	15.3 (±10.2)	26.9 (±19.5)[a,b,c]	<0.001
Serum C3 (mg/dl)	29.1 (±19.8)	22.8 (±21.9)	35.4 (±34.3)	93.2 (±26.1)[a,b,c]	<0.001
Serum C4 (mg/dl)	18.7 (±6.7)	16.8 (±11.3)	20.6 (±8.7)	21 (±10)	0.11
Plasma sC5b-9 (ng/ml)	1378 (±1255)[c,d]	1861 (±1357)[c,d]	540 (±604)[a,b]	302 (±145)[a,b]	<0.001
Low serum C3	100%	100%	93.9%	49%[a,b,c]	<0.001
Low serum C4	7.7%	28.6%[a,c,d]	6.2%	10%	0.004
Low serum C3 and normal C4	92.3%[b]	71.4%[a]	87.5%	44%[a,c]	<0.001
High plasma sC5b-9	76.3%	83%	32.3%[a,b]	20%[a,b]	<0.001
RV carriers	27.3%	22.4%	15.2%	3.9%[a,b]	0.01
C3NeF positive	53.4%	62%	79.3%[a]	5.9%[a,b,c]	<0.001
RV carriers and/or C3NeF	71.6%	74.5%	82.8%	9.8%[a,b,c]	<0.001
FH levels (mg/L)	301.9 (±70.9)	284.5 (±72)	275.7 (±65)	322.5 (±76)[b,c]	0.02
Low FH levels	6.9%	8.7%	6.9%	0%	0.26
Anti-FH antibodies	1.7%	6.5%	10.3%	0%	0.08

Quantitative variables are expressed as mean (±SD).
Abbreviations and limit of normal range:
C3: 90–180 mg/dl.
C4: 10–40 mg/dl;
Normal plasma sC5b-9 levels: ≤400 ng/ml;
Normal serum/plasma FH levels: ≥193 mg/L;
RV, rare variant defined as genetic variant in coding and splicing regions of complement genes already related to C3G- IC-MPGN (CFH, CFI, CD46, CFB, C3, and THBD) with minor allele frequency (MAF) in the gnomAD database <0.001 and with Combined Annotation Dependent Depletion (CADD) phred score ≥10;
p-values were corrected for multiple tests.
[a]Significant different from cluster 1; [b]significant different from cluster 2; [c]significant different from cluster 3; [d]significant different from cluster 4.

Using the published three-step algorithm (Iatropoulos et al., 2018), patients were assigned to clusters 1 (*n* = 66), 2 (*n* = 49) and 3 (*n* = 33) or cluster 4 (*n* = 51) (**Table 1**). As expected, plasma sC5b-9 levels were significantly higher in clusters 1 and 2 than in cluster 3.

Serum Factor H Abnormalities

Circulating factor H (FH) levels were measured in 181 patients and were lower than normal (reference ≥193 mg/L) in 9 patients (5%; C3G, *n* = 5; IC-MPGN, *n* = 4), all from clusters 1, 2, and 3. Six of the nine patients carried *CFH* RVs (N-terminal RV, *n* = 5; SCR15 -mid-region of FH-, *n* = 1) compared to 4 of 172 patients with normal FH levels (**Table 2**).

Screening for FH autoantibodies (FHAAs) identified 7 (4%) positive patients, all from clusters 1, 2, and 3 (IC-MPGN, *n* = 6; C3G, *n* = 1) (**Table 2**). All 7 patients had low C3 and normal FH levels, although 2 had low C4 levels. In addition to FHAAs, 5 patients were co-positive for C3NeFs. One patient negative for C3NeFs carried a RV in *C3* (p.Ser1063Asn; gnomAD global MAF = 6.9×10^{-5}). Six of the seven patients experienced childhood disease onset (ranging from 4.8 to 10.6 years).

CFH-CFHR Copy Number Variations

Common CNVs, namely the *CFHR3-CFHR1* (*CFHR3-CFHR1* del) and/or the *CFHR1-CFHR4* (*CFHR1-CFHR4* del) deletions, were identified in 32.8% of patients and 36.9% of controls (**Figure 5** and **Table 3**). Although there was no difference in the prevalence of the homozygous *CFHR3-CFHR1* del when patients and controls were compared, across patient groups, the homozygous *CFHR3-CFHR1* del was more frequently observed in cluster 3 than in cluster 1 (**Table 3**). This relationship remained when we also included two patients (one in cluster 1 and one in cluster 3) who were compound heterozygotes for *CFHR3-CFHR1* del and *CFHR1-CFHR4* del. There was no association between the homozygous *CFHR3-CFHR1* del and FHAAs.

Seven patients (3.6%) carried novel or rare CNVs that included a hybrid gene, two gene deletions, and four gene duplications. The new or rare CNVs were distributed among all clusters (**Figure 5**). Histologic, biochemical and genetic data of these patients are reported in **Table 4**.

CFHR3$_{1-5}$-*CFHR4*$_{10}$ Hybrid Gene

A new deletion involving *CFHR3*, *CFHR1* and *CFHR4* genes was identified in 1 patient (cluster 3; DDD; Patient #1678; **Table 4**) who presented with proteinuria (2 g/day) and low C3 levels (C3 = 45 mg/dl) at the age of 26. Her renal impairment progressed from the age of 32, reaching end-stage renal disease (ESRD) by age 38. She has received 3 kidney transplants, losing the first and second allografts to disease recurrence. Prior to her third transplant, she had slightly reduced C3 (72.5 mg/dl) but

TABLE 2 | List of patients with low FH levels and/or genetic or acquired FH abnormalities.

Pat ID	Histol. group	Algor. cluster	Age of onset (y)	Rare variant	Zyg.	gnomAD global freq.	CADD	Variant classif.	SVs	C3NeF	Serum C3 (mg/dl)	Serum C4 (mg/dl)	Plasma sC5b-9 (ng/ml)	FH levels (mg/dl)	FHAAs
1073	IC	2	28	p.FH: C494R[a] (SCR8)	Het	0	24.0	LPV*	Normal	Neg	70	28	1332	180	Neg
1304	IC	2	30.6	No	/	/	/	/	Heterozygous *CFHR3-CFHR1* del	Pos	45	9	249	91	Neg
1773	DDD	3	11.8	No	/	/	/	/	Normal	Pos	9	24	267	91	Neg
2032	C3GN	1	24	p.FH: R78G[a,b,c,d] (SCR1)	Hom	0	16	P	Normal	Neg	15	27.5	1530	133	Neg
2082	C3GN	1	0.75	p.FH: R127C[a,b] (SCR2)	Het	0	33	LPV	Normal	NA	72	26	2789	151	Neg
2158	C3GN	1	41.7	p.FH: R78G[a,b,c,d] (SCR1)	Hom	0	16	P	Normal	Neg	14	35	4571	154	Neg
2192	C3GN	1	9	p.FH: G133R[a] (SCR2)	Het	8E-06	31	LPV[§]	Normal	Neg	55	20	355	178	Neg
2585	IC	2	8.8	p.FH: G879R[b] (SCR15)	Het	4E-06	28	LPV	Normal	Pos	49	5	1209	119.5	Neg
2888	IC	2	16	No	/	/	/	/	Heterozygous *CFH-CFHR3-CFHR1* del	Neg	8.5	23	253	156	Neg
1026	IC	2	7	No	/	/	/	/	Normal	Pos	5	6	1080	222	Pos
1837	DDD	3	10.6	No	/	/	/	/	Homozygous *CFHR3-CFHR1* del	Pos	9	29	545	264	Pos
1967	IC	3	22	No	/	/	/	/	Heterozygous *CFHR3-CFHR1* del	Pos	84	28.4	257	265	Pos
2047	IC	2	10	p.C3: S1063N[a,b] (TED domain)	Het	6.9E-05	10	VUS	Normal	Neg	70	6	235	315	Pos
2081	IC	2	8.5	No	/	/	/	/	Normal	Pos	16	13.1	1643	313	Pos
2163	IC	3	6.5	No	/	/	/	/	Normal	Pos	18	19	277	376	Pos
2557	IC	1	4.8	No	/	/	/	/	Normal	Neg	33	11	605	267	Pos
1101	DDD	3	48.7	p.FH: R1210C[a,b,e,f] (SCR20)	Het	1.5E-04	12	P	Normal	Neg	154	14	368	394	Neg
1284	IC	2	0.4	p.FH: P88T[a,b] (SCR2)	Hom	0	29	LPV	Normal	Neg	5.4	24.7	3596	NA*	NA
1287	IC	2	0.3	p.FH: P88T[a,b] (SCR2)	Hom	0	29	LPV	Normal	Neg	47.7	45.3	2074	NA*	NA
1549	DDD	3	24.7	p.FH: R2I[a] (Signal peptide)	Het	0	11	VUS	1 copy of *CFHR3* + 3 copies of *CFHR4*	Pos	54	18	286	216	Neg

FH abnormalities are highlighted with gray color.
Abbreviations and limit of normal range:
Pat. ID, patient ID; Histol. Group, histologic group according to the current classification; Algor. Cluster, cluster group assigned using three-step algorithm;
y, years; Zyg, zygosity;
Rare variant is defined as genetic variant in coding and splicing regions of complement genes with minor allele frequency (MAF) in the gnomAD (genome aggregation database) <0.001 and with CADD (Combined Annotation Dependent Depletion) phred score ≥10.
gnomAD Freq., MAF in all subjects of the gnomAD database (v2.1.1);
Variant Classif., variant Classification reported in the Database on complement gene variant (https://www.complement-db.org/home.php) based on guidelines from ACMG (Richards et al., 2015) and from the KDIGO conference on aHUS and C3G (Goodship et al., 2017). When variant classification was not available in the database, the variant was classified using in silico predictions on the basis of KDIGO guidelines. P, pathogenic; LPV, likely pathogenic variant; VUS, variant of uncertain significance.
SVs, structural variants defined as genomic rearrangements longer than 1 kb.
C3NeF, C3 nephritic factor;
C3, 90–180 mg/dl;
C4, 10–40 mg/dl;
Normal plasma sC5b-9 levels: ≤400 ng/ml;
Normal serum/plasma FH levels: ≥193 mg/L;
FHAAs, anti-FH antibodies.
[a] *Iatropoulos et al. (2018).*
[b] *Osborne et al. (2018).*
[c] *Pechtl et al. (2011).*
[d] *Caprioli et al. (2003).*
[e] *Servais et al. (2012).*
[f] *Maga et al. (2010).*
**Pathogenic in 11 of 11 in silico tools;*
[§] *Pathogenic in 10 of 11 in silico tools.*

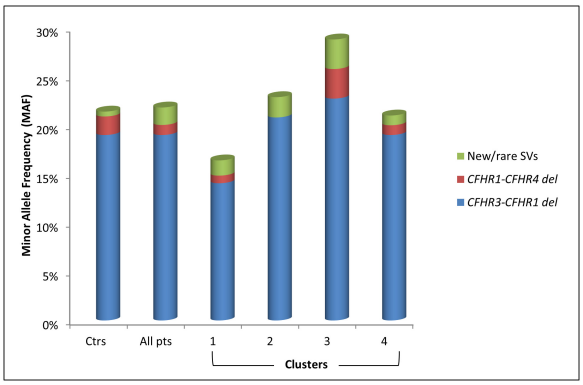


FIGURE 5 | Histogram representing the Minor Allele Frequency (MAF) and distribution of common (*CFHR3-CFHR1* del and *CFHR1-CFHR4* del) and new/rare structural variants (SVs) in patients (pts) and in healthy controls (ctrs).

normal C4 (23 mg/dl), sC5b-9 (269 ng/ml), and FH (323 mg/L) levels. C3NeFs and FHAAs were absent and genetic screening failed to identify any RVs in *CFH*, *C3*, *CD46*, *CFI*, *CFB*, and *THBD*. CNV analysis was remarkable for one copy of *CFHR3* that lacked exon 6, zero copies of *CFHR1*, and two copies of *CFHR4*, one of which carried a large deletion (**Figure 6A**). Long PCR and Sanger sequencing confirmed a deletion extending from exon 6 of *CFHR3* to exon 9 of *CFHR4*, predicting a novel $CFHR3_{1-5}$-$CFHR4_{10}$ hybrid gene. The breakpoint region was mapped between chr1:196760556 (intron 5 of *CFHR3*) and chr1:196886396 (intron 9 of *CFHR4*). Within the breakpoint region, we identified an insertion of 305 bp with sequence similarity to the two Alu Repeats located in intron 5 of *CFHR3* and in intron 9 of *CFHR4* (**Supplementary Figure 3**).

Because the *CFH-CFHR1-5* genomic region has several duplicated regions and a large number of Alu repeats that represent a strong limitation for sequence characterization of *CFH-CFHR* genomic rearrangements, we used SMRT, a DNA sequencing long-read approach. SMRT correctly identified the $CFHR3_{1-5}$-$CFHR4_{10}$ hybrid gene on one allele and distinguished it from the *CFHR3-CFHR1* del present on the other allele in the positive control (patient #1678 DNA), and confirmed the breakpoint region identified by Sanger sequencing (**Figure 7A**).

It is noteworthy that patient #1678, who belongs to cluster 3, is completely deficient in *CFHR1*.

The same $CFHR3_{1-5}$-$CFHR4_{10}$ hybrid gene was detected in this patient's two unaffected sons (**Figure 6B**) and in one of 214 healthy controls. WB using a polyclonal anti-human FHR3 antibody showed that the $CFHR3_{1-5}$-$CFHR4_{10}$ hybrid gene generates a $FHR3_{1-4}$-$FHR4_9$ hybrid protein (II-1; **Figure 6C**).

To search for additional genetic abnormalities in *CFHR* genes that may contribute to the disease phenotype in the patient, we performed targeted sequencing using CasCADE and identified two heterozygous nonsense RVs on the same allele in *CFHR2* (p.Gln211Ter – rs41299605 – and p.Arg254Ter – rs41313888 -; gnomAD global MAF: 6.5×10^{-5} and 7.5×10^{-4}, respectively) that were not transmitted to her healthy sons (**Figure 6B**).

TABLE 3 | Frequency of common structural variants (SVs) in patients and in controls.

Common SVs	Ctrs	All patients (*n* = 195)			Cluster 1 (*n* = 64)			Cluster 2 (*n* = 48)			Cluster 3 (*n* = 33)			Cluster 4 (*n* = 50)		
		Freq (*n*)	OR (95% CI)	*P*-value[a]	Freq (*n*)	OR (95% CI)	*P*-value[a]	Freq (*n*)	OR (95% CI)	*P*-value[a]	Freq (*n*)	OR (95% CI)	*P*-value[a]	Freq (*n*)	OR (95% CI)	*P*-value[a]
Het*CFHR3-CFHR1*del	32% (32/100)	27.7% (54)	0.81 (0.5–1.4)	0.44	26.6% (17)	0.77 (0.4–1.5)	0.46	31.2% (15)	0.97 (0.46–2)	0.93	18.2% (6)	0.47 (0.2–1.3)	0.13	32% (16)	1 (0.5–2.2)	1
Hom*CFHR3-CFHR1*del	3% (3/100)	3.1% (6)	1.03 (0.2–4.2)	0.97	0	0.22 (0–4.2)	0.31	4.2% (2)	1.41 (0.2–8.7)	0.71	9.1% (3)	2.7 (0.6–16.9)	0.16	2% (1)	0.85 (0.1–6.5)	0.72
*CFHR3-CFHR1*del+ *CFHR1-CFHR4*del	0% (0/214)	1% (2)	5.54 (0.3–116.2)	0.27	1.6% (1)	10.14 (0.4–251.8)	0.16	0	4.42 (0.1–225.7)	0.46	3% (1)	19.8 (0.8–496.5)	0.07	0	4.25 (0.1–216.6)	0.47
Het*CFHR1-CFHR4*del	1.9% (4/214)	1% (2)	1.84 (0.3–10.2)	0.48	0	2.75 (0.2–51.9)	0.5	0	2.07 (0.1–39.2)	0.62	3% (1)	0.61 (0.1–5.6)	0.66	2% (1)	0.93 (0.1–8.5)	0.95

[a]*P-values were calculated vs. controls (ctrs).*
Patients carrying both rare/new and common SVs were excluded.

TABLE 4 | List of patients carrying new/rare SVs.

Family ID	Patient ID	Gender	Histologic diagnosis	Algorithm-based cluster	Age of onset (years)	New/rare SVs	Serum C3 levels (mg/dl)	Serum C4 levels (mg/dl)	Plasma sC5b-9 levels (ng/ml)	RV	C3NeF	FHAAs	FH levels (mg/dl)
913	1678	Female	DDD	3	26	*CFHR3-CFHR4* hybrid gene + *CFHR3-CFHR1* del	72.5	20	269	No	Neg	Neg	323
1876	2870	Female	C3GN	4	50	*partial CFHR3 deletion* + *CFHR3-CFHR1* del	94.7	65.3	470	No	Neg	Neg	573
1892	2888	Male	IC	2	16	*CFH-CFHR3-CFHR1* deletion	8.5	23	253	No	Neg	Neg	156
1866	2856	Male	C3GN	1	11	3 copies of *CFHR1* + 3 copies of *CFHR4*	9	17.1	1930	No	Pos	Neg	454
1970	2979	Male	C3GN	1	5	3 copies of *CFHR1* + 3 copies of *CFHR4*	51	19	NA	No	NA	NA	NA
950	1726	Female	IC	2	25	1 copy of *CFHR3* + 3 copies of *CFHR4*	20	6	291	*CFB:* p. R679W	Neg	Neg	393.5
811	1549	Female	DDD	3	25	1 copy of *CFHR3* + 3 copies of *CFHR4*	54	18	286	*CFH:* p.R2I	Pos	Neg	216

Abbreviations and limit of normal range:
SVs, structural variants defined as genomic rearrangements resulting in duplications, deletions and inversions larger than 1 kb;
NA, not available;
C3, 90–180 mg/dl;
C4, 10–40 mg/dl;
Normal plasma sC5b-9 levels: ≤400 ng/ml;
Normal serum/plasma FH levels: ≥193 mg/L;
RV, rare variant defined as genetic variant in coding and splicing regions of complement genes already related to C3G- IC-MPGN (CFH, CFI, CD46, CFB, C3, and THBD) with MAF in the gnomAD database <0.001 and with Combined Annotation Dependent Depletion (CADD) phred score ≥10;
C3NeF, C3 nephritic factor;
FHAAs, anti-FH antibodies.

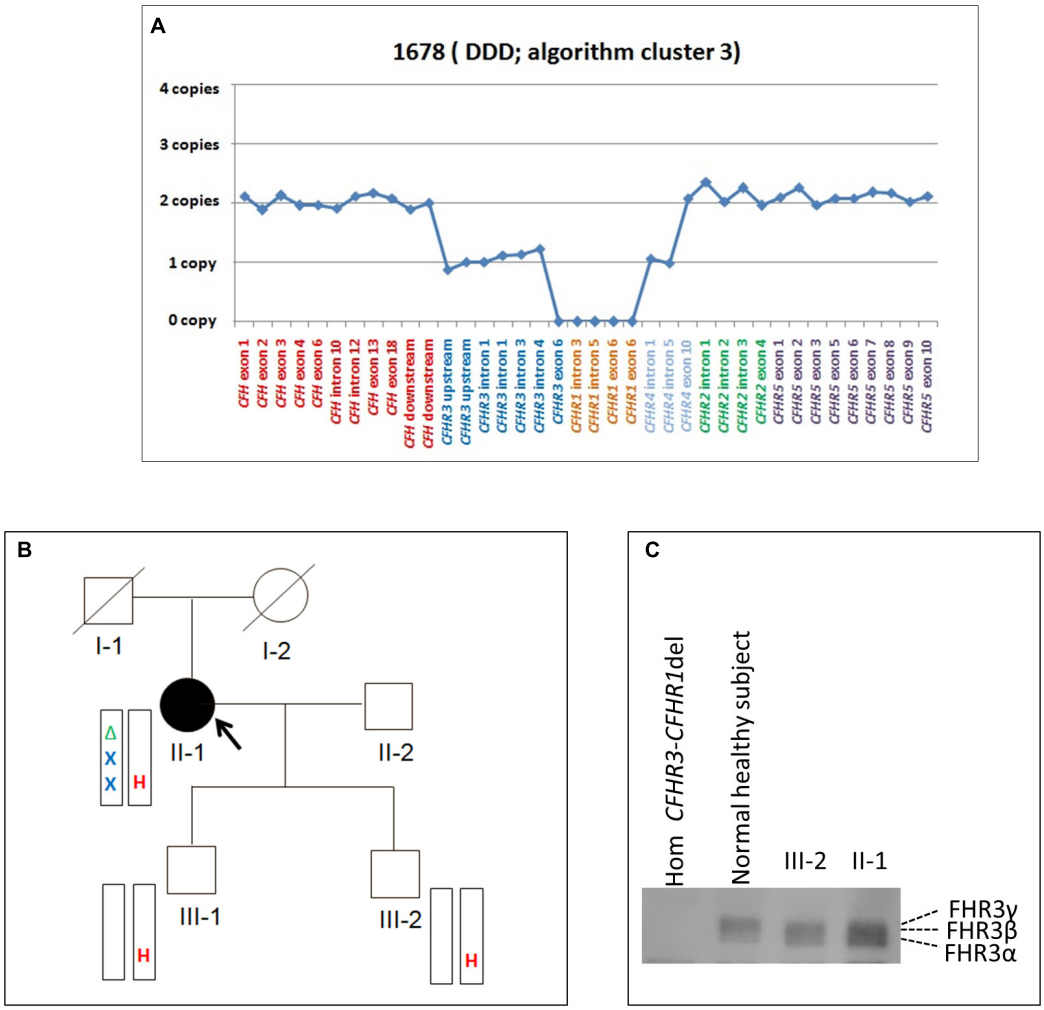


FIGURE 6 | The *CFHR3*$_{1-5}$-*CFHR4*$_{10}$ hybrid gene identified in a DDD patient in cluster 3. **(A)** Results of MLPA showing in patient #1678 two normal copies of *CFH*, only one copy of *CFHR3* lacking exon 6, zero copies of *CFHR1*, one normal and one partially deleted copy of *CFHR4* and two copies of *CFHR5*. **(B)** Pedigree (#913) of the DDD patient (II-1; indicated by the black circle) carrying the *CFHR3*$_{1-5}$-*CFHR4*$_{10}$ hybrid gene on one allele and the *CFHR3-CFHR1* del on the other allele. The *CFHR3*$_{1-5}$-*CFHR4*$_{10}$ hybrid gene is indicated in red (H) and the *CFHR3-CFHR1* del is indicated in green (Δ). The patient also carries two heterozygous nonsense rare variants of unknown significance (VUS) in the *CFHR2* (p.Gln211Ter - rs41299605 – and p.Arg254Ter – rs41313888 -; gnomAD global MAF: 6.5 × 10^{-5} and 7.5 × 10^{-4}, respectively), indicated in blue (X).The *CFHR3*$_{1-5}$-*CFHR4*$_{10}$ hybrid gene, but not the *CFHR2* rare variants (RVs) and the *CFHR3-CFHR1* del, was transmitted to the two healthy patients' sons (III-1 and III-2). **(C)** Western Blot (WB) of FHR3 was performed using an anti-FHR3 polyclonal antiserum (diluted 1:2,000), under non-reducing conditions, using the sera from the proband (II-1), her healthy son (III-2), a healthy control with normal CNVs (positive control) and a patient carrying the homozygous *CFHR3-CFHR1* del (negative control). The presence of 3 bands in the proband, corresponding to the different glycosylated variants of FHR3, indicates that the FHR3$_{1-4}$-FHR4$_9$ hybrid protein is secreted, since she is *CFHR3-CFHR1* deleted on the other allele.

CFHR3 Deletion

In a patient from cluster 4 (#2870; **Table 4**), MLPA revealed 1 copy of *CFHR3* to intron 3, 0 copies of *CFHR3* from intron 4 to exon 6, and 1 copy of *CFHR1* (**Figure 8A**). The patient, who had a family history of nephropathy, developed disease heralded by microhaematuria and proteinuria at 50 years of age. Because proteinuria persisted (0.6–1.0 g/day for at least 8 years), at age 58, a kidney biopsy was performed and a diagnosis of C3GN was made. Serum protein electrophoresis was normal and the patient was negative for C3NeFs and FHAAs. Six years later, proteinuria increased to the nephrotic-range (3.7 g/day), renal function declined (creatinine 1.3 mg/dl), and treatment with diuretics, angiotensin-converting-enzyme (ACE) inhibitors and angiotensin II receptor blockers (ARBs) was initiated. At last follow-up, creatinine was 1.1 mg/dl, C3 (94.7 mg/dl) and C4 (65.3 mg/dl) were normal, and sC5b-9 was slightly increased (470 ng/ml).

We were not able to identify the deletion breakpoints by long PCR and Sanger sequencing; however, SMRT sequencing showed that the abnormal MLPA pattern derived from both the *CFHR3-CFHR1* del on one allele and a novel deletion from *CFHR3*-intron 3 to *CFHR3*-3′UTR on the other allele (**Figure 7B**). SMRT data also identified the two genomic breakpoints (hg19: chr1:196756789 at *CFHR3* intron 3 and

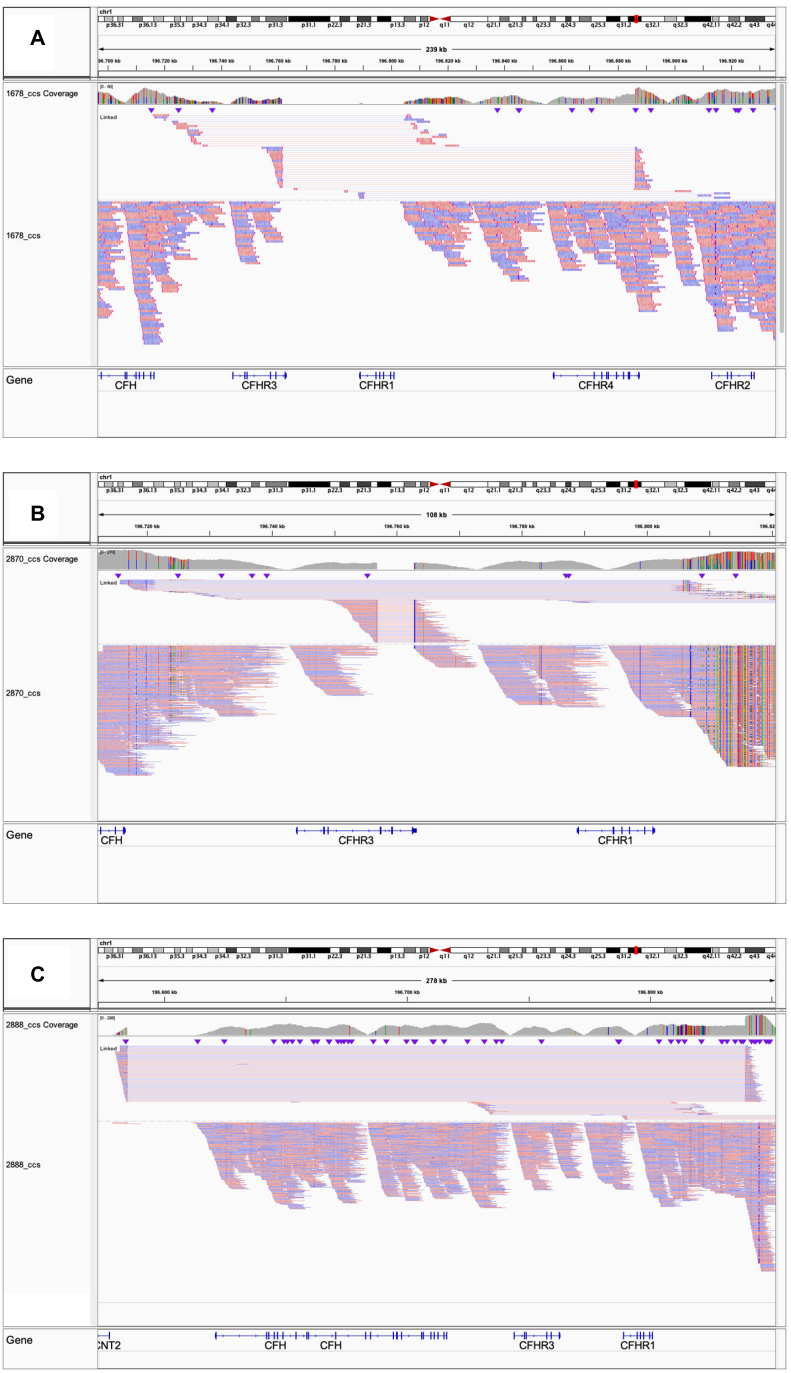


FIGURE 7 | Screenshot from IGV (Integrative Genomics Viewer) showing reads from SMRT sequencing. **(A)** Patient #1678 carrying the *CFHR3*$_{1-5}$-*CFHR4*$_{10}$ and the *CFHR3-CFHR1* del. **(B)** Patient #2870 carrying the *CFHR3*-intron 3 to *CFHR3*-3′UTR deletion and the *CFHR3-CFHR1* del. **(C)** Patient #2888 carrying the heterozygous deletion of *CFH*, *CFHR3*, and *CFHR1*.

chr1:196762816 at *CFHR3* 3′UTR), which were confirmed by long PCR and Sanger sequencing using primers targeting the breakpoint region (**Supplementary Table 2**). These data indicate the presence of a shorter *CFHR3* gene comprised of only exons 1, 2, and 3. In addition, NGS identified a heterozygous RV in *CFHR4* (p.Val438Gly; rs766466004; gnomAD global MAF: 4×10^{-6}; II-4, **Figure 8B**). Both the partial *CFHR3* deletion and the *CFHR4* rare variant were identified in a maternal female cousin (II-7; **Figure 8B**) with a history of proteinuria from the age of 15 and a biopsy diagnosis of MPGN (IF and EM data are not available). She developed progressive chronic renal failure and received a

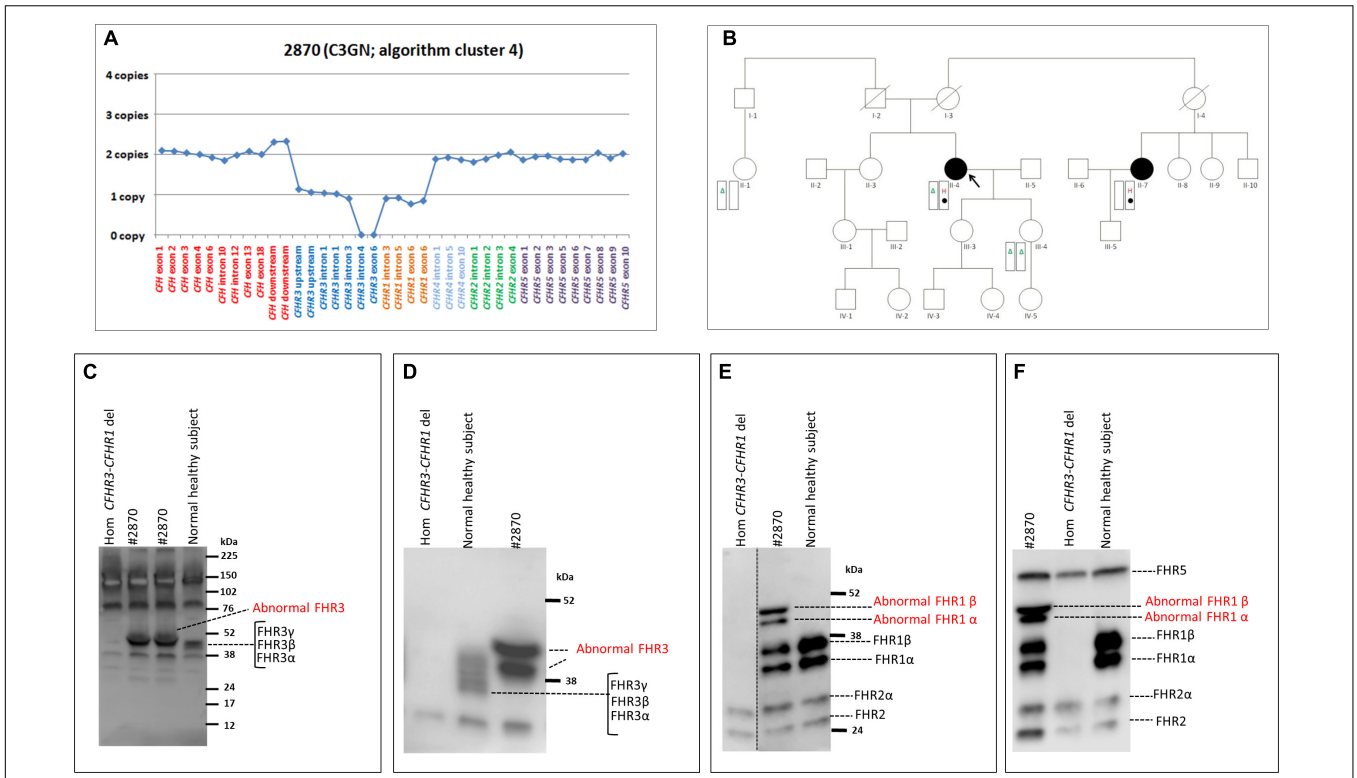


FIGURE 8 | The *CFHR3* deletion identified in the C3GN patient in cluster 4. **(A)** Results of MLPA showing two normal copies of *CFH*, one copy of *CFHR3* until intron 3, zero copies of *CFHR3* from intron 4 to exon 6, one copy of *CFHR1* and two normal copies of *CFHR4*, *CFHR2* and *CFHR5*. **(B)** Pedigree (#1876) of the C3GN patient (II-4; indicated by the black circle) carrying the *CFHR3* SV (H, indicated in red) on one allele and the *CFHR3-CFHR1* del (Δ, indicated in green) on the other allele. The patient also carries a variant of unknown significance (VUS; indicated with a filled circle) in *CFHR4* (p.Val438Gly; rs766466004; gnomAD global MAF: 4×10^{-6}). Both the *CFHR3* SV and the *CFHR4* VUS were also found in the maternal cousin (II-7, indicated by the black circle) who has an MPGN diagnosis but, not in the patient's healthy sons. **(C–F)** Western Blot (WB) analyses were performed under non-reducing conditions using the sera from the proband (#2870), a healthy control with normal CNVs (positive control) and a patient carrying the homozygous *CFHR3-CFHR1* del (negative control). Using the rabbit anti-FHR3 polyclonal antiserum (diluted 1:2000; panel **C**,**D**) we did not observe the predicted band of the shorter FHR3 at 16 kDa (**C**; predicted MW based on the partial *CFHR3* deletion). Instead we observed two bands with a MW (around 50 kDa) higher than normal FHR3, which are better evidenced in **(D)**, obtained after a longer run. Using an anti-FHR1 antibody (JHD; diluted 1:1,000) we found both the two bands corresponding to normal glycosylated isoforms of FHR1 (around 37 and 41 kDa, respectively) and two abnormal bands around 50 kDa, identical to those observed with the anti-FHR3 antiserum **(E)**. The same WB pattern was confirmed using the anti-FHR1-2-5 antibody (2C6; **F**). Altogether the WB findings indicate the presence in the proband of both the normal FHR1 and a fusion protein encompassing FHR3 and FHR1.

kidney transplantation 33 years after onset. Neither the *CFHR3* genomic abnormality nor the *CFHR4* variant were identified in the unaffected patient's daughter (III-4; **Figure 8B**) or in a healthy paternal female cousin (II-1; **Figure 8C, D**) or in 100 healthy controls.

The predicted MW of the protein encoded by the partially deleted *CFHR3* gene is about 16 kDa. However, WB analyses of patient serum using an anti-FHR3 antibody showed two bands with a MW around 50 kDa and no bands at 16 kDa (**Figures 8C, D**). Western blot with an anti-FHR1 antibody revealed two bands corresponding to normal glycosylated isoforms of FHR1 and two additional bands with MWs (about 50 kDa; **Figure 8E**) identical to the bands observed with the anti-FHR3 antibody. The same results were observed with an anti-FHR1-2-5 antibody (**Figure 8F**). These results suggest the presence of 1) a hybrid protein between the shorter FHR3 and the full FHR1 (likely $FHR3_{1-3}$-FHR1); 2) a normal FHR1.

CFH-CFHR3-CFHR1 Gene Deletion

Heterozygosity for a large deletion that included *CFH*, *CFHR3* and *CFHR1* was identified by MLPA analysis (**Figure 9A**) in a patient in cluster 2 with histologic diagnosis of IC-MPGN (#2888; **Table 4**). At the age of 16, the patient presented with nephrotic syndrome, haematuria, low C3 levels (8.5 mg/dl), normal C4 and hypertension. No family history of nephropathy was reported. After 25 years, renal function deteriorated and the patient underwent a pre-emptive kidney transplantation (the donor was his father). Two years later, the patient lost the allograft due to rejection and started dialysis. At that time, biomarkers showed low C3 (47 mg/dl) and normal C4 (27 mg/dl). No C3NeF or FHAAs were detected and genetic screening did not reveal RVs in complement genes. Consistent with the deletion of one copy of *CFH*, FH levels were low (156 mg/dl). SMRT sequencing confirmed a 254 kb long deletion from chr1:196584749 (between *KCNT2*

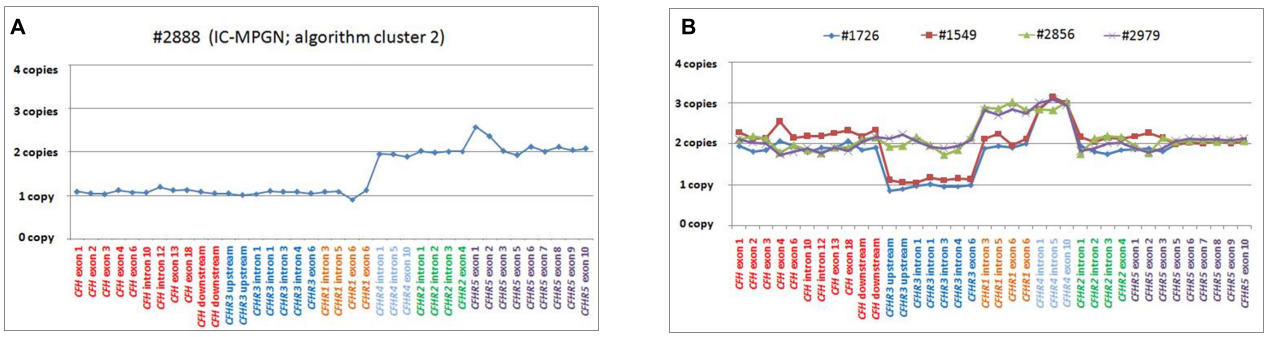


FIGURE 9 | Graphic representation of MLPA results from the patient carrying the *CFH-CFHR3-CFHR1* deletion and patients with *CFHR4* duplication. **(A)** MLPA results showing the genomic deletion, including the entire copy of *CFH*, *CFHR3*, and *CFHR1* in a patient with IC-MPGN (#2888; cluster 2). **(B)** Analysis of MLPA showing three copies of *CFHR1* and *CFHR4* in two C3GN patients (#2856 and #2979; both from cluster 1) and one copy of *CFHR3*, two copies of *CFHR1* and 3 copies of *CFHR4* in two patients with IC-MPGN and DDD (#1726 cluster 2; #1549, cluster 3), respectively.

and *CFH*) and extending to chr1:196839345 (in the *CFHR1-CFHR4* intergenic region) (**Figure 7C**). Breakpoints were confirmed by Sanger sequencing (primers are reported in **Supplementary Table 2**). This deletion was not identified in any controls.

Gene Duplications

Two young male patients from cluster 1 (#2856, #2979; **Table 4**) carried a duplication of *CFHR1-CFHR4* and therefore had 3 copies of both *CFHR1* and *CFHR4* (**Figure 9B**).

The first, patient #2856, presented with proteinuria and haematuria at age 11 and had biopsy-confirmed C3GN. In the following years, he experienced progressive proteinuria, peaking at 11.8 g/day at the age of 25. C3 levels were low (9 mg/dl), sC5b-9 levels were high (1,930 ng/ml) and he was C3NeF positive. Mycophenolate mofetil (MMF; 2g/day) and prednisone (PDN; 1 mg/kg) were initiated, with an associated reduction in proteinuria (1.1 g/day) and at last follow-up (at 26 years of age), C3 levels had improved, sC5b-9 levels had normalized (328 ng/ml), and C3NeF was absent.

The second case, patient #2979, presented with proteinuria (0.26 g/day), haematuria and low C3 (51 mg/dl) at the age of 5; one year later, because of the persistence of proteinuria, he underwent a kidney biopsy, which showed C3GN. At last follow-up, one year later, proteinuria had increased (0.69 g/day), renal function was normal (creatinine 0.34 mg/dl), and C3 levels remained low (66 mg/dl).

One patient from cluster 2 (#1726) with IC-MPGN also carried 3 copies of *CFHR4* (but at variance with the first two cases, she had only two copies of *CFHR1* and one copy of *CFHR3;* **Figure 9B**). Disease developed during pregnancy when she presented at age 25 with proteinuria, microhaematuria, low C3 (20 mg/dl) and C4 (6 mg/dl) but normal renal function. Post-pregnancy treatment included chronic immunosuppression (corticosteroids, cyclophosphamide, MMF) and antihypertensive therapies (ACE inhibitors and ARBs), and at 45 years of age, C3 and C4 levels were normal and proteinuria and haematuria resolved. At last follow-up (at the age of 46), creatinine was 0.9 mg/dl, C3 and C4 were 140 mg/dl and 12 mg/dl, respectively, and morning urine spot was negative for microhaematuria and slightly positive for proteinuria (155 mg/g creatinine, normal values < 200 mg/g). C3NeFs and FHAAs were absent. Segregation analysis showed that the patient inherited an allele with zero copies of *CFHR3*, one copy of *CFHR1* and two copies of *CFHR4* from the unaffected father (allele A, **Supplementary Figure 4**). The other allele is normal. Of the two unaffected sons, one has inherited the maternal abnormal allele A and the paternal *CFHR3-CFHR1* deletion allele (**Supplementary Figure 4**). NGS also identified homozygosity for the *CFB* RV (p.Arg679Trp, gnomAD global MAF: 0), inherited from the consanguineous healthy parents. The patient's sons are heterozygous for this variant (**Supplementary Figure 4**).

The same MLPA pattern seen in #1726 was also identified in a patient in cluster 3, who was diagnosed with DDD at 25 years of age when he developed nephrotic range proteinuria (6.2 g/day) in the face of low C3 levels (54 mg/dl) (#1549; **Table 4** and **Figure 9B**). Renal function and blood pressure remained normal and conservative therapy with statins and ACE inhibitors was initiated, resulting in progressive reduction of proteinuria to below the nephrotic range. The patient was C3NeF positive. The patient has remained stable and at last follow-up (at the age of 34) had sub-nephrotic range proteinuria (1.8 g/24 h) and normal renal function (creatinine 0.55 mg/dl). C3 remained low (64 mg/dl) but sC5b-9 was normal (142 ng/ml) and C3NeFs had resolved. Segregation analysis showed that the abnormal allele (allele A) was maternally inherited. NGS studies identified a heterozygous RV in *CFH* (p.Arg2Ile; gnomAD global MAF: 0) that does not appear to impact FH levels (216 mg/dl); this variant was also maternally inherited.

Notably, we were not able to discriminate between carriers (#2856, #1726, and #1549) and non-carriers of the *CFHR1-CFHR4* duplication with CCS reads, likely due to the fact that the breakpoints of this SV are in the r1/r2 duplicated regions, which can lead to erroneous mapping (**Supplementary Figure 5**), as described in the Section "Materials and Methods." No controls had more than 2 copies of *CFHR1* and/or *CFHR4*.

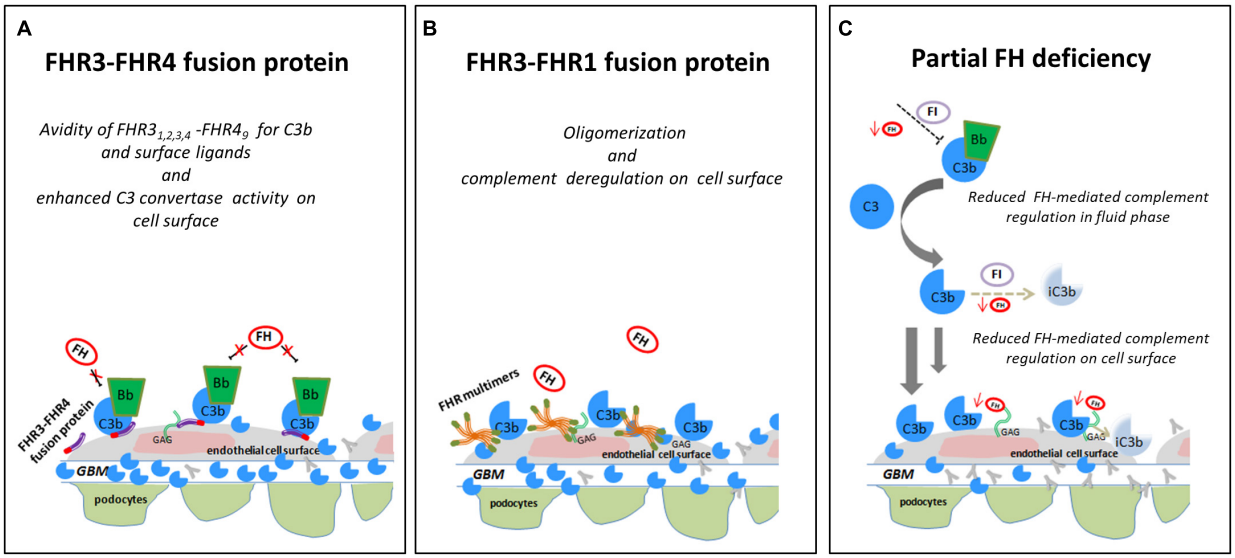


FIGURE 10 | Hypothesis of the effects of FHR SVs on FH-complement regulation. **(A)** The hypothesis is that the FHR3$_{1-4}$-FHR4$_9$ fusion protein identified in a DDD patient (cluster 3) binds GAGs and C3b on glomerular cells, favouring the formation of an active AP C3 convertase that is resistant to FH-mediated decay, promoting the formation of highly electron-dense deposits in the glomerular basement membrane (GBM). **(B)** The FHR3-FHR1 fusion protein identified in a C3GN patient (cluster 4), through FHR1 portion, may generate multimeric complexes and through FHR3 domains increase the affinity of multimers for FH ligands and C3b, preventing FH-complement regulation (this process is known "FH deregulation"). The final effect is the bright C3 glomerular staining in the face of normal circulating C3. **(C)** In an IC-MPGN patient (cluster 2) we identified a heterozygous deletion of *CFH-CFHR3-CFHR1*. Low FH serum levels caused by the heterozygous deletion of the entire *CFH* gene may result in impaired FH-complement regulation both in the fluid phase and on the glomerular surface. The consequence is the deposition of C3b molecules on endothelial cells that promote glomerular chronic complement activation caused by immune-complexes.

DISCUSSION

Here we performed a comprehensive analysis to characterize genetic and acquired FH-FHR abnormalities in a large cohort of 199 C3G/IC-MPGN patients, classified into four clusters, with the main focus on *CFH-CFHR* CNVs.

Low FH levels and genetic and acquired FH abnormalities were identified only in patients in clusters 1–3, which are characterized by fluid-phase complement activation. Specifically, 7% of cluster 1–3 patients had *CFH* RVs, consistent with our results in a smaller cohort (Iatropoulos et al., 2018). FHAAs were also found in 5% of cluster 1–3 patients, all with childhood onset. All but 1 FHAA-positive patient were diagnosed with IC-MPGN, suggesting a possible link between FHAAs and immune-complexes in the glomeruli. In addition, the majority of patients with FHAAs were co-positive for another autoantibody, C3NeF, consistent with other reports (Blanc et al., 2015). These findings suggest a cumulative or synergistic effect of FHAAs and C3NeF in inducing fluid-phase AP overactivation although the specific contribution of each autoantibody remains unclear.

At variance with aHUS patients, we did not observe a correlation between the presence of FHAAs and homozygosity for *CFHR1* del in C3G/IC-MPGN patients, consistent with previous data (Blanc et al., 2015; Valoti et al., 2019; Zhang et al., 2020). In addition, the prevalence of the common SVs (*CFHR3-CFHR1* del or *CFHR1-CFHR4* del) did not differ between patients and healthy controls, indicating that common SVs are not risk factors for C3G/IC-MPGN. However, the finding that total deficiency for *CFHR1* was more frequent in patients in cluster 3 compared to patients in cluster 1, may indicate that FHR1 deficiency plays a role in driving the disease phenotype characteristic of cluster 3 patients.

To date, with the exception of the fusion protein (FHR5$_{1,2}$-FHR5) identified in Greek Cypriot patients with C3GN, rare SVs in the *CFH-CFHR* region have been described in only a few familial cases of C3G. They have not been implicated in IC-MPGN. This knowledge gap reflects, in part, the high degree of similarity within the *CFH-CFHR* region, which is a strong limitation in designing specific probes for copy number variation (CNV) analysis and leads to an incomplete investigation of this locus.

To optimize the CNV analysis in this region, we used available and custom MLPA probes to provide an overview of SVs, which we then further resolved through PacBio long-read sequencing (SMRT, Single- Molecule Real-Time, sequencing). Using this protocol, we identified rare *CFH-CFHR* SVs in patients with IC-MPGN and an overall prevalence of 4% of new and rare *CFH-CFHR* SVs in C3G/IC-MPGN patients.

We detected a duplication of *CFHR1-CFHR4* in 2% of patients distributed amongst clusters 1–3 but not in cluster 4, often in combination with other complement RVs and/or the common *CFHR3-CFHR1* del. This duplication has also been identified in patients with aHUS and AMD (Bu et al., 2014; Cantsilieris et al., 2018). We verified segregation in healthy relatives indicating that, alone, the *CFHR1-CFHR4* duplication is not sufficient to induce

disease and that other risk factors are required to determine the ultimate phenotype.

Interestingly, another genomic rearrangement altering *CFHR4* was identified in a DDD patient from cluster 3, namely a *CFHR3*$_{1-5}$-*CFHR4*$_{10}$ hybrid gene that encodes the fusion protein FHR3$_{1,2,3,4}$-FHR4$_9$. SCRs 1-3 of FHR3 have high sequence similarity with FH SCRs 6-8, which form a second FH heparan-sulfate binding site on cell surfaces and the glomerular basement membrane (GBM) (Borza, 2017). Hebecker and Jozsi have shown that FHR4 favors the assembly of the AP C3 convertase through its C-terminal region, which contains a C3b binding sites (Hebecker and Jozsi, 2012). These data suggest that the FHR3$_{1,2,3,4}$-FHR4$_9$ fusion protein may compete with FH for binding to both glycosaminoglycans/sialic acid and C3b fragments in the GBM, thereby enhancing C3 convertase activity and favoring the formation of the high electron-dense deposits, a characteristic feature of cluster 3-patients (**Figure 10A**). This hypothesis warrants testing.

In addition to the above *CFHR4* CNVs, in a familial case of C3GN in cluster 4 we identified a shorter *CFHR3*$_{1-3}$ gene caused by a deletion spanning intron 3 to 3′UTR, followed by a normal copy of *CFHR1*, which leads to a fusion protein likely consisting of the 2 N-terminal SCRs of FHR3 and the entire FHR1 (FHR3$_{1-2}$-FHR1). A comparable fusion protein generated by a different genomic rearrangement has been described by Malik et al. (2012) in a familial C3GN case. In both cases, C3 levels are normal, suggesting that complement dysregulation occurs primarily in the glomeruli microenvironment. The likely mechanism of action is secondary to multimeric complexes of FH-related proteins that outcompete FH for binding to the glomerular glycomatrix (Goicoechea de Jorge et al., 2013; Medjeral-Thomas and Pickering, 2016; Csincsi et al., 2017) (**Figure 10B**). Functional studies, however, would be required to elucidate the functional effects of the identified genomic *CFHR* abnormalities and their pathogenetic role in C3G/IC-MPGN.

A final important finding of this study is the identification of a large deletion encompassing *CFH*, *CFHR3*, and *CFHR1* in a IC-MPGN patient in cluster 2 with low C3 and FH serum levels. The deletion causes FH haplodeficiency. As a consequence, fluid-phase AP regulation is impaired, which thereby sustains chronic complement activation initiated through the CP by immune-complexes in the glomeruli, a feature typical of cluster 2 patients (**Figure 10C**).

CONCLUSION

In this study we have used established and innovative techniques to characterize SVs over the *CFH-CFHR* genomic region in a large cohort of C3G/IC-MPGN patients. We have demonstrated that while common *CFH-CFHR* SVs are not risk factors for disease, rare SVs do predispose to disease, but typically in combination with RVs in complement genes or acquired drivers of disease like autoantibodies. Our findings support the overarching concept that C3G/IC-MPGN are genetically complex, with the ultimate phenotype reflecting the delicate balance of serum levels of FH and the FHR proteins. Our results also illustrate the value of SMRT sequencing methodology as a tool for resolving the complexity of SVs in this genomic region.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: EBI European Nucleotide Archive, accession no: PRJEB44176.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of Bergamo. 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

RP, MN, and GR designed research, interpreted data, and wrote the manuscript. RP, MB, MA, EV, PI, CM, RD, and PC performed the research and analyzed the data. EB provided detailed clinical information of patients. AB and RS analyzed the data and critically revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was partially supported by the Kidneeds Foundation (Kidneeds projects 735/8215 and 698/7603), the European Union Framework Programme 7 (FP7)-EURenOmics project 305608, the Italian Ministero della Salute (RF-2016-02361720), the National Institutes of Health R01 DK110023 (RS), and Fondazione Regionale per la Ricerca Biomedica (Regione Lombardia), Project ERAPERMED2020-151, GA 779282. RP, MB, and EV were recipients of a research contract from Progetto DDD Onlus – Associazione per la lotta alla DDD (Milan, Italy). PC was the recipient of a fellowship from Fondazione Aiuti per la Ricerca sulle Malattie Rare ARMR ONLUS (Bergamo, Italy). The funding sources had no role in study design, or in the collection, analysis and interpretation of data, nor in the writing of the report or in the decision to submit the paper for publication.

ACKNOWLEDGMENTS

We thank Gregorini, Scolari, D'Amico, Pecoraro, Fisher, Hirt-Minkowski, and Forster for providing clinical and histology data of patients, the laboratory of clinical chemistry at the Mario Negri Institute for biochemical evaluation of C3 and C4, Serena Bettoni for assistance with WB analysis, Sara Gamba

for the collection of biological samples, Miriam Rigoldi for managing and reviewing clinical data of patients, Paola Rizzo for histological images, Nicole Meyer, Bertha Martin, Nicolò Ghiringhelli Borsa and Carla Nishimura for NGS support at University of Iowa, Kerstin Mierke for editing the manuscript. We would like to thank Ave Tooming-Klunderud for SMRT sequencing service, which was provided by the Norwegian Sequencing Centre (www.sequencing.uio.no), a national technology platform hosted by the University of Oslo and supported by the "Functional Genomics" and "Infrastructure" programmes of the "Research Council of Norway and the Southeastern Regional Health Authorities", and David Stucki and Deborah Moine from PacBio for bioinformatic assistance. We thank Peter Zipfel (from Friedrich Schiller University Jena, Jena) and Santiago Rodriguez de Córdoba (from Centro de Investigaciones Biologicas, Madrid) for kindly providing us with the anti-FHR antibodies. Finally, we acknowledge all clinicians and patients for their membership in and support to the MPGN/C3G Registry.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.670727/full#supplementary-material

REFERENCES

Bhakdi, S., and Tranum-Jensen, J. (1988). Damage to cell membranes by pore-forming bacterial cytolysins. *Prog. Allergy* 40, 1–43. doi: 10.1159/000414943

Blanc, C., Togarsimalemath, S. K., Chauvet, S., Le Quintrec, M., Moulin, B., Buchler, M., et al. (2015). Anti-factor H autoantibodies in C3 glomerulopathies and in atypical hemolytic uremic syndrome: one target, two diseases. *J. Immunol.* 194, 5129–5138. doi: 10.4049/jimmunol.1402770

Borza, D. B. (2017). Glomerular basement membrane heparan sulfate in health and disease: a regulator of local complement activation. *Matrix Biol.* 5, 299–310. doi: 10.1016/j.matbio.2016.09.002

Bu, F., Maga, T., Meyer, N. C., Wang, K., Thomas, C. P., Nester, C. M., et al. (2014). Comprehensive genetic analysis of complement and coagulation genes in atypical hemolytic uremic syndrome. *J. Am. Soc. Nephrol.* 25, 55–64. doi: 10.1681/ASN.2013050453

Cantsilieris, S., Nelson, B. J., Huddleston, J., Baker, C., Harshman, L., Penewit, K., et al. (2018). Recurrent structural variation, clustered sites of selection, and disease risk for the complement factor H (CFH) gene family. *Proc. Natl. Acad. Sci. U.S.A.* 115, E4433–E4442. doi: 10.1073/pnas.1717600115

Caprioli, J., Castelletti, F., Bucchioni, S., Bettinaglio, P., Bresin, E., Pianetti, G., et al. (2003). Complement factor H mutations and gene polymorphisms in haemolytic uraemic syndrome: the C-257T, the A2089G and the G2881T polymorphisms are strongly associated with the disease. *Hum. Mol. Genet.* 12, 3385–3395. doi: 10.1093/hmg/ddg363

Chen, Q., Wiesener, M., Eberhardt, H. U., Hartmann, A., Uzonyi, B., Kirschfink, M., et al. (2014). Complement factor H-related hybrid protein deregulates complement in dense deposit disease. *J. Clin. Invest.* 124, 145–155. doi: 10.1172/jci71866

Cook, H. T., and Pickering, M. C. (2015). Histopathology of MPGN and C3 glomerulopathies. *Nat. Rev. Nephrol.* 11, 14–22. doi: 10.1038/nrneph.2014.217

Csincsi, A. I., Szabo, Z., Banlaki, Z., Uzonyi, B., Cserhalmi, M., Karpati, E., et al. (2017). FHR-1 binds to C-reactive protein and enhances rather than inhibits complement activation. *J. Immunol.* 199, 292–303. doi: 10.4049/jimmunol.1600483

Diaz-Guillen, M. A., Rodriguez de Cordoba, S., and Heine-Suner, D. (1999). A radiation hybrid map of complement factor H and factor H-related genes. *Immunogenetics* 49, 549–552. doi: 10.1007/s002510050534

Donadelli, R., Pulieri P., Piras, R., Iatropoulos, P., Valoti, E., Benigni, A., et al. (2018). Unravelling the molecular mechanisms underlying complement dysregulation by nephritic factors in C3G and IC-MPGN. *Front. Immunol.* 9:2329. doi: 10.3389/fimmu.2018.02329

, Pulieri P., Piras, R., Iatropoulos, P., Valoti, E., Benigni, A., et al.

Feuk, L., Carson, A. R., and Scherer, S. W. (2006). Structural variation in the human genome. *Nat. Rev. Genet.* 7, 85–97. doi: 10.1038/nrg1767

Fremeaux-Bacchi, V., Weiss, L., Brun, P., and Kazatchkine, M. D. (1994). Selective disappearance of C3NeF IgG autoantibody in the plasma of a patient with membranoproliferative glomerulonephritis following renal transplantation. *Nephrol. Dial Transplant.* 9, 811–814.

Gale, D. P., de Jorge, E. G., Cook, H. T., Martinez-Barricarte, R., Hadjisavvas, A., McLean, A. G., et al. (2010). Identification of a mutation in complement factor H-related protein 5 in patients of Cypriot origin with glomerulonephritis. *Lancet* 376, 794–801. doi: 10.1016/s0140-6736(10)60670-8

Gharavi, A. G., Kiryluk, K., Choi, M., Li, Y., Hou, P., Xie, J., et al. (2011). Genome-wide association study identifies susceptibility loci for IgA nephropathy. *Nat. Genet.* 43, 321–327.

Goicoechea de Jorge, E., Caesar, J. J., Malik, T. H., Patel, M., Colledge, M., Johnson, S., et al. (2013). Dimerization of complement factor H-related proteins modulates complement activation in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4685–4690. doi: 10.1073/pnas.1219260110

Goodship, T. H., Cook, H. T., Fakhouri, F., Fervenza, F. C., Fremeaux-Bacchi, V., Kavanagh, D., et al. (2017). Atypical hemolytic uremic syndrome and C3 glomerulopathy: conclusions from a "kidney disease: improving global outcomes" (KDIGO) controversies conference. *Kidney Int.* 91, 539–551.

Hebecker, M., and Jozsi, M. (2012). Factor H-related protein 4 activates complement by serving as a platform for the assembly of alternative pathway C3 convertase via its interaction with C3b protein. *J. Biol. Chem.* 287, 19528–19536. doi: 10.1074/jbc.M112.364471

Holmes, L. V., Strain, L., Staniforth, S. J., Moore, I., Marchbank, K., Kavanagh, D., et al. (2013). Determining the population frequency of the CFHR3/CFHR1 deletion at 1q32. *PLoS One* 8:e60352. doi: 10.1371/journal.pone.0060352

Hou, J., Markowitz, G. S., Bomback, A. S., Appel, G. B., Herlitz, L. C., Barry Stokes, M., et al. (2014). Toward a working definition of C3 glomerulopathy by immunofluorescence. *Kidney Int.* 85, 450–456. doi: 10.1038/ki.2013.340

Hughes, A. E., Orr, N., Esfandiary, H., Diaz-Torres, M., Goodship, T., and Chakravarthy, U. (2006). A common CFH haplotype, with deletion of CFHR1 and CFHR3, is associated with lower risk of age-related macular degeneration. *Nat. Genet.* 38, 1173–1177. doi: 10.1038/ng1890

Iatropoulos, P., Daina, E., Curreri, M., Piras, R., Valoti, E., Mele, C., et al. (2018). Cluster analysis identifies distinct pathogenetic patterns in C3 glomerulopathies/immune complex-mediated membranoproliferative GN. *J. Am. Soc. Nephrol.* 29, 283–294. doi: 10.1681/asn.2017030258

Iatropoulos, P., Noris, M., Mele, C., Piras, R., Valoti, E., Bresin, E., et al. (2016). Complement gene variants determine the risk of immunoglobulin-associated MPGN and C3 glomerulopathy and predict long-term renal outcome. *Mol. Immunol.* 71, 131–142. doi: 10.1016/j.molimm.2016.01.010

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

Lupski, J. R., and Stankiewicz, P. (2005). Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet.* 1:e49. doi: 10.1371/journal.pgen.0010049

Maga, T. K., Nishimura, C. J., Weaver, A. E., Frees, K. L., and Smith, R. J. (2010). Mutations in alternative pathway complement proteins in American patients with atypical hemolytic uremic syndrome. *Hum. Mutat.* 31, E1445–E1460. doi: 10.1002/humu.21256

Malik, T. H., Lavin, P. J., Goicoechea de Jorge, E., Vernon, K. A., Rose, K. L., Patel, M. P., et al. (2012). A hybrid CFHR3-1 gene causes familial C3 glomerulopathy. *J. Am. Soc. Nephrol.* 23, 1155–1160. doi: 10.1681/asn.2012020166

Marinozzi, M. C., Roumenina, L. T., Chauvet, S., Hertig, A., Bertrand, D., Olagne, J., et al. (2017). Anti-factor B and anti-C3b autoantibodies in C3

glomerulopathy and ig-associated membranoproliferative GN. *J. Am. Soc. Nephrol.* 28, 1603–1613. doi: 10.1681/asn.2016030343

Medjeral-Thomas, N., Malik, T. H., Patel, M. P., Toth, T., Cook, H. T., Tomson, C., et al. (2014). A novel CFHR5 fusion protein causes C3 glomerulopathy in a family without Cypriot ancestry. *Kidney Int.* 85, 933–937. doi: 10.1038/ki.2013.348

Medjeral-Thomas, N., and Pickering, M. C. (2016). The complement factor H-related proteins. *Immunol. Rev.* 274, 191–201. doi: 10.1111/imr.12477

Moore, I., Strain, L., Pappworth, I., Kavanagh, D., Barlow, P. N., Herbert, A. P., et al. (2010). Association of factor H autoantibodies with deletions of CFHR1, CFHR3, CFHR4, and with mutations in CFH, CFI, CD46, and C3 in patients with atypical hemolytic uremic syndrome. *Blood* 115, 379–387. doi: 10.1182/blood-2009-05-221549

Morgan, B. P. (1999). Regulation of the complement membrane attack pathway. *Crit. Rev. Immunol.* 19, 173–198.

Muller-Eberhard, H. J. (1986). The membrane attack complex of complement. *Annu. Rev. Immunol.* 4, 503–528. doi: 10.1146/annurev.iy.04.040186.002443

Noris, M., Caprioli, J., Bresin, E., Mossali, C., Pianetti, G., Gamba, S., et al. (2010). Relative role of genetic complement abnormalities in sporadic and familial aHUS and their impact on clinical phenotype. *Clin. J. Am. Soc. Nephrol.* 5, 1844–1859. doi: 10.2215/CJN.02210310

Noris, M., and Remuzzi, G. (2015). Glomerular diseases dependent on complement activation, including atypical hemolytic uremic syndrome, membrano-proliferative glomerulonephritis, and C3 glomerulopathy: core curriculum 2015. *Am. J. Kidney Dis.* 66, 359–375. doi: 10.1053/j.ajkd.2015.03.040

Osborne, A. J., Breno, M., Borsa, N. G., Bu, F., Fremeaux-Bacchi, V., Gale, D. P., et al. (2018). Statistical validation of rare complement variants provides insights into the molecular basis of atypical hemolytic uremic syndrome and C3 glomerulopathy. *J. Immunol.* 200, 2464–2478. doi: 10.4049/jimmunol.1701695

Pechtl, I. C., Kavanagh, D., McIntosh, N., Harris, C. L., and Barlow, P. N. (2011). Disease-associated N-terminal complement factor H mutations perturb cofactor and decay-accelerating activities. *J. Biol. Chem.* 286, 11082–11090. doi: 10.1074/jbc.M110.211839

Pickering, M. C., D'Agati, V. D., Nester, C. M., Smith, R. J., Haas, M., Appel, G. B., et al. (2013). C3 glomerulopathy: consensus report. *Kidney Int.* 84, 1079–1089.

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

Servais, A., Noel, L. H., Roumenina, L. T., Le Quintrec, M., Ngo, S., Dragon-Durey, M. A., et al. (2012). Acquired and genetic complement abnormalities play a critical role in dense deposit disease and other C3 glomerulopathies. *Kidney Int.* 82, 454–464. doi: 10.1038/ki.2012.63

Sethi, S., and Fervenza, F. C. (2011). Membranoproliferative glomerulonephritis: pathogenetic heterogeneity and proposal for a new classification. *Semin. Nephrol.* 31, 341–348. doi: 10.1016/j.semnephrol.2011.06.005

Skerka, C., Chen, Q., Fremeaux-Bacchi, V., and Roumenina, L. T. (2013). Complement factor H related proteins (CFHRs). *Mol. Immunol.* 56, 170–180. doi: 10.1016/j.molimm.2013.06.001

Togarsimalemath, S. K., Sethi, S. K., Duggal, R., Le Quintrec, M., Jha, P., Daniel, R., et al. (2017). A novel CFHR1-CFHR5 hybrid leads to a familial dominant C3 glomerulopathy. *Kidney Int.* 92, 876–887. doi: 10.1016/j.kint.2017.04.025

Tortajada, A., Yebenes, H., Abarrategui-Garrido, C., Anter, J., Garcia-Fernandez, J. M., Martinez-Barricarte, R., et al. (2013). C3 glomerulopathy-associated CFHR1 mutation alters FHR oligomerization and complement regulation. *J. Clin. Invest.* 123, 2434–2446. doi: 10.1172/jci68280

Valoti, E., Alberti, M., Iatropoulos, P., Piras, R., Mele, C., Breno, M., et al. (2019). Rare functional variants in complement genes and anti-FH autoantibodies-associated aHUS. *Front. Immunol.* 10:853. doi: 10.3389/fimmu.2019.00853

Xiao, X., Ghossein, C., Tortajada, A., Zhang, Y., Meyer, N., Jones, M., et al. (2016). Familial C3 glomerulonephritis caused by a novel CFHR5-CFHR2 fusion gene. *Mol. Immunol.* 77, 89–96. doi: 10.1016/j.molimm.2016.07.007

Zhang, Y., Ghiringhelli Borsa, N., Shao, D., Dopler, A., Jones, M. B., Meyer, N. C., et al. (2020). Factor H autoantibodies and complement-mediated diseases. *Front. Immunol.* 11:607211. doi: 10.3389/fimmu.2020.607211

Zhang, Y., Meyer, N. C., Wang, K., Nishimura, C., Frees, K., Jones, M., et al. (2012). Causes of alternative pathway dysregulation in dense deposit disease. *Clin. J. Am. Soc. Nephrol.* 7, 265–274. doi: 10.2215/cjn.07900811

Zhao, J., Wu, H., Khosravi, M., Cui, H., Qian, X., Kelly, J. A., et al. (2011). Association of genetic variants in complement factor H and factor H-related genes with systemic lupus erythematosus susceptibility. *PLoS Genet.* 7:e1002079. doi: 10.1371/journal.pgen.1002079

Conflict of Interest: MN has received honoraria from Alexion Pharmaceuticals for giving lectures, and for participating in advisory boards and research grants from Omeros, ChemoCentryx, Inception Science Canada, BioCryst Pharmaceuticals. AB has received honoraria from Boehringer Ingelheim, Alexion Phamaceuticals, Janssen Pharmaceuticals, Akebia Therapeutics, Inception Science Canada, BioCryst Pharmaceuticals. GR has consultancy agreements with Boehringer Ingelheim, Janssen Pharmaceuticals, Akebia Therapeutics, Alexion Pharmaceuticals, Alnylam, Inception Science Canada, BioCryst Pharmaceuticals. RS directs the Molecular Otolaryngology and Renal Research Laboratories, which provides genetic testing for complement-mediated renal diseases. None of these activities have had any influence on the results or interpretations in this article.

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 30 July 2021
doi: 10.3389/fgene.2021.645595

Deep Phenotyping and Genetic Characterization of a Cohort of 70 Individuals With 5p Minus Syndrome

Julián Nevado[1,2,3†], Cristina Bel-Fenellós[4], Ana Karen Sandoval-Talamantes[1,2,3,4,5], Adolfo Hernández[6†], Chantal Biencinto-López[4], María Luisa Martínez-Fernández[7], Pilar Barrúz[1], Fernando Santos-Simarro[1,2,3], María Ángeles Mori-Álvarez[1,2,3], Elena Mansilla[1,2,3], Fé Amalia García-Santiago[1,2,3], Isabel Valcorba[1,2,3], Belén Sáenz-Rico[8], María Luisa Martínez-Frías[7] and Pablo Lapunzina[1,2,3†]*

[1] CIBERER, Centro de Investigación Biomédica en Red de Enfermedades Raras, Instituto de Salud Carlos (ISCIII), Madrid, Spain, [2] Instituto de Genética Médica y Molecular (INGEMM)-IdiPAZ, Hospital Universitario La Paz, Madrid, Spain, [3] ITHACA-European Reference Network-Hospital la Paz, Madrid, Spain, [4] Departamento de Investigación y Psicología en Educación, Facultad de Educación, Universidad Complutense de Madrid (UCM), Madrid, Spain, [5] Servicio de Genética, Centro de Rehabilitacion Infantil Teleton (CRIT), Guadalajara, Mexico, [6] Departamento de Economía Financiera y Actuarial y Estadística, Facultad de Comercio y Turismo, Universidad Complutense de Madrid, Madrid, Spain, [7] Spanish Collaborative Study of Congenital Malformations (ECEMC), Research Unit on Congenital Anomalies (UIAC), Instituto de Salud Carlos III (ISCIII), Madrid, Spain, [8] Departamento Estudios Educativos, Facultad de Educación, Universidad Complutense de Madrid, Madrid, Spain

Chromosome-5p minus syndrome (5p-Sd, OMIM #123450) formerly known as *Cri du Chat* syndrome results from the loss of genetic material at the distal region of the short arm of chromosome 5. It is a neurodevelopmental disorder of genetic cause. So far, about 400 patients have been reported worldwide. Individuals affected by this syndrome have large phenotypic heterogeneity. However, a specific phenotype has emerged including global developmental delay, microcephaly, delayed speech, some dysmorphic features, and a characteristic and monochromatic high-pitch voice, resembling a cat's cry. We here describe a cohort of 70 patients with clinical features of 5p- Sd characterized by means of deep phenotyping, SNP arrays, and other genetic approaches. Individuals have a great clinical and molecular heterogeneity, which can be partially explained by the existence of additional significant genomic rearrangements in around 39% of cases. Thus, our data showed significant statistical differences between subpopulations (simple 5p deletions versus 5p deletions plus additional rearrangements) of the cohort. We also determined significant "functional" differences between male and female individuals.

Keywords: 5p-minus syndrome, intellectual disabilities, Cri du chat, subtelomeric deletion, behavior problems

OPEN ACCESS

Edited by:
Katalin Komlosi, Medical Center – University of Freiburg, Germany

Reviewed by:
Shabeesh Balan, RIKEN Center for Brain Science (CBS), Japan
Emanuela Volpi, University of Westminster, United Kingdom
Rincic Martina, University of Zagreb, Croatia

***Correspondence:**
Julián Nevado
jnevado@salud.madrid.org

†ORCID:
Julián Nevado
orcid.org/0000-0001-5611-2659
Adolfo Hernández
orcid.org/0000-0003-1078-2328
Pablo Lapunzina
orcid.org/0000-0002-6324-4825

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *23 December 2020*
Accepted: *14 May 2021*
Published: *30 July 2021*

Citation:
Nevado J, Bel-Fenellós C, Sandoval-Talamantes AK, Hernández A, Biencinto-López C, Martínez-Fernández ML, Barrúz P, Santos-Simarro F, Mori-Álvarez MÁ, Mansilla E, García-Santiago FA, Valcorba I, Sáenz-Rico B, Martínez-Frías ML and Lapunzina P (2021) Deep Phenotyping and Genetic Characterization of a Cohort of 70 Individuals With 5p Minus Syndrome. Front. Genet. 12:645595. doi: 10.3389/fgene.2021.645595

INTRODUCTION

The syndrome of 5p- (5p- Sd) is caused by partial deletion of the short arm of chromosome 5. The size of the deletion is variable ranging from 500 kb or less to 45 Mb (Simmons et al., 1995; Gu et al., 2013; Elmakky et al., 2014). This syndrome is a rare chromosomal disease, with an incidence between 1 in 15,000 and 1 in 50,000 live births (Niebuhr, 1978; Higurashi et al., 1990; Cerruti Mainardi, 2006). The prevalence is higher among females (66%) than males, but the reason

is unclear. No differences in prevalence between races or geographical areas have been found or related to prenatal events or age of the parents. In Spain, it is estimated that there are around 500–700 patients (Rodríguez-Caballero et al., 2012, and unpublished data from patient's Associations). It has been suggested that the great phenotypic variability observed among individuals with this syndrome is related to both the size and the location of the deletion (between 5p15.3 and 5p15.2 bands), since it is a chromosomal region with a large gene content (Nguyen et al., 2015; Correa et al., 2019).

Ninety percent of cases are *de novo*, and 10% are inherited, due to a rearrangement in the parents (unbalanced segregation of translocations, or recombination involving a pericentric inversion, rarely a parental mosaicism, or an inherited terminal deletion). In *de novo* cases, between 80% and 90% are of paternal origin possibly due to chromosome breakage during the formation of male gametes (Cerruti Mainardi et al., 2001). Prenatal diagnoses in 5p- Sd (at 12–16 weeks of gestation) are common because fetuses frequently show abnormal ultrasound signs (∼65–90%) including cerebral abnormalities, cerebellar hypoplasia, absent/hypoplastic nasal bone, hydrops fetalis, ascites or encephalocele, hypospadias, lung dysplasia, IUGR, microcephaly, and micro/retrognatia (Mak et al., 2019; Su et al., 2019; Peng et al., 2020).

Over the past decade, the accuracy of genetic diagnosis and the advances of analytical techniques have allowed to expand the genetic information associated with the short arm of chromosome 5. However, a full map of the involved genes in this syndrome is not completely established, nor the consequences of their haploisufficiency for subjects with 5p- Sd. In this sense, Nguyen et al. (2015) established a role for 11 dose-sensitive genes within the 5p- arm. In five of them, losses may lead to haploinsufficiency (*TERT*, *SEMA5A*, *MARCH6*, *CTNND2*, and *NPR3*), and in the remaining six genes their haploinsufficiency is conditioned by an additional environmental factor (*SLC6A3*, *CDH18*, *CDH12*, *CDH10*, *CDH9*, and *CDH6*). In addition, two additional genes have been suggested to have haplolethal effects (*RICTOR* and *DAB2*).

We here describe the clinical and molecular data of a cohort of 70 unrelated patients with a cytogenetic and/or molecular diagnosis of 5p- Sd. High-resolution single-nucleotide polymorphism (SNP) array, cytogenetic, fluorescence *in situ* hybridization (FISH), and multiplex ligation-probe amplification (MLPA) techniques were applied to most patients, in order to delineate the size, extent, gene content, and additional rearrangements. Genotype–phenotype relationship analyses were also established. A comparison of the clinical features with published patients in the literature and relevant findings that all patients share in this series were also discussed.

MATERIALS AND METHODS

Subjects

During the period between 2017 and 2020, around 100 patients with 5p minus syndrome, formerly called Cri du chat syndrome (CDCS), were recruited for this study in our center. At this moment, around 30 cases had incomplete either clinical or molecular data and were finally not included in this study. The final cohort is constituted by 70 individuals (see **Figure 1** and **Supplementary Data**). Most of the DNA samples from these patients were extracted and analyzed by SNP arrays at INGEMM (Madrid, Spain), and standard cytogenetic studies were made at the Spanish Collaborative Study of Congenital Malformations Centre (ECEMC) and INGEMM. Clinical information of patients was obtained from the referring physicians by two standardized questionnaires (INGEMM and ECEMC), completed with data of the medical reports, and interviewing most of the parents. Parents or guardians provided informed consent and the Institutional Review Board of our Hospital approved the study (HULP, Madrid, Spain).

Methods

Karyotyping and FISH

Cytogenetic analyses were performed on GTG-banded metaphases at a resolution of about 550 bands according to standard laboratory protocol using Chromosome Kit P (Euroclone, Siziano PV, Italy). FISH was performed according to standard laboratory protocols using commercial subtelomeric 5pter probes, LPU 013SA (covering CTNND2, 5p15.2 and UBE2QL1, 5p15.31, with control at 5q35) and probe FLJ25076 (CytoCell Ltd., Tarrytown, NY, United States) and probe CTNND2 (from Kreatech, Leica, Wetzlar, Germany).

Multiplex Ligation-Probe Amplification (MLPA)

We used MLPA Salsa kits P036 and P070 (subtelomeric probes for all chromosomes) and/or P096 and P358 (specific telomeric probes for the 5p arm) to characterize patients with 5p- Sd (MRC Holland, Amsterdam, Netherland). Data analyses were performed according to the protocols supplied by the provider defining relative probe signals by dividing each measured peak area by the sum of all peak areas of the control probes of that sample. The ratio of each peak's relative probe area was then compared versus a DNA control sample (Promega, United Kingdom), using Coffalyser v.9.4 (MRC Holland).

SNP-Array Analysis

A genome-wide scan of 850,000 tag SNPs (Infinium CytoSNP-850k BeadChip, Illumina, San Diego, CA, United States) was performed at INGEMM, in the majority of the patients, but three (analyzed by array-CGH at ECEMC). They were analyzed by using the Chromosome Viewer tool contained in the Genome Studio package (Illumina). In Chromosome Viewer, gene call scores <0.15 at any locus were considered "no calls." In addition, an allele frequency analysis was applied for all SNPs. All genomic positions were established according to the 2009 human genome build 19 (GRCh37/NCBI build 37.1). Deletion sizes were plotted on the genome browser (**Figure 2** and **Supplementary Data**) using the University of California at Santa Cruz Genome Browser[1].

[1]http://genome.ucsc.edu/; Kent et al., 2002.

FIGURE 1 | (A) Facial features of a patient with 5p minus syndrome at age of 6 days, 7 years, 11 years, and 21 years. **(B)** Details of different ear alterations. **(C)** Details of several dental anomalies, and of a wide mouth. **(D)** Details of some hand and finger anomalies. **(E)** Details of several foot and toe anomalies.

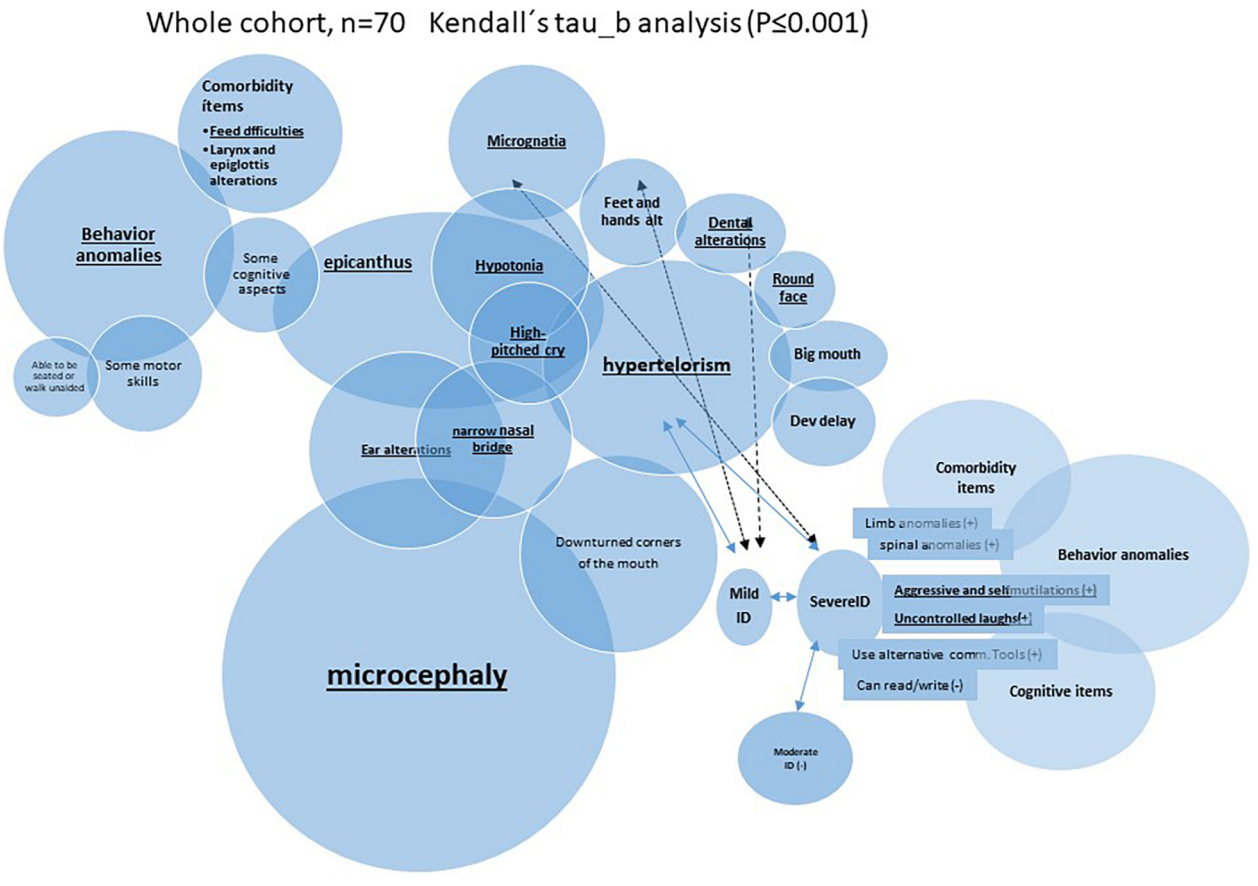


FIGURE 2 | Schematic representation of very significant ($P \leq 0.01$) inter-correlation among microcephaly and other categoric variables (Kendall's tau_b analysis was performed).

TABLE 1 | Descriptive statistics for continuous variables in the whole cohort.

	N	Mean	Standard deviation	Median	Range
Age at evaluation (years)	70	8.99	8.94	7.00	0.1–45
Gestational age at birth (weeks)	68	38.29	2.59	39.00	30–42
Weight at birth (g)	68	2,602.13	677.55	2,600.00	1,170–4,500
OFC at birth (cm)	68	32.20	2.42	32.00	27–37
Height at birth (cm)	68	45.89	3.90	46.75	32–52
Number of surgeries	70	0.71	1.37	0.0001	0–7
Size of deletion (Mb)	70	20.22	9.29	22.55	0.62–35.01

N, number of patients evaluated.

Validation of Global Functional Assessment of the Patients (GFAP)

We estimated individual functional assessment in our cohort by using different features taken from the questionnaires and weighed them by Human Phenotype Ontology (HPO) term frequencies in a numerical scale of five "nuclear" items in the syndrome, based on our clinical experience. A final patient assessment (GFAP) was constructed by the summatory of items "(i) to (v)" as is indicated in **Supplementary Table 1**, and its validation is explained in the "Results" section.

Statistical Analysis

Statistical analysis was performed with SPSS version 25 (IBM Corporation, United States). Descriptive analysis included mean ± SD for continuous variables and frequency tables for categorical variables. These categorical variables were expressed as 1 or 0, indeed grouped as *"ever"* having a given condition compared to *"never"* having the condition, taken from the two questionnaires and curated from medical records. Correlation associations were calculated using Pearson's linear correlation coefficient (continuous variables) or Spearman's Rho and Kendall's tau_b (categorical variables). Comparisons between two groups (as based on sex or to have additional rearrangements) were performed either by Student's *t*-test (for continuous variables) or by chi-square tests (for categorical ones). For more than two groups, ANOVA analysis (and Bonferroni's *post hoc* tests) was run for continuous variables, and *z*-tests between column proportions for categorical variables. PCA (principal component analysis) was used to validate our GFAP construct, containing Kaiser–Meyer–Olkin's measure and Barlett's test. Ward's minimum variance method was the criterion used in hierarchical cluster analysis, and the number of clusters was selected using the Bayesian information criterion (BIC) or Akaike information criterion (AIC). A *P*-value (observed significance level) lower than 0.05 or 0.01 was considered to indicate a statistically significant or very significant difference, respectively.

RESULTS

Clinical Findings

We evaluate 70 unreported individuals. All but three were from Spain (see **Figure 1** and **Supplementary Data**). The female/male ratio (2.04:1; 47/23) was very similar to previously described cohorts, and ages ranged from birth to 45 years (see **Supplementary Table 2**). The highest number of individuals with 5p- Sd in our cohort are individuals in the pediatric age (between 0 and 12 years, 77.23%). Descriptive statistics (for continuous variables) and frequencies (for categoric items) are shown in **Tables 1**, **2**, respectively. The mean and median age at evaluation were 8 years and 9 months, and 7 years old, respectively (**Table 1**).

Perinatal and Neonatal Data

Regarding neonatal data, the average gestational age of our cohort was 38.28 ± 2.59 weeks (**Table 1**). Grossly, 53% (37 subjects) were born between weeks 39th and 40th. Nineteen individuals were born before the 38th week of gestation, three of them below week 32th, and 27 after week 40th. The average birth weight is 2602.13 ± 677.50 g (centile below 5%; Marinescu et al., 2000), which corresponds to the average weight of a neonate of 35th–36th weeks (at centile 50%), and the average length, 45.89 ± 3.90 cm (centile below 5%; Marinescu et al., 2000). Finally, the mean of the cephalic perimeter (OFC) at birth was 32.20 ± 2.42 cm (centile below 5%; Marinescu et al., 2000). More than one-third of subjects were hospitalized at birth. The main causes were prematurity, low weight, and suspected chromosomal abnormality. During the first months of life, several individuals also had feeding difficulties.

Postnatal Clinical Findings

The frequencies of clinical features observed in this cohort were recorded using the HPO terms and are listed in **Table 3**. In **Table 3**, we also listed data from previous published series of 5p- Sd individuals (Cerruti Mainardi et al., 2006; Van Buggenhout et al., 2000; Rodríguez-Caballero et al., 2012; Espirito Santo et al., 2016; Rodrigues de Medeiros, 2017; Honjo et al., 2018; Chehimi et al., 2020). **Figure 1A** shows that facial features are not always typical of the syndrome and that a specific *gestalt* is not always present. Nonetheless, microcephaly, large nose bridge, epicanthal folds, hypertelorism, high arched palate, downturned corners of the mouth, round face, ear anomalies (**Figure 1B**), dental alterations (**Figure 1C**), short philtrum, micrognathia, and feeding difficulties were present in around or higher than 60% of patients. These should considered, in addition to hypotonia, typical cry/acute voice, breathing problems, and behavior anomalies, as the commonest features in this syndrome (**Table 3**). On the other hand, alterations of the hands or feet (see **Figures 1D,E**), hyperlaxity, divergent/convergent strabismus, down-slanting

TABLE 2 | Frequencies for categorical variables in the whole cohort.

Categorical variables	Female	Male
Gender	47 (67.1%)	23 (32.9%)
	0/"never"/condition not present	1/"ever"/condition present
IUGR	43 (61.4%)	27 (38.6%)
Postnatal growth failure	37 (52.9%)	33 (47.1%)
Microcephaly	11 (15.7%)	59 (84.3%)
Facial asymmetry	61 (87.1%)	99 (12.9%)
Round face	38 (54.3%)	32 (45.7%)
Enlarged face	47 (67.1%)	23 (32.9%)
Hearing problems	40 (57.1%)	30 (42.9%)
Ear alterations	32 (45.7%)	38 (54.3%)
Epicanthus	37 (52.9%)	33 (47.1%)
Ophtalmic anomalies	38 (54.3%)	32 (45.7%)
Prominent superciliary arches	65 (92.9%)	5 (7.1%)
Downslanted palpebral fissures	56 (80.0%)	14 (20.0%)
Hypertelorism	29 (41.4%)	41 (58.6%)
Palpebral fissures size anomalies	64 (91.4%)	6 (8.6%)
Nasal defects	56 (80.0%)	14 (20.0%)
Narrow nasal bridge	26 (37.1%)	44 (62.9%)
Short philtrum	60 (85.7%)	10 (14.3%)
Downturned corners of the mouth	62 (88.6%)	8 (11.4%)
Lip and palate anomalies	63 (90.0%)	7 (10.0%)
Micrognathia	40 (57.1%)	30 (42.9%)
Thick lower lip	54 (77.1%)	16 (22.9%)
Big mouth	52 (74.3%)	18 (25.7%)
Teeth alterations	36 (51.4%)	34 (48.6%)
Neck anomalies	57 (81.4%)	13 (18.6%)
Single palmar crease	58 (82.9%)	12 (17.1%)
Breath problems	43 (61.4%)	27 (38.6%)
Cardiac anomalies	46 (65.7%)	24 (34.3%)
Difficult to feed	42 (60.0%)	28 (40.0%)
Larynx and epiglottis alterations	47 (67.1%)	23 (32.9%)
Gastrointestinal anomalies	31 (44.3%)	39 (55.7%)
Renal anomalies	61 (87.1%)	9 (12.9%)
Genital anomalies	54 (77.1%)	16 (22.9%)
Anal anomalies	65 (92.9%)	5 (7.1%)
Limb anomalies	62 (88.6%)	8 (11.4%)
Alterations in hands or feet	39 (55.7%)	31 (44.3%)
Spinal anomalies	52 (74.3%)	18 (25.7%)
Scoliosis	45 (64.3%)	25 (35.7%)
Joint dislocation includes hip	55 (78.6%)	15 (21.4%)
Joint laxity	39 (55.7%)	31 (44.3%)
Pes cavus	57 (81.4%)	13 (18.6%)
MRI images	18 (25.7%)	52 (74.3%)
Anomalies in MRI images	32 (61.5%)	20 (38.5%)
Hypotonia	21 (30.0%)	49 (70.0%)
Hypertonia	63 (90.0%)	7 (10.0%)
Seizures	66 (94.3%)	4 (5.7%)
Developmental delay	4 (5.7%)	66 (94.3%)
Mild ID	62 (88.6%)	8 (11.4%)
Moderate ID	56 (80.0%)	14 (20.0%)
Severe ID	39 (55.7%)	31 (44.3%)

(Continued)

TABLE 2 | Continued

Categorical variables	Female	Male
Behavior anomalies	50 (71.4%)	20 (28.6%)
Autism (ASD)	61 (87.1%)	9 (12.9%)
Hyperactivity	53 (75.7%)	17 (24.3%)
Aggressive and self-mutilation	43 (61.4%)	27 (38.6%)
Stereotypes and repetitive movements	39 (55.7%)	31 (44.3%)
Frustration intolerance	44 (62.9%)	26 (37.1%)
Uncontrolled laughs	50 (71.4%)	20 (28.6%)
Sleeping problems	32 (45.7%)	38 (54.3%)
Cephalic support	19 (27.1%)	51 (72.9%)
Able to stay seated	22 (31.4%)	48 (68.6%)
Able to stay seated unaided	24 (34.3%)	46 (65.7%)
Able to walk unaided	29 (41.4%)	41 (58.6%)
Able to walk with help	23 (32.9%)	47 (67.1%)
Use diapers	38 (55.7%)	31 (44.3%)
Interact with the environment	20 (28.6%)	50 (71.4%)
Can read/write	56 (83.4%)	12 (17.6%)
Use alternative communicative tools	39 (55.7%)	29 (41.4%)
No words at all	44 (65.7%)	26 (34.3%)
Use less than 10 words	43 (63.3%)	25 (36.7%)
Short understandable sentences	52 (76.5%)	16 (23.5%)
High-pitched or horsed cry	31 (44.3%)	39 (55.7%)
Cry w/o sound	68 (97.2%)	2 (2.8%)
Family member no longer work for care	38 (54.3%)	32 (45.7%)
Surgery	46 (67.1%)	24 (32.9%)
Suspicion of pathology prior to diagnosis	48 (68.6%)	22 (31.4%)
Normal electro encephalogram	54 (77.1%)	16 (22.9%)
Normal metabolic screening	53 (75.7%)	17 (24.3%)
Additional duplication	43 (61.4%)	27 (38.6%)

"1" means "ever" having a given condition compared to 0, "never" having the condition, taken from either of our two questionnaires, and curated from medical records.

palpebral fissures, stereotypies, gastrointestinal anomalies, short neck, scoliosis, cardiac anomalies, and speech delay were present in 25–59% of the cases and should be considered frequent findings in the syndrome (**Table 3**).

It is remarkable that many of those called "nuclear clinical features" (the most frequent findings) seemed to be interrelated among them. Indeed, it showed significant positive correlation among them when a Kendall's tau_b analysis was performed. For example, microcephaly presented in more than 65% of the cases correlated with epicanthus, narrow nasal bridge, or ear alterations. As an example, **Figure 2** summarizes some of those very significant intra-correlations ($P \leq 0.01$), e.g., microcephaly. The expandend analyses of these correlations are summarized in **Supplementary Table 3**.

Brain MRI studies were performed in almost 75% of the individuals, though only 28.6% showed some kind of alterations (**Table 2**), including cerebellar amygdala herniation, abnormalities of the corpus callosum (ranging from thinness to agenesis), frontal horn ectasia, brainstem hypoplasia, dilated ventricular system, cysts, or hydrocephalus. Electroencephalograms showed normal results in only a reduced number of individuals (12/70, 22.90%) (**Table 2**).

Speech abilities (evaluated only in patients aged ≥3 years; $n = 56$) showed severe abnormalities in the majority of patients (40/56; ∼71.50%). In fact, 30.35% of patients (17/56) had no speech at all, 41.07% (23/56) had an elementary vocabulary of 10 words or less, and 28.57% (16/56) were reported to have a mild vocabulary and the ability to use limited phrases for a short and comprehensible conversation (**Tables 2**, **3**).

As examples of comorbidities, almost 33% of the cohort undertook at least one surgery (ranging from one to seven, **Table 1**) and include ventricular septal defect (VSD), percutaneous closure of the patent ductus arteriosus, closure of open foramen oval, duodenal atresia, strabismus, and inguinal hernia (the most frequent).

Genetic Findings

Breakpoint Data Analysis

SNP-array analysis was performed in most cases except three patients who had comparative genomic hybridization (CGH array). Genomic coordinates for microdeletions affecting the short arm of chromosome 5 and other genomic rearrangements are listed in **Table 4**. A graphic representation of the deletions is shown in **Supplementary Figure 2**. Briefly, the average size of the losses was 20.21 ± 9.28 Mb (range 0.62–35.01 Mb). SNP arrays established the existence of other clinically significant genomic rearrangements in almost 39% of the patients (**Table 4**);

TABLE 3 | Phenotypic, comorbidities, and global developmental features of our cohort compared to previous published work.

Items	This study	Cerruti Mainardi et al., 2006	Espirito Santo et al., 2016	Honjo et al., 2018	Chehimi et al., 2020	Van Buggenhout et al., 2000	Rodrigues de Medeiros, 2017	Rodríguez-Caballero et al., 2012	Total (factored)
N	*N* = 70	*N* = 220	*N* = 6	*N* = 73	*N* = 14	*N* = 7	*N* = 3	*N* = 32	432
Range of age	0.1–45 years	0.8–61 years	6–38 years	9.5–40 years	2–38 years		17–23 years	2–35 years	0.1–45 years
Mean age	8.80 years		16.80 years	13.80 years	13.30 years		19.66 years	14.65 ± 10.19 years	12.32 years
Developmental delay (HP:0001263)	91.40			100.00	100.00		100.00	89.47	95.11
Hypotonia (HP:0003808)	70.00	72.20	100.00				100.00	87.50	99.38
Micrognathia (HP:0000347)	42.90	96.70	100.00		71.00		100.00	90.62	84.26
Epicanthal folds (HP:0000286)	47.10	90.20	83.30			85.71	100.00	93.75	81.67
Large nose bridge (HP:0000446)	62.90	87.20	100.00		57.00	71.42	100.00	93.75	81.78
Typical cry/acute voice (HP:0200046)	55.70	95.90	100.00	94.40	93.00		33.33	93.75	88.25
Hypertelorism (HP:0000316)	58.60	81.40	83.30		71.00	57.14	100.00	93.75	77.16
Aggressive and self- mutilation (HP:0000718)	84.60							65.63	78.65
Behavior anomalies (HP:0012433)	71.40							68.75	70.57
Round face (HP:0000311)	45.70	83.50	100.00		29.00		100.00	25.00	68.62
High arched palate (HP:0000218)	10.00	83.80	100.00		64.00			16/29 (56.17)	65.58
Independent walking	58.60			72.20					65.54
Low-set ears (HP:0000369)	54.30	69.80	33.30			14.28	100.00	78.12	65.20
Microcephaly (HP:0000252)	84.30		66.70		91.00	85.71		6.25	64.93
Use Diapers	44.30			84.00					64.56
Difficult to feed (HP:0011968)	40.00		60.00	80.30				71.87	62.55
Downturned corners of the mouth (HP:0002714)	11.40						81.00		64.20
Dental anomalies (HP:0000164)	48.60	75.00					100.00	13/23 (56.52)	58.30
Short philtrum (HP:0000322)	14.30	60.50			86.00			96.87	55.40
Hyperlaxity (HP:0002761)	44.30							78.12	54.91
Downslanting palpebrals fissures (HP:0200005)	20.00	56.90	83.30						48.71
Strabismus divergent/convergent (HP:0000486)	45.70	47.50	100.00			42.86		17/31 (54.83)	48.67
Short neck (HP:0000470)	18.60	56.20	33.30				100.00		47.38
Stereotypies (HP:0000733) (HP:0008762)	44.30			40.30					42.25
No words at all (HP:0001344)	35.30			47.20					41.37
Hyperactivity (HP:0000752)	24.30							71.87	39.22
Scoliosis (HP:0002944)	*35.70*		33.30	28.80			*42.60*		38.41
Breath problems (HP:0002098)	*38.60*						100.00		38.38
Cardiac anomalies (HP:0115080)	*34.30*	35.80		31.50			100.00	14/29 (48.27)	36.23
Gastro-intestinal anomalies (HP:0011024)	*55.70*						*21.40*	65.62	33.25

(Continued)

TABLE 3 | Continued

Items	This study	Cerruti Mainardi et al., 2006	Espirito Santo et al., 2016	Honjo et al., 2018	Chehimi et al., 2020	Van Buggenhout et al., 2000	Rodrigues de Medeiros, 2017	Rodríguez-Caballero et al., 2012	Total (factored)
Facial asymmetry (HP:0000324)	12.90					57.14		68.75	31.28
Larynx and epiglottis alterations (HP:0001600)	32.90							28.00	31.36
Enlarged face (HP:0100729)	32.90							13.79	26.90
Alterations of the hands or feets (HP:0011297) (HP:0002813)	44.30		83.30				19.50		26.66
Short understandable sentences	28.50			18.10					22.74
Single palmar crease (HP:0000954)	17.10						100.00		20.51
Autism (HP:0000717)	12.90							25.80	19.95
Genital anomalies (HP:0000078)	22.90			8.20					15.40
Hypertonia (HP:0001276)	*10.00*							15.62	11.76
Renal anomalies (HP:0004742)	*12.90*			5.50					12.67

Data are expressed in percentages. All columns are weighed based on the number of total cases in each item. N, is number of patient evaluated.

most of them were not previously detected by cytogenetic studies (see **Table 4** and **Supplementary Data**). Most subjects had terminal deletions (65/70, 92.85%), and five individuals carried interstitial deletions (represented in **Supplementary Figure 2B**). Among the terminal deletions, 16 of them (22.85%), had an additional terminal duplication in other chromosome, which could result from a possible translocation (*de novo* or inherited). Cytogenetic analysis of the parents allowed us to establish whether the rearrangements were familiar (6 cases) or *de novo* (10 cases). In one case, the 5p deletion was inherited from a maternal mosaicism (6.5% of cells in blood) unknown until the moment of diagnosis of the child. We found patients with additional terminal deletions in other chromosomes (two cases, 2.85%) and additional rearrangements at chromosome-5 nearby the deletions (seven cases, 10%). Finally, three children inherited from their mothers a simple, isolated terminal deletion.

"Functional" Findings

Individual GFAP

The great heterogeneity observed in patients with this syndrome together with the high number of other significant genomic rearrangements (besides the 5p deletions) raised the question whether the presence of these additional rearrangements may modulate functionally the clinical features in this syndrome and to explain the high intra-cohort variability. We proposed a graduation of the individual global assessment of functionality (GFAP), constructing one continuous variable, based on the frequency of the different "nuclear" clinical items (i to v, see section "Materials and Methods"), and our clinical experience in the syndrome.

To verify this GFAP scale construction, a statistical combined analysis of Kaiser–Meyer–Olkin's, Bartlett's, and principal component analysis (PCA) test were performed to detect the best way of association between these grouped clinical features. Indeed, the first principal component (PCA 1) from PCA weighed the major score of the variance, supporting that PCA1 can be written as a weighted average of the five original variables. Finally, Pearson correlation analysis showed that PCA1 and the item GFAP are very significantly correlated (Pearson correlation value = 0.846; P = 0.001). The dispersion plot shows the strong linear correlation among them and therefore justifies GFAP as a valid construct (**Figure 3**).

Table 5 shows the median and mean ± SD values for GFAP and its intermediates "functional" components for the whole cohort, and both subpopulations: simple deletions (47 cases) and patients with deletions and additional rearrangements (mainly duplications, 23 cases).

Comparative Analysis

Among Subpopulations With and Without Additional Rearrangements

A chi-square test was performed to compare categoric variables in both groups: simple (isolated) 5p deletions and those including 5p deletions and additional rearrangements

TABLE 4 | Genomic coordinates from all the rearrangements (GRCh37, hg19).

	Deletions				Duplications			
Individual	**Chromosome**	**Coordinates start**	**Coordinates end**	**Size (Mb)**	**Chromosome**	**Coordinates start**	**Coordinates end**	**Size (Mb)**
1	5p15.33-p15.1	25328	17981563	17.98	Xp22.33	169805	2123990	2.12
2	5p15.33-p14.3	25328	22446214	22.44	10q25.3-q26.3	115402474	135434319	20.08
3	5p15.33-p15.2	25328	12995938	12.99				
4	5p15.33-p14.1	25328	28779357	28.77				
5	5p15.33-p14.3	25328	22956970	22.95				
6	5p15.33-p15.31	25328	9438756	9.43	Xq28	154933691	155236712	3.72
7	5p15.33-p13.2	25328	34602269	34.60				
	5p13.2**	34602654	36816661	2.21				
8	5p15.33-p13.3	25328	31853346	31.85				
9	5p15.33-p13.2	25328	35015297	35.02				
10	5p15.33-p14.1	25328	25290077	25.30				
11*	5p15.32-p15.1	4928318	15418957	10.49				
12	5p15.33-p14.2	25328	24430251	24.43				
13	5p15.33-p14.1	25328	28783716	28.78				
14	5p15.33-p15.32	25328	4938756	4.94				
15	5p15.33-p13.3	25328	34986724	34.98				
16	5p15.33-p15.1	25328	17665529	17.60	12p11.21	32875287	33056330	0.18
17	5p15.33-p14.1	25328	25027051	25.02				
18	5p15.33-p15.1	25328	15913112	15.91	8p23.3p-23.1	176617	11860710	11.86
19	5p15.33-p14.1	25328	25396006	25.40	5p14.1	25409917	28435493	3.025
20	5p15.33-p15.1	25328	15808138	15.81				
21	5p15.33-p15.2	25328	12978580	12.98	10q-26.11-q26.3	121556072	135425341	13.87
22	5p15.33-p15.2	25328	11037420	11.037				
23	5p15.33-p15-32	25328	4356789	4.35	5p15.33-p15.32	4355708	4969019	0.60
					5p15.31	6325532	6642356	0.30
24	5p15.33-p14.1	25328	27108052	27.10				
25	5p15.33-p14.3	25328	21872896	21.88	9p24.3-p22.1	46587	19713500	19.7
26	5p15.33-p14.3	25328	22658970	22.65	8p23.2-p11.23	2061877	34908297	34.94
27	5p15.33-p15.2	25328	14360436	14.36				
28	5p15.33-p13.3	25328	29485091	29.48				
29	5p15.33-p13.3	25328	29292854	29.29	1p13.1-p12	117594464	117989275	0.38
30	5p15.33-p13.3	25328	32130401	32.13				
31	5p15.33-p14.1	25328	27708038	27.71	18p11.32	13034	2656248	2.65
32	5p15.33-p14.1	25328	28147535	28.15				
33	5p15.33-p14.1	25328	26622073	26.62	5p13.3-p13.2	26695268	34019038	7.67
34	5p15.33-p14.1	25328	28796749	28.79				
35*	5p15.33-p15.1	560000	17509888	16.95				
36	5p15.33-p14.1	25328	25821865	25.82				
37	5p15.33-p14.3	25328	21504581	21.50	8p23.3-p23.2	164984	752709	0.75
					22q11.21	25661725	25914593	0.25
38	5p15.33-p15.32	25328	4610206	4.61	5q35.1	169708691	169893751	0.18
					5q35.1-q35.3	171656863	180693344	9.03
39	5p15.33-p15.1	25328	15922302	15.92				
40	5p15.33-p14.2	25328	24438467	24.43				
41	5p15.33-p14.1	25328	25135494	25.13	11q22.1	100578089	100870339	0.29
42	5p15.33-p15.1	25328	15022112	15.02	9p24.3-p21.3	162931	23232287	23.24
43	5p15.33-p14.1	25328	28464893	28.46	18p11.32-p11.31	141896	6785383	6.78
44	5p15.33-p14.2	25328	24247673	24.24				
45	5p15.33-p15.1	25328	17704161	17.70	5p14.3	19970119	20370847	0.4

(Continued)

TABLE 4 | Continued

	Deletions				Duplications			
Individual	**Chromosome**	**Coordinates start**	**Coordinates end**	**Size (Mb)**	**Chromosome**	**Coordinates start**	**Coordinates end**	**Size (Mb)**
46	5p15.33-p13.2	25328	28929082	28.92	5p13.3-p13.2	28968268	34497445	5.53
47	5p15.33-p15.1	25328	16410000	16.41				
48	5p15.33-p13.3	25328	31131409	31.13				
49	5p15.33-p15.2	25328	14270500	14.27				
50	5p15.33-p15.2	25328	14061184	14.06	7p22.3-p11.3	44935-	11857504	11.86
51	5p15.33-p15.1	25328	16661282	16.67				
52	5p15.33	25328	618586	0.62	2q36.3-2q27.3	22990625	243048760	13.14
	2q36.3	228782976-	229669531	0.89				
53	5p15.33-p15.31	25328	7125022	7.12				
54	5p15.33-p15.31	25328	7125022	7.12				
55	5p15.33-p13.3	25328	30210500	30.21				
56	5p15.33-p15.32	25328	4610206	4.61	2p15.3-p25.2	14238	4698068	4.68
57*	5p15.1-p13.3	17509888	32677299	15.17				
58	5p15.33-p13.3	25328	34119847	34.11	Xp22.31	6447911-	8135053	
59	5p15.33-p13.3	25328	30480030	30.48				
60	NA	NA	NA	NA				
61	5p15.33-p15.1	25328	15652433	15.65	11p15.5-p15.4	75328-	10525251	10.5
62	NA	NA	NA	NA				
63	5p15.33-p15.13.3	25328	30445734	30.44				
64	5p15.33-p15.32	25328	4610206	4.61	2p15.3-p25.2	14238	4698068	4.68
65	5p15.33-p15.32	25328	4610206	4.61	2p15.3-p25.2	14238	4698068	4.68
66*	5p142-p13.2	23383424-	36609355	13.22				
67	5p15.33-p15.31	25328	7125022	7.12				
68**	5p15.33-p13.3	25328	29292854	29.29				
69	5p15.33-p15.32	25328	5014883	5.01				
70*	5p15.2-p13.3	9860050	33760050	23.09				

**Means, interstitial cases. **Means, mosaics.*

(mainly duplications). Interestingly, the presence of additional rearrangements may exert significant differences on prenatal and postnatal growth delay findings, cardiac anomalies, and *speech abilities* in the expressive language (**Table 6**, $P \leq 0.05$, at CI 95%). Remarkably, other findings became significant at CI of 90%, such as cleft lip/palate, renal anomalies, autistic spectrum disorders (ASD), or breathing difficulties (**Table 6**).

Ward cluster analysis allowed us to compare the frequencies of these variables in both subpopulations. We denote that better figures (low percentages) were more represented in the simple 5p deletion group, but with motor items, slightly better than in the group with additional rearrangements (**Table 5** and **Supplementary Data**). Although the simple deletion group had a higher size of 5p deletions on average (see **Supplementary Table 5**), no statistical significant differences could be observed between the two different subpopulations (*Student t-test*, see **Figure 4**).

We also performed an association analysis among categoric variables in both subpopulations by Kendall's tau_b analysis (expanded analysis for the whole cohort and subgroups is presented in **Supplementary Table 3**). Interestingly, some of the observed correlations in the simple 5p deletion group disappeared in the group with additional rearrangements (**Figure 5**). A more specific example for three of these categoric variables is presented in **Supplementary Figure 3**.

Genotype–Phenotype Correlations

We made Ward's hierarchical cluster analysis using the item "size of deletion" as unique variable, in order to verify how individuals (initially, from the whole cohort) group according to their deletion size. At the end, individuals were grouped in four clusters (the number was established by BIC and AIC algorithms), as follows: 4.97 ± 1.83 Mb, 14.64 ± 2.31 Mb, 24.01 ± 1.38 Mb, and 29.95 ± 2.93 Mb (**Figures 6A,B**). ANOVA analysis discarded a significant correlation between the size of the deletion and the functional item, GFAP, or any of its intermediates ($P = 0.07$ at CI 95%, data not shown). However, ANOVA analysis for continuous variables or by chi-square test for categoric variables shows the existence of significant differences between clusters in a few variables, mostly related to perinatal parameters, some dysmorphic features, behavior, and cognitive features (**Figure 6C**). Further, Bonferroni's and *z*-tests for previous significant variables revealed that cluster 3

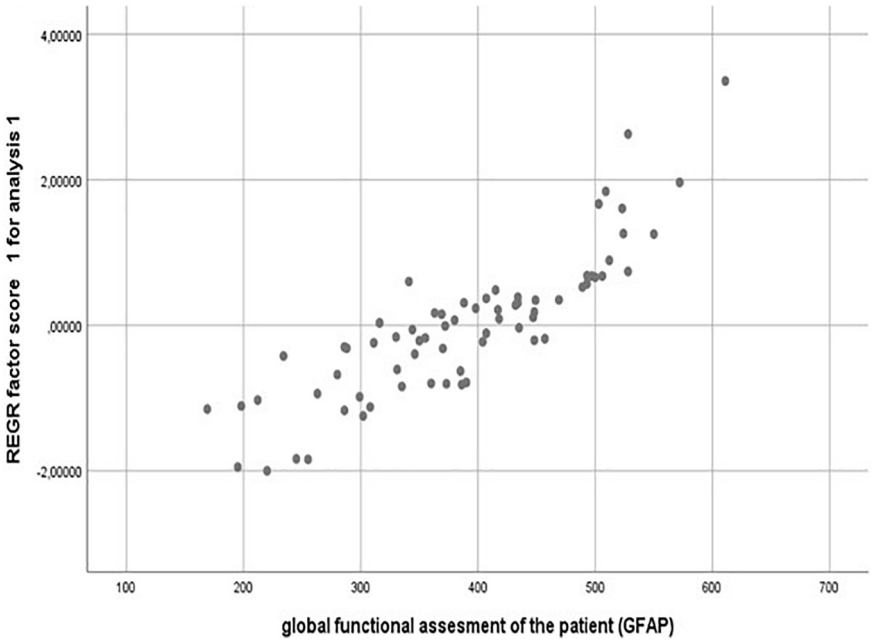


FIGURE 3 | Validation of a GFAP construct by PCA (principal component analysis) statistical approach. Pearson correlation value = 0.846; *P* = 0.001.

TABLE 5 | Mean (±SD) and median of GFAP (Global Functional Assessment of the Patient) and its intermediates (items "i" to "v") from the whole cohort and subpopulations of 5p- individuals.

	Whole cohort		Single 5p deletions		5p deletions plus additional rearrangements	
"Functional" variable	**Median**	**Mean ± *SD***	**Median**	**Mean ± *SD***	**Median**	**Mean ± *SD***
GFAP	387.60	388.40 ± 100.00	361.50	362.89 ± 98.50	404.00	398.70 ± 92.00
(i) Developmental delay items corrected by age	251.00	244.60 ± 66.40	228.50	233.80 ± 59.50	249.00	229.10 ± 73.00
(ii) Behavioral alteration items	7.50	13.40 ± 15.40	7.00	10.80 ± 12.10	7.00	15.00 ± 19.60
(iii) Dysmorphic items	24.00	21.10 ± 11.60	24.00	20.30 ± 11.80	24.00	19.00 ± 0.60
(iv) Communication skills	48.00	54.10 ± 25.80	45.00	50.50 ± 25.80	60.00	58.20 ± 23.80
(v) Comorbidity items	47.00	55.20 ± 47.00	47.00	48.80 ± 38.80	67.00	61.50 ± 37.50

Values are expressed in arbitrary units. Higher values are associated with a worse functional prognosis.

TABLE 6 | Comparison between subpopulations in 5p- individuals, regarding categorical variables taken.

Chi-square test

	Simple 5p dels *N* = 47	5p dels + addt rearrangement *N* = 23	Value	df	Sig. asymptotic (bilateral)
IUGR	14	13	1.701	1	0.019*
Postnatal growth failure	16	17	3.716	1	0.036*
Cardiac anomalies	10	14	6.020	1	0.014*
Short understandable sentences	14	2	5.299	1	0.021*
Cleft lip-palate anomalies	2	5	3.543	1	0.060$
Renal anomalies	3	6	3.441	1	0.064$
ASD	3	6	3.441	1	0.064$
Prominent superciliary arches	5	0	3.381	1	0.066$
Breath problems	13	14	3.272	1	0.070$
No words at all	10	11	3.187	1	0.074$
Interact with the environment	34	16	3.189	1	0.074$

** means significant p-value ≤ 0.05; $ means possible tendency, significant at CI 90% (data not shown). N is the number of patients evaluated.*

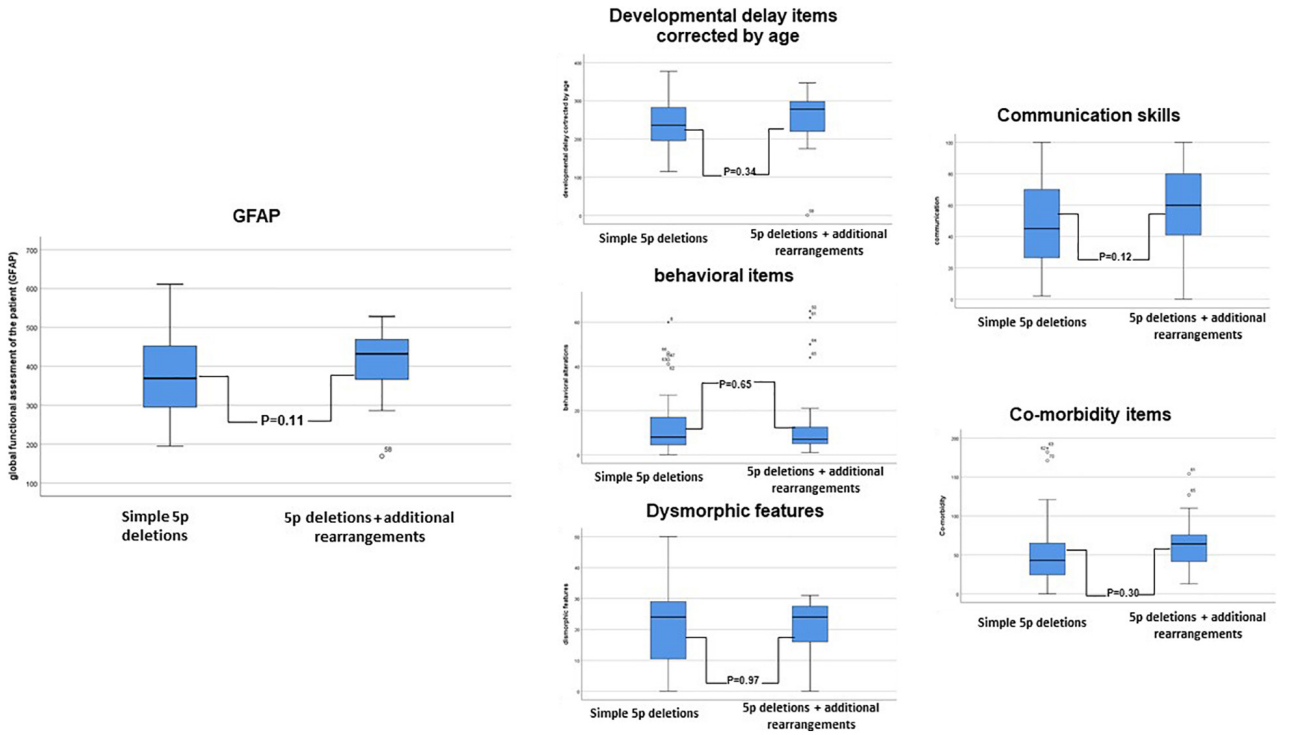


FIGURE 4 | Schematic representation of the comparison between subpopulations in 5p- individuals (5p deletions vs. 5p deletion plus additional rearrangements) in the "functional" constructed GFAP (Global functional Assessment of the Patient) and its intermediates (developmental delay items corrected by age, behavioral items, dysmorphic items, communication skills, and comorbidity items). A Student *t*-test was performed.

(size 24.01 ± 1.38 Mb; 5p15.1–p14.1) is the most represented among the cluster pairs with significant differences among them (**Figure 6C**).

When Ward's clusters were dissected by item frequencies (in percentages), higher percentages (normally associated with a worse prognostic) seemed to be mapped, in cluster 3 too (**Table 6** and **Supplementary Data**). However, expressive language (followed by item, *the ability to make short sentences*) or the ability to write or read was associated preferentially with clusters 1 and 2 (11/16 individuals, 69%). **Figure 7** shows how the four clusters integrate into suggested functional areas of chromosome-5p of several previously published data. We observed some relevant mapping findings such as the item speech delay, which was mapped at the beginning of the telomere.

We further analyzed these possible differences among clusters (by size of deletion) in the two subpopulations of 5p- Sd individuals (simple, isolated 5p deletions vs. 5p deletions plus additional rearrangements), using the same statistical approach presented above. **Supplementary Figure 4A** showed a similar result for simple deletions. We also found intra-cluster significant differences for some variables, with cluster 3 again as the most representative cluster for significant differences in the pair of cluster comparisons (**Figure 4A** and **Supplementary Data**). Remarkably, one of these variables that showed differences among clusters was GFAP. For analysis of the group with additional rearrangements, we generated only two clusters for comparison (due to the number of individuals) but also denoted significant differences between clusters "A" and "B" (now, cluster "A" aggregates clusters 1 and 2 and "B" clusters 3 and 4, **Supplementary Figure 4B**).

Finally, Pearson correlation analysis established that the size of the deletion inversely correlated with some neonatal parameters, such as weight or OFC ($P \leq 0.001$), and almost with birth length ($P = 0.061$). However, the most significant genotype/phenotype correlation was observed between size of the deletions and gender (males, 15.79 ± 8.79 vs. females, 22.38 ± 8.84. Student *t*-test, $P = 0.004$).

Male vs. Female Comparative Analysis

A chi-square test was performed for the whole cohort and two of the subpopulations. **Table 7i** shows the statistic significant differences between males and females in the whole cohort. These differences were mostly related to growth delay (prenatal and postnatal), dysmorphic features, some spinal comorbidities, and behavioral and cognitive aspects. In addition, Ward cluster analysis between males and females showed the worst frequencies (in percentages) in females (**Table 7** and **Supplementary Data**). As we expected, neonatal data at birth showed also significant differences among gender and weight and OFC ($P \leq 0.01$ and $P \leq 0.05$, respectively, Student *t*-test) or with length at birth ($P = 0.074$, Student *t*-test). Most remarkably, there were also significant differences at the functional GFAP ($P = 0.05$, Student *t*-test). These differences showed higher values of frequencies (mainly, a worse prognosis) in females. Similarly, we

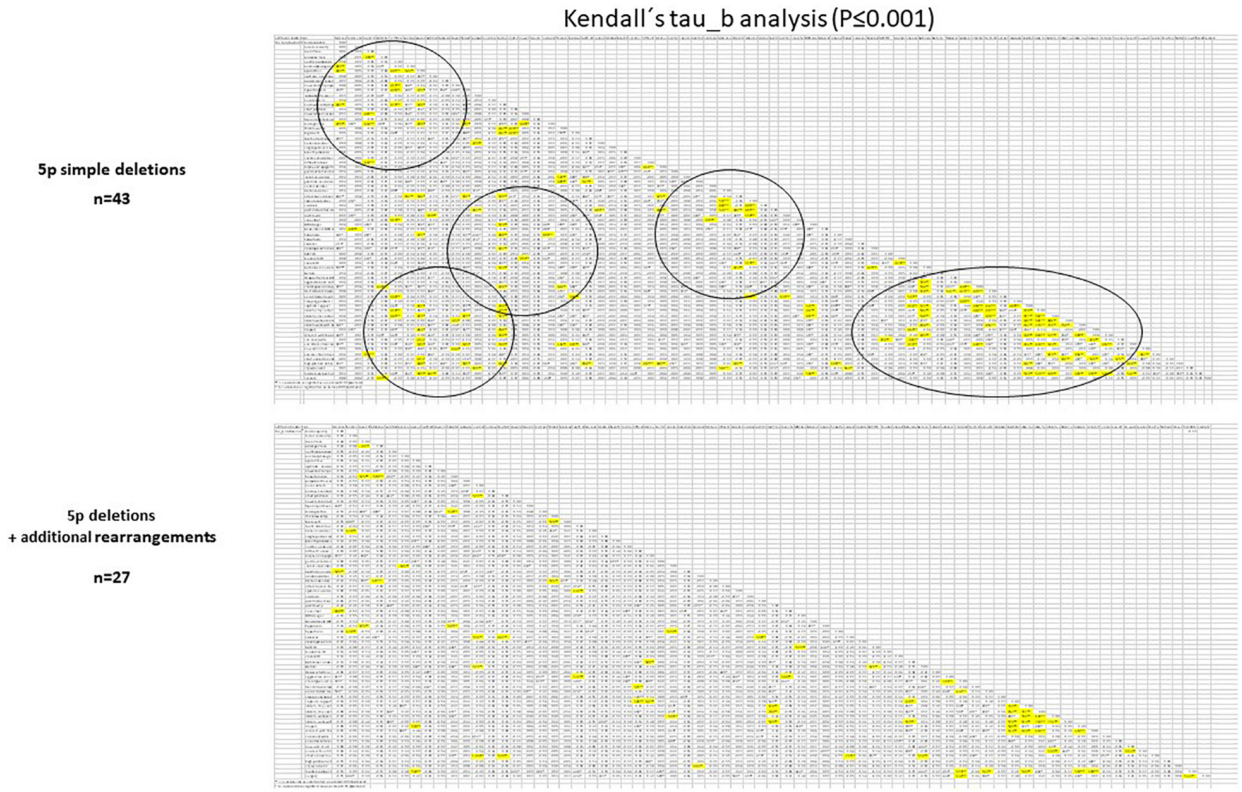


FIGURE 5 | Schematic representation of the comparison between subpopulations in 5p- individuals (5p deletions vs. 5p deletion plus additional rearrangements) in categoric variables using Kendall's tau_b statistical analysis. Very significant differences ($P \leq 0.01$) were denoted in bold. Circles denote significant differences among variables observed in the 5p deletion group and absent in the 5p deletion + additional rearrangement group.

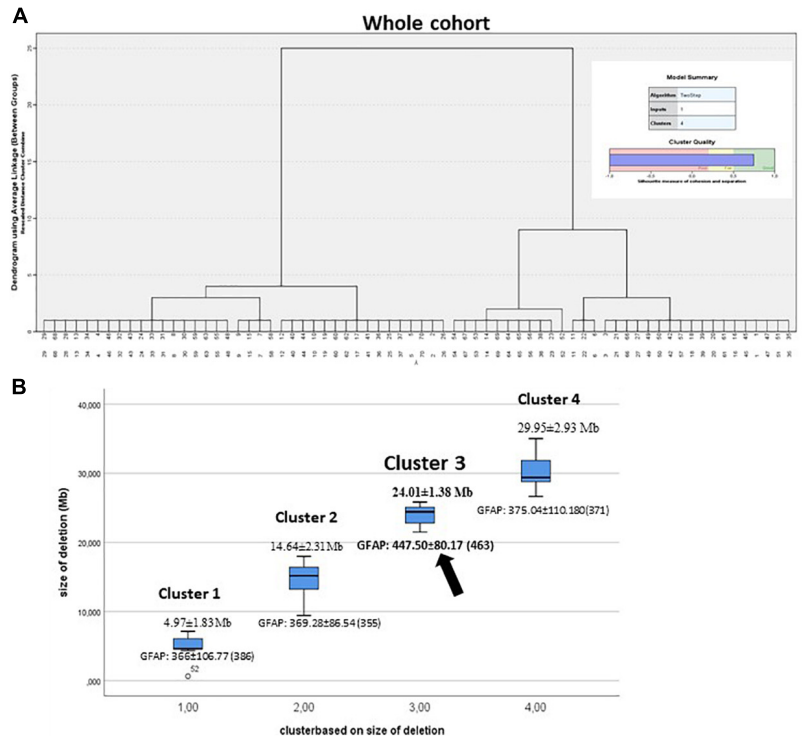
compared male vs. female significant differences for all categoric variables (in both isolated deletions and deletions + additional rearrangements) (**Tables 7ii,iii**). Significant correlations were found among gender, independently of the group. The only significant difference in common was intrauterine growth retardation (IUGR). Again, the most remarkable finding with significant differences in the simple 5p deletion group was GFAP (**Table 7ii**), but not in the group with additional rearrangements (**Table 7iii**). On the other hand, patients with additional rearrangements also showed significant differences in neonatal data, such as weight or OFC at birth, again as the whole cohort, showing better numbers in males than in females. Expanded Student *t*-test analysis is shown in **Supplementary Table 8**.

DISCUSSION

In this work, we describe the largest cohort of Spanish patients with 5p- Sd and one of the largest series of these patients so far, characterized by means of CMA and other genetic approaches, such as cytogenetics, MLPA, and FISH. Although its prevalence is still unknown, it was estimated around 1:15,000–50,000 (Niebuhr, 1978; Higurashi et al., 1990; Cerruti Mainardi et al., 2006). Our data showed that 5p- individuals may have a high clinical variability that is accompanied also by a high genetic heterogeneity. In fact, individuals with 5p- syndrome do not always carry a single rearrangement. In our cohort, around 39% of the individuals presented an additional clinically significant genomic rearrangement, mainly a duplication in other chromosomes. In other cases, additional deletions and duplications can be observed nearby the main 5p deletion (seven cases), probably as a result of a complex rearrangement, as it has been previously suggested (Gu et al., 2013). These additional rearrangements raised the question of whether additional genomic rearrangements may have a role in the syndrome, and thus, it may explain part of its variability, or if individuals with additional rearrangements should be considered as having 5p minus syndrome.

We described and compared our cohort with other previously reported series in terms of clinical features. Some limitations of this study come from information taken from the questionnaires filled up by parents or caregivers. This could explain part of these differences among subjects. We strongly recommend systematic codification of clinical features using the HPO system.

We think that frequency-weighed HPO terms grouped in five main nuclear features of the syndrome will help clinicians to describe 5p- Sd patients (**Table 3**). We built a quotation scale called GFAP (see sections "Materials and Methods" and "Results"). We compared this "functional" GFAP and its intermediate components in order to establish



C

Variable	p-value	Statistical test	Pairs with significant differences
gestational age at birth	0.029	ANOVA	(1.3)
Weight at birth	0.017	ANOVA	(1.3)
epicanthus	0.042	Chi square	(1.4)
hypertelorism	0.002	Chi square	(1.3) (1.4)
spinal anomalies	0.004	Chi square	(1.3) (2.3)
hypotonia	0.009	Chi square	(1.3)
severe ID	0.0001	Chi square	(1.3) (1.4) (2.3)
sterotypes and repetitive movements	0.018	Chi square	(1.3)
can read/write	0.001	Chi square	(1.3) (1.4)
use alternative communicative tools	0.031	Chi square	(1.3)
short understandable sentences	0.021	Chi square	(1.4)

FIGURE 6 | (A) Ward's hierarchical cluster of the whole cohort by size of the deletions. BIC and AIC determined to be grouped by four clusters. **(B)** Plot segregation ordered by deletion size in Mb. Every cluster showed the GFAP value. **(C)** ANOVA analysis for continuous variables or by chi-square test for categoric variables was performed to establish the existence of significant differences between clusters. Further, Bonferroni's test and *z*-test for previous significant variables revealed cluster pairs with significant differences among them.

putative significant differences between both subpopulations: simple, isolated 5p deletions and 5p deletions with additional rearrangements. However, no statistical significant differences (*Student t-test*) could be observed between the two different subpopulations for the GFAP variable, although several significant differences could be denoted among other clinical features. The most relevant were cardiac anomalies and speech delay and the presence of additional rearrangements. Regarding behavioral aspects, there were significant differences among subpopulations in sleeping problems, stereotypical or aggressive behavior, and number of behavioral problems, being more common in the group with an additional genomic rearrangement. Thus, the latter showed better numbers in some cognitive items than simple 5p deletions. Altogether and based on statistic analysis, the presence of additional duplications did not have any significant representation over the whole phenotype of the 5p- patient, but it might have specific contributions for some clinical findings such as growth delay (either prenatal or postnatal) as well as cardiac anomalies.

Genotype–Phenotype Correlations

Some authors have previously stated that the severity of the phenotype and the cognitive delay of 5p- Sd were associated with an increased size of the deletion at chromosome 5p (Wilkins et al., 1983; Cornish et al., 1999; Cerruti Mainardi et al., 2001). However, this fact was not confirmed by others (Marinescu et al., 1999; Espirito Santo et al., 2016). Thus, this aspect is still controversial. We used our "functional" construct GFAP to validate this hypothesis. Our data supported these genotype–phenotype correlations only in simple deletions. Since there is a scant number of publications in this syndrome incorporating microarray data, it cannot be discussed whether this fact has occurred in other cohorts. For instance, Cerruti Mainardi et al. (2006) analyzed genotype–phenotype correlations but only in patients with isolated 5p terminal deletions (151/185 cases).

Furthermore, if a part of the huge phenotypic variability observed among 5p- individuals was not related to the size of the deletion, the other possibility may be established by the location of the deletion, since it is a chromosomal region with an important gene content. Our data supported that specific regions at chromosome 5p may have more significant roles in the syndrome than others. Our analysis of clusters (by size of the deletion) showed that cluster 3 was the most relevant among the cluster pairs with statistically significant differences, both in the whole cohort and subpopulation groups. In fact, the worst frequencies of most categorical items, as well as GFAP and its intermediates in cluster 3, seem to support this observation. Cluster 3 mapped at 18–25 Mb from the telomere (chromosomal bands 5p15.1–5p14.1). Among the genes mapping in this area were the cadherin (CDH) cluster, including *CDH10*, *CDH9*, *CDH12*, *CDH18*, and *CDH6*, strongly associated with this syndrome. This CDH cluster has been described to be

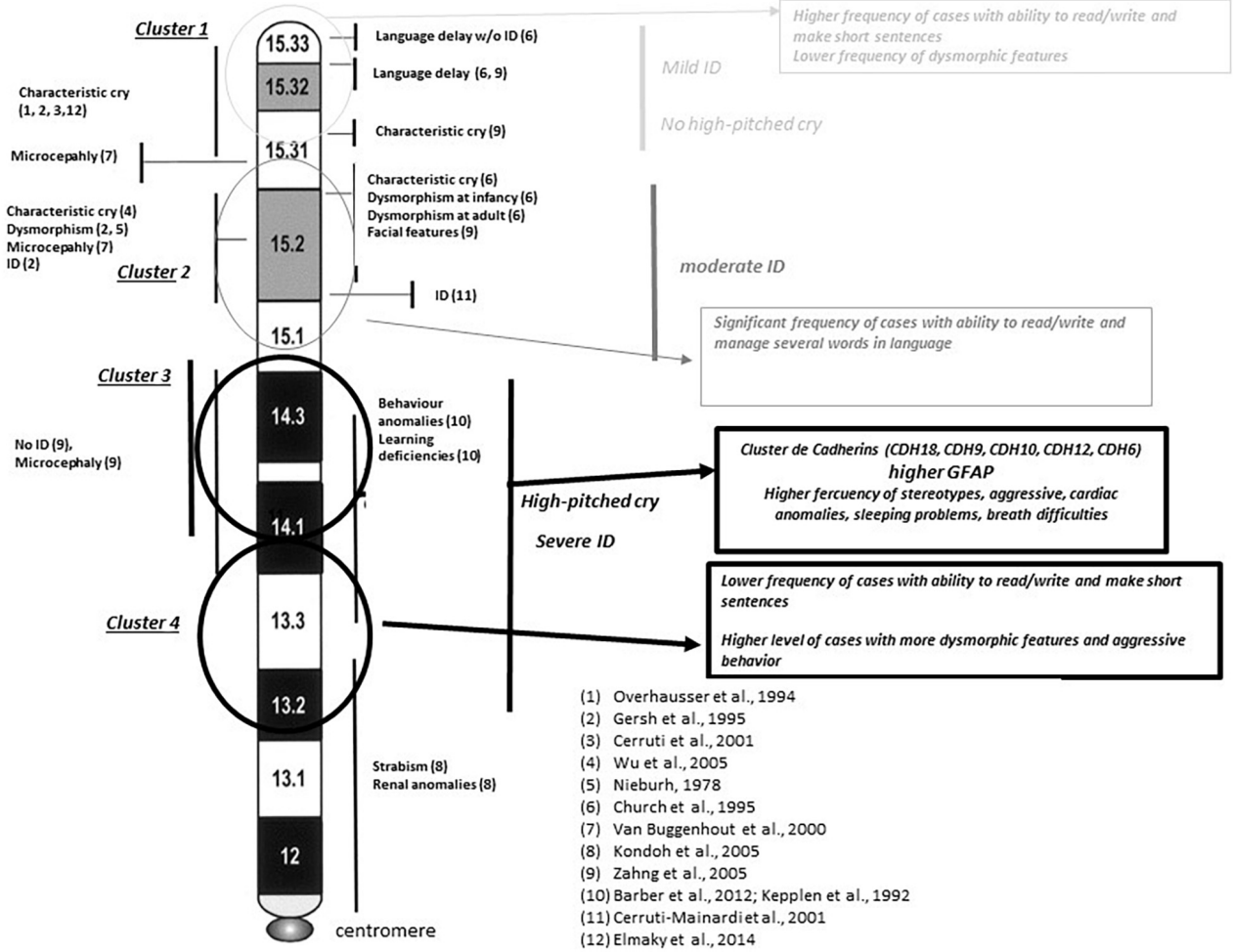


FIGURE 7 | Integrative map for clusters shown in **Figure 6**. Comparisons with previously reported critical regions for phenotype sings at 5p minus syndrome (references refer to clinical symptoms reported in individual families with interstitial deletions in different reported works). Circles represent the areas mapped for the different clusters obtained in Ward's analysis. Major findings observed in our study are in italic text. Chromosome bands are reported according to ISCN. ID, intellectual disability.

conditionally haploinsufficient and depend on other genetic or environmental factors leading to an abnormal phenotype. This is an interesting fact and could also explain part of the variability observed in 5p-Sd. Other genes in this region are *FBXL7*, *MARCH11*, *FAM134B*, *MYO10*, *DROSHA*, *PDZD2*, *GOLPH3*, *MTMR12*, *ZFR*, *SUB1*, *NPR3*, and *TARS*. All genes have a significant level of haploinsufficiency (see **Supplementary Table 9**). However, we cannot rule out a role for other genes such as *CTNND2*, *TERT*, and *MED10*, commonly deleted in 5p- Sd and associated with neuronal development/function and cellular death. The smallest region of overlap patients with interstitial deletions pointed out to two potential regions, one mapping at these genes and the other in the cadherin cluster. Additional interstitial cases and functional assays are needed to unreveal the role of all these genes.

Speech skills (evaluated only in patients aged >3 years) yielded that the potentially affected region is near to the telomere (5p15.33–5p15.31) supporting previous findings (Church et al., 1995; Zhang et al., 2005). High frequencies for most of the behavioral findings seemed to be associated with clusters 3 and 4 in our Ward's cluster analysis and supported previous studies for its hypothetical mapping (Barber et al., 2011). On the other hand, our data also showed some discrepancies with previous studies. In our study, high-pitched cry seemed to map at p14.3–p13.2 bands versus bands p15.33–p15.31 (Overhauser et al., 1994; Gersh et al., 1995; Cerruti Mainardi et al., 2001), p15.31 (Zhang et al., 2005), p15.31–p15.2 (Church et al., 1995), or p15.2 (Wu et al., 2005) in other previous reports.

Gender as a Differentiating Factor: Correlations Depending on Gender

A suspicion of putative cognitive and "functional" differences between males and females patients has been constantly suggested to us by parents, caregivers, and several clinical specialists. This is the first report showing "functional"

TABLE 7 | Comparison between male and female 5p- individuals, regarding categorical and continuous variables taken by means of the chi-square and Student *t*-test, respectively, in **(i)** the whole cohort, **(ii)** simple 5p deletions, and **(iii)** 5p deletions plus additional rearrangements.

(i) The whole cohort

Chi-square

	Male	Female	Value	df	Sig. asymptotic (bilateral)
IUGR	3	27	9.42	1	0.002**
Postnatal growth failure	7	26	3.838	1	0.050*
Round face	6	26	5.318	1	0.021*
Enlarged face	12	11	5.794	1	0.016*
Neck anomalies	1	12	4.583	1	0.032*
Alterations of the fingers or toes	6	25	4.598	1	0.032*
Spinal anomalies	2	16	5.194	1	0.023*
Scoliosis	4	21	5.009	1	0.035*
Severe ID	5	26	7.058	1	0.008**
Aggressive and self-mutilation	4	23	6.486	1	0.011*
Sleeping problems	8	30	5.255	1	0.022*
Can read/write	7	5	3.911	1	0.048*
Short understandable sentences	9	7	4.701	1	0.030*
Mild ID	5	3	3.598	1	0.058$

Student *t*

	Male		Female		Sig.
Weight at birth (g)	2925.91	689.24	2447.28	621.33	0.006**
OFC at birth (cm)	33.17	2.27	31.74	2.37	0.021*
GFAP	358.87	73.39	402.94	108.60	0.050*
Height at birth (cm)	47.11	3.79	45.31	3.86	0.074$

(ii) Simple 5p deletions

Chi-square

	Male	Female	Value	df	Sig. asymptotic (bilateral)
IUGR	2	12	4.669	1	0.0031**
Faillure to thrive	3	13	3.716	1	0.054
Larynx and epiglottis alterations	1	10	5.002	1	0.025*
Severe ID	3	14	4.605	1	0.032*
Aggressive and self-mutilation	3	14	4.605	1	0.032*
Sleeping problems	4	16	4.740	1	0.029*
Spinal anomalies	1	8	3.318	1	0.069$

Student *t*

	Male		Female		Sig.
GFAP	336.31	61.36	396.37	123.10	0.040*
Behavioral item, as component of GFAP	7.73	11.72	22.63	13.77	0.063$

(iii) 5p deletions + additional rearrangements

Chi-square

	Male	Female	Value	df	Sig. asymptotic (bilateral)
IUGR	1	12	4.34	1	0.037*
Round face	1	12	1.73	1	0.037*
Enlarged face	6	4	9.60	1	0.002**

(Continued)

TABLE 7 | Continued

(iii) 5p deletions + additional rearrangements					
Chi-square					
	Male	**Female**	**Value**	**df**	**Sig. asymptotic (bilateral)**
Ophthalmological anomalies	1	12	4.34	1	0.037*
Short understandable sentences	2	0	5.59	1	0.018*
Auditive problems	5	6	3.68	1	0.055$
Student *t*					
	Male		**Female**		**Sig.**
Weight at birth (g)	3314.29	686.24	2469.60	586.85	0.004**
OFC at birth (cm)	34.79	1.55	31.44	2.11	0.001**
Size of the deletions	10.25	6.81	21.60	8.69	0.004**

** means significant p-value ≤ 0.05. ** means significant p-value ≤ 0.01. $ means possible tendency, significant at CI 90% (data not shown).*

differences between males and females in 5p- Sd individuals. We found that some of the clinical features analyzed showed statistically significant differences among males and females, for instance in the GFAP variable. Thus, we denoted worse functional scores and higher deletion sizes in females than in males using Ward's cluster analysis. Additional efforts with systematic cognitive–behavioral evaluations of the patients must be performed in order to assign more precise differences.

The reason why the ratio female–male is 2:1 is still unknown. One of the most relevant differences between genders is the mean value for size of the deletions. Interestingly, Ward's cluster analysis allowed us to observe how the female/male ratio was modulated by the different sizes of the deletions in the clusters (**Tables 5**, **7** and **Supplementary Data**). The number of males in these clusters decreased drastically when the size of the deletion increased over 15 Mb. This fact may suggest a different, possibly lethal, effect of deletions over 15 Mb in males and might explain the differences among the female/male ratio in this syndrome. In fact, miscarriages are frequent in this syndrome. This is not an unusual effect because other genes at 5p13.1, such as *RICTOR* and *DAB2*, have been suggested to be haplolethals (Peng et al., 2020) and may explain how deletions do not expand in size, more than 39 Mb from the telomere. However, we cannot rule out any other additional genetic or epigenetic effect in males, affecting chromosomal bands 5p15.1–p13.2. In fact, an aberrant DNA methylation in Cri du chat syndrome related to development conditions has been already suggested (Naumonova et al., 2018).

CONCLUSION

Summing up, we here report a large series of patients with 5p minus syndrome emphasizing some phenotype–genotype correlations. Remarkably, we found statistically significant "functional" differences among males and females. We also dissected subpopulations in 5p- Sd based on the presence/absence of clinical significant additional rearrangements, besides losses at the 5p arm. The presence of these additional rearrangements may have a role modulating part of the phenotype in the syndrome.

Finally, we recommend combining typical karyotyping with CMA as the definitive method for a precise diagnosis of 5p- Sd, in order to provide a more accurate genetic counseling for these families.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found in the DECIPHER database (https://www.deciphergenomics.org/) with the following accession numbers: 436269 to 436336 corresponding to cases 5pIMG01-5pIMG70.

ETHICS STATEMENT

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

AUTHOR CONTRIBUTIONS

JN conceived the presented idea, completed the data analysis, and wrote the manuscript. JN and PL designed the study. JN coordinated the data acquisition and collected the data from the second questionnaires. AS-T created and managed the Final 5p- database. PL, AH, CB-L, and BS-R assisted with the data management and statistics. CB-L analyzed the conductual and cognitive profiles of the patients. JN, PB, and MM-Á managed the SNP and CGH microarrays at the INGEMM. EM, IV, and FG-S provided the FISH and karyotyping studies at INGEMM. FS-S and PL provided and explored several patients at INGEMM. MM-Fr and MM-Fe created the first questionnaire and managed

some of the patients' cytogenetic analysis. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We would like to thank all the individuals and families for their participation in this study, especially Viviana Alonso, Felix Casado, Pilar Castaño, Carolina Nicolás, Esther Pérez, María Joséfa Porras, Margarita Ruether, and Sonia Saíz, for helping us develop this project. In addition, we also thank local 5p- patients' associations: Adais 5p, Asimaga, Fundación Síndrome5p-, and Planet 5p.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.645595/full#supplementary-material

REFERENCES

Barber, J. C., Huang, S., Bateman, M. S. Y., and Collins, A. L. (2011). Transmitted deletions of medial 5p and learning difficulties; does the cadherin cluster only become penetrant when flanking genes are deleted? *Am. J. Med. Genet. Part A* 155A, 2807–2815. doi: 10.1002/ajmg.a.34241

Cerruti Mainardi, P. (2006). Cri du chat syndrome. *Orphanet J. Rare Dis.* 1:33.

Cerruti Mainardi, P., Pastore, G., Castronovo, C., Godi, M., Guala, A., Tamiazzo, S., et al. (2006). The natural history of Cri du Chat Syndrome. a report from the Italian Register. *Eur. J. Med. Genet.* 49, 363–383. doi: 10.1016/j.ejmg.2005.12.004

Cerruti Mainardi, P., Perfumo, C., Cali, A., Coucourde, G., Pastore, G., Cavani, S., et al. (2001). Clinical and molecular characterisation of 80 patients with 5p deletion: genotipe-phenotipe correlation. *J. Med. Genet.* 38, 151–158. doi: 10.1136/jmg.38.3.151

Chehimi, S. N., Zanardo, V. A., Ceroni, J. R. M., Nascimento, A. M., Madia, F. A. R., Dias, A. T., et al. (2020). Breakpoint delineation in 5p- patients leads to new insights about microcephaly and the typical high-pitched cry. *Mol. Genet. Genomic Med.* 8:e957. doi: 10.1002/mgg3.957

Church, D. M., Bengtsson, U., Nielsen, K. V., Wasmuth, J. J., and Niebuhr, E. (1995). Molecular definition of deletions of different segments of distal 5p that result in distinct phenotypic features. *Am. J. Hum. Genet.* 56, 1162–1172.

Cornish, K., Bramble, K., Munir, F. Y., and Pigram, J. (1999). Cognitive functioning in children with typical cri du chat (5p-) syndrome. *Dev. Med. Child Neurol.* 41, 263–266. doi: 10.1017/s0012162299000559

Correa, T., Feltes, B. C., and Riegel, M. (2019). Integrated analysis of the critical region 5p15.3-p15.2 associated with cri-du-chat syndrome. *Genet. Mol. Biol.* 42, 186–196. doi: 10.1590/1678-4685-GMB-2018-0173

Elmakky, A., Carli, D., Lugli, L., Torelli, P., Guidi, B., Falcinelli, C., et al. (2014). A three-generation family with terminal microdeletion involving 5p15.33-32 due to a whole-arm 5:15 chromosomal translocation with a steady phenotype of atypical cri du chat syndrome. *Eur. J. Med. Genet.* 57, 145–150. doi: 10.1016/j.ejmg.2014.02.005

Espirito Santo, L. D., Moreira, L. M., and Riegel, M. (2016). Cri-Du-Chat syndrome: clinical profile and chromosomal microarray analysis in six patients. *BioMed Res. Int.* 5467083.

Gersh, M., Goodart, S. A., Pasztor, L. M., Harris, D. J., Weiss, L., and Overhauser, J. (1995). Evidence for a distinct region causing a cat-like cry in patients with 5p deletions. *Am. J. Hum. Genet.* 56, 1404–1410.

Gu, H., Jiang, J. H., Li, J. Y., Zhang, Y. N., Dong, X. S., Huang, Y. Y., et al. (2013). A familial cri-du-chat/5p deletion syndrome resulted from rare maternal complex chromosomal rearrangements (ccrs) and/or possible chromosome 5p chromothripsis. *PLoS One* 8:e76985. doi: 10.1371/journal.pone.0076985

Higurashi, M., Oda, M., Iijima, K., Iijima, S., Takeshita, T., Watanabe, N., et al. (1990). Livebirths prevalence and follow-up of malformation síndromes. *Brain Dev.* 12, 770–773. doi: 10.1016/s0387-7604(12)80004-0

Honjo, R. S., Mello, C. B., Pimenta, L. S. E., Nuñes-Vaca, E. C., Benedetto, L. M., Khoury, R., et al. (2018). Cri du Chat syndrome: characteristics of 73 Brazilian patients. *J. Intellect. Disabil. Res.* 62, 467–473. doi: 10.1111/jir.12476

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., et al. (2002). The human genome browser at UCSC. *Genome Res.* 12, 996–1006. doi: 10.1101/gr.229102.

Kondoh, T., Shimokawa, O., Harada, N., Doi, T., Yun, C., Gohda, Y., et al. (2005). Genotype-phenotype correlation of 5p-syndrome: pitfall of diagnosis. *J. Hum. Genet.* 50, 26–29. doi: 10.1007/s10038-004-0213-9

Mak, A., Ma, T., Chan, K., Kan, A., Tang, M., and Leung, K. Y. (2019). Prenatal diagnosis of 5p deletion syndrome: report of five cases. *J. Obstet. Gynaecol. Res.* 45, 923–926. doi: 10.1111/jog.13911

Marinescu, R. C., Cerruti, P., Collins, M. R., Kouahou, M., Coucourde, G., Pastore, G., et al. (2000). Growth charts for cri-du-chat syndrome: an international collaborative study. *Am. J. Med. Genet.* 94, 153–162. doi: 10.1002/1096-8628(20000911)94:2<153::aid-ajmg8>3.0.co;2-#

Marinescu, R. C., Johnson, E. I., Dykens, E. M., Hodapp, R. M. Y., and Overhauser, J. (1999). No relationship between the size of the deletion and the level of developmental delay in cri-du-chat syndrome. *Am. J. Med. Genet.* 86, 66–70. doi: 10.1002/(sici)1096-8628(19990903)86:1<66::aid-ajmg13>3.0.co;2-n

Naumonova, O. Y., Rychkov, S. Y., Kuznetzova, T. V., Odintsova, V. V., Kornilov, S. A., and Grigorenko, E. L. (2018). DNA methylation alterations in the genome of a toddler with cri-du-chat syndrome. *Clin. Case Rep.* 6, 14–17. doi: 10.1002/ccr3.1274

Nguyen, J., Qualmann, K., Okashah, R., Reilly, A., Alexeyev, M., and Campbell, D. (2015). 5p deletions: current knowledge and future directions. *Am. J. Med. Genet. Part C, Sem. Med. Genet.* 169, 224–238. doi: 10.1002/ajmg.c.31444

Niebuhr, E. (1978). The cri du chat syndrome. epidemiology, cytogenetics and clinical features. *Hum. Genet.* 44, 227–275. doi: 10.1007/bf00394291

Overhauser, J., Huang, X., Gersh, M., Wilson, W., McMahon, J., Bengtsson, U., et al. (1994). Molecular and phenotypic mapping of the short arm of chromosome 5: sublocalization of the critical region for the cri-du-chat syndrome. *Hum. Mol. Genet.* 3, 247–252. doi: 10.1093/hmg/3.2.247

Peng, Y., Pang, J., Hu, J., Jia, Z., Xi, H., Ma, N., et al. (2020). Clinical and molecular characterization of 12 prenatal cases of Cri-du-chat syndrome. *Mol. Genet. Genomic Med.* 8:e1312.

Rodrigues de Medeiros, J. B. (2017). *Cases Report: Cri-du-chat Syndrome*. Master Degree thesis, Uberlândia: Universidade Federal de Uberlândia. Faculdade de Odontologia.

Rodríguez-Caballero, A., Torres-Lagares, D., Yáñez-Vico, R. M., Gutiérrez-Pérez, J. L. G., and Machuca-Portillo, G. (2012). Assessment of orofacial characteristics and oral pathology associated with cri-du-chat syndrome. *Oral Dis.* 18, 191–197. doi: 10.1111/j.1601-0825.2011.01864.x

Simmons, A. D., Goodart, S. A., Gallardo, T. D., Overhauser, J., and Lovett, M. (1995). Five novel genes from the cri-du-chat critical region isolated by direct selection. *Hum. Mol. Genet.* 4, 295–302. doi: 10.1093/hmg/4.2.295

Su, J., Fu, H., Xie, B., Lu, W., Li, W., Wei, Y., et al. (2019). Prenatal diagnosis of cri-du-chat syndrome by SNP array: report of twelve cases and review of the literature. *Mol. Cytogenet.* 12:49.

Van Buggenhout, G. J., Pijkels, E., Holvoet, M., Schaap, C., Hamel, B. C., and Fryns, J. P. (2000). Cri du chat syndrome: changing phenotype in older patients. *Am. J. Med. Genet.* 90, 203–215. doi: 10.1002/(sici)1096-8628(20000131)90:3<203::aid-ajmg5>3.0.co;2-a

Wilkins, L. E., Brown, J. A., Nance, W. E., and Wolf, B. (1983). Clinical heterogeneity in 80 home-reared children with cri-du-chat syndrome. *J. Pediatr.* 102, 528–533. doi: 10.1016/s0022-3476(83)80179-6

Wu, Q., Niebuhr, E., Yang, H., and Hansen, L. (2005). Determination of the 'critical region' for cat-like cry of Cri-du-chat syndrome and analysis of candidate genes by quantitative PCR. *Eur. J. Hum. Genet. EJHG* 13, 475–485. doi: 10.1038/sj.ejhg.5201345

Zhang, A., Snijders, A., Segraves, R., Zhang, X., Niebuhr, A., Albertson, D., et al. (2005). High resolution mapping of genotype-phenotype relationships in cri du chat syndrome using array comparative genomic hybridization. *Am. J. Hum. Genet.* 76, 312–326. doi: 10.1086/427762

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 Pediatrics

CASE REPORT
published: 20 August 2021
doi: 10.3389/fped.2021.664548

Case Report: Expressive Speech Disorder in a Family as a Hallmark of 7q31 Deletion Involving the *FOXP2* Gene

Orsolya Nagy[1], *Judit Kárteszi*[2], *Beatrix Elmont*[3] *and Anikó Ujfalusi*[1]*

[1] Division of Clinical Genetics, Department of Laboratory Medicine, Faculty of Medicine, University of Debrecen, Debrecen, Hungary, [2] Hospital of Zala County, Zalaegerszeg, Hungary, [3] Department of Pediatrics, Hospital of Zala County, Zalaegerszeg, Hungary

Pathogenic variants of *FOXP2* gene were identified first as a monogenic cause of childhood apraxia of speech (CAS), a complex disease that is associated with an impairment of the precision and consistency of movements underlying speech, due to deficits in speech motor planning and programming. *FOXP2* variants are heterogenous; single nucleotide variants and small insertions/deletions, intragenic and large-scale deletions, as well as disruptions by structural chromosomal aberrations and uniparental disomy of chromosome 7 are the most common types of mutations. *FOXP2*-related speech and language disorders can be classified as "*FOXP2*-only," wherein intragenic mutations result in haploinsufficiency of the *FOXP2* gene, or "*FOXP2*-plus" generated by structural genomic variants (i.e., translocation, microdeletion, etc.) and having more likely developmental and behavioral disturbances adjacent to speech and language impairment. The additional phenotypes are usually related to the disruption/deletion of multiple genes neighboring *FOXP2* in the affected chromosomal region. We report the clinical and genetic findings in a family with four affected individuals having expressive speech impairment as the dominant symptom and additional mild dysmorphic features in three. A 7.87 Mb interstitial deletion of the 7q31.1q31.31 region was revealed by whole genome diagnostic microarray analysis in the proband. The *FOXP2* gene deletion was confirmed by multiplex ligation-dependent probe amplification (MLPA), and all family members were screened by this targeted method. The *FOXP2* deletion was detected in the mother and two siblings of the proband using MLPA. Higher resolution microarray was performed in all the affected individuals to refine the extent and breakpoints of the 7q31 deletion and to exclude other pathogenic copy number variants. To the best of our knowledge, there are only two family-studies reported to date with interstitial 7q31 deletion and showing the core phenotype of *FOXP2* haploinsufficiency. Our study may contribute to a better understanding of the behavioral phenotype of *FOXP2* disruptions and aid in the identification of such patients. We illustrate the importance of a targeted MLPA analysis suitable for the detection of *FOXP2* deletion in selected cases with a specific phenotype of expressive speech disorder. The "phenotype first" and targeted diagnostic strategy can improve the diagnostic yield of speech disorders in the routine clinical practice.

Keywords: 7q31 deletion, *FOXP2*, expressive speech disorder, MLPA, case report

OPEN ACCESS

Edited by:
Katalin Komlosi, Medical Center University of Freiburg, Germany

Reviewed by:
Maria Isabel Melaragno, Federal University of São Paulo, Brazil
Agnieszka Stembalska, Wroclaw Medical University, Poland
Leslie Domenici Kulikowski, University of São Paulo, Brazil

***Correspondence:**
Anikó Ujfalusi
ujfalusi.aniko@med.unideb.hu

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

Received: *05 February 2021*
Accepted: *22 July 2021*
Published: *20 August 2021*

Citation:
Nagy O, Kárteszi J, Elmont B and Ujfalusi A (2021) Case Report: Expressive Speech Disorder in a Family as a Hallmark of 7q31 Deletion Involving the FOXP2 Gene. Front. Pediatr. 9:664548. doi: 10.3389/fped.2021.664548

INTRODUCTION

In the general population, childhood speech disorders are common clinical conditions affecting 1 in 20 preschool children (1). In regard to its possible causes, hearing impairment, autism, intellectual or psychomotor developmental disorders, and genetic syndromes are all in contention (2). Although the starting point and quality of speech in the first years of life may vary greatly, psychomotor alteration with or without the other clinical signs are indications for the child-neurologists to consult a clinical geneticist. Speech delay is one of the most frequent reason for genetic workup in early childhood. The lack of etiological diagnosis causes difficulty in giving proper genotype-phenotype correlations, in providing counseling with respect to possible outcome and treatment opportunities, and furthermore, in cases with a genetic background assessing the risk of recurrence. The genetic background of abnormal speech development is very heterogenous, therefore, phenotypical sub-characterization comes in handy for the clinical geneticist (3, 4). The type of the speech delay (global, expressive, receptive) and the presence of accompanying phenotypic signs and organ developmental disorders can orient the clinical geneticist. It is common that routine brain MRI scans, as part of the diagnostic workup, give negative results in childhood speech disorders, suggesting that brain abnormalities may be present at the sub-macroscopic level (3, 5). As such, the genetic workup is determined by the presence of the additional clinical symptoms of the patient. In case of a syndromic form, the characteristic symptoms of the assumed syndrome define the diagnostic methodology (e.g., Fragile-X syndrome–*FMR1* gene mutation analysis). In the majority of patients, the speech delay is not syndromic or the phenotype of the patient is not specific, as such, the currently available genome wide molecular (cyto)genetic techniques are chosen first. Chromosomal microarray is used to detect chromosomal copy number variations (CNVs), while whole exome sequencing (WES) is the main approach to identify mutations in the protein-coding genes.

Childhood apraxia of speech (CAS) is an uncommon motor speech disorder, defined as a higher-order motor system deficit of motor planning and programming of speech (3). The patients usually have an impaired speech development from infancy manifested as poor feeding, lack of babbling, delayed onset of first words, and limited number of spoken words. According to the American Speech-Language-Hearing Association consensus statement, there are three core diagnostic symptoms of CAS: (i) inconsistent errors on consonants and vowels; (ii) lengthened and disrupted coarticulatory transitions between sound and syllables; and (iii) inappropriate prosody (https://www.asha.org/policy/PS2007-00277). Differentiation from the other types of speech disorders such as articulation or phonological disorders, dysarthria, and stuttering is crucial for providing prognostic information to patients (3). This is a lifelong speech disorder. Many children with CAS also have language problems and literacy impairments that can influence their educational and employment outcomes (6).

FOXP2 (Forkhead box protein P2) (MIM # 605317) was the first gene to be associated predominantly with speech and language disorders (7). It was recognized in a multigenerational "KE" family, wherein the affected members carried a point mutation at a highly conserved position (R553H) within the forkhead DNA binding domain of the protein. The *FOXP2* gene encodes a conserved transcription factor that is important in the development and functioning of the motor cortex, striatum, and cerebellum responsible for fine motor control (2, 8). Depending on the underlying genetic mechanism of *FOXP2* insufficiency, the individuals with *FOXP2*-related speech and language disorders can be categorized as "*FOXP2*-only" and "*FOXP2*-plus." Patients with inactivating intragenic mutations in the *FOXP2* gene have only a speech and language disorder, therefore, these cases are classified as "*FOXP2*-only." In "*FOXP2*-plus" individuals, the genetic background can be large copy number variants (i.e., contiguous gene deletions) (52% of affected individuals), structural variants (i.e., chromosome translocation or inversion) (8% of patients), or maternal uniparental disomy of chromosome 7 (UPD7) involving the *FOXP2* gene (11% of patients) (9). Microdeletions of different chromosomal regions (7q31, 2p15p16.1, 12p13.33, 16p11.2, 17q21.31) and other gene mutations (i.e., *GRIN2A*, *SETBP1*) have also been reported in CAS (7, 10–15).

The 7q31 microdeletion syndrome is an ultra-rare chromosomal anomaly (<1:1,000,000) characterized by a speech and language disorder. Individuals with larger deletions in this region have also been reported to display intellectual disability, dysmorphic features, developmental delay, and autism. About 52% of the *FOXP2*-related speech and language disorders are caused by 7q31q33 deletions that encompass the *FOXP2* gene and flanking DNA. In total, 80% of these deletions are *de novo* and the remaining cases are inherited in an autosomal dominant manner (9).

We report herein the clinical and genetic characterization of a large family with four affected individuals having speech impairment with variable severity and mild dysmorphic features, where the patients carry 7q31.1q31.31 deletion involving the *FOXP2* gene.

CASE REPORT

The female proband (III-4) was examined first from her family. She was born after an uneventful pregnancy at 38th weeks of gestation with a birth weight of 3,000 g from non-consanguineous parents (**Figure 1A**). Perinatal anamnesis was negative. At 10 months of age, a complicated febrile convulsion was observed and antiepileptic treatment was received for a while. She remained symptom free after finishing the therapy. Brain MRI was negative. No feeding difficulty was observed. Early motor development in infancy was almost normal. She walked independently at 15 months of age. However, the parents noticed that the overall movement was somewhat slow and sluggish, and she fell quite a lot of times. The neurologist detected muscle hypotonia and imbalance, and as a result, she received physiotherapy from the age of 2 years. Later, she got physiotherapy for scoliosis. She never reached harmonic walking at all and she walked with bent knees. Her speech development

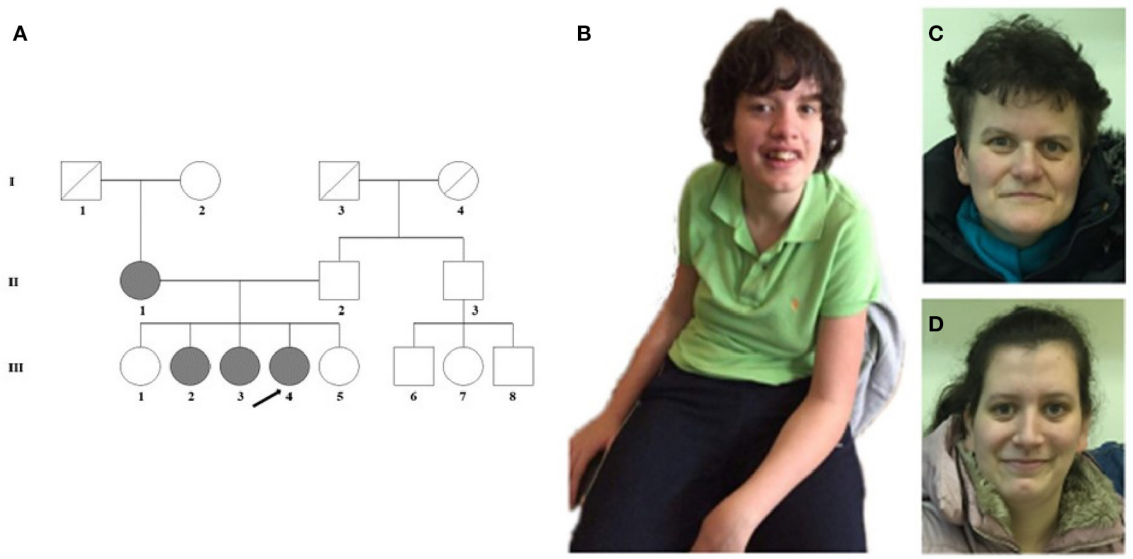


FIGURE 1 | (A) Pedigree of the family. **(B)** The proband (III-4) shows scoliosis and mild facial dysmorphic features such as low forehead, slight hypertelorism, prominent nose, and wide lips. **(C)** The minor facial anomalies of the mother (II-1): maxillary hypoplasia, prominent nose, and pointed chin. **(D)** Face of the sister (III-3) with almost no dysmorphic features.

was delayed and it was retracted even further. She participated in special education considering her speaking disability that was mainly a motor speech impairment. Additionally, her receptiveness was also much below average. Hearing impairment was not detected. She was diagnosed to have a moderately severe psychomotor retardation based on the regular follow-ups of the pedagogical services that regularly check and analyse the development of children globally in different fields of mental and motor development. A child neurologist and a special needs teacher together took parts in every evaluation process performed regularly in this case. **Table 1** summarizes the retracted development of the patient till school age. In spite of the profound delay, she presented for initial genetic evaluation at the age of 16 years. The delay was due to the attitude of the family who lived in a small village, where the mother and two sisters of the patient also suffered from speech impairment. At presentation, physical examination of the proband revealed normal body anthropometric data but several minor anomalies were described as follows: low forehead, slight hypertelorism, prominent nose, wide lips, long and slender fingers, and long toes (**Figure 1B**). Internal organs and external genitalia were normal. Neurological examination found very slow movements, thoracolumbar scoliosis, slight muscular atrophy with sluggish reflexes, convergent strabismus, and intellectual disability, but no abnormal reflexes or ataxia. The patient received regular and specific educational treatment by trained experts for speech and mental impairments with limited results.

The genetic examination of this patient was commenced with peripheral blood cytogenetics using the standard procedure, where the G-banded chromosome analysis showed a normal female karyotype (46,XX). Based on her speech delay and mild dysmorphic features, targeted FISH analysis was done to confirm or exclude the DiGeorge/Velo-cardio-facial syndrome using a locus specific FISH probe (DiGeorge/VCFS TUPLE1) (Cytocell, Rainbow Scientific Inc., Windsor, CT). The result of this test was normal. Array CGH was performed as the next routine diagnostic step using oligonucleotide microarray composed of ∼60,000 probes, distributed through all the genome (*qChip Post*) (Quantitative Genomic Medicine Laboratories, S.L., Barcelona, Espana). Microarray analysis identified a pathogenic interstitial deletion on the long arm of chromosome 7 from 7q31.1 to 7q31.31 cytoband with the following coordinates: arr[GRCh37] 7q31.1q31.31(109745411_117482692)×1. It was an ∼7.73 Mb deletion, that alters the dosage of multiple reference genes, including the morbid gene *FOXP2*.

Detailed phenotype analysis of the family revealed speech difficulties among family members (**Table 1**; **Figure 1A**). The mother (II-1) of the proband had speech impairment since her childhood. She attended a normal primary school and worked in a factory as a manual laborer. She has some distinctive minor anomalies with maxillary hypoplasia, prominent nose, and pointed chin (**Figure 1C**). The older sister (III-2) of the proband, aged 27 years, has severe articulation problems and difficulties in social interactions and communication. She has abilities necessary for daily living, but she does not work. She is aware of her limitations and thus inhibited with unfamiliar people. The other affected sister (III-3) of the proband, aged 25 years, has a borderline IQ (81) with nasal speech and almost no dysmorphic features (**Figure 1D**). She has regular work and no communication difficulties besides the speech dyspraxia. The unaffected, completely healthy sisters (III-1 and III-5) and father (II-2) of the proband have an average intellectual status. No other

TABLE 1 | Clinical features of the affected family members and detailed developmental evaluation of the proband.

	II-1 (Mother)	**III-2 (Sister)**	**III-3 (Sister)**	**III-4 (Proband)**
Age (years)	49	27	25	16
Type of speech disorder	CAS	CAS	Dyspraxia of speech	CAS
Psychomotor development	Normal	Normal	Normal (IQ: 81)	Moderately severe psychomotor retardation (IQ: 31)
Minor anomalies	+	+	-	+
Education	Normal + speech therapy	Normal + speech therapy	Normal + speech therapy	Special school
Regular work	+	–	+	–
Sociability	Normal	Some difficulties	Normal	Severe delay
Developmental evaluation of the Proband (III-4)				
Age of evaluation (year; month)	**Walking***	**Hand manipulation***	**Speech***	**Sociability***
2; 11	19 months (broad base)	10 months (grasps objects for a few minutes)	6 months (gestures and voices)	22 months (hardly accepts unfamiliar people)
5; 3	22 months (some imbalance)	15 months (plays upon being instructed but not spontaneously)	9 months (some syllables)	12 months (cooperates for only a few minutes with the therapist)
7; 4	24 months (disharmonic)	18 months (plays with blocks)	9 months (slight improvement in understanding, says some words)	12 months (activity in cooperation is very simple)
17; 0	Wide based gait	Plays a lot with blocks, constructs complicated objects	Understands, says some words and simple sentences	Friendly, open for cooperation with the therapist

**Age in months is equivalent to the indicated developmental age.*
CAS, childhood apraxia of speech; IQ, intelligence quotient.

family members are known to have any intellectual disability. It was obvious that our proband had the most severe clinical phenotype in the family; nonetheless, analysis of the other family members seemed reasonable.

Based on the array CGH result of the proband, targeted MLPA (Multiplex Ligation-dependent Probe Amplification) analysis was applied as a first-tier test on all family members to examine the deletion status of the *FOXP2* gene. For this analysis, SALSA MLPA Probemix P475-A1 FOXP1-FOXP2 (MRC-Holland, Amsterdam, The Netherlands) that contains probes for all exons of the *FOXP1* and *FOXP2* genes was used. This analysis identified the deletion of all exons of the *FOXP2* gene and no alteration in the *FOXP1* gene in the affected family members (II-1, III-2, III-3, III-4). The clinically healthy sisters (III-1, III-5) and the father (II-2) had normal MLPA results. According to the family history, the grandparents (I-1, I-2) did not have speech and language disturbances, therefore, we assumed that the 7q31 deletion was a *de novo* event in the mother. The grandparents were not available for genetic testing.

After all the aforementioned genetic investigations, we performed array CGH for all the affected family members using a higher resolution CytoScan 750K Array (Affymetrix, Thermo Fisher Scientific, Waltham, MA) to refine the extent of the 7q31 deletion, allowing an accurate determination of the breakpoints. In addition, we wanted to check the presence of concomitant CNVs that can explain the clinical heterogeneity of the affected family members. The higher resolution microarray revealed a heterozygous 7.87 Mb deletion with the following breakpoints: arr[GRCh37] 7q31.1q31.31(109708675_117578862)×1 (**Figure 2A**). The size of the deletion and the breakpoints were exactly the same in every patient. The extent of the 7q31 deletion was slightly bigger than the one detected by the diagnostic microarray without altering the gene content (40 RefSeq genes, among them 20 OMIM morbid genes). No other concomitant pathogenic CNV was detected in the affected individuals. **Figure 2B** shows the genome map of the 7q22-7q32 region depicting the previously reported deleted cases and result of this study.

The observed microdeletion in the family has a great impact on family planning concerning the two fertile sisters with the CAS as they have a 50% chance to transmit the microdeletion to their children with an uncertain impact on speech development. The clinical geneticist has also emphasized that they are carriers of a *CFTR* gene deletion and, therefore, carrier testing of their partners is recommended. It was mentioned that targeted prenatal diagnostics is available in case of future pregnancies.

DISCUSSION

In this study, we present a large family with four affected individuals having an expressive speech impairment caused by the interstitial deletion of the 7q31 region involving the *FOXP2* gene. The 7q31 deletion was maternally inherited resulting in a distinct severity of speech disorder, mild behavioral alteration, and dysmorphic features in the affected family members.

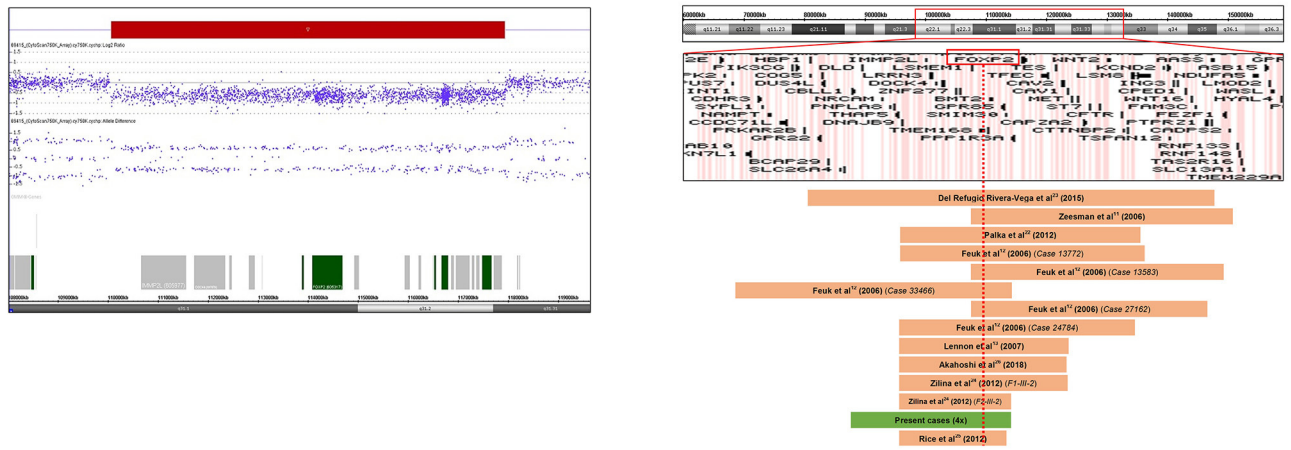


FIGURE 2 | (A) Result of chromosome microarray analysis. A 7.87 Mb deletion at chromosome 7q31.1q31.31 (109708675-117578862). Signal intensity is plotted on a log2 scale, the deleted segment is shown as a red bar. Allele differences confirm the deletion. **(B)** Schematic representation of the 7q22q32 chromosomal region with gene content based on the UCSC Genome Browser. Deleted regions reported in previous cases are represented as brown bars, and the deletion of our patients is depicted as a green bar. *FOXP2* gene is highlighted by a red rectangle. The red dotted line shows that all reported deleted cases involve the *FOXP2* gene. Genomic positions refer to build19.

Known genetic background of CAS is highly heterogeneous. Apart from the first-described *FOXP2* gene mutations, numerous other genes and chromosomal loci were discovered as causative factors in motor speech disorders. Highly penetrant variants usually affect common transcriptional pathways suggesting the essential role of transcriptional regulation in the normal speech development (4, 16). The *FOXP2* gene plays an important role in the speech and language development. Point mutations and deletions of *FOXP2* lead to verbal dyspraxia with impaired expressive and receptive language and are common in most individuals having CAS. Some patients may have a mild developmental delay as well (7). *FOXP2* is a member of the forkhead family of transcription factors, and homologs to other members of this family (*FOXP1*, *FOXP4*), having highly conserved domains (17). The gene is expressed in several structures in the brain including the cortical plate, basal ganglia, thalamus, inferior olives, and cerebellum where the FOXP2 protein may regulate the expression of other genes (18). The expression pattern is specific to subpopulations of neuronal cells in different structures (e.g., Purkinje cells in the cerebellum, deep layers of the cortex, and medium spiny neurons of the striatum). Disruptions of these cells during embryogenesis and postnatal development are risk factors for speech disorders. The human brain imaging studies indicate that *FOXP2* mutations alter the structure and function of the aforementioned brain structures (19, 20). The pathogenic variants of *FOXP2* are heterozygous and predicted to be loss-of-function changes, but the dominant negative effect of the mutant allele has also been suggested (2). Haploinsufficiency of the *FOXP2* gene results in impairments in the sequencing of movement and procedural learning leading to "Speech-language disorder 1" (*FOXP2*, MIM # 602081) in the affected individuals. In "*FOXP2*-only"-related disorders, non-verbal (performance) IQ is typically more preserved compared to verbal IQ. Core features of the disorder are childhood apraxia of speech and patients show difficulties in performing sequential orofacial movements, both linguistic and non-linguistic (9). They have an inability to generate syntactic grammar rules, impaired processing, and expressive language (4, 21).

The 7q31 deletion is a very rare chromosomal abnormality and familiar cases are even more unique. Zeesman et al. (11) first suggested that patients with chromosomal deletions involving 7q31 may define a new contiguous gene deletion syndrome characterized by developmental verbal dyspraxia. Speech and language deficits, articulation problems, and limited oral vocabulary are observed in all patients with haploinsufficiency of *FOXP2*. To date, <30 cases have been reported with interstitial 7q31 deletion encompassing the *FOXP2* gene. These cases usually carry different sized 7q31 deletion and consequently, they differ in clinical manifestation. Most of the 7q31 deletions reported are larger than 10 Mb and present a more complex clinical phenotype (22). The reported symptoms beside the speech impairment were the following: developmental delay, mild intellectual disability, and dysmorphic features (**Table 2**) (23–25). Intellectual disability, paranoid schizophrenia, and unilateral sensorineural hearing loss were described as CAS-associated symptoms only in single cases emphasizing their unknown genetic background (13, 26, 27). To the best of our knowledge, there are only two family studies with interstitial 7q31 deletion reported to date with one or two affected individuals and showing the core phenotype of *FOXP2* haploinsufficiency. Rice et al. (25) reported a detailed clinical assessment (speech, language, cognition, motor functions) of a moderately affected mother and her son with a severe apraxia of speech. Both of them carried a very small, 1.57 Mb deletion on chromosome 7q31 detected by array CGH. The deleted region involved only three genes: *FOXP2*, *MDFIC*, and *PPP1R3A*. Because the last

TABLE 2 | Cytogenetic, molecular, and clinical data of patients with 7q deletion encompassing the *FOXP2* gene.

References	Chromosomal region	Deletion size (Mb)	Speech and language	Other features
Del Refugio Rivera-Vega et al. (23)	7q22.3q32.1	23.1	Language delay	Short stature, motor delay, craniofacial dysmorphism, microcephaly, hand anomalies, intellectual disability
Zeesman et al. (11)	7q31.2q32.2	16	Dyspraxia	Hypotonia, malformed ears, down-turned mouth, brachycephaly
Palka et al. (22)	7q31.1q31.3	14.8 (mosaic)	Dyspraxia	Mild psychomotor delay, high arched palate, lordosis
Feuk et al. (12)	7q31.1q31.3 (*Case 13772*)	15	Dyspraxia	Psychomotor delay, cognitive impairment, behavioral disturbances
	7q31.2q32.3 (*Case 13583*)	15	Dyspraxia	Psychomotor delay, cognitive impairment, ASD-like
	7q22q31.3 (*Case 33466*)	15	Dyspraxia	Psychomotor delay, cognitive impairment
	7q31.2q32 (*Case 27162*)	13	Dyspraxia	Psychomotor delay, cognitive impairment, ASD
	7q31.1q31.3 (*Case 24784*)	11	Dyspraxia	Autism, craniostenosis
Lennon et al. (13)	7q31.1q31.31	9.1	Dyspraxia	Ptosis, plagiocephaly, hypertelorism, bulbous nose, moderate psychomotor delay
Akahoshi et al. (26)	7q31.1q31.31	8	NA	Mild intellectual disability, minor anomalies, paranoid schizophrenia
Zilina et al. (24)	7q31.1q31.31 (*F1-III-2*)	8.3	Dyspraxia	Developmental delay, failure to thrive, dysmorphic phenotype, positive Graefe symptom, urinary tract anomalies, autistic features
	7q31.1q31.2 (*F2-III-2*)	6.5	Dyspraxia	Developmental delay, slightly dysmorphic phenotype, mild ataxia, occasional agressive behavioral
Present cases	7q31.1q31.31 (*III-4*)	7.87	CAS	Muscle hypotonia and imbalance, thoracolumbal scoliosis, moderately severe psychomotor retardation, minor anomalies
	7q31.1q31.31 (*III-3*)	7.87	Dyspraxia	Borderline IQ (81), no dysmorphic features
	7q31.1q31.31 (*III-2*)	7.87	CAS	Difficulties in social interactions and communication
	7q31.1q31.31 (*II-1*)	7.87	CAS	Minor anomalies with maxillary hypoplasia, prominent nose and pointed chin
Rice et al. (25)	7q31.1q31.2	1.57	CAS	No major congenital anomalies or dysmorphic features

ASD, autism spectrum disorder; CAS, childhood apraxia of speech; FOXP2, forkhead box protein P2; IQ, intelligence quotient; Mb, megabase; NA, not applicable.

two genes have not been associated with speech or language disorders, the clinical assessment of these patients provided informative phenotypic data on *FOXP2* haploinsufficiency. Their findings confirmed that *FOXP2* haploinsufficiency can disrupt development in cognition, speech, language, and sensorimotor domains. In the second family study, the authors described the clinical and molecular characterization of two familial cases with speech impairment, developmental delay, and congenital anomalies. They compared the phenotype of the affected patients with deletions of *FOXP2* inherited paternally and maternally (24). The authors did not find a significant difference due to the parental origin of the 7q31 deletion in the investigated two families. They could not confirm the hypothesis published earlier by Feuk et al. (12) that the loss of maternal *FOXP2* should be relatively benign while the loss of paternal *FOXP2* yields severe speech problems because of the differential parent-of-origin expression of the *FOXP2* locus. The clinical findings of our presented cases could not support this assumption either.

It is noteworthy that the 7.87 Mb deletion detected in our proband (III-4) by array CGH covers the two-thirds of the 7q31 region that is known as the autism susceptibility locus 9 (AUTS9, MIM # 611015) as well. A meta-analysis of genome studies on autism or autism spectrum disorders (ASD) found a significant linkage to 7q31 suggesting that this chromosomal region is likely to harbor a susceptibility gene for autism. Although there are contradictory findings on the direct correlation between *FOXP2* variants and ASD, new data emphasize the misregulation of the target genes controlled by *FOXP2* in the downstream signaling pathways as the possible explanation of the autistic features of some of the *FOXP2* mutated patients (28). Autistic features were not observed in the affected members in the presented family, although the oldest patient (III-2) shows a more severe speech impairment and limited communication with unfamiliar people.

The clinical heterogeneity among the affected individuals in this family remains to be elucidated. According to literature data, the mechanism of the phenotypic manifestation of the CNVs and their incomplete penetrance remain largely unclear (29). Recently, it has been reported that differentially methylated regions inside CNVs may be one of the mechanisms of incomplete penetrance of inherited CNVs associated with neurodevelopmental disorders (30). Vasilyev et al. reported differential DNA methylation

of intragenic CpG sites of the *IMMP2L* gene located in a critical region for the autism susceptibility locus on chromosome 7q (AUTS1). The authors suggest a partial compensation of *IMMP2L* gene haploinsufficiency in healthy CNV carriers by reducing the DNA methylation level (31).

The rearrangements of the *FOXP2* gene are considered rare events, probably because of the limitations of the targeted investigation used. Our results contribute to the better understanding of the behavioral phenotype of *FOXP2* disruptions and can aid in the identification of patients. We emphasize the importance of the careful evaluation of speech and language disturbances, focusing on the discrepancy between verbal and non-verbal abilities, lack of behavioral problems, hyperactivity, and autistic features that are frequently associated with speech delay. Our results also emphasize the importance of a targeted MLPA analysis suitable for the detection of *FOXP2* deletion and can improve the diagnostic yield of speech impairment in routine practice. Early molecular diagnosis is highly beneficial for patients as it can help in the assessment of the possible outcome and risk of recurrence.

CONCLUSION

To the best of our knowledge, this is the first report of a family with four affected individuals carrying 7q31 deletion involving the *FOXP2* gene and presenting phenotypic variability both in speech impairment and in other symptoms. The maternally inherited *FOXP2* deletion provides additional support to the previously described role of *FOXP2* haploinsufficiency as a causative factor in speech disorder.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by GINOP Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ON performed the MLPA and the high-resolution microarray analysis and wrote the first draft of the manuscript. JK collected the clinical data and contributed to writing of the manuscript. BE contributed to clinical data collection and phenotypic description. AU performed the supervision and edited the writing. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Ministry of National Economy, Hungary (GINOP-2.3.2-15-2016-00039).

ACKNOWLEDGMENTS

We would like to thank the patients and parents for their kind cooperation and consent to this study.

REFERENCES

1. Reilly S, McKean C, Morgan A, Wake M. Identifying and managing common childhood language and speech impairments. *BMJ.* (2015) 350:h2318. doi: 10.1136/bmj.h2318
2. Vernes SC, Nicod J, Elahi FM, Coventry JA, Kenny N, Coupe AM, et al. Functional genetic analysis of mutations implicated in a human speech and language disorder. *Hum Mol Genet.* (2006) 15:3154–67. doi: 10.1093/hmg/ddl392
3. Morgan AT, Webster R. Aetiology of childhood apraxia of speech: a clinical practice update for paediatricians. *J Paediatr Child Health.* (2018) 54:1090–5. doi: 10.1111/jpc.14150
4. Hildebrand MS, Jackson VE, Scerri TS, Van Reyk O, Coleman M, Braden RO, et al. Severe childhood speech disorder: gene discovery highlights transcriptional dysregulation. *Neurology.* (2020) 94:e2148–67. doi: 10.1212/WNL.0000000000009441
5. Liégeois FJ, Morgan AT. Neural bases of childhood speech disorders: lateralization and plasticity for speech functions during development. *Neurosci Biobehav Rev.* (2012) 36:439–58. doi: 10.1016/j.neubiorev.2011.07.011
6. Gillon GT, Moriarty BC. Childhood apraxia of speech: children at risk for persistent reading and spelling disorder. *Semin Speech Lang.* (2007) 28:48–57. doi: 10.1055/s-2007-967929
7. Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature.* (2001) 413:519–23. doi: 10.1038/35097076
8. Spiteri E, Konopka G, Coppola G, Bomar J, Oldham M, Ou J, et al. Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. *Am J Hum Genet.* (2007) 81:1144–57. doi: 10.1086/522237
9. Morgan A, Fisher SE, Scheffer I, Hildebrand M. FOXP2-related speech and language disorders. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, et al., editors. *GeneReviews(®).* Seattle, WA: University of Washington, Seattle. Copyright © 1993-2020, University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved (1993).
10. Hannula K, Lipsanen-Nyman M, Kristo P, Kaitila I, Simola KO, Lenko HL, et al. Genetic screening for maternal uniparental disomy of chromosome 7 in prenatal and postnatal growth retardation of unknown cause. *Pediatrics.* (2002) 109:441–8. doi: 10.1542/peds.109.3.441
11. Zeesman S, Nowaczyk MJ, Teshima I, Roberts W, Cardy JO, Brian J, et al. Speech and language impairment and oromotor dyspraxia due to deletion of 7q31 that involves FOXP2. *Am J Med Genet A.* (2006) 140:509–14. doi: 10.1002/ajmg.a.31110
12. Feuk L, Kalervo A, Lipsanen-Nyman M, Skaug J, Nakabayashi K, Finucane B, et al. Absence of a paternally inherited FOXP2 gene in developmental verbal dyspraxia. *Am J Hum Genet.* (2006) 79:965–72. doi: 10.1086/508902
13. Lennon PA, Cooper ML, Peiffer DA, Gunderson KL, Patel A, Peters S, et al. Deletion of 7q31.1 supports involvement of FOXP2 in language impairment: clinical report and review. *Am J Med Genet A.* (2007) 143a:791–8. doi: 10.1002/ajmg.a.31632
14. Liégeois F, Morgan AT, Connelly A, Vargha-Khadem F. Endophenotypes of FOXP2: dysfunction within the human articulatory network. *Eur J Paediatr Neurol.* (2011) 15:283–8. doi: 10.1016/j.ejpn.2011.04.006
15. Turner SJ, Hildebrand MS, Block S, Damiano J, Fahey M, Reilly S, et al. Small intragenic deletion in FOXP2 associated with childhood apraxia of speech and dysarthria. *Am J Med Genet A.* (2013) 161a:2321–6. doi: 10.1002/ajmg.a.36055

16. Eising E, Carrion-Castillo A, Vino A, Strand EA, Jakielski KJ, Scerri TS, et al. A set of regulatory genes co-expressed in embryonic human brain is implicated in disrupted speech development. *Mol Psychiatry.* (2019) 24:1065–78. doi: 10.1038/s41380-018-0020-x
17. Takahashi H, Takahashi K, Liu FC. FOXP genes, neural development, speech and language disorders. *Adv Exp Med Biol.* (2009) 665:117–29. doi: 10.1007/978-1-4419-1599-3_9
18. Enard W. FOXP2 and the role of cortico-basal ganglia circuits in speech and language evolution. *Curr Opin Neurobiol.* (2011) 21:415–24. doi: 10.1016/j.conb.2011.04.008
19. Liégeois FJ, Hildebrand MS, Bonthrone A, Turner SJ, Scheffer IE, Bahlo M, et al. Early neuroimaging markers of FOXP2 intragenic deletion. *Sci Rep.* (2016) 6:35192. doi: 10.1038/srep35192
20. Morgan AT, Su M, Reilly S, Conti-Ramsden G, Connelly A, Liégeois FJ. A brain marker for developmental speech disorders. *J Pediatr.* (2018) 198:234–9.e1. doi: 10.1016/j.jpeds.2018.02.043
21. Co M, Anderson AG, Konopka G. FOXP transcription factors in vertebrate brain development, function, and disorders. *Wiley Interdiscip Rev Dev Biol.* (2020) 9:e375. doi: 10.1002/wdev.375
22. Palka C, Alfonsi M, Mohn A, Cerbo R, Guanciali Franchi P, Fantasia D, et al. Mosaic 7q31 deletion involving FOXP2 gene associated with language impairment. *Pediatrics.* (2012) 129:e183–8. doi: 10.1542/peds.2010-2094
23. Del Refugio Rivera-Vega M, Gómez-Del Angel LA, Valdes-Miranda JM, Pérez-Cabrera A, Gonzalez-Huerta LM, Toral-López J, et al. A novel 23.1 Mb interstitial deletion involving 7q22.3q32.1 in a girl with short stature, motor delay, and craniofacial dysmorphism. *Cytogenet Genome Res.* (2015) 145:1–5. doi: 10.1159/000381234
24. Zilina O, Reimand T, Zjablovskaja P, Männik K, Männamaa M, Traat A, et al. Maternally and paternally inherited deletion of 7q31 involving the FOXP2 gene in two families. *Am J Med Genet A.* (2012) 158a:254–6. doi: 10.1002/ajmg.a.34378
25. Rice GM, Raca G, Jakielski KJ, Laffin JJ, Iyama-Kurtycz CM, Hartley SL, et al. Phenotype of FOXP2 haploinsufficiency in a mother and son. *Am J Med Genet A.* (2012) 158a:174–81. doi: 10.1002/ajmg.a.34354
26. Akahoshi K, Yamamoto T. Interstitial deletion within 7q31.1q31.3 in a woman with mild intellectual disability and schizophrenia. *Neuropsychiatr Dis Treat.* (2018) 14:1773–8. doi: 10.2147/NDT.S168469
27. Zhao J, Noon SE, Krantz ID, Wu Y. A de novo interstitial deletion of 7q31.2q31.31 identified in a girl with developmental delay and hearing loss. *Am J Med Genet C Semin Med Genet.* (2016) 172:102–8. doi: 10.1002/ajmg.c.31488
28. Bowers JM, Konopka G. The role of the FOXP family of transcription factors in ASD. *Dis Markers.* (2012) 33:251–60. doi: 10.1155/2012/456787
29. Kirov G, Rees E, Walters JT, Escott-Price V, Georgieva L, Richards AL, et al. The penetrance of copy number variations for schizophrenia and developmental delay. *Biol Psychiatry.* (2014) 75:378–85. doi: 10.1016/j.biopsych.2013.07.022
30. Barbosa M, Joshi RS, Garg P, Martin-Trujillo A, Patel N, Jadhav B, et al. Identification of rare de novo epigenetic variations in congenital disorders. *Nat Commun.* (2018) 9:2064. doi: 10.1038/s41467-018-04540-x
31. Vasilyev SA, Skryabin NA, Kashevarova AA, Tolmacheva EN, Savchenko RR, Vasilyeva OY, et al. Differential DNA methylation of the IMMP2L gene in families with maternally inherited 7q31.1 microdeletions is associated with intellectual disability and developmental delay. *Cytogenet Genome Res.* (2021) 161:105–19. doi: 10.1159/000514491

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 | Frontiers in Genetics

ORIGINAL RESEARCH
published: 12 April 2022
doi: 10.3389/fgene.2022.652454

Variability in Phelan-McDermid Syndrome in a Cohort of 210 Individuals

Julián Nevado[1,2,3]†, Sixto García-Miñaúr[1,2,3], María Palomares-Bralo[1,2,3], Elena Vallespín[1,2,3], Encarna Guillén-Navarro[4], Jordi Rosell[5], Cristina Bel-Fenellós[6,7], María Ángeles Mori[1,2,3], Montserrat Milá[8], Miguel del Campo[9], Pilar Barrúz[1], Fernando Santos-Simarro[1,2,3], Gabriela Obregón[10], Carmen Orellana[11], Harry Pachajoa[12], Jair Antonio Tenorio[1,2,3], Enrique Galán[13], Juan C. Cigudosa[14], Angélica Moresco[10], César Saleme[15], Silvia Castillo[16,17], Elisabeth Gabau[18], Luis Pérez-Jurado[2,19]†, Ana Barcia[20], Maria Soledad Martín[1], Elena Mansilla[1,2,3], Isabel Vallcorba[1,2,3], Pedro García-Murillo[21], Franco Cammarata-Scalisi[22], Natálya Gonçalves Pereira[23], Raquel Blanco-Lago[24], Mercedes Serrano[25], Juan Dario Ortigoza-Escobar[25], Blanca Gener[26], Verónica Adriana Seidel[27], Pilar Tirado[28], Pablo Lapunzina[1,2,3]† and Spanish PMS Working Group*

[1]Instituto de Genética Médica y Molecular (INGEMM)-IdiPAZ, Hospital Universitario La Paz, Madrid, Spain, [2]CIBERER, Centro de Investigación Biomédica en Red de Enfermedades Raras, ISCIII, Madrid, Spain, [3]ITHACA-European Reference Network, Hospital La Paz, Madrid, Spain, [4]Hospital Virgen de la Arrixaca, Murcia, Spain, [5]Hospital Son Espases, Palma de Mallorca, Spain, [6]Departamento de Investigación y Psicología en Educación, Facultad de Educación, UCM, Madrid, Spain, [7]CEE Estudio-3, Afanias, Madrid, Spain, [8]Hospital Clinic, Barcelona, Spain, [9]Hospital Vall D'Hebron, Barcelona, Spain, [10]Hospital Juan P. Garrahan, Buenos Aires, Argentina, [11]Hospital La Fé, Valencia, Spain, [12]Universidad Icesi, Cali, Colombia, [13]Hospital Materno-Infantil Infanta Cristina, Badajoz, Spain, [14]NIM-Genetics Madrid, Alcobendas, Spain, [15]Maternity Nuestra Señora de la Merced, Tucumán, Argentina, [16]Sección Genética, Hospital Clínico Universidad de Chile, Santiago, Chile, [17]Clínica Alemana, Santiago, Chile, [18]Corporación Sanitaria Parc Taulí, Barcelona, Spain, [19]Servicio de Genética, Instituto de Investigaciones Médicas Hospital del Mar (IMIM)/Universitat Pompeu Fabra, Barcelona, Spain, [20]Hospital Universitario Virgen del Rocío, Sevilla, Spain, [21]Unidad de Genética, Hospital Virgen de la Salud, Toledo, Spain, [22]Servicio de Pediatría, Hospital Regional de Antofagasta, Antofagasta, Chile, [23]Clínica Reprodutiva NIDUS, Juiz de Fora, Brasil, [24]Servicio de Neuropediatría, Hospital Universitario Central de Asturias, Oviedo (Asturias), Spain, [25]Unidad de Neuropediatría, Hospital San Joan de Deu, Barcelona, Spain, [26]Hospital Universitario de Cruces, Bilbao, Spain, [27]Servicio de Genética Clínica, Hospital Universitario Gregorio Marañon, Madrid, Spain, [28]Servicio de Neuropediatría, Hospital Universitario La Paz, Madrid, Spain

Phelan-McDermid syndrome (PMS, OMIM# 606232) results from either different rearrangements at the distal region of the long arm of chromosome 22 (22q13.3) or pathogenic sequence variants in the *SHANK3* gene. *SHANK3* codes for a structural protein that plays a central role in the formation of the postsynaptic terminals and the maintenance of synaptic structures. Clinically, patients with PMS often present with global developmental delay, absent or severely delayed speech, neonatal hypotonia, minor dysmorphic features, and autism spectrum disorders (ASD), among other findings. Here, we describe a cohort of 210 patients with genetically confirmed PMS. We observed multiple variant types, including a significant number of small deletions (<0.5 Mb, 64/189) and *SHANK3* sequence variants (21 cases). We also detected multiple types of rearrangements among microdeletion cases, including a significant number with post-zygotic mosaicism (9.0%, 17/189), ring chromosome 22 (10.6%, 20/189), unbalanced translocations (*de novo* or inherited, 6.4%), and additional rearrangements at 22q13 (6.3%, 12/189) as well as other copy number variations in other chromosomes, unrelated to 22q deletions (14.8%, 28/189). We compared the

OPEN ACCESS

Edited by:
Katalin Komlosi,
Medical Center University of Freiburg, Germany

Reviewed by:
Andreas Martin Grabrucker, University of Limerick, Ireland
Catalina Betancur, Institut National de la Santé et de la Recherche Médicale (INSERM), France

***Correspondence:**
Julián Nevado
jnevado@salud.madrid.org

†ORCID:
Julian Nevado
orcid.org/0000-0001-5611-2659
Pablo Lapunzina
orcid.org/0000-0002-6324-4825
Luis Pérez-Jurado
orcid.org/0000-0002-1988-3005

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: 12 January 2021
Accepted: 16 February 2022
Published: 12 April 2022

Citation:
Nevado J, García-Miñaúr S, Palomares-Bralo M, Vallespín E, Guillén-Navarro E, Rosell J, Bel-Fenellós C, Mori MÁ, Milá M, Campo Md, Barrúz P, Santos-Simarro F, Obregón G, Orellana C, Pachajoa H, Tenorio JA, Galán E, Cigudosa JC, Moresco A, Saleme C, Castillo S, Gabau E, Pérez-Jurado L, Barcia A, Martín MS, Mansilla E, Vallcorba I, García-Murillo P, Cammarata-Scalisi F, Gonçalves Pereira N, Blanco-Lago R, Serrano M, Ortigoza-Escobar JD, Gener B, Seidel VA, Tirado P, Lapunzina P and Spanish PMS Working Group (2022) Variability in Phelan-McDermid Syndrome in a Cohort of 210 Individuals. Front. Genet. 13:652454. doi: 10.3389/fgene.2022.652454

clinical and genetic characteristics among patients with different sizes of deletions and with *SHANK3* variants. Our findings suggest that *SHANK3* plays an important role in this syndrome but is probably not uniquely responsible for all the spectrum features in PMS. We emphasize that only an adequate combination of different molecular and cytogenetic approaches allows an accurate genetic diagnosis in PMS patients. Thus, a diagnostic algorithm is proposed.

Keywords: autistic behavior, 22q13 deletion syndrome, Phelan-McDermid syndrome (PMS), *SHANK3*, subtelomeric deletion syndrome, intellectual disabilities (ID)

INTRODUCTION

In the past 15–20 years, the increasing use of genome-wide telomere screening by fluorescence *in situ* hybridization (FISH), multiplex ligation-dependent probe amplification (MLPA, Schouten et al., 2002), and more recently chromosome microarrays (CMA) has provided evidence of the presence of subtle abnormalities involving telomeres in around 5% (range, 2%–30%) of patients with intellectual disability (ID) (Anderlid et al., 2002a; Shao et al., 2008). In the evaluation of ID patients, deletion of 22q13.3, also known as Phelan-McDermid syndrome (PMS; OMIM#:606232), is one of the most common subtelomeric deletions after 1p36.3 deletion syndrome (Heilstedt et al., 2003; Delahaye et al., 2009). PMS usually results from either the loss of genetic material at the distal region of the long arm of chromosome 22 (including *SHANK3*) or pathogenic sequence variants in *SHANK3*.

SHANK3 plays a central role in forming the postsynaptic environment, integrating the protein network of glutamate receptors at postsynaptic density and the maintenance of synaptic structures (Boeckers, 2006; Durand et al., 2007). Deletion sizes vary considerably among PMS individuals, ranging from intragenic deletions in the *SHANK3* gene (~13 Kb) to around 9 Mb (Bonaglia et al., 2011; Phelan et al., 2018). The deletion occurs with similar frequency in male and female. *SHANK3* haploinsufficiency is proposed to be responsible for the major neurological features of the 22q13 deletion syndrome (Bonaglia M. C. et al., 2001; Anderlid et al., 2002b; Wilson et al., 2003; Durand et al., 2007; Phelan et al., 2018) and recently has also been shown to be involved in additional clinical features of the syndrome in humans (De Rubeis et al., 2018) and mice (Sauer et al., 2019). However, interstitial deletions disrupting the 22q13.3 band, not including *SHANK3* (Wilson et al., 2008; Disciglio et al., 2014; Ha et al., 2017), are also reported. The clinical features in these patients overlap those of PMS, raising debate about whether they can be diagnosed as having PMS.

Although many PMS patients have been diagnosed worldwide, most of the individuals included in previous genotype-phenotype analyses had microdeletions (Cusmano-Ozog et al., 2007; Dhar et al., 2010; Sarasua et al., 2011; Soorya et al., 2013; Sarasua et al., 2014a,b; Tabet et al., 2017; Samogy-Costa et al., 2019). Indeed, the proportion of patients with *SHANK3* variants in previous data is 3%–25% (Phelan et al., 2018; De Rubeis et al., 2018; and ClinVar, Varsome, LOVD databases) or 8.6% in the PMS International Registry (among genetically confirmed cases; Kolevzon et al., 2019). Thus, PMS seems to be underdiagnosed, and its exact prevalence in is unknown.

Here, we describe the clinical and molecular data of one of the largest cohorts of patients with confirmed genetic diagnosis of PMS, most of them with microdeletions (189/210, 90%) and 21 with *SHANK3* sequence variants (10%). High-resolution CMA, cytogenetic, and MLPA techniques were necessary to delineate the size and gene content of the deletions and to identify additional rearrangements. Exome and/or target panel sequencing analysis of *SHANK3* were preferentially applied for *SHANK3* sequence variant analysis.

MATERIAL AND METHODS

Subjects

Between 2008 and 2020, 242 patients with confirmed PMS, mostly nonrelated (except for four individuals from two families), were recruited for this study in collaboration with the Spanish PMS Association and the Argentinean PMS Group. Twenty-eight of these had incomplete clinical or molecular data and were not included in this study. Three were excluded because they carried deletions at 22q13.33 nearby to *SHANK3* but not including this gene, and one had an intragenic *SHANK3* duplication and was also excluded because, at this time, we are not able to confirm that the duplication is in tandem and disrupts *SHANK3*. Thus, 210 individuals constituted the final cohort (**Supplementary Figure S1**).

Most of the DNA samples from these patients were extracted and analyzed at INGEMM (Madrid, Spain). A minority of them had been previously analyzed outside of our institution by high-resolution CMA or next generation sequencing (NGS). The patients' clinical information was obtained from the referring physicians and/or their clinical geneticists and compiled in two questionnaires. Data were completed by reviewing medical records and parents' interviews. Parents or guardians provided informed consent. The Institutional Review Board of Hospital Universitario La Paz approved the study (PI: 2735 HULP, Madrid. Spain).

METHODS

Karyotyping and FISH

Cytogenetic analyses were performed on GTG-banded metaphases at a resolution of about 550 bands according to

standard laboratory protocols using Chromosome Kit P (Euroclone, Siziano PV, Italy). FISH was performed according to standard laboratory protocols using the subtelomeric 22q13 probe (D22S1056, Kreatech Biotechnology B.V, Amsterdam, Netherlands) or the DiGeorge/VCFS probe mixture (Vysis Inc., IL, United States), containing a control probe in *ARSA* that maps to the 22q13.3 region. In some cases, the probe N25/N85A3 (Cytocell, Cambridge, United Kingdom) within the *SHANK3* locus was also used.

Parental Origin Analysis

We used highly polymorphic short tandem repeats (D22S1169, D22S1149, D22S444, D22S1170, D22S295, and D22S1141) mapping within the *SHANK3* gene and around it to evaluate parental segregation. The forward primers were synthesized and labeled with fluorescein-amidite (Sigma-Aldrich, St. Louis, MO, United States), whereas the reverse primers were not labeled (primer sequences are available upon request). The region amplified by these primers depended on the number of repeats. Capillary electrophoresis (Applied Biosystems Genetic Analyzer System 3130) was used to detect the length of the fragments (Thermo Fisher, CA, United States).

MLPA Probe Kits

We used several MLPA-Salsa kits in this study (MRC-Holland, Amsterdam, Netherlands). For patients referred to rule out subtelomeric rearrangements in the first years of the study, MLPA kits P036 and P070 were used. DNA samples of all patients with 22q13 deletions were further characterized with the specific MLPA P188 and P339 probe mixes for PMS (MRC-Holland). Both kits contain 34 sequence probes on chromosome 22q13 and control ones for other chromosomes (12 and 9, respectively). The majority of the 22q13 probes (22/34) are in the 1 Mb terminal region of the long arm (P188) and include multiple probes within *SHANK3* (P339). Data analyses were performed according to the protocols supplied by the providers defining relative probe signals by dividing each measured peak area by the sum of all peak areas of the control probes of that sample. Each peak's relative probe area ratio was then compared to a DNA control sample (Promega, United Kingdom), using Coffalysser.net (MRC- Holland).

Chromosome Microarray Analysis (CMA)

Different array platforms were used in this study: *1*) a clinical 60K-array CGH (INGEMM, KaryoArray-®, Vallespin et al., 2013) in 72 of 189 patients; *2*) a high-resolution customized- 60K aCGH (INGEMM custom design, not published) at 22q13.3 in 30 of 189 patients; *3*) different custom or commercial CGH-microarrays with a variety of resolutions in 59 patients (**Supplementary Figure S2A**); *4*) a genome-wide scan of 850,00 tag SNPs (Illumina Infinium CytoSNP-850K BeadChip) in 56 patients (Supplemental Data, **Supplementary Figure 2B**) at INGEMM; and *5*) a genome-wide scan of 750,00 tag SNPs (Affymetrix, ThermoFisher Scientific, Waltham, MA, United States) in 11 patients. Arrays in 1–3 were analyzed with Cytogenomics software (Agilent Corporation; Santa Clara, CA, United States). Image data from 4 were analyzed using the Chromosome Viewer tool contained in the Genome Studio package (Illumina, San Diego, CA, United States). In Chromosome Viewer, gene call scores <0.15 at any locus were considered "no calls." In addition, allele frequency analysis was applied for all SNPs. For the analysis of 5, the ChAS software (Affymetrix, Thermo-Fisher Scientific, Waltham, MA, United States) was used.

All genomic coordinates were established according to the 2009 human genome build 19 (GRCh37/NCBI build 37.1). Deletion coordinates were plotted using the University of California at Santa Cruz Genome Browser (http://genome.ucsc.edu/).

SHANK3 Sequencing Analysis

These studies were performed either at INGEMM or outside of our institution, using different NGS approaches, all under the manufacturer's guidelines: *1*) exome sequencing by trio analysis using the Agilent SureSelect XT clinical research exome (Agilent Tech) and IDT Technologies (Coralville, IA, United States); *2*) singleton exome sequencing CentoXome Gold®, and NOVAGENE (Agilent all exon V6) and MedExome, Q-Genomics (Barcelona, Spain); and *3*) a customized gene panel of specific genes related to ID or/and autism (Agilent-based Technologies). Most samples (98%) were run in Illumina instruments (such as Nextseq500; Miseq, Hiseq 2000/4000; Illumina, San Diego, CA, United States). Classification of the variants follows ACMG/AMP criteria (Richards et al., 2015), using VarSome 10.2 as a web source.

Validation of Global Functional Assessment of the Patients (GFAP)

We estimated an individual severity score in our cohort using different features taken from the questionnaires and weighed them by Human Ontology Phenotype (HPO) term frequencies on a numerical scale of core features of the syndrome. The GFAP was constructed as follows: items with a frequency between 0% and 20%, 1 point; between 20% and 35%, 2 points; between 35% and 70%, 5 points, and >70%, 10 points. Principal component analysis (PCA) was used to validate the GFAP construct, containing Kaiser-Meyer-Olkin's measure and Barlett's test.

Statistical Analysis

Statistical analyses were performed with SPSS version 25 (IBM Corporation, Chicago, IL, United States). Descriptive analysis included mean ± SD for continuous variables and frequency tables for categorical variables (**Table 1**). The categorical variables were taken from our two questionnaires curated from medical records and were expressed as "1" (condition present at some point) or "0" (condition not present at any time). Correlation associations were calculated using Pearson's linear correlation coefficient (continuous variables) or Spearman's Rho and Kendall's tau_b (categorical variables). Comparisons between two groups were performed either with Student's *t*-test (for continuous variables) or chi-square tests (for categorical ones). For more than two groups, ANOVA (followed by Bonferroni's or T3-Dunnett *post hoc* tests) were run for continuous variables and *z*-tests between column

TABLE 1 | Descriptive statistics and frequencies of variables used in the study of 22q13.3 microdeletions and *SHANK3* variants.
a) Categorical variables

		Deletions		*SHANK3* variants	
		Frequency	**Percentage**	**Frequency**	**Percentage**
Sex	Male	85	44.7	13	61.9
	Female	105	55.3	8	38.1
	Total	190	100	21	100
Growth	Centile ≤3	23	12.1	0	0
	Normal	105	56.3	16	88.9
	Centile ≥95	60	31.6	2	11.1
	Total	188	100	18	100
Walk independently	≤15 months	50	26.3	13	72.2
	>15 months	139	73.7	5	27.8
	Total	189	100	18	100
Delayed/absent speech	No words	65/181	36	5/18	27.7
	Some words, 10–20	70/181	39.6	8/18	44.4
	Many words, and ability to make sentences	46/181	25.4	5/18	27.7
	Total	181	100	18	100
Hypotonia	No	45	24.1	8	38
	Yes	142	75.9	13	62
	Total	187	100	21	100
Behavior abnormalities (e.g., stereotypies, manic behavior)	No	39	20.9	1	5.9
	Yes	148	79.1	20	94.1
	Total	187	100	21	100
Regressions	No	98	52.1	10	52.6
	Yes	90	47.9	9	47.4
	Total	188	100	19	100
Seizures	No	129	69	16	84.2
	Yes	58	31	3	15.8
	Total	187	100	19	100
High pain threshold	No	62	33.2	4	21
	Yes	125	66.8	15	79
	Total	187	100	19	100
Decreased perspiration	Yes	99	52.7	5	31.2
	Normal	77	42.4	11	68.8
	Increased	11	5.9	0	0
	Total	187	100	16	100
Microcephaly	Normal	151	81.1	16	88.9
	Yes	37	18.9	2	11.1
	Total	188	100	18	100
Macrocephaly	Normal	139	73.9	14	76.8
	Yes	49	26.1	4	22.2
	Total	188	100	18	100
	No	150	79.8	16	94.1
Dolicocephaly	Yes	38	20.2	1	5.9
	Total	188	100	17	100
	No	161	86.1	16	88.9
Flat midface	Yes	26	13.9	2	11.1
	Total	187	100	18	100
	No	134	71.3	16	88.9
Epicanthal folds	Yes	54	28.7	2	11.1
	Total	188	100	18	100
	No	138	73.8	16	88.9
Strabismus	Yes	49	26.2	2	11.1
	Total	187	100	18	100
	No	153	81.8	15	88.2
Ptosis	Yes	34	18.2	2	11.8
	Total	187	100	17	100
	No	80	42.8	8	44.4
Long eyelashes	Yes	107	57.2	10	55.6
	Total	187	100	18	100
	No	113	60.4	16	94.1
Full eyebrow	Yes	74	39.6	1	5.9
	Total	187	100	17	100
	No	144	75.8	17	94.5

(Continued on following page)

TABLE 1 | (*Continued*) Descriptive statistics and frequencies of variables used in the study of 22q13.3 microdeletions and *SHANK3* variants.
a) Categorical variables

		Deletions		*SHANK3* variants	
		Frequency	**Percentage**	**Frequency**	**Percentage**
Full/puffy eyelids	Yes	43	22.6	1	5.5
	Total	187	100	18	100
	No	143	77	17	94.5
Deep set eyes	Yes	44	23	1	5.5
	Total	187	100	18	100
	No	82	43.9	12	60
Wide nasal bridge	Yes	105	56.1	8	40
	Total	187	100	20	100
	No	79	42.2	11	61.1
Bulbous nose	Yes	108	57.8	7	38.9
	Total	187	100	18	100
	No	102	54	11	57.9
Ear anomalies	Yes	86	46	8	42.1
	Total	188	100	19	100
	No	145	77.5	15	79
Full/puffy cheeks	Yes	42	22.5	4	21
	Total	187	100	19	100
	No	99	52.9	15	83.3
Widely spaced teeth/malocclusion	Yes	88	47.1	3	16.7
	Total	187	100	18	100
	No	78	41.7	7	38.9
Pointed chin	Yes	109	58.3	11	61.1
	Total	187	100	18	100
	No	136	72.7	17	94.5
Toe syndactyly	Yes	51	27.3	1	5.5
	Total	187	100	18	100
	No	86	45.3	9	52.9
Large, fleshy hands	Yes	101	53.2	8	47.1
	Total	187	100	17	100
	No	152	80	16	94.1
Fifth finger clinodactyly	Yes	35	18.4	1	5.9
	Total	187	100	17	100
	No	111	59.4	11	61.1
Hypoplastic/dysplastic nails	Yes	76	40.6	7	38.9
	Total	187	100	18	100
Main reason for genetic consultation	DD	102	54	5	29.4
	ASD	26	13.8	8	47.1
	Dysmorphic features	7	3.7	0	0
	ID	17	9.0	1	5.9
	Hypotonia	16	8.5	0	0
	Language problems	8	4.2	3	17.6
	Other	13	6.8	0	0
	Total	189	100	17	100
	DD	66	34.9	2	11.8
Second reason for genetic consultation	ASD	26	13.8	4	23.5
	Dysmorphic features	9	4.6	1	5.9
	ID	27	14.3	1	5.9
	Hypotonia	19	10.1	0	0
	Language problems	26	13.8	9	52.9
	Other	16	8.5	0	0
	Total	189	100	17	100
	No	159	84.1	16	94.1
Cardiac anomalies	Yes	30	15.9	1	5.9
	Total	189	100	17	100
	No	146	78.1	13	76.5
Ophthalmologic anomalies	Yes	41	21.9	4	23.5
	Total	187	100	17	100
	No	161	86.1	8	47.1
Sphincter control	Yes	26	13.9	9	52.9
	Total	187	100	17	100
	No	146	77.7	16	94.1

(Continued on following page)

TABLE 1 | (*Continued*) Descriptive statistics and frequencies of variables used in the study of 22q13.3 microdeletions and *SHANK3* variants.
a) Categorical variables

		Deletions		*SHANK3* variants	
		Frequency	**Percentage**	**Frequency**	**Percentage**
Renal and urogenital anomalies	Yes	42	22.2	1	5.9
	Total	188	100	17	100
	No	171	91.0	16	88.9
Lip/palate abnormalities	Yes	17	9.0	2	11.1
	Total	188	100	18	100
	No	142	75.9	8	42.1
Sleeping disorders	Yes	45	24.1	11	57.9
	Total	187	100	19	100
	No	146	77.7	13	76.5
Skin anomalies	Yes	42	22.2	4	23.5
	Total	188	100	17	100
	No	157	83.9	13	72.2
Recurrent infections	Yes	30	16.1	5	27.8
	Total	187	100	18	100
	No	175	93.6	17	100
Herniae	Yes	12	6.4	0	0
	Total	187	100	17	100
	No	184	97.9	17	100
Obesity	Yes	4	2.1	0	0
	Total	188	100	17	100
	No	167	89.3	14	82.2
Hearing problems	Yes	20	10.7	3	17.8
	Total	187	100	17	100
	No	169	90.4	17	100
Lymphedema	Yes	18	9.6	0	0
	Total	187	100	17	100
	No	153	81.8	13	72.2
Gastrointestinal problems	Yes	34	18.2	5	27.8
	Total	187	100	18	100
	Not performed	92	49.2	8	38.1
Brain MRI	Normal	59	31.6	11	52.4
	With abnormalities	36	19.2	2	9.5
	Total	187	100	21	100
	No	81	43.3	9	50
Poor visual contact	Yes	106	56.7	9	50
	Total	187	100	18	100
	No	117	62.6	11	61.1
Biting	Yes	70	37.4	7	38.9
	Total	186	100	18	100
	No	126	67.4	6	33.3
Very sensitive to touch	Yes	61	32.6	12	66.7
	Total	187	100	18	100
	No	118	63.1	11	61.1
Uncontrolled laughter	Yes	69	36.9	7	38.9
	Total	187	100	18	100
	No	90	48.1	10	52.6
Impulsive	Yes	97	51.9	9	47.4
	Total	187	100	19	100
	No	117	63.1	13	72.2
Excessive yelling	Yes	69	36.9	5	27.8
	Total	186	100	18	100
	No	145	77.2	13	76.5
Hair pulling	Yes	42	22.6	4	23.5
	Total	187	100	17	100
	No	144	77	13	76.5
Skin picking	Yes	43	23	4	23.5
	Total	187	100	17	100
	No	161	86.1	13	76.5
Nonstop crying	Yes	26	13.9	4	23.5
	Total	187	100	17	100
	No	151	80.8	17	89.5

(Continued on following page)

TABLE 1 | (*Continued*) Descriptive statistics and frequencies of variables used in the study of 22q13.3 microdeletions and *SHANK3* variants.
a) Categorical variables

		Deletions		*SHANK3* variants	
		Frequency	**Percentage**	**Frequency**	**Percentage**
Aggressive behavior	Yes	36	19.2	2	10.5
	Total	187	100	19	100
	No	125	66.9	12	66.7
Tongue thrusting, sticking out	Yes	62	33.1	6	33.3
	Total	187	100	18	100
	No	89	47.6	4	22.2
Abnormal emotional response	Yes	98	52.4	14	77.8
	Total	187	100	18	100
Formal ASD evaluation	Not performed	152	81.3	13	62
	Normal	13	6.9	1	4.8
	ASD diagnosis[a]	22	11.8	7	33.2
	Total	187	100	21	100

[a]*ASD diagnosis according to the psychiatrists of the referring institutions.*

TABLE 1a | b) Continuous variables

	Deletions					*SHANK3* variants				
	N	**Mean**	**Standard error**	**Standard deviation**	**Median**	***N***	**Mean**	**Standard error**	**Standard deviation**	**Median**
Age at evaluation (years)	189	12.44	0.63	8.67	10.30	21	10.99	1.41	5.95	9.10
Age at diagnosis (months)	184	71.40	6.42	86.90	36	19	94.40	13.06	53.85	84
Size (Mb)	189	3.54	0.21	2.85	3.29	19	—	—	—	—
GFAP (arbitrary units)	187	109.54	2.81	33.70	11.50	21	86.11	11.20	46.25	83.30
Final *N* (per list)	189					21				

Descriptive analysis included mean ± SD for continuous variables and frequency/percentages for categorical variables. The categorical variables were taken from our two questionnaires curated from medical records and were expressed as "1" (condition present at some point) or "0" (condition not present at any time). ASD, autism spectrum disorder; DD, developmental delay; ID, intellectual disability; GFAP, global functional assessment of the patients; MRI, magnetic resonance image.

proportions for categorical variables. Ward's minimum variance method was the criterion used in hierarchical cluster analysis, and the number of clusters was selected using the Bayesian information criterion (BIC) or Akaike information criterion (AIC). A p-value lower than .05 was considered to indicate a statistically significant difference.

RESULTS

Cohort

Individuals (n = 210), all previously nonreported, are mostly from Spain, all over the country (n = 178), and from South America (n = 32), mainly from Argentina (**Supplementary Figure S1**). The female/male ratio, 1.12:1 (111/99), was similar to previous reports, and ages ranged from birth to 62 years. Descriptive statistics (for continuous variables) and frequencies (for categorical items) are shown in **Table 1**. The majority of individuals with PMS in our cohort are of pediatric age (between 0 and 16 years old, 146 patients; 69.5%). The mean age at diagnosis was around 6 years old for deletions (**Table 1b**) and around 8 years for the group with sequence variants in *SHANK3*. The mean age at evaluation were 12.44 ± 8.7 years and 10.99 ± 5.95 years for deletions and *SHANK3* sequence variants, respectively (**Table 1b**).

Clinical Findings

The clinical features observed in this cohort by weighed-HPO terms are listed in **Table 2a** for 22q13.3 microdeletions. **Table 2a** also shows the frequencies of clinical features observed in other representative studies with deletion cases (Sarasua et al., 2014a; Tabet et al., 2017, Samogy-Costa et al., 2019). **Table 2b** shows the frequencies of clinical features observed in patients with *SHANK3* variants, and data from De Rubeis et al. (2018) and other previously published cases (Gauthier et al., 2009; Boccuto et al., 2013; Leblond et al., 2014; O'Roak et al., 2014; Bramswig et al., 2015; Nemirovsky et al., 2015; Zhang et al., 2015; Holder & Quach 2016; Bowling et al., 2017; Lim et al., 2017; Yuen et al., 2017).

Figure 1 shows that facial features are neither typical nor specific for PMS. Patients presented a high degree of facial variability even among individuals with similar deletion size. Significant facial differences can be observed when comparing bigger deletions (>5 Mb) with either small deletions (≤0.5 Mb) or sequence variants in *SHANK3* (**Figure 1**). Facial features such as bulbous nose, pointed chin, ear anomalies, full eyebrows, long eyelashes, and wide nasal bridge were observed in around 35%–80% of the individuals (**Table 2a**). These facial features, together with hypotonia, high pain threshold, developmental delay, speech delay, ID, behavior abnormalities, large/fleshly hands, hypoplastic/dysplastic nails, decreased perspiration, and ASD, should be considered as core features of this syndrome (at

TABLE 2 | Frequency of clinical features observed in this cohort.
a) Microdeletions at 22q13.3

HPO clinical features frequencies	This study (189 cases)	Sarasua et al., 2014a (201 cases)	Tabet et al., 2017 (78 cases)	Samogy-Costa et al., 2019 (34 cases)
≥70 Intellectual disability	95.8% (181/189)	NA	100% (66/66)	NA
≥70 Speech delay	97.4% (184/189)	86.0% (37/43)	100% (65/65)	88.9 (24/27)
≥70 Developmental delay	74.3% (139/187)	88.0% (44/50)	NA	NA
≥70 Hypotonia	75.9% (142/187)	74.5% (82/110)	42.1% (32/76)	84.8% (28/33)
≥70 Behavior abnormalities	79.1%(148/187)	65.3% (83/127)	77.3% (34/44)	NA
≥70 High pain threshold	66.8% (125/187)	77.1% (131/170)	NA	80.0%(24/30)
35–60% ASD diagnosis[a]	62.9% (22/35)	NA	NA	NA
35–60% Pointed chin	58.3% (109/187)	52.3% (58/111)	6.6% (5/76)	NA
35–60% Wide nasal bridge	56.1% (105/187)	NA	2.6% (2/76)	42.3% (11/26)
35–60% Decreased perspiration	52.9% (99/187)	36% (18/50)	NA	NA
35–60% Ear anomalies	45.7% (86/188)	NA	15.8% (12/76)	73.1% (19/26)
35–60% Full brow	39.6% (74/187)	NA	NA	NA
35–60% Impulsive	51.9% (97/187)	40% (78/166)	NA	NA
35–60% Long eyelashes	57.2% (107/187)	84% (95/113)	2.6% (2/76)	11.5% (3/26)
35–60% Bulbous nose	57.8% (108/187)	NA	2.6% (2/76)	15.4% (4/26)
35–60% Large/fleshly hands	54.0% (101/187)	63.4%(71/112)	6.6%(5/76)	NA
40–60% Abnormal emotional response	52.4% (98/187)	NA	NA	NA
35–60% Regressions	47.9% (90/188)	NA	9.2% (6/65)	NA
35–60% Widely spaced teeth/malocclusion	47.1% (88/187)	NA	11.8% (9/76)	7.7% (2/26)
35–60% Hypoplastic/dysplastic nails	40.6% (76/187)	73% (81/111)	3.9% (3/76)	7.7% (2/26)
35–60% Abnormal brain MRI	37.9% (36/95)	NA	NA	NA
40–60% Biting	37.6% (70/186)	45.8% (82/179)	NA	NA
35–60% Excessive yelling	37.1% (69/186)	31% (54/174)	NA	NA
35–60% Uncontrolled laughter	36.9%(69/187)	NA	3.1% (2/65)	NA
20–35% Play frequently with tongue thrusting/sticking out	33.2% (62/187)	NA	NA	NA
20–35% Very sensitive to touch	32.6% (61/187)	NA	NA	NA
20–35% Growth centile >95%	31.9% (60/188)	9.4% (9/96)	4.6% (3/65)	NA
20–35% Seizures	31% (58/187)	54.3% (82/151)	18.5% (12/65)	NA
20–35% Epicanthus	28.7% (54/188)	46.8% (52/111)	10.5% (8/76)	7.7% (2/26)
20–35% 2/3 toe syndactyly	27.3% (51/187)	48.2%(53/110)	10.5% (8/76)	7.7% (2/26)
20–35% Strabismus	26.2% (49/187)	26.6% (29/109)	30.3% (23/76)	11.5% (3/26)
20–35% Macrocephaly	26.1% (49/188)	18.2% (20/110)	1.7% (1/60)	NA
20–35% Sleep disorders	24.1% (45/187)	46.2% (12/26)	5.7% (3/53)	42.4% (14/33)
20–35% Ability to make sentences	25.4% (46/181)	NA	NA	NA
20–35% Deep set eyes	23.5% (44/187)	28.8% (32/111)	NA	NA
20–35% Skin picking	23% (43/187)			
20–35% Hair pulling	22.5% (42/187)	25.5% (48/188)	NA	NA
20–35% Full/puffy cheeks	22.5% (42/187)	NA	NA	NA
20–35% Renal and urogenital anomalies	22.3% (42/188)	26.4% (39/148)	7.5% (4/53)	30.3% (10/33)
20–35% Skin anomalies	22.3% (42/188)	NA	NA	NA
20–35% Ophthalmological anomalies	21.9% (41/187)	NA	NA	NA
20–35% Dolichocephaly	20.2% (38/188)	31.9% (36/113)	NA	NA
<20% Aggressive behavior	19.3% (36/187)	38.6% (49/127)	10.8% (7/65)	NA
<20% Microcephaly	19.7% (37/188)	10.9% (12/110)	6.6% (5/76)	NA
<20% Gastrointestinal problems	18.2% (34/187)	41.6% (62/149)	18.5%(12/65)	56.7%(17/30)
<20% Recurrent infections	16.0% (30/187)	NA	13.2% (7/53)	60.6% (20/33)
<20% Growth centile <3%	12.2% (23/188)	11.5% (11/96)	16.9% (11/65)	NA

[a]*ASD diagnosis according to the psychiatrists of the referring institutions.*

least in patients with microdeletions; **Table 2a**). On the other hand, patients with variants in *SHANK3* seemed to have fewer dysmorphic features than patients with microdeletions (**Figure 1** and **Table 2b**).

Interestingly, many of these core features seem to be interrelated among them. Significant positive correlations were observed when Kendall's tau_b analysis was performed between categorical variables (**Supplementary Table S1**). An example with three of these categorical variables is schematized in **Figure 2**.

Brain MRI studies were performed in 51% (95/187) of individuals in the microdeletion group and 62% (13/21) in the *SHANK3* sequence variant group with abnormal findings found in 38% (36/95) and 15% (2/13), respectively (**Table 1a**). Abnormal findings included hypoplasia/atrophy of the cerebellar vermis, abnormalities of the corpus callosum (ranging from thinness to agenesis or dysgenesis), abnormalities of the white matter, arachnoid cysts, and hydrocephalus. We also found other abnormalities, such as ventriculomegaly, enlarged cisterna-magna and vermis, prominent metopic suture, cerebral

TABLE 2b | b) *SHANK3* variants

HPO Frequencies	Clinical features	This study (21 cases)	De Rubeis et al. (17 cases)	Other cases (33 cases)
≥70	Intellectual disability	95.2% (20/21)	100% (17/17)	100% (33/33)
≥70	Speech delay	85.7% (18/21)	82.4% (14/17)	95.7% (22/23)
≥70	ASD diagnosis[b]	100% (7/8)	68.8% (11/16)	93.9% (31/33)
≥70	Behavior anomalies	95.2% (20/21)	94.1% (16/17)	71.4% (15/21)
≥70	High pain threshold	79.0% (15/19)	94.1% (16/17)	100% (1/1)
≥70	Hypotonia	65% (13/21)	94.1% (16/17)	66.7% (8/12)
≥70	Abnormal emotional response	77.7% (14/18)	NA	NA
≥70	Developmental delay	66.6% (14/21)	82.4 (14/17)	54.5% (6/11)
≥70	Very sensitive to touch	66.7% (12/18)	NA	NA
≥70	Long eyelashes	55.6% (10/18)	72.7% (8/11)	100% (5/5)
35–60%	Sleep disorders	57.9% (11/19)	58.8% (10/17)	100% (6/6)
35–60%	Wide nasal bridge	55.6% (12/20)	55.5% (6/11)	85.7% (6/7)
35–60%	Pointed chin	61.1% (11/18)	63.6% (7/11)	57.1% (8/14)
35–60%	Regressions	47.4% (9/19)	64.7% (11/17)	66.7% (16/24)
35–60%	Hypoplastic/dysplastic nails	38.9% (7/18)	63.6% (7/11)	100% (2/2)
35–60%	Ear anomalies	42.1% (8/19)	36.4% (4/11)	87.5% (7/8)
35–60%	Uncontrolled laughter	38.9% (7/18)	NA	NA
35–60%	Biting her/himself or others	38.9% (7/18)	NA	NA
35–60%	Impulsive	47.4% (9/19)	NA	NA
35–60%	Recurrent infections	27.8% (5/18)	52.9% (9/17)	50% (1/2)
35–60%	Gastrointestinal problems	27.8% (5/18)	29.4% (5/17)	75% (6/8)
35–60%	Seizures	15.8% (3/19)	29.4% (5/17)	56.7% (17/30)
35–60%	Head size anomalies	33.3% (6/18)	28.6% (4/14)	57.1% (4/7)
35–60%	Dental anomalies	16.7% (3/18)	63.6% (7/11)	100% (1/1)
35–60%	Decreased perspiration	38.9% (7/18)	16.7% (2/12)	50% (1/2)
20–35%	Poor visual contact	50.1% (9/18)	29.4%(5/17)	NA
20–35%	Fifth finger clinodactyly	5.6% (1/18)	81.8% (9/11)	NA
20–35%	Lip/palate anomalies	11.2% (2/18)	NA	50% (1/2)
20–35%	Tongue thrusting, sticking out	33.3% (6/18)	NA	NA
20–35%	Excessive yelling	27.8% (5/18)	NA	NA
20–35%	Decrease perspiration/heat intolerance	31.3% (5/16)	16.7% (2/12)	NA
20–35%	Deep set eyes	5.6% (1/18)	45.5% (5/11)	75% (3/4)
20–35%	Abnormal brain MRI	15.4% (2/13)	33.3% (5/15)	25% (2/8)
20–35%	Bulbous nose	38.9% (7/18)	54.5% (6/11)	85.7% (6/7)
20–35%	Epicanthus	11.1% (2/18)	45.5% (5/11)	50% (1/2)
20–35%	Macrocephaly	22.2% (4/18)	21.4% (3/14)	28.6% (2/7)
20–35%	Hair pulling	23.5% (4/17)	NA	NA
20–35%	Full/puffy cheeks	21.1% (4/19)	18.2% (2/11)	0.0% (0/1)
20–35%	2/3 toe syndactyly	5.6% (1/18)	45.5% (5/11)	0.0% (0/1)
20–35%	Strabismus	11.2% (2/18)	11.8% (2/17)	50% (3/6)
20–35%	Aggressive behavior	10.5% (2/19)	47.1%(8/17)	9.1%(2/22)
20–35%	Verbally fluent	27.8% (5/18)	17.6% (3/17)	4.3% (1/23)
20–35%	Flat midface	11.2% (2/18)	NA	50% (1/2)

[b]*ASD diagnosis according to the psychiatrists of the referring institutions.*

dysplasia with lateral ventricular dilatation, and frontal cerebral hypertrophy.

Speech abilities (evaluated only in patients ≥3 years old; n = 199/210, 94.8%) showed severe abnormalities in most of the patients evaluated (148/199, 74.4%). Thirty-five percent of patients (70/199) had no speech at all, around 39% (78/199) had an elementary vocabulary of 10 words or less, and around 26% (51/199) were reported to have a significant vocabulary and the ability to use limited phrases for a short and comprehensible conversation (**Table 1a**). **Table 1** segregates the numbers by deletions and *SHANK3* variants. Remarkably, most of the verbally fluent individuals in the microdeletion group have small deletions.

The main reason for referral to a genetic consultation in patients with microdeletions was developmental delay, whereas in individuals with sequence variants, ASD and language delay were the most frequent reasons for referral (**Figure 3** and **Table 1**). Similarly, ASD and delayed or absent speech were the main cause of genetic consultation among patients with smaller deletions (≤0.25 Mb). We compared these groups by Chi-square test and z-test (*post hoc*, corrected by Bonferroni). We choose 0.25 Mb as the size of the deletions with the minimal telomeric lost segment, including the *SHANK3* gene. The chi-square test revealed differences between groups constituted by large deletions (>0.25 Mb, 153 cases), small deletions (<0.25 Mb, 36 cases), and *SHANK3* variants (21 cases) for the first- and second-main reasons for referral to genetic consultation (p = .0001, F = 43.491 and p = .0001, F = 37.491, respectively). These differences were mainly observed between deletions >0.25 Mb and both smaller deletions and

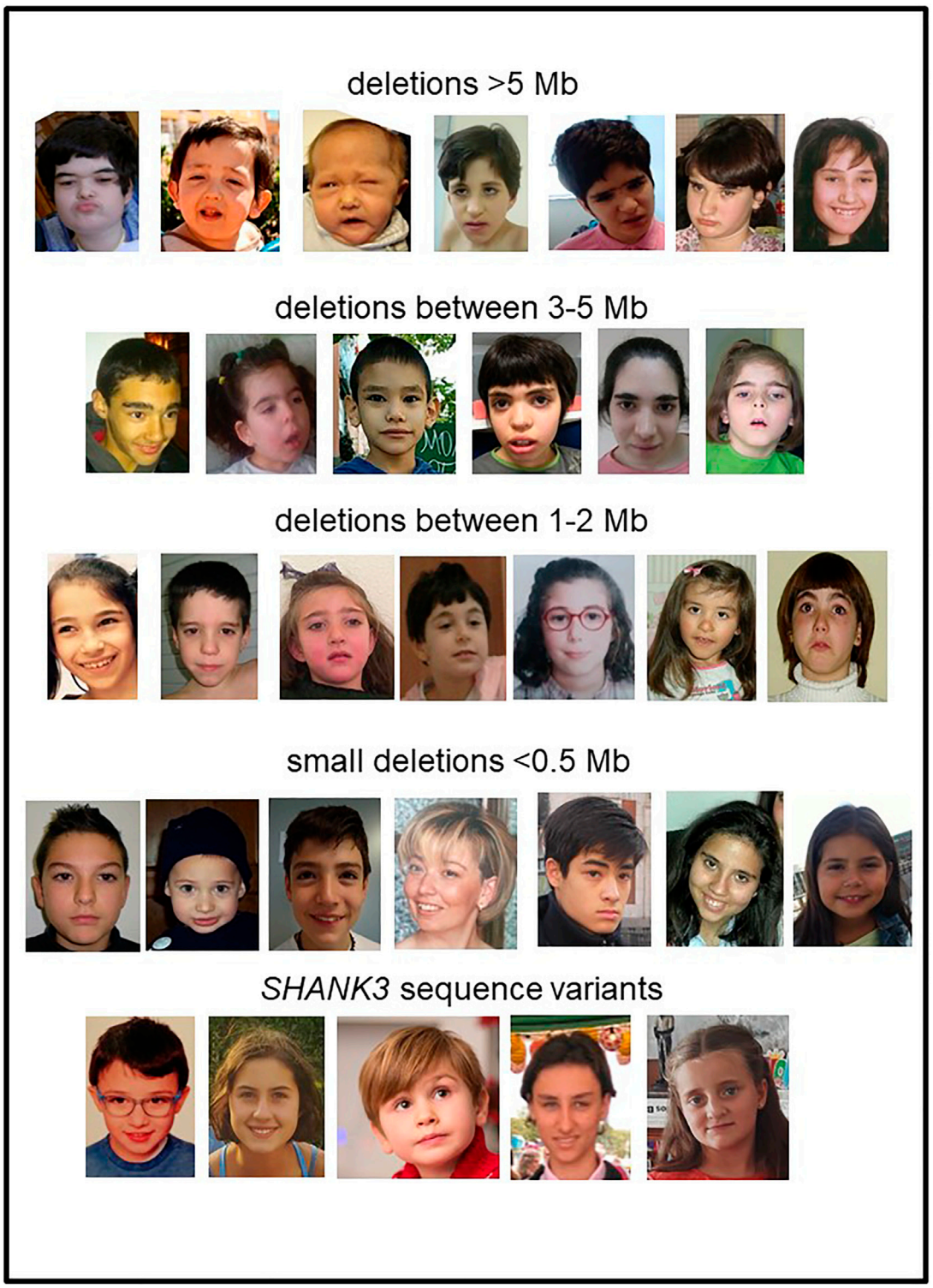


FIGURE 1 | Facial views of individuals with PMS with 22q13.3 deletions or *SHANK3* sequence variants.

variants in *SHANK3* (**Figure 3**). In addition, hypotonia and dysmorphic features were the main reasons for referral in individuals with medium-size deletions (2.5–5.0 Mb). In patients with deletions ≥5 Mb, the main reason for genetic consultation was severe ID and developmental delay with other severe comorbidities (data not shown).

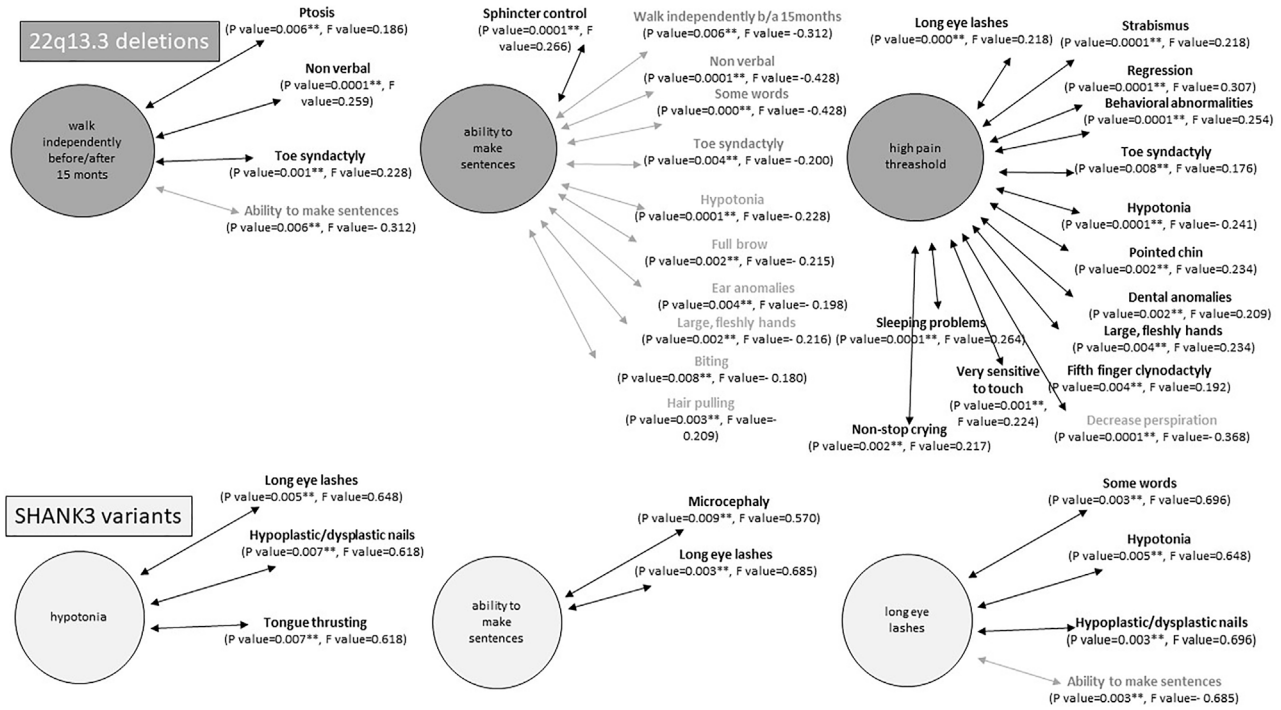


FIGURE 2 | Examples of statistically significant correlations ($p < 0.001$) between intercategorical variables in individuals with 22q13.3 deletions (top) or *SHANK3* variants (bottom). Statistical analyses were performed using Kendal tau_b correlation coefficient. In bold, positive correlations and in gray negative correlations.

Genetic Findings

Analysis of 22q13.3 Deletion Breakpoints

We applied different CMA platforms and MLPA approaches to confirm and establish the size of the deletions. **Figure 4** illustrates the need to use MLPA for a complete characterization of patients with deletions. This is explained by the lack of probes at the end of the 22q13.33 band in commercial microarrays versus customized microarrays (**Supplementary Figure S2**). A compilation of additional examples is shown in **Supplementary Figure S3**.

One-hundred eighty-nine out of 210 individuals carried deletions at 22q13.3 (90%), all of them including *SHANK3* (**Table 3**). **Table 3** also summarizes how the different genomic rearrangements were distributed in the cohort. The number of individuals with ring chromosome 22 (r(22), 20 cases), post-zygotic mosaicism (17 cases), or additional genomic rearrangements (40 cases, including variants of uncertain significance (VUS) and clinically relevant variants in other chromosomes as well as 12 cases with other rearrangements at chromosome 22), is remarkable. **Supplementary Table S2** shows the genomic coordinates of the 22q13 deletions and other CNVs identified in the cohort. The mean 22q13 deletion size was 3.52 ± 2.83 Mb (median: 3.29 Mb), ranging from 12 Kb within the last exon of *SHANK3* (individual 51) to 10.30 Mb (individual 170) from the telomere. To our knowledge, the latter is the largest deletion reported so far and was likely not lethal because it is in mosaic form. Cytogenetic data of most of these individuals are shown in **Supplementary Table S2**.

The use of combined SNP arrays and MLPA allowed finding different degrees of post-zygotic mosaicism in microdeletion cases. We found 17 patients with mosaicism ranging from 10% to 82% (**Figure 5**). In addition, the finding of two siblings with the same deletion (a 48 Kb-interstitial microdeletion with breakpoints within genes *SHANK3* and *RABL2B*, **Supplementary Figure S4**) suggests parental germinal mosaicism, which was later confirmed as paternal after haplotype analysis using SNP arrays (CytoScan 850K, Illumina).

Breakpoint analyses showed a recurrent 5′breakpoint hot spot, apparently the same described by Bonaglia M. C. et al. (2001). We observed a similar breakpoint in 22 individuals with smaller deletions (coordinates 51123505 to telomere, GCRh37, **Supplementary Figure S5**). This region is rich in SINEs and LINEs, such as Alu sequences, which could be involved in causing these rearrangements by various mechanisms (Bonaglia et al., 2011; Cooper et al., 2011; Oberman et al., 2015). Our data also point out two additional 3′ recurrent breakpoints (**Supplementary Figure S5**), which are also extremely rich in Alu sequences. The first recurrent breakpoint was located between coordinates 51146663 and 51175872 (GCRh37; patients 94, 99, and 117) and the second one was located between intron 19 and the end of the last exon of *SHANK3* (NM_001372044.1; patients 31, 57, 75, and 77). Both hypothetical breakpoints were close to the one predicted in a patient reported by Bonaglia M. C. et al. (2001). Additional cases are needed to confirm these new hot spot breakpoints.

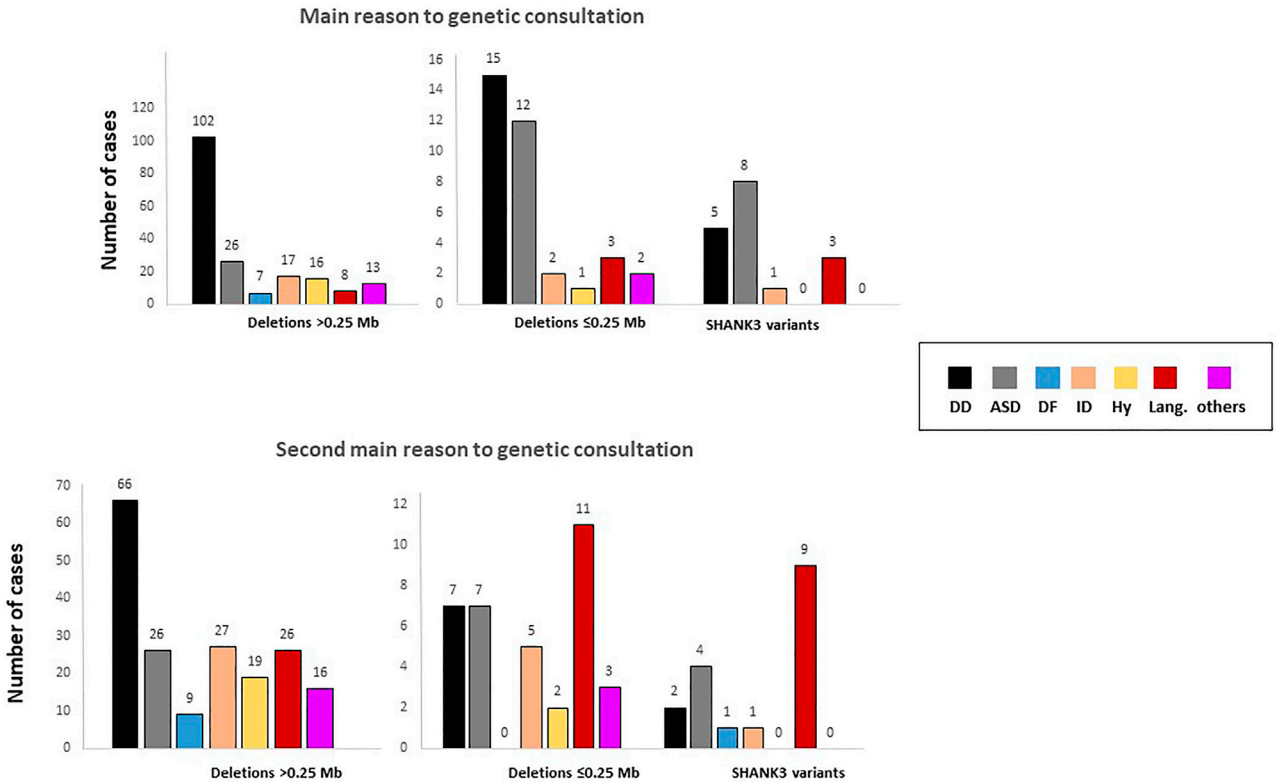


FIGURE 3 | Reasons for referral for genetic evaluation stratified according to the type of genetic defect. Analyses were performed by one-way ANOVA. DD, developmental delay; ASD, autism spectrum disorder; DF, dysmorphic features; ID, intellectual disability; Hy, hypotonia; Lang., language.

Parental Origin of the Deletions

We tested six highly polymorphic short tandem repeats (STR) to identify the parental origin of the deleted chromosome in 86 trios. In 35 cases (40.7%), the results were noninformative. Among 51 trios with informative findings, we found that deletions originated from the paternally inherited chromosome in 76.5% (39/51) and the maternally inherited chromosome in 23.5% of cases (12/51).

Sequence Variants in *SHANK3*

In this cohort, we also evaluated 21 patients (10%) carrying *SHANK3* variants (**Table 4**). All of them were *de novo*; 19 variants were within the penultimate exon (NM_001372044.2), one affected the canonical splicing site at exon 24, and one was located in exon 20. There were 17 frameshift, one nonsense, one splice site, and two missense variants. Some of the variants (**Table 4**) have been previously described in public databases (ClinVar, LOVD, Varsome) and several publications and are recurrent in our patients (Leblond et al., 2014; Bramswig et al., 2015; Holder & Quach 2016; Thevenon et al., 2016; Yuen et al., 2017; De Rubeis et al., 2018; Zhou et al., 2019; Kaplanis et al., 2020; Feliciano et al., 2019; Lelieveld et al., 2016; Retterer et al., 2016; O´Roak et al., 2014; Farwell et al., 2015; Durand et al., 2007), suggesting several hot spots for *de novo* variants.

The interpretation of these two missense variants within *SHANK3* remains difficult (**Table 4**). We classified them as VUS-likely pathogenic by following ACMG/AMP criteria based on *de novo* condition, the individuals' clinical features, their absence in European non-Finnish population in gnomAD, the domain of the protein affected, *in silico* pathogenicity scores, and its medium-high level of conservation position in the evolution. However, the missense *SHANK3* variant c.3673C>T(p.Pro1225Ser) was observed in two independent individuals of African descent (total allele frequency 7×10^{-6}; gnomAD v2.1.1), a finding that may question its association with the clinical features observed in the patient.

Finally, the presence of the same *SHANK3* variant in male monochorionic dizygotic twins suggested potential gonadal mosaicism in one of the parents (data not shown). Haplotype analysis using SNP array suggested a paternal origin of the variant. We also have the suspicion for another case with parental mosaicism in a family with two affected twins.

Genotype-Phenotype Analysis

Individual GFAP

The significant clinical and genetic heterogeneity observed in patients with PMS suggests the type of genetic defect modulates

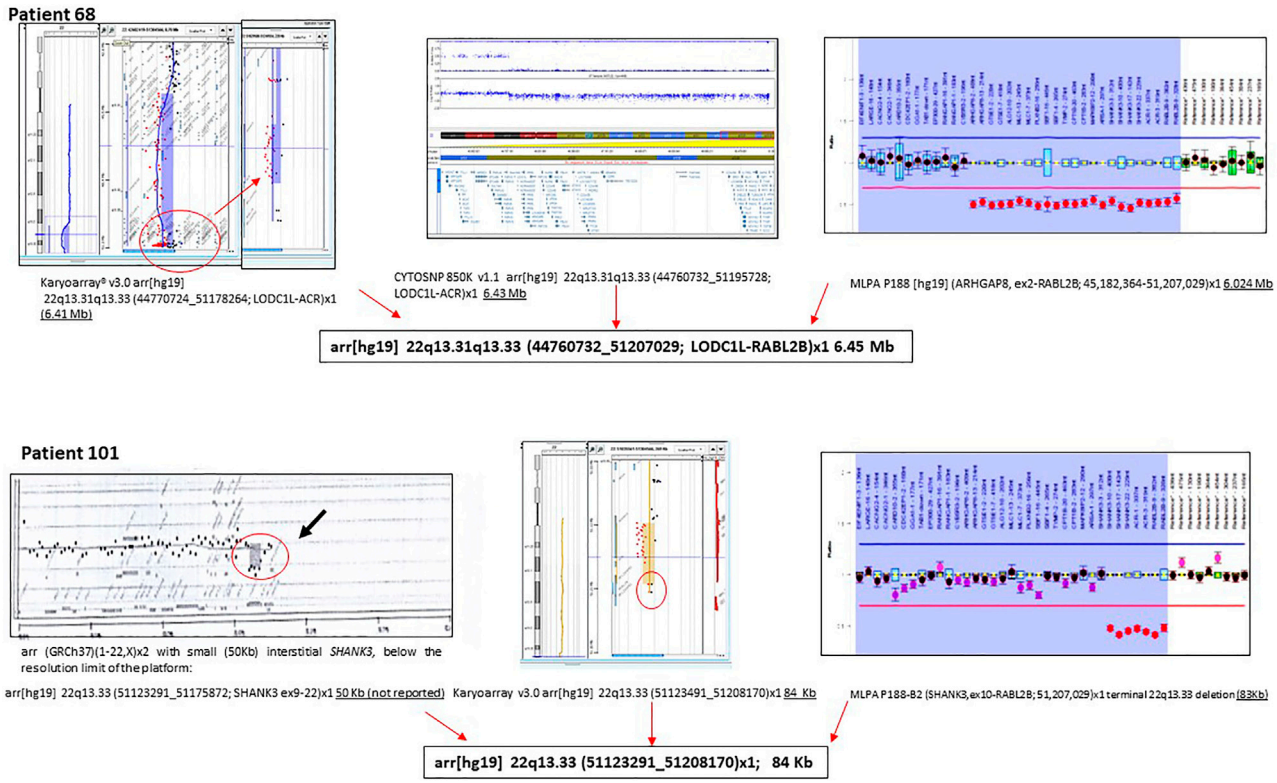


FIGURE 4 | Examples of molecular characterization of two individuals with PMS. Different molecular approaches were used, including CGH-array, SNP array, and MLPA.

TABLE 3 | Summary of genetic findings from the cohort.

Type of genetic alteration	Number of cases
Deletions	189/210 (90%)
Simple terminal deletions	144/189 (76.9%)
Ring 22	20/189 (10.6%)
Mosaic	8/20 (40%)
Unbalanced translocations	13/189 (6.9%)
Inherited	5
De novo	8
Postzygotic mosaic deletions	17/189 (9.0%)
Parental germinal mosaicism	1
Interstitial deletions (including *SHANK3*)	12/189 (6.3%)
Additional genomic rearrangements	40/189 (21.1%)
At chromosome 22	12
In other chromosomes	28
SHANK3 sequence variants	21/210 (10%)
Parental germinal mosaicism	1

the clinical features. Thus, we propose a numerical score of the GFAP, constructing a continuous variable based on a prioritization array of different "core" clinical weighted-HPO items (*see* **Methods**). These variables were based on comorbidity items, developmental delay, speech delay, dysmorphic features, and behavior items. **Figure 6A** shows the median values for GFAP for the whole cohort and different types of genetic defects. **Figures 6B–D** shows median values for other continuous variables (age at diagnosis and evaluation and size of deletions) in the different groups.

Comparative Analysis Between Genetic Subgroups

We compared 10 subgroups of individuals with different types of genetic defects: *1*) large deletions (>0.25 Mb; mean size ± SD, 4.29 ± 2.50), *2*) smaller deletions (≤0.25 Mb, 0.10 ± 0.05), *3*) interstitial deletions (1.94 ± 3.55 Mb), *4*) *SHANK3* sequence variants, *5*) ring 22 (3.53 ± 2.44 Mb), *6*) unbalanced translocations (3.69 ± 1.61 Mb), *7*) mosaic deletions (3.5 ± 3.48 Mb), *8*) additional rearrangement at chromosome 22 (3.32 ± 2.02 Mb), *9*) additional rearrangement in other chromosomes (2.62 ± 2.26 Mb), and *10*) all cases with additional rearrangements (2.99 ± 2.26 Mb) (**Table 5**). Bonferroni or T3-Dunnett *post hoc* tests reveal that the significant differences in the variable "size of deletion" were mainly due to differences between large (>0.25 Mb) and small (≤0.25 Mb) or interstitial deletions, and between small deletions and ring 22 or unbalanced translocations (**Table 5**).

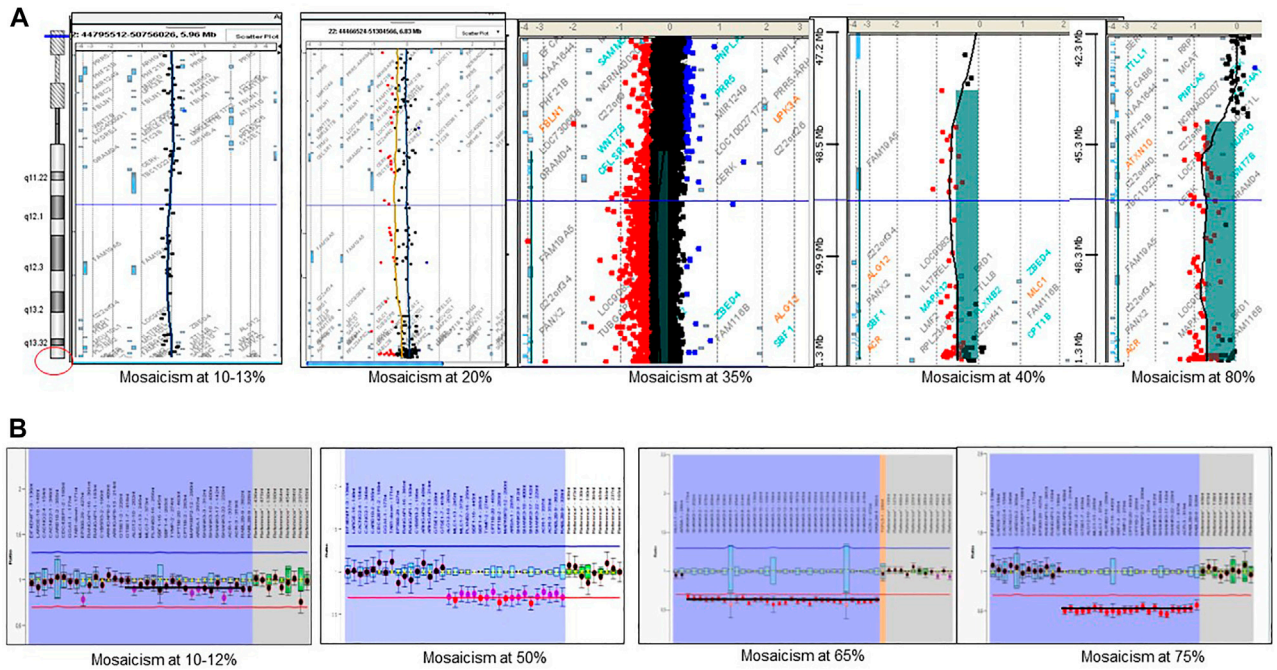


FIGURE 5 | Detection of post-zygotic mosaicism in PMS by using microarrays and MLPA. **(A)** examples of mosaicism detected by CGH-array; **(B)** examples of mosaicism detected by MLPA.

Using the GFAP, we observed significant differences mainly between patients with large deletions compared with patients with small deletions, interstitial deletions, and sequence variants (**Table 5**). Remarkably, no significant differences were detected between small deletions and individuals with sequence variants in *SHANK3* (**Table 5**).

Pearson statistical analysis was performed to explore correlations between these continuous variables. We observed significant direct correlations between size of the deletion and GFAP (Pearson value = 0.33, p = .0001) as well as inverse correlations between age at diagnosis and size of the deletions (Pearson value = −0.240, p = .001) and GFAP (Pearson value = −0.133, p = .03). Altogether, our data suggest that the age at diagnosis seems to be inversely related to the degree of difficulty at diagnosis. Indeed, patients with small deletions (below 0.25 Mb; mean 0.10 ± 0.05 Mb) were diagnosed later (mean 7.61 ± 4.47 years) than those with large-size deletions (˃0.25 Mb, 4.35 ± 2.62 Mb, mean age at diagnosis: 5.52 ± 7.87 years). This fact was also observed in patients with interstitial deletions (mean age at diagnosis 9.75 ± 8.07 years and 1.91 ± 3.51 Mb for deletion size) and *SHANK3* gene variants (mean age at diagnosis 7.86 ± 4.49 years).

Individuals with r(22), mosaic deletions, and unbalanced translocations affecting the 22q13 band were diagnosed significantly earlier than the average (mean ages 5.59, 4.41, and 3.57 years, respectively) even though the mean deletion size in those cases was 3–4 Mb (3.19, 3.24, and 3.91 Mb, respectively) similar to the average of the cohort (median 3.08 Mb).

Although individuals with small deletions and *SHANK3* variants showed similar findings in most of the categorical variables (**Table 5**), a remarkable difference was observed in "the ability to make sentences" between the two groups, with 30/65 (46.2%, **Supplementary Table S3**) of individuals with deletions below 0.25 Mb able to make sentences compared with 5/18 (27.7%, **Table 1a**) among those with *SHANK3* variants. Interestingly, we also found significant differences in the variable "parental origin" between groups with additional rearrangements (at chromosome 22 vs. other chromosomes). As expected, significant differences were found between all deletions and individuals with *SHANK3* variants, mostly affecting dysmorphic features (**Table 5**).

No statistically significant differences were detected between gender and continuous variables (size of the deletion, age of diagnosis, age of evaluation or GFAP, Student's t-test, data not shown). However, significant differences were observed between gender and several categorical variables (seizures, decreased perspiration, microcephaly, fifth finger clinodactyly, and lymphedema; chi-square test, p = 0.023, 0.056, 0.008, 0.029 and 0.001, respectively; data not shown), with higher frequencies in females.

Finally, we observed significant differences between parental origin and GFAP (p = 0.048, Student's t-test) and two categorical variables, high pain threshold and lymphedema (chi-square test, p = 0.039 and 0.027, respectively, n = 51). In all cases, maternal origin (n = 12) was associated with higher GFAP values and with a worse prognosis (**Table 5**).

Genotype-Phenotype Correlations

We applied Ward's hierarchical cluster analysis using deletion size as the unique variable to test how individuals with microdeletions group according to their deletion size. Individuals were grouped into four clusters (the number was

TABLE 4 | *SHANK3* sequence variants identified in this study.

Case	Exon/ total exons	Genomic change NC_000022.1(GCRh37/hg19)	Nucleotide change NM_001372044.2	Amino acid change	Effect	ACMG/AMP classification; others
PMS209	20/25	g.51144533dupC	c.2249dupC	p.Leu751ThrfsTer11	frameshift	P (PVS1, PS2, PM2,PP3, PP4)
PMS187[o]	ivs22/ ivs24	g.51153476G>A	c.2451+1G>A[a]	?	splice site	P (PVS1, PS2, PM2, PP3,PP5); ClinVar (P, LP)
PMS207	24/25	g.51158717delC	c.2643delC	p.Ala882ArgfsTer73	frameshift	P (PVS1, PS2, PM2, PP4)
PMS124	24/25	g.51159024delG	c.2949delG	p.Pro984ArgfsTer34	frameshift	P (PVS1, PS2, PP4)
PMS213	24/25	g.51159481_51159497delGTGTCTGCCCTGAAGCC	c.3408_3424del	pSer1137GlyfsTer215	frameshift	P (PVS1, PS2, PM2,PP3)
PMS146[o]	24/25	g.51159685_51159686delCT	c.3610_3611delCT[b,c,d,e]	p.Leu1204ValfsTer153	frameshift	P (PVS1, PS2, PM2, PP3, PP5) ClinVar (P, LP)
PMS180[o]	24/25	g.51159685_51159686delCT	c.3610_3611delCT[b,c,d,e]	p.Leu1204ValfsTer153	frameshift	P (PVS1, PS2, PM2, PP3, PP5) ClinVar (P, LP)
PMS208[o]	24/25	g.51159685_51159686delCT	c.3610_3611delCT[b,c,d,e]	p.Leu1204ValfsTer153	frameshift	P (PVS1, PS2, PM2, PP3, PP5) ClinVar (P, LP)
PMS181[m,o]	24/25	g.51159685_51159686delCT	c.3610_3611delCT[b,c,d,e]	p.Leu1204ValfsTer153	frameshift	P (PVS1, PS2, PM2, PP3, PP5) ClinVar (P, LP)
PMS182[m,o]	24/25	g.51159685_51159686delCT	c.3610_3611delCT[b,c,d,e]	p.Leu1204ValfsTer153	frameshift	P (PVS1, PS2, PM2, PP3, PP5) ClinVar (P, LP)
PMS175	24/25	g.51159748C>T	c.3673C>T[n]	p.Pro1225Ser	missense	VUS-LP? (PS2, PM2)
PMS211	24/25	g.51159787delG	c.3712delG	p.Glu1238Argfster19	frameshift	P (PVS1, PS2, PM2, PP3)
PMS185[o]	24/25	g.51159940dupG	c.3865dupG[c,d,f,g,h,i,j,k]	p.Ala1289GlyfsTer69	frameshift	P (PVS1, PS2, PM2,PP3, PP5); ClinVar (P)
PMS212[o]	24/25	g.51159940dupG	c.3865dupG[c,d,f,g,h,i,j,k]	p.Ala1289GlyfsTer69	frameshift	P (PVS1, PS2, PM2,PP3, PP5); ClinVar (P)
PMS165	24/25	g.51160025_51160037del GGGCCCAGCCCCC	c.3950_3962del	p.Arg1317LeufsTer25	frameshift	P (PVS1, PS2, PM2, PP3, PP5); ClinVar (P) CClinPP5)
PMS198[o]	24/25	g.51160025dupG	c.3952dupG	p.Ala1318GlyfsTer40	frameshift	P (PVS1, PS2, PM2, PP3; PP5); ClinVar (P)
PMS214[o]	24/25	g.51160025dupG	c.3952dupG	p.Ala1318GlyfsTer40	frameshift	P (PVS1, PS2, PM2, PP3; PP5); ClinVar (P)
PMS137	24/25	g.51160235dupG	c.4160dupG	p.Ser1391LeufsTer16	frameshift	LP (PVS1, PS2, PM2)
PMS177	24/25	g.51160291_51160312delGAGCCACCCCCTGCCCCTGAGT	c.4216-4237del	p.Glu1406LeufsTer35	frameshift	P (PVS1, PS2, PM2, PP3)
PMS201	24/25	g.51160349G>A	c.4274G>A	p.Arg1425His	missense	VUS-LP (PS2, PM2, PP3)
PMS145[o]	24/25	g.51160594C>T	c.4519C>T[l]	p.Gln1507Ter	nonsense	P (PVS1,PS2, PM2, PP3)

[a]Bramswig et al. (2015), Holder and Quach (2016), Yuen et al. (2017); [b]Leblond et al. (2014); [c]De Rubeis et al. (2018); [d]Zhou et al. (2019); [e]Kaplanis et al. (2020); [f]Feliciano et al. (2019); [g]Lelieveld et al. (2016); [h]Retterer et al. (2016); [i]O'Roak et al. (2014); [j]Farwell et al. (2015); [k]Durand et al. (2007); [l]Thevenon et al. (2016); [m]Individuals PMS181 and PMS182 are siblings; [n]The variant c.3673C>T(p.Pro1225Ser) has been previously described in two individuals of African descent (gnomAD v2.1.1.), a fact that may question its association with the clinical features observed in the patient; [o]Variants described previously in unrelated individuals or recurrent in our cohort. P, pathogenic; LP, likely pathogenic; VUS, variant of uncertain significance.

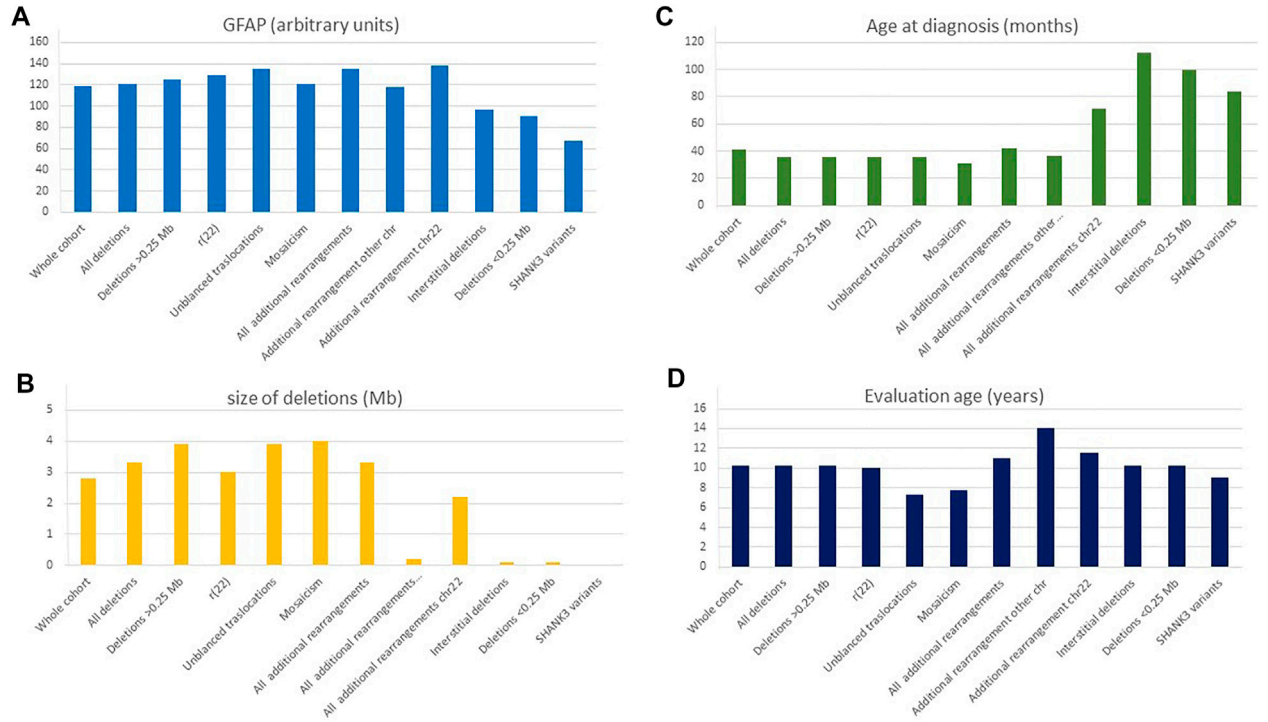


FIGURE 6 | Distribution of continuous variables according to the type of genetic defect. **(A)** GFAP, global function assessment of the patient (arbitrary values); **(B)** Size of the deletions (Mb); **(C)** Age at diagnosis (months); **(D)** Age at evaluation (years). *ASD diagnosis according to the psychiatrists of the referring institutions.

established by BIC and AIC algorithms) as follows: cluster 1: 0.52 ± 0.51 Mb (64 individuals), cluster 2: 3.39 ± 0.77 Mb (66 individuals), cluster 3: 6.10 ± 0.69 Mb (29 individuals), and cluster 4: 8.27 ± 0.74 Mb (28 individuals). Extended variable frequencies in each cluster are shown in **Supplementary Table S3**. One-way ANOVA followed by a *post hoc* test (Bonferroni or T3-Dunnett) revealed statistically significant differences between age at diagnosis, GFAP, and size of deletions in different clusters (p = 0.009, 0.0001, and 0.0001, respectively, **Table 6**). **Supplementary Figure S6** shows that some clinical findings, such as "ability to make sentences" or "walk independently before/after 15 months," were preferentially associated with cluster 1. In fact, in cluster 1 (deletions 0.52 ± 0.51 Mb), 53.8% of these individuals were able to make sentences (35/65), followed by 15.6% (10/64) in cluster 2 and only 3.7% (1/27) in clusters 3 and 4. The chi-square test followed by z *post hoc* test with Bonferroni correction showed significant differences among clusters for several categorical variables (**Table 6**).

When Ward's clusters were dissected by frequencies of these variables (in percentages), we observed higher frequencies of several core features, considered as a better prognosis, in cluster 1 than in other clusters (**Supplementary Table S3**). On the other hand, higher percentages of other core items, reflecting comorbidity (normally associated with a worse prognosis; renal and urogenital abnormalities, hearing problems, lymphedema, no words, or growth above the 95th percentile, **Supplementary Table S3**) mapped preferentially in cluster 4, which is associated with large deletions. Finally, other items seemed to correlate directly (toe syndactyly, ear anomalies, GFAP, MRI anomalies, abnormal emotional response, or renal and urogenital anomalies) or inversely (age at diagnosis) to the size of the deletions (**Supplementary Table S3**).

Linear regression was used to obtain a coefficient of correlation to deletion size at 22q13 for each feature (**Table 7**). The coefficient of correlation ranged between 0 and 0.7. "F value" was examined to determine if the coefficient of correlation was significant. For most features, no correlation to deletion size was found. However, several clinical features were found to have a statistically significant correlation with the size of the deletion (**Table 7**), including the ability to make sentences, lymphedema, macrocephaly, renal and urogenital anomalies, and brain MRI anomalies. At a significance level of 0.05, one would expect 1 in 20 significant correlations by chance, whereas 14/61 (23%) correlations for the size of deletion were obtained. With a similar approach, we identified 6/61 (9.8%) correlations with age at diagnosis and 8/61 (13.1%) with age at evaluation.

DISCUSSION

We describe one of the largest series of patients with PMS characterized by CMA and other genetic approaches, including karyotype, MLPA, and FISH. We also explored the high genetic, and phenotypic variability observed in PMS individuals. Although the true prevalence of this rare disease is still unknown, it is among the most common subtelomeric

TABLE 5 | Comparison between groups with different types of genetic alterations.

Variable	*p*-value/F-value	Statistical test	Pairs of groups with test significant differences after *post hoc* test
Size of deletion[a]	0.0001/40.46	ANOVA	(1.2) (1.3) (2.5) (2.6)
Age at evaluation	0.556/0.59	ANOVA	none
Age at diagnosis	0.008/4.03	ANOVA	(1.2)
GFAP	0.0001/11.24	ANOVA	(1.2) (1.3) (1.4) (1.6) (2.6) (4.5) (4.6) (4.7) (4.8) (4.9) (4.10)
Walk independently before/after 15 months	0.0001/27.51	Chi square	(1.2) (1.4) (2.4) (4.8) (4.9) (4.10)
Single words	0.005/12.89	Chi square	(1.2)
Ability to make sentences	0.0001/27.11	Chi square	(1.2) (1.3) (2.4) (2.6)
Full brow	0.010/11.34	Chi square	(1.4)
Dental anomalies	0.0044/8.09	Chi square	(1.4) (2.4) (4.9) (4.10)
Deep set eyes	0.037/11.84	Chi square	(1.4) (4.6) (4.10)
Toe syndactyly	0.001/17.15	Chi square	(1.2) (1.4)
Large fleshy hands	0.001/17.93	Chi square	(1.2) (1.3)
Sphincter control	0.0001/17.93	Chi square	(1.4) (2.4) (4.8) (4.9) (4.10) (7.8) (7.9)
Very sensitive to touch	0.0001/11.52	Chi square	(1.4) (2.4) (3.4) (4.6) (4.9)
Parental origin	0.021/11.60	Chi square	(8.9)
Recurrent infections	0.016/12.21	Chi square	(1.8)
Hair pulling	0.013/12.74	Chi square	(1.7) (7.9)
Gastrointestinal anomalies	0.002/17.22	Chi square	(1.7) (1.8) (1.10)
Sleeping problems	0.020/11.69	Chi square	(1.4) (1.10) (4.10)
Epicanthus	0.046^{FET}/5.14	Chi square	(1.4)
Full/puffy eyelids	0.026/7.82	Chi square	(1.4)
Poor visual contact	0.043^{FET}/4.22	Chi square	(1.4)
Formal ASD evaluation	0.040^{FET}/4.22	Chi square	(1.4)
Abnormal emotional response	0.047^{FET}/3.67	Chi square	(1.4)
Growth, centile >95th	0.054^{FET}/3.29	Chi square	(1.4)
Hypotonia	0.059^{FET}/3.55	Chi square	(1.4)

*Group 1 (deletions >0.25 Mb, mean size ± SD, 4.29 ± 2.50); group 2 (smaller deletions ≤0.25 Mb, 0.10 ± 0.05); group 3 (interstitial deletions, 1.94 ± 3.55 Mb); group 4 (*SHANK3 *variants); group 5 (ring 22, 3.53 ± 2.44 Mb); group 6 (unbalanced translocations, 3.69 ± 1.61 Mb); group 7 (mosaic deletions, 3.5 ± 3.48 Mb); group 8 (additional rearrangement at chromosome 22, 3.32 ± 2.02 Mb); group 9 (additional rearrangement in other chromosomes, 2.62 ± 2.26 Mb), and group 10 (all cases with additional rearrangements, 2.99 ± 2.26 Mb). FET, corrected by Fisher's exact test; GFAP, global functional assessment of the patient.*

[a]*Group 4 (*SHANK3 *variants) was not included in the analysis of deletion size.*

TABLE 6 | Comparison between Ward's clusters obtained using deletion size.

Variable	*p*-value/F-value	Statistical test	Pairs of clusters with significant differences after *post hoc* test
Size of deletion	0.001/7.509	ANOVA	(1.2) (1.3) (1.4) (2.3) (2.4)
Age at diagnosis	0.009/3.861	ANOVA	(1.2)
GFAP	0.001/7.509	ANOVA	(1.2) (1.3) (2.3) (2.4)
Age at evaluation	0.086/1.951	ANOVA	none
Walk independently before/after 15 months	0.0001/18.996	Chi square	(1.2) (1.4)
Growth, percentile >95th	0.020/9.867	Chi square	(2.4)
Ability to make sentences	0.0001/27.996	Chi square	(1.2) (1.3) (1.4)
Some words	0.0001/17.906	Chi square	(1.2) (1.3)
Hypotonia	0.003/13.726	Chi square	(1.3)
Microcephaly	0.012/10.897	Chi square	(1.3)
Macrocephaly	0.004/13.512	Chi square	(1.4) (2.4)
Sphincter control	0.009/11.604	Chi square	(1.3)
Renal and urogenital anomalies	0.009/11.504	Chi square	(1.4)
Lymphedema	0.0001/26.883	Chi square	(1.4) (2.4)
Ear anomalies	0.009/11.504	Chi square	(1.4)
Biting	0.037/8.494	Chi square	(1.2)
Nonstop crying	0.044/8.116	Chi square	(3.4)

Mean deletion size cluster 1 (0.52 ± 0.51 Mb), cluster 2 (3.39 ± 0.77 Mb), cluster 3 (6.10 ± 0.69 Mb), and cluster 4 (8.27 ± 0.74 Mb). GFAP, global functional assessment of the patients.

TABLE 7 | Comparison of clinical features and the size of the 22q13 deletion, age at diagnosis and age at evaluation using linear regression to obtain a coefficient of correlation.

Clinical feature	Coefficient of correlation	Significance F
Dependent variable: size of deletion		
Ability to make sentences	0.37	0.0001
Lymphedema	0.49	0.0001
Macrocephaly	0.53	0.002
Renal and urogenital anomalies	0.55	0.010
Seizures	0.57	0.014
Other genomic rearrangements	0.59	0.021
Sphincter control	0.61	0.011
Abnormal brain MRI	0.63	0.013
Deep set eyes	0.65	0.011
Growth, percentile >95th	0.66	0.037
Herniae	0.67	0.024
Abnormal emotional response	0.68	0.037
Toe syndactyly	0.69	0.036
Epicanthal folds	0.70	0.046
Dependent variable: age at diagnosis		
Sphincter control	0.23	0.003
Biting	0.29	0.017
Seizures	0.33	0.024
Dolichocephaly	0.37	0.020
Lip/palate anomalies	0.41	0.026
Nonverbal	0.43	0.046
Dependent variable: age at evaluation		
Brain MRI	0.27	0.0001
Sphincter control	0.35	0.002
ASD diagnosis[a]	0.39	0.015
Dolichocephaly	0.43	0.012
Ability to make sentences	0.46	0.010
Seizures	0.50	0.004
Obesity	0.52	0.025
Poor visual contact	0.54	0.029

[a]*ASD diagnosis according to the psychiatrists of the referring institutions.*

microdeletion syndromes (Delahaye et al., 2009). Previous findings show that PMS is diagnosed in around 0.5% of individuals with ASD and ID (Cooper et al., 2011; Betancur and Buxbaum, 2013; Leblond et al., 2014; Chen et al., 2017; Samogy-Costa et al., 2019). Previous data suggest that the prevalence of this syndrome remains underestimated worldwide due to several reasons:

a) The lack of a distinctive phenotype without significant dysmorphic features (**Figure 1**). In most cases, individuals carrying *SHANK3* variants and small deletions do not have a distinctive facial appearance.
b) High genetic and clinical variability. We observed marked intracohort variability. Analysis of GFAP revealed significant differences depending on the type of genetic defect and the type of rearrangements found in individuals. We found additional rearrangements in 21.2% of the cases. Some of them involved other OMIM-related syndromes (**Supplementary Table S2**), including hereditary neuropathy with liability to pressure palsies (OMIM#162500), affecting *PMP22*; Chromosome 15q11.2 deletion syndrome BP1-BP2 (OMIM#615656), affecting *NIPA1-NIPA2*; 15q13.3 deletion syndrome (OMIM#2612001), affecting *CHRNA7*, and 16p11.2 microdeletion syndrome (OMIM#611913), which may contribute partially to the variability of some individuals. Previous studies also report the presence of additional rearrangements with putative clinical relevance in individuals with PMS (Tabet et al., 2017; Samogy-Costa et al., 2019). Interestingly, our data show that individuals with additional rearrangements and, in particular, those with small 22q13 deletions had higher values of GFAP (associated with worse prognosis) than cases with simple small deletions. In our series, some of the patients carried the same additional CNVs reported by Tabet and others (2017), in most cases inherited from a reportedly healthy parent. We do not know the consequences of these findings or if it is just a coincidence. Most of these and other similar CNVs (15q11.2 deletions and duplications, 15q13.3 deletions and duplications, 16p13.11 deletions, 16p12.1 deletions, 16p11.2 proximal and distal deletions, 17q12 deletions and duplications, and 22q11.21 duplications) are linked to susceptibility loci for a variety of pediatric diseases (Girirajan and Eichler, 2010; Cooper et al., 2011). For some of these CNVs, the enrichment in affected individuals (mainly ID, ASD, or DD cases) in comparison with healthy controls seems to give them a putative pathogenic classification (Rosenfeld et al., 2013).

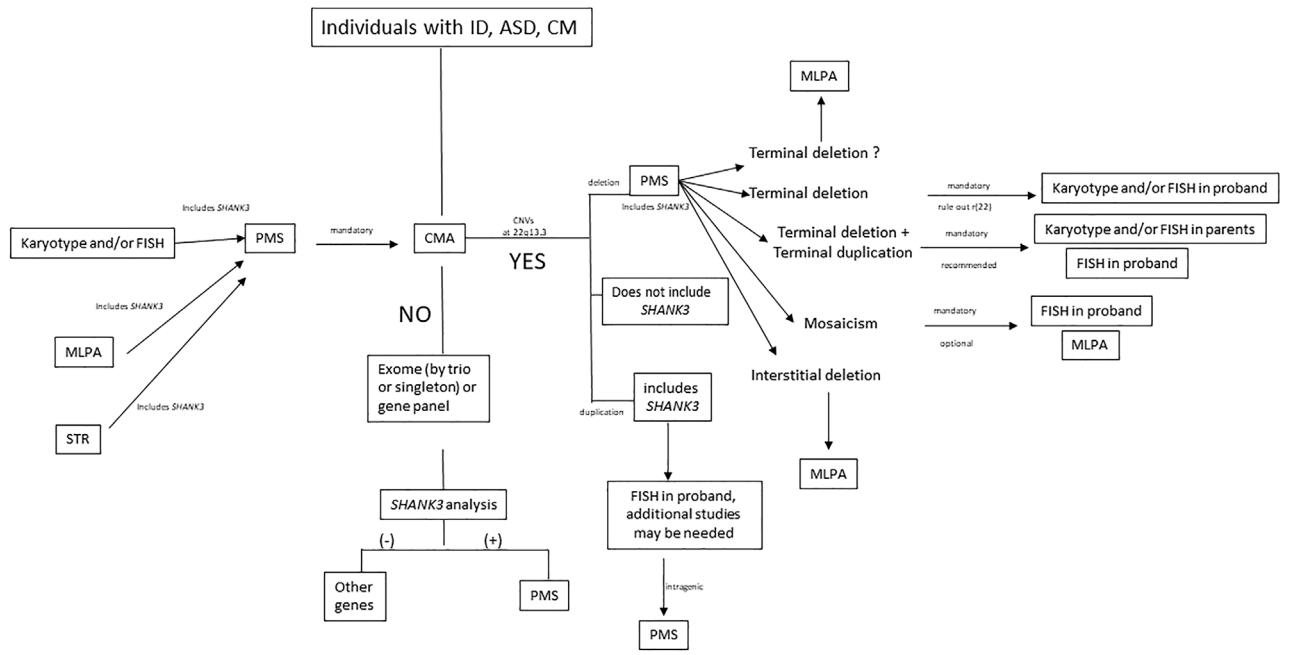


FIGURE 7 | Laboratory algorithm for management of samples suspected of PMS. ID, intellectual disability; ASD, autism spectrum disorder; CM, congenital malformations; PMS, Phelan-McDermid syndrome; CMA, chromosome microarray analysis; MLPA, multiplex ligation-dependent probe amplification; STR, short tandem repeat; CNV, copy number variation; r(22), ring chromosome 22.

c) The difficulty in detecting chromosome 22 microdeletions in routine cytogenetic analysis even at the 550–850 band level of resolution. Our data show that small terminal deletions, interstitial deletions, and *SHANK3* variants were diagnosed later than those carrying other type of rearrangements, such as ring chromosomes, mosaic deletions, or unbalanced translocations. Thus, most cases were diagnosed in tertiary hospitals that applied CMA testing as a first-tier test through its laboratory routines for individuals with ID, ASD, and congenital malformations, following international guidelines (Miller et al., 2010). Misdiagnosis or underdiagnosis of mosaicism could be observed when using CMA as a unique tool. Mosaicism lower than 15% cannot be easily detected by CMA (**Figure 3**) owing to the variability of the assay and the fact that most of the commercial CMA platforms do not have a significant number of probes at the end of the telomere of chromosome 22 (**Supplementary Figure S5**). FISH or MLPA combined with CMA must be applied in suspected patients. We found an unexpectedly high number of post-zygotic mosaicism (17/189; 9.0%) in patients with microdeletions when compared with a previous report, which established a mosaic frequency of around 2.5%–5.8% for deletions at 22q13.3 (Samogy-Costa et al., 2019). It is not easy to predict the expected clinical features in patients with mosaicism though patients with <10% of mosaicism in blood can present a complete manifestation of the disease (Phelan et al., 2018). We also found two independent families with suspected gonadal mosaicism. This aspect is important because it complicates genetic counseling. Germinal mosaicism in PMS is not frequent, but it has been described in a few families (Tabolacci et al., 2005; Durand et al., 2007; Gauthier et al., 2009; Nemirovsky et al., 2015; Zwanenburg et al., 2016).

In PMS individuals with terminal deletions diagnosed with CMA, it is essential to rule out the presence of r(22). Confirmation of r(22) has significant implications for clinical management because individuals with r(22) have an increased risk of tumors in the nervous system due to biallelic loss of the *NF2* (neurofibromatosis type 2) gene (Lyons-Warren et al., 2017; Ziats et al., 2020). We observed three out of 20 patients with r(22) with neurofibromatosis type 2; these three individuals were included in the series reported by Zyats and others (2020). The prevalence of tumors associated with r(22) is unknown. Thus, we recommend follow-up of PMS patients carrying r(22) and highlight the importance of karyotyping individuals with terminal deletions of the long arm of chromosome 22.

d) Difficulties in testing *SHANK3* variants. Implementing exome or panels to analyze *SHANK3* variants was rare and expensive during the period of recruitment of this cohort in our country. However, in recent years (2019–2020), we recruited 18 patients with *SHANK3* variants.

We propose an algorithm for laboratory management of individuals with PMS (**Figure 7**). We recommend CMA as a first-tier test for patients with ID and ASD to determine the exact deletion size, define the deletion breakpoints, and detect

additional genomic rearrangements, such as terminal duplications in other chromosomes. Most patients also need other molecular approaches, such as MLPA or FISH, for accurate laboratory characterization (**Supplementary Figure S5**). Terminal deletions need karyotyping to rule out a r(22), and FISH is mandatory in parents when suspicion of unbalanced or balanced translocation is suspected. Low-grade mosaicism may be detected by applying FISH in the proband. When other techniques, such as FISH or MLPA, established the diagnosis of PMS as the first test (**Figure 7**), CMA is still mandatory to complete the diagnosis of individuals (to determine the affected genes, deletion size, other rearrangements, etc.). Finally, when all cytogenetic and molecular approaches are negative in individuals with ID or ASD with other clinical features of PMS, we recommend an exome-analysis (trio or singleton) with extensive analysis of *SHANK3* sequence variants (**Figure 7**).

It is also remarkable that, although formal ASD studies were only performed in 20% (43/210) of the cohort, 29/43 (67%) of them have an ASD diagnosis according to the psychiatrists of the referring institutions. Thus, for PMS individuals, formal ASD evaluation is mandatory. Sixty individuals of this cohort are included in a recent study of the behavioral profile in PMS performed by our colleagues (Burdeus-Olavarrieta et al., 2021).

Genotype-Phenotype Correlations

It is suggested that the haploinsufficiency of *SHANK3* is the most significant contributor to PMS. We believe that *SHANK3* is a major contributor to the neurocognitive features of the syndrome, but not the only one. Other genes may contribute to the PMS phenotype by modulating *SHANK3* action. Several authors review a possible effect of different genes in the PMS phenotype (Tabet et al., 2017; Mitz et al., 2018; Ziats et al., 2019; Li et al., 2020; Ricciardello et al., 2021), but how those genes contribute is still unknown.

Only a few studies investigate putative relations between the size of the deletions and clinical features of PMS, and the causality remains unclear (Cusmano-Ozog et al., 2007; Dhar et al., 2010; Sarasua et al., 2011; Soorya et al., 2013; Sarasua et al., 2014a; Sarasua et al., 2014b; Tabet et al., 2017; Samogy-Costa et al., 2019). The clinical features of patients with pathogenic variants in *SHANK3* overlap with those of individuals with deletions, giving this gene an important role in the spectrum of clinical features of PMS.

We found that speech skills, one of the main features of the syndrome, might be directly associated with the size and/or mapping of the deletion. Indeed, most individuals who can make sentences (aged older than 3 years) had smaller deletions, supporting previously described observations (Sarasua et al., 2014a; Samogy-Costa et al., 2019; Brignell et al., 2021). In addition, among individuals with *SHANK3* variants, 27% (5/18) of patients in this study were able to maintain short conversations, compared with 18% (3/17) and 38% (3/8) of individuals verbally fluent reported by De Rubeis et al. (2018) and Xu et al. (2020), respectively.

Our data also support significant differences between individuals with *SHANK3* variants and small deletions in the ability to make sentences. Thus, other genes or some interaction nearby could modulate language abilities. In fact, a recent study also showed that *SHANK3* seemed necessary but not exclusive for expressive language in PMS individuals (Brignell et al., 2021).

Additional differences between individuals with *SHANK3* variants and those with small deletions were also observed for cognitive features, such as sleeping anomalies or sphincter control, with higher frequencies in individuals with *SHANK3* variants than in the smaller deletion group. As expected, differences in several facial dysmorphic features were observed between individuals with deletions and *SHANK3* variants.

The cluster analysis showed a positive correlation between deletion size and GFAP, brain MRI abnormalities, ear anomalies, and toe syndactyly as well as a negative correlation between deletion size and age at diagnosis and abnormal emotional response. It is also clear that several clinical features mapped preferentially in specific regions of the clusters. Indeed, two clear genomic regions can be associated with the size of the cranium. Whereas medium- and large-size deletions seem to be associated with macrocephaly, microcephaly seems to be present only in patients with small deletions. We established an interval between 0.40 and 3.4 Mb linked to microcephaly and between 4.50 and 8 Mb from the telomere related to macrocephaly. This fact suggests the contribution of at least two independent genes for alterations in the cranium size. Interestingly, there are no more than 10 high dosage-sensitive genes (ClinGen, http://www.clinicalgenome.org) in the latter interval. Among them is *GRAMD4*, which has been established experimentally to have protein-protein interaction with PIAS1 (**Supplementary Figure S7**). PIAS1 is a member of the ubiquitin protein family, like PIAS4. The *PIAS4* gene has been involved in macro/microcephaly in distal 19p13.3 microdeletion/microduplication syndrome (Nevado et al., 2015; Tenorio et al., 2020).

The existence of interstitial deletions not including *SHANK3* (Wilson et al., 2008; Disciglio et al., 2014; Ha et al., 2017; this study), which partly overlap some clinical features of PMS (**Supplementary Table S4**), may also indirectly support a role for additional genes in the clinical spectrum of PMS. At this point, we cannot rule out a positional/regulating effect on *SHANK3* in all these cases, nor global alteration of topological chromatin organization (TAD; topological association domains) as is been suggested by others (Kurtas et al., 2018; Srikanth et al., 2021) rather than simply by the deletion of dosage-sensitive genes. This hypothesis needs to be explored in future studies.

Correlations by Age

A previous large cohort study reported a small but significant increase with age of several clinical findings in PMS, including sensory dysfunction, reduced response to pain, epilepsy, and lymphedema (Sarasua et al., 2014b). Similarly, the risk of psychiatric disorders in PMS increases with age (Denayer et al., 2012; Verhoeven et al., 2012; Kolevzon et al., 2019). Regarding the correlation of clinical features with age, our data cannot support any solid conclusion about the contribution of age to the clinical features of PMS. We found in our cohort six and eight items of 61 that rejected this null

hypothesis (~10% and 13%) for age at diagnosis and age at evaluation, respectively. This is twice the number expected by chance.

CONCLUSIONS

Here, we report a large series of Spanish and South American patients with PMS, focusing on phenotype-genotype correlations. The analysis of individuals with sequence variants and their comparison with patients with small deletions support the notion that *SHANK3* is essential in most core phenotypic findings of PMS but is not the unique one. Additional genes may modulate the whole phenotype in PMS individuals with microdeletions.

The existence of different types of rearrangements and genomic variations may explain the high variability observed in PMS individuals. Finally, an accurate laboratory approach for PMS individuals using a diagnostic algorithm is proposed to offer appropriate management, follow-up, and genetic counselling to these families.

SPANISH PMS WORKING GROUP

INGEMM, Madrid, Spain (Rocío Mena, Roser Lleguer, Victoria Fernández-Montaño, Rubén Martín, Blanca Fernández, Fé García-Santiago, Victoria Gómez del Pozo, Carolina Peña; Spanish PMS Association (Norma Alhambra, Carlos García, Juan Ramón Rodríguez); Servicio de Neuropediatría, Hospital Universitario La Paz, Madrid, Spain (Antonio Martínez-Bermejo); Hospital Central de Asturias, Oviedo, Spain (Ignacio Málaga); Hospital San Joan de Deu, Barcelona, Spain (Antonio Federico Martínez-Monseny, Judith Armstrong, Jennifer Anticona, Cristina Hernando-Davalillo, Adrián Alcalá San Martí, Loreto Martorell, Delia Yubero, Tania Nunes, Mar O´Callaghan, Carlos Ortez, Xenia Alonso, Federico Ramos, Jesús Casas López); Hospital Virgen de la Arrixaca, Murcia, Spain (Vanesa López-González, M. Juliana Ballesta); Q- Genomics Laboratory Barcelona, Spain (Lluís Armengol); Hospital Virgen del Rocío, Sevilla, Spain (Antonio González-Meneses; Salud Borrego); Hospital Universitario la Fé, Valencia, Spain (Mónica Roselló); NIM-Genetics, Madrid, Spain (Javier Suela); Hospital Son Espases, Palma de Mallorca, Spain (Ángeles Pérez-Granero); Hospital Clinic, Barcelona, Spain (Laia Rodríguez-Revenga).

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: DECIPHER Genomics, accession no: 432868 - 433079, IMMGPMS1 - IMMGPMS211.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board of Hospital la Paz, Madrid. Spain PI: 2735 HULP. 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

JN: Conceived the presented idea, completed data analysis and wrote the manuscript; JN and PL: Designed the study; JN Coordinated data acquisition and collected the data; EV, MP-B, JT, MM and PL: corrected the manuscript; PB assisted with data management; CB statistical analysis; EV, MP-B, JN, PB, SM and MM: managed miocroarrays at INGEMM; EM and IV: performed FISH and karyotyping studies at INGEMM; SG-M, EG-N, JR, MM, MC, FS-S, GO, CO, HP, EG, JC, AM, CS, SC, EG, LP-J, AB, PG-M, FC-S, NG-P, RB-L, MS, JO-E, BG, VS, PT and the Spanish PMS working group: contributed patients from their institutions.

FUNDING

REDES/FIBHULP08. FIBHULP PI: 2735. FIBHULP Auchan Reserch Project. FIBHULP. Raregenomics (B2017/BMD-3721).

ACKNOWLEDGMENTS

We would like to thank all individuals and families for their participation in this study, as well as Norma Alhambra, Carlos García, and Juan Ramón Rodríguez (Spanish PMS Association) for helping us to develop this project. We thank Rocío Mena, who coordinates the PMS working group. In memoriam of Sol Martin for her valuable support work in the laboratory. This work was supported by a grant from REDES/FIBHULP08 of the Fundación para la Investigación Biomédica Hospital Universitario La Paz (FIBHULP); PI: 2735 from FIBHULP, and AUCHAN Research Project on PMS.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.652454/full#supplementary-material

REFERENCES

Anderlid, B.-M., Schoumans, J., Annerén, G., Sahlén, S., Kyllerman, M., Vujic, M., et al. (2002a). Subtelomeric Rearrangements Detected in Patients with Idiopathic Mental Retardation. *Am. J. Med. Genet.* 107 (4), 275–284. doi:10.1002/ajmg.10029

Anderlid, B.-M., Schoumans, J., Annerén, G., Tapia-Paez, I., Dumanski, J., Blennow, E., et al. (2002b). FISH-mapping of a 100-kb Terminal 22q13 Deletion. *Hum. Genet.* 110 (5), 439–443. doi:10.1007/s00439-002-0713-7

Betancur, C., and Buxbaum, J. D. (2013). SHANK3 Haploinsufficiency: a "common" but Underdiagnosed Highly Penetrant Monogenic Cause of Autism Spectrum Disorders. *Mol. Autism* 4 (1), 17. doi:10.1186/2040-2392-4-17

Boccuto, L., Lauri, M., Sarasua, S. M., Skinner, C. D., Buccella, D., Dwivedi, A., et al. (2013). Prevalence of *SHANK3* Variants in Patients with Different Subtypes of Autism Spectrum Disorders. *Eur. J. Hum. Genet.* 21, 310–316. doi:10.1038/ejhg.2012.175

Boeckers, T. M. (2006). The Postsynaptic Density. *Cell Tissue Res* 326 (2), 409–422. doi:10.1007/s00441-006-0274-5

Bonaglia, M. C., Giorda, R., Beri, S., De Agostini, C., Novara, F., Fichera, M., et al. (2011). Molecular Mechanisms Generating and Stabilizing Terminal 22q13 Deletions in 44 Subjects with Phelan/McDermid Syndrome. *Plos Genet.* 7, e1002173. doi:10.1371/journal.pgen.1002173

Bonaglia, M. C., Giorda, R., Borgatti, R., Felisari, G., Gagliardi, C., Selicorni, A., et al. (2001b). Disruption of the ProSAP2 Gene in a t(12;22)(q24.1;q13.3) Is Associated with the 22q13.3 Deletion Syndrome. *Am. J. Hum. Genet.* 69 (2), 261–268. doi:10.1086/321293

Bonaglia, M. C., Giorda, R., Mani, E., Aceti, G., Anderld, B. M., Baroncini, A., et al. (2001a). Identification of a Recurrent Breakpoint within the SHANK3 Gene in the 22q13.3 Deletion Syndrome. *Am. J. Hum. Genet.* 69 (2), 261–268.

Bowling, K. M., Thompson, M. L., Amaral, M. D., Finnila, C. R., Hiatt, S. M., Engel, K. L., et al. (2017). Genomic Diagnosis for Children with Intellectual Disability And/or Developmental Delay. *Genome Med.* 9, 43. doi:10.1186/s13073-017-0433-1

Bramswig, N. C., Lüdecke, H.-J., Alanay, Y., Albrecht, B., Barthelmie, A., Boduroglu, K., et al. (2015). Exome Sequencing Unravels Unexpected Differential Diagnoses in Individuals with the Tentative Diagnosis of Coffin-Siris and Nicolaides-Baraitser Syndromes. *Hum. Genet.* 134 (6), 553–568. doi:10.1007/s00439-015-1535-8

Brignell, A., Gu, C., Holm, A., Carrigg, B., Sheppard, D. A., Amor, D. J., et al. (2021). Speech and Language Phenotype in Phelan-McDermid (22q13.3) Syndrome. *Eur. J. Hum. Genet.* 29 (4), 564–574. doi:10.1038/s41431-020-00761-1

Burdeus-Olavarrieta, M., San José-Cáceres, A., García-Alcón, A., González-Peñas, J., Hernández-Jusdado, P., and Parellada-Redondo, M. (2021). Characterisation of the Clinical Phenotype in Phelan-McDermid Syndrome. *J. Neurodevelop Disord.* 13, 26–40. doi:10.1186/s11689-021-09370-5

C Yuen, R. K., Merico, D., Bookman, M., L HoweThiruvahindrapuram, J. B., Thiruvahindrapuram, B., Patel, R. V., et al. (2017). Whole Genome Sequencing Resource Identifies 18 New Candidate Genes for Autism Spectrum Disorder. *Nat. Neurosci.* 20, 602–611. doi:10.1038/nn.4524

Chen, C.-H., Chen, H.-I., Liao, H.-M., Chen, Y.-J., Fang, J.-S., Lee, K.-F., et al. (2017). Clinical and Molecular Characterization of Three Genomic Rearrangements at Chromosome 22q13.3 Associated with Autism Spectrum Disorder. *Psychiatr. Genet.* 27 (1), 23–33. doi:10.1097/ypg.0000000000000151

Cooper, G. M., Coe, B. P., Girirajan, S., Rosenfeld, J. A., Vu, T. H., Baker, C., et al. (2011). A Copy Number Variation Morbidity Map of Developmental Delay. *Nat. Genet.* 43 (9), 838–846. doi:10.1038/ng.909

Cusmano-Ozog, K., Manning, M. A., and Hoyme, H. E. (2007). 22q13.3 Deletion Syndrome: A Recognizable Malformation Syndrome Associated with Marked Speech and Language Delay. *Am. J. Med. Genet.* 145C, 393–398. doi:10.1002/ajmg.c.30155

De Rubeis, S., Siper, P. M., Durkin, A., Weissman, J., Muratet, F., Halpern, D., et al. (2018). Delineation of the Genetic and Clinical Spectrum of Phelan-McDermid Syndrome Caused by SHANK3 point Mutations. *Mol. Autism* 9, 31. doi:10.1186/s13229-018-0205-9

Delahaye, A., Toutain, A., Aboura, A., Dupont, C., Tabet, A. C., Benzacken, B., et al. (2009). Chromosome 22q13.3 Deletion Syndrome with a De Novo Interstitial 22q13.3 Cryptic Deletion Disrupting SHANK3. *Eur. J. Med. Genet.* 52 (5), 328–332. doi:10.1016/j.ejmg.2009.05.004

Denayer, A., Van Esch, H., de Ravel, T., Frijns, J. P., Van Buggenhout, G., Vogels, A., et al. (2012). Neuropsychopathology in 7 Patients with the 22q13 Deletion Syndrome: Presence of Bipolar Disorder and Progressive Loss of Skills. *Mol. Syndromol* 3 (1), 14–20. doi:10.1159/000339119

Dhar, S. U., del Gaudio, D., German, J. R., Peters, S. U., Ou, Z., Bader, P. I., et al. (2010). 22q13.3 Deletion Syndrome: Clinical and Molecular Analysis Using Array CGH. *Am. J. Med. Genet.* 152A, 573–581. doi:10.1002/ajmg.a.33253

Disciglio, V., Rizzo, C. L., Mencarelli, M. A., Mucciolo, M., Marozza, A., Di Marco, C., et al. (2014). Interstitial 22q13 Deletions Not Involving SHANK3 Gene: a New Contiguous Gene Syndrome. *Am. J. Med. Genet.* 164 (7), 1666–1676. doi:10.1002/ajmg.a.36513

Durand, C. M., Betancur, C., Boeckers, T. M., Bockmann, J., Chaste, P., Fauchereau, F., et al. (2007). Mutations in the Gene Encoding the Synaptic Scaffolding Protein SHANK3 Are Associated with Autism Spectrum Disorders. *Nat. Genet.* 39 (1), 25–27. doi:10.1038/ng1933

Farwell, K. D., Shahmirzadi, L., El-Khechen, D., Powis, Z., Chao, E. C., Tippin Davis, B., et al. (2015). Enhanced Utility of Family-Centered Diagnostic Exome Sequencing with Inheritance Model-Based Analysis: Results from 500 Unselected Families with Undiagnosed Genetic Conditions. *Genet. Med.* 17 (7), 578–586. doi:10.1038/gim.2014.154

Feliciano, P., Zhou, X., Astrovskaya, I., Turner, T. N., Wang, T., Brueggeman, L., et al. (2019). Exome Sequencing of 457 Autism Families Recruited Online Provides Evidence for Autism Risk Genes. *NPJ Genom Med.* 4, 19. doi:10.1038/s41525-019-0093-8

Gauthier, J., Spiegelman, D., Piton, A., Lafrenière, R. G., Laurent, S., St-Onge, J., et al. (2009). Novel De Novo SHANK3 Mutation in Autistic Patients. *Am. J. Med. Genet.* 150B, 421–424. doi:10.1002/ajmg.b.30822

Girirajan, S., and Eichler, E. E. (2010). Phenotypic Variability and Genetic Susceptibility to Genomic Disorders. *Hum. Mol. Genet.* 19 (R2), R176–R187. doi:10.1093/hmg/ddq366

Ha, J. F., Ahmad, A., and Lesperance, M. M. (2017). Clinical Characterization of Novel Chromosome 22q13. *Microdeletions Int. J. Pediatr. Otorhinolaryngol.* 95, 121–126.

Heilstedt, H. A., Ballif, B. C., Howard, L. A., Lewis, R. A., Stal, S., Kashork, C. D., et al. (2003). Physical Map of 1p36, Placement of Breakpoints in Monosomy 1p36, and Clinical Characterization of the Syndrome. *Am. J. Hum. Genet.* 72 (5), 1200–1212. doi:10.1086/375179

Holder, J. L., Jr, and Quach, M. M. (2016). The Spectrum of Epilepsy and Electroencephalographic Abnormalities Due to SHANK3 Loss-Of-Function Mutations. *Epilepsia* 57 (10), 1651–1659. doi:10.1111/epi.13506

Kaplanis, J., Samocha, K. E., Wiel, L., Zhang, Z., Arvai, K. J., Eberhardt, R. Y., et al. (2020). Evidence for 28 Genetic Disorders Discovered by Combining Healthcare and Research Data. *Nature* 586 (7831), 757–762. doi:10.1038/s41586-020-2832-5

Kolevzon, A., Delaby, E., Berry-Kravis, E., Buxbaum, J. D., and Betancur, C. (2019). Neuropsychiatric Decompensation in Adolescents and Adults with Phelan-McDermid Syndrome: a Systematic Review of the Literature. *Mol. Autism* 10, 50. doi:10.1186/s13229-019-0291-3

Kurtas, N., Arrigoni, F., Errichiello, E., Zucca, C., Maghini, C., D'Angelo, M. G., et al. (2018). Chromothripsis and Ring Chromosome 22: a Paradigm of Genomic Complexity in the Phelan-McDermid Syndrome (22q13 Deletion Syndrome). *J. Med. Genet.* 55 (4), 269–277. doi:10.1136/jmedgenet-2017-105125

Leblond, C. S., Nava, C., Polge, A., Gauthier, J., Huguet, G., Lumbroso, S., et al. (2014). Meta-analysis of SHANK Mutations in Autism Spectrum Disorders: A Gradient of Severity in Cognitive Impairments. *Plos Genet.* 10, e1004580. doi:10.1371/journal.pgen.1004580

Lelieveld, S. H., Reijnders, M. R. F., Pfundt, R., Yntema, H. G., Kamsteeg, E.-J., de Vries, P., et al. (2016). Meta-analysis of 2,104 Trios Provides Support for 10 New Genes for Intellectual Disability. *Nat. Neurosci.* 19, 1194–1196. doi:10.1038/nn.4352

Li, S., Xi, K.-w., Liu, T., Zhang, Y., Zhang, M., Zeng, L.-d., et al. (2020). Fraternal Twins with Phelan-McDermid Syndrome Not Involving the SHANK3 Gene: Case Report and Literature Review. *BMC Med. Genomics* 13, 146–153. doi:10.1186/s12920-020-00802-0

Lim, E. T., Uddin, M., Uddin, M., De Rubeis, S., Chan, Y., Kamumbu, A. S., et al. (2017). Rates, Distribution and Implications of Postzygotic Mosaic Mutations in Autism Spectrum Disorder. *Nat. Neurosci.* 20 (9), 1217–1224. doi:10.1038/nn.4598

Lyons-Warren, A. M., Cheung, S. W., and Holder, J. L., Jr (2017). Clinical Reasoning: A Common Cause for Phelan-McDermid Syndrome and Neurofibromatosis Type 2. *Neurology* 89, e205–e209. doi:10.1212/wnl.0000000000004573

Miller, D. T., Adam, M. P., Aradhya, S., Biesecker, L. G., Brothman, A. R. R., Carter, N. P., et al. (2010). Consensus Statement: Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies. *Am. J. Hum. Genet.* 86, 749–764. doi:10.1016/j.ajhg.2010.04.006

Mitz, A. R., Philyaw, T. J., Boccuto, L., Shcheglovitov, A., Sarasua, S. M., Kaufmann, W. E., et al. (2018). Identification of 22q13 Genes Most Likely to Contribute to Phelan McDermid Syndrome. *Eur. J. Hum. Genet.* 26, 293–302. doi:10.1038/s41431-017-0042-x

Nemirovsky, S. I., Cordoba, M., Zaiat, J. J., Completa, S. P., Vega, P. A., Gonzalez-Moron, D., et al. (2015). Whole-genome Sequencing Reveals a De Novo SHANK3 Mutation in Familial Autism Spectrum Disorder. *PLoS One* 10, e0116358. doi:10.1371/journal.pone.0116358

Nevado, J., Rosenfeld, J. A., Mena, R., Palomares-Bralo, M., Vallespín, E., Ángeles Mori, M., et al. (2015). PIAS4 Is Associated with Macro/microcephaly in the Novel Interstitial 19p13.3 Microdeletion/microduplication Syndrome. *Eur. J. Hum. Genet.* 23 (12), 1615–1626. doi:10.1038/ejhg.2015.51

O'Roak, B. J., Stessman, H. A., Boyle, E. A., Witherspoon, K. T., Martin, B., Lee, C., et al. (2014). Recurrent De Novo Mutations Implicate Novel Genes Underlying Simplex Autism Risk. *Nat. Commun.* 5, 5595. doi:10.1038/ncomms6595

Oberman, L. M., Boccuto, L., Cascio, L., Sarasua, S., and Kaufmann, W. E. (2015). Autism Spectrum Disorder in Phelan-McDermid Syndrome: Initial Characterization and Genotype-Phenotype Correlations. *Orphanet J. Rare Dis.* 10, 105. doi:10.1186/s13023-015-0323-9

Phelan, K., Rogers, R. C., and Boccuto, L. (2018). *22q13.3 Deletion Syndrome, Chromosome 22q13.3 Deletion Syndrome, Deletion 22q13 Syndrome. Genereviews*.

Retterer, K., Juusola, J., Cho, M. T., Vitazka, P., Millan, F., Gibellini, F., et al. (2016). Clinical Application of Whole-Exome Sequencing across Clinical Indications. *Genet. Med.* 18, 696–704. doi:10.1038/gim.2015.148

Ricciardello, A., Tomaiuolo, P., and Persico, A. M. (2021). Genotype-phenotype Correlation in Phelan-McDermid Syndrome: A Comprehensive Review of Chromosome 22q13 Deleted Genes. *Am. J. Med. Genet.* 185, 2211–2233. doi:10.1002/ajmg.a.62222

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

Rosenfeld, J. A., Coe, B. P., Eichler, E. E., Cuckle, H., and Shaffer, L. G. (2013). Estimates of Penetrance for Recurrent Pathogenic Copy-Number Variations. *Genet. Med.* 15 (6), 478–481. doi:10.1038/gim.2012.164

Samogy-Costa, C. I., Varella-Branco, E., Monfardini, F., Ferraz, H., Fock, R. A., Barbosa, R. H. A., et al. (2019). A Brazilian Cohort of Individuals with Phelan-McDermid Syndrome: Genotype-Phenotype Correlation and Identification of an Atypical Case. *J. Neurodevelop Disord.* 11 (1), 13. doi:10.1186/s11689-019-9273-1

Sarasua, S. M., Boccuto, L., Sharp, J. L., Dwivedi, A., Chen, C.-F., Rollins, J. D., et al. (2014a). Clinical and Genomic Evaluation of 201 Patients with Phelan-McDermid Syndrome. *Hum. Genet.* 133 (7), 847–859. doi:10.1007/s00439-014-1423-7

Sarasua, S. M., Dwivedi, A., Boccuto, L., Chen, C.-F., Sharp, J. L., Rollins, J. D., et al. (2014b). 22q13.2q13.32 Genomic Regions Associated with Severity of Speech Delay, Developmental Delay, and Physical Features in Phelan-McDermid Syndrome. *Genet. Med.* 16, 318–328. doi:10.1038/gim.2013.144

Sarasua, S. M., Dwivedi, A., Boccuto, L., Rollins, J. D., Chen, C.-F., Rogers, R. C., et al. (2011). Association between Deletion Size and Important Phenotypes Expands the Genomic Region of Interest in Phelan-McDermid Syndrome (22q13 Deletion Syndrome). *J. Med. Genet.* 48, 761–766. doi:10.1136/jmedgenet-2011-100225

Sauer, K. A., Bockmann, J., Steionestel, K., Boeckers, T. M., and Grabucker, A. M. (2019). Altered Intestinal Morphology and Microbiota Composition in the Autism Spectrum Disorders Associated SHANK3 Mouse Model. *Int. J. Mol. Sci.* 20 (09), 2134. doi:10.3390/ijms20092134

Schouten, J. P., McElgunn, C. J., Waaijer, R., Zwijnenburg, D., Diepvens, F., and Pals, G. (2002). Relative Quantification of 40 Nucleic Acid Sequences by Multiplex Ligation-dependent Probe Amplification. *Nucleic Acids Res.* 30 (12), e57. doi:10.1093/nar/gnf056

Shao, L., Shaw, C. A., Lu, X.-Y., Sahoo, T., Bacino, C. A., Lalani, S. R., et al. (2008). Identification of Chromosome Abnormalities in Subtelomeric Regions by Microarray Analysis: a Study of 5,380 Cases. *Am. J. Med. Genet.* 146A (17), 2242–2251. doi:10.1002/ajmg.a.32399

Soorya, L., Kolevzon, A., Zweifach, J., Lim, T., Dobry, Y., Schwartz, L., et al. (2013). Prospective Investigation of Autism and Genotype-Phenotype Correlations in 22q13 Deletion Syndrome and SHANK3 Deficiency. *Mol. Autism* 4 (4), 18. doi:10.1186/2040-2392-4-18

Srikanth, S., Jain, L., Zepeda-Mendoza, C., Cascio Kelly Jones, L., Pauly, R., DuPont, B., et al. (2021). Position Effects of 22q13 Rearrangements on Candidate Genes in Phelan- McDermid Syndrome. *PLoS One* 16 (7), e0253859. doi:10.1371/journal.pone.0253859

Tabet, A.-C., Rolland, T., Ducloy, M., Lévy, J., Buratti, J., Mathieu, A., et al. (2017). A Framework to Identify Contributing Genes in Patients with Phelan-McDermid Syndrome. *Npj Genom. Med.* 2, 32. doi:10.1038/s41525-017-0035-2

Tabolacci, E., Zollino, M., Lecce, R., Sangiorgi, E., Gurrieri, F., Leuzzi, V., et al. (2005). Two brothers with 22q13 Deletion Syndrome and Features Suggestive of the Clark???Baraitser Syndrome. *Clin. Dysmorphol.* 14, 127–132. doi:10.1097/00019605-200507000-00004

Tenorio, J., Nevado, J., González-Meneses, A., Arias, P., Dapía, I., Venegas-Vega, C. A., et al. (2020). Further Definition of the Proximal 19p13.3 Microdeletion/microduplication Syndrome and Implication of PIAS4 as the Major Contributor. *Clin. Genet.* 97 (3), 467–476. doi:10.1111/cge.13689

Thevenon, J., Duffourd, Y., Masurel-Paulet, A., Lefebvre, M., Feillet, F., El Chehadeh-Djebbar, S., et al. (2016). Diagnostic Odyssey in Severe Neurodevelopmental Disorders: toward Clinical Whole-Exome Sequencing as a First-Line Diagnostic Test. *Clin. Genet.* 89 (6), 700–707. doi:10.1111/cge.12732

Vallespín, E., Palomares Bralo, M., MoriÁ., M. Á., Martín, R., García-Miñaúr, S., Fernández, L., et al. (2013). Customized High Resolution CGH-Array for Clinical Diagnosis Reveals Additional Genomic Imbalances in Previous Well-Defined Pathological Samples. *Am. J. Med. Genet.* 161 (8), 1950–1960. doi:10.1002/ajmg.a.35960

Verhoeven, W. M. A., Egger, J., Willemsen, M. H., de Leijer, G. J., and Kleefstra, T. (2012). Phelan-McDermid Syndrome in Two Adult brothers: Atypical Bipolar Disorder as its Psychopathological Phenotype? *Ndt* 8, 175–179. doi:10.2147/ndt.s30506

Wilson, H. L., Crolla, J. A., Walker, D., Artifoni, L., Dallapiccola, B., Takano, T., et al. (2008). Interstitial 22q13 Deletions: Genes Other Than SHANK3 Have Major Effects on Cognitive and Language Development. *Eur. J. Hum. Genet.* 16, 1301–1310. doi:10.1038/ejhg.2008.107

Wilson, H. L., Wong, A. C., Shaw, S. R., Tse, W. Y., Stapleton, G. A., Phelan, M. C., et al. (2003). Molecular Characterisation of the 22q13 Deletion Syndrome Supports the Role of Haploinsufficiency of SHANK3/PROSAP2 in the Major Neurological Symptoms. *J. Med. Genet.* 40, 575–584. doi:10.1136/jmg.40.8.575

Xu, N., Lv, H., Yang, T., Du, X., Sun, Y., Xiao, B., et al. (2020). A 29 Mainland Chinese Cohort of Patients with Phelan-McDermid Syndrome: Genotype-Phenotype Correlations and the Role of SHANK3 Haploinsufficiency in the Important Phenotypes. *Orphanet J. Rare Dis.* 15 (1), 335. doi:10.1186/s13023-020-01592-5

Zhang, Y., Kong, W., Gao, Y., Liu, X., Gao, K., Xie, H., et al. (2015). Gene Mutation Analysis in 253 Chinese Children with Unexplained Epilepsy and Intellectual/developmental Disabilities. *PLoS One* 10, e0141782. doi:10.1371/journal.pone.0141782

Zhou, W. Z., Zhang, J., Li, Z., Lin, X., Li, J., Wang, S., et al. (2019). Targeted Resequencing of 358 Candidate Genes for Autism Spectrum Disorder in a

Chinese Cohort Reveals Diagnostic Potential and Genotype-Phenotype Correlations. *Hum. Mutat.* 40 (6), 801–815. doi:10.1002/humu.23724

Ziats, C. A., Grosvenor, L. P., Sarasua, S. M., Thurm, A. E., Swedo, S. E., Mahfouz, A., et al. (2019). Functional Genomics Analysis of Phelan-McDermid Syndrome 22q13 Region during Human Neurodevelopment. *PLoS One* 14, e0213921. doi:10.1371/journal.pone.0213921

Ziats, C. A., Jain, L., McLarney, B., Vandenboom, E., DuPont, B. R., Rogers, C., et al. (2020). Neurofibromatosis Type 2 in Phelan-McDermid Syndrome: Institutional Experience and Review of the Literature. *Eur. J. Med. Genet.* 63 (11), 104042. doi:10.1016/j.ejmg.2020.104042

Zwanenburg, R. J., Ruiter, S. A. J., van den Heuvel, E. R., Flapper, B. C. T., and Van Ravenswaaij-Arts, C. M. A. (2016). Developmental Phenotype in Phelan-McDermid (22q13.3 Deletion) Syndrome: a Systematic and Prospective Study in 34 Children. *J. Neurodevelop Disord.* 8, 16–28. doi:10.1186/s11689-016-9150-0

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.

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