NEXT GENERATION SEQUENCING (NGS) FOR RARE DISEASES DIAGNOSIS

EDITED BY: Xiu-An Yang, Hu Hao, Yanling Yang and Can Liao

PUBLISHED IN: Frontiers in Genetics

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88974-203-5
DOI 10.3389/978-2-88974-203-5

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

NEXT GENERATION SEQUENCING (NGS) FOR RARE DISEASES DIAGNOSIS

Topic Editors:
Xiu-An Yang, Chengde Medical College, China
Hu Hao, The Sixth Affiliated Hospital of Sun Yat-Sen University, China
Yanling Yang, Peking University First Hospital, China
Can Liao, Guangzhou Women and Children's Medical Center, China

Citation: Yang, X.-A., Hao, H., Yang, Y., Liao, C., eds. (2022). Next Generation Sequencing (NGS) for Rare Diseases Diagnosis. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-203-5

Table of Contents

frontiers in Genetics

EDITORIAL
published: 23 December 2021
doi: 10.3389/fgene.2021.808042

Editorial: Next Generation Sequencing (NGS) for Rare Diseases Diagnosis

Xiu-An Yang[1,2]*

[1]Laboratory of Genetic Engineering and Genomics, School of Basic Medical Sciences, Chengde Medical University, Chengde, China, [2]Hebei Key Laboratory of Nerve Injury and Repair, Chengde Medical University, Chengde, China

Keywords: next generation sequencing, rare diseases diagnosis, whole-exome sequencing, copy number variation, bioinformatics

Editorial on the Research Topic

Next Generation Sequencing (NGS) for Rare Diseases Diagnosis

INTRODUCTION

The rapidly increasing amount of genome data provides wide opportunities to address new research questions and to translate new knowledge in clinical practice (Koriath et al., 2021). Currently, most mature areas for clinical applications of genomics include rare disease diagnostics, oncology, infectious diseases, and pharmacogenomics. We expect that the next wave is seen in improving our understanding of chronic diseases that change traditional diagnostic categories. These new disease classifications hopefully result in new, more tailored treatment and prediction options (Holzinger et al., 2015). Genomic medicine publishes high-quality research that facilitates the use of genomic information in medicine, including treatment, diagnostics, and prevention (Franks et al., 2021). With the goal to explore studies regarding clinical work that benefited from Next Generation Sequencing (NGS), this issue entitled "Next Generation Sequencing (NGS) for Rare Diseases Diagnosis" is published.

In a hemophilia A family with two male patients, Bai et al. had used NGS for preimplantation genetic testing to screen chromosome copy number variation and mutation sites. Sanger Sequencing was enrolled to verify the identified mutation sites and single nucleotide polymorphisms (SNP). Finally, a novel missense mutation, p (Phe690Leu)/c.2070C > A, occurring in exon 13 of factor VIII gene (*F8/FVIII*) was identified as a potential pathogenic mutation. Defects in F8 are documented to be the pathogenic mechanism of an X-linked recessive bleeding disorder called hemophilia A. After an F8 normal euploid blastocyst was transferred, the family had a healthy fetus, indicating the importance of NGS in assisted reproduction technology. It is worth noting that structural modeling was used for speculating the possible pathogenic mechanism of the mutation both in this study and other studies in this issue. Due to the limitation of diagnostic conditions and time, it is often difficult to verify the functions of the mutations found in possible therapeutic gene sites. Therefore, *in silico* analysis such as structural modeling plays an important role in determining the pathogenicity of mutation sites. This illustrates the importance of combining clinical practice with basic practice.

MOCS1 (Molybdenum cofactor synthesis 1), *MOCS2* (Molybdenum cofactor synthesis 2), and *GPHN* (Gephyrin) are shown to be the pathogenic genes of molybdenum cofactor (Moco) deficiency in humans (Mayr et al., 2021). These genes, together with *MOCS3* (Molybdenum cofactor synthesis 3) are involved in Moco biosynthesis and providing cofactors to Moco-dependent enzymes (Mechler et al., 2015). However, variants in *MOCS3* have not been reported in molybdenum cofactor (Moco)

OPEN ACCESS

Edited and reviewed by:
Thomas Liehr,
Friedrich Schiller University Jena, Germany

***Correspondence:**
Xiu-An Yang
yangxiuan07@mails.ucas.edu.cn

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

Received: *02 November 2021*
Accepted: *06 December 2021*
Published: *23 December 2021*

Citation:
Yang X-A (2021) Editorial: Next Generation Sequencing (NGS) for Rare Diseases Diagnosis. Front. Genet. 12:808042. doi: 10.3389/fgene.2021.808042

deficiency in patients so far. Using trio whole-exome sequencing (Trio-WES) accompanied with Sanger sequencing validation, Tian et al. identified compound heterozygous variants in NM_014484.49 (MOCS3): c.1375C > T; p. Gln459Ter (chr20:49576754-49576754) and NM_014484.49 (MOCS3): c.325C > G; p. Leu109Val (chr20:49575704-49575704) in a proband. The identified variants were inherited from parents respectively. According to the authors, this is the first case of *MOCS3* variants causing Moco deficiency. It shows the power of NGS in the diagnosis of rare diseases, and also indicates the importance of familial segregation analysis.

Due to its rarity and sporadic characteristics, the diagnosis of rare diseases is usually quite complicated. In light of this, retrospectively investigation for a cohort of individuals who are suspected having molecular deficiency is effective for expanding both the genetic and clinic spectrum of rare diseases. Patients with disorders of nervous system (67 cases), kidney (65 cases and 69 cases respectively), and hematology (25 cases) were retrospectively studied (Ferese et al.; Kristan et al.; Stippel et al.; Wang et al.). To reveal the role of rare variants in the genotype-phenotype correlation, 67 patients with Charcot-Marie-Tooth (CMT) disease were enrolled in the study performed by (Ferese et al.). The cohort had taken copy number variation (CNV) detection and NGS panel for sequencing of 47 genes known to be associated with CMT and routinely screened in medical genetics. Both novel and previously reported causative variants were identified, broadening the phenotype expression produced by either pathogenic or undefined variants. Similarly, the other three studies are of great significance for the improvement of diagnosis of patients with unexplained rare disorders.

With the development of high-throughput sequencing technology, a large amount of sequencing data have been uploaded to relevant databases, and the mining of these public data has played an important role in the research of diseases pathogenesis including rare diseases and tumors. By means of bioinformatics, He et al. investigated the novel key genes of non-obstructive azoospermia affect spermatogenesis using RNA-seq and scRNA-seq data. Finally, 430 differentially expressed genes were identified. Of these 430 genes, C22orf23, TSACC, and TTC25, which have not been reported related to spermatogenesis are identified as hub genes. This study showed the importance of bioinformatics in disease pathogenesis.

In summary, new genes and rare variants were identified in the studies enrolled in this special topic, showing the important role of NGS in providing a tremendous opportunity to better serve our patients. The studies highlighted the association of phenotypes and the genetic heterogeneity in rare disease diagnosis. Lohmann & Klein have pointed out the importance role of NGS in the diagnosis of both rare diseases and common but heterogeneous disorders as early as in 2014 (Lohmann and Klein 2014). NGS is now widely accepted for diagnostic purposes of rare diseases. It should be noted that for most studied cases Sanger sequencing would have been as good to identify the mutations found as NGS. However, NGS can be performed in a short time at reasonable costs comparing with Sanger sequencing and can provide a large-scale potential variants (Lohmann and Klein 2014).

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This study was supported by Natural Science Foundation of Hebei Province (Grant Number H2020406049); Scientific and Technological Research Projects of Hebei Higher Education (Grant Number ZD2019084); and Hebei Key Laboratory of Nerve Injury and Repair, Technology Innovation Guidance Project-Science and Technology Work Conference.

REFERENCES

Franks, P. W., Melén, E., Friedman, M., Sundström, J., Kockum, I., Klareskog, L., et al. (2021). Technological Readiness and Implementation of Genomic-driven Precision Medicine for Complex Diseases. *J. Intern. Med.* 290, 602–620. doi:10.1111/joim.13330

Holzinger, D., Kessel, C., Omenetti, A., and Gattorno, M. (2015). From Bench to Bedside and Back Again: Translational Research in Autoinflammation. *Nat. Rev. Rheumatol.* 11, 573–585. doi:10.1038/nrrheum.2015.79

Koriath, C. A. M., Kenny, J., Ryan, N. S., Rohrer, J. D., Schott, J. M., Houlden, H., et al. (2021). Genetic Testing in Dementia - Utility and Clinical Strategies. *Nat. Rev. Neurol.* 17, 23–36. doi:10.1038/s41582-020-00416-1

Lohmann, K., and Klein, C. (2014). Next Generation Sequencing and the Future of Genetic Diagnosis. *Neurotherapeutics* 11, 699–707. doi:10.1007/s13311-014-0288-8

Mayr, S. J., Mendel, R.-R., and Schwarz, G. (2021). Molybdenum Cofactor Biology, Evolution and Deficiency. *Biochim. Biophys. Acta (Bba) - Mol. Cel Res.* 1868, 118883. doi:10.1016/j.bbamcr.2020.118883

Mechler, K., Mountford, W. K., Hoffmann, G. F., and Ries, M. (2015). Ultra-orphan Diseases: a Quantitative Analysis of the Natural History of Molybdenum Cofactor Deficiency. *Genet. Med.* 17, 965–970. doi:10.1038/gim.2015.12

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 26 February 2021
doi: 10.3389/fgene.2021.608629

The Novel Key Genes of Non-obstructive Azoospermia Affect Spermatogenesis: Transcriptomic Analysis Based on RNA-Seq and scRNA-Seq Data

Haihong He[1†], Fan Yu[1†], Wang Shen[2], Keyan Chen[1], Lijun Zhang[1], Shuang Lou[1], Qiaomin Zhang[1], Siping Chen[1], Xinhua Yuan[1], Xingwang Jia[1] and Yiwen Zhou[1]*

[1]Department of Emergency Laboratory, Clinical Laboratory Medical Center, Shenzhen Hospital, Southern Medical University, Shenzhen, China, [2]Department of Clinical Laboratory, Affiliated Jiangmen TCM Hospital of Ji'nan University, Jiangmen, China

Non-obstructive azoospermia (NOA) is one of the most important causes of male infertility. It is mainly characterized by the absence of sperm in semen repeatedly or the number of sperm is small and not fully developed. At present, its pathogenesis remains largely unknown. The goal of this study is to identify hub genes that might affect biomarkers related to spermatogenesis. Using the clinically significant transcriptome and single-cell sequencing data sets on the Gene Expression Omnibus (GEO) database, we identified candidate hub genes related to spermatogenesis. Based on them, we performed Gene Ontology (GO) functional enrichment analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment pathway analyses, protein-protein interaction (PPI) network analysis, principal component analysis (PCA), cell cluster analysis, and pseudo-chronological analysis. We identified a total of 430 differentially expressed genes, of which three have not been reported related to spermatogenesis (C22orf23, TSACC, and TTC25), and the expression of these three hub genes was different in each type of sperm cells. The results of the pseudo-chronological analysis of the three hub genes indicated that TTC25 was in a low expression state during the whole process of sperm development, while the expression of C22orf23 had two fluctuations in the differentiating spermatogonia and late primary spermatocyte stages, and TSACC showed an upward trend from the spermatogonial stem cell stage to the spermatogenesis stage. Our research found that the three hub genes were different in the trajectory of sperm development, indicating that they might play important roles in different sperm cells. This result is of great significance for revealing the pathogenic mechanism of NOA and further research.

Keywords: non-obstructive azoospermia, spermatogenesis, scRNA-seq, differentially express genes, TSSK6

OPEN ACCESS

Edited by:
Xiu-An Yang,
Chengde Medical College, China

Reviewed by:
Qiu-Ning Liu,
Yancheng Teachers University, China
Xiaona Wang,
Children's Hospital Affiliated to Zhengzhou University, China

****Correspondence:***
Yiwen Zhou
zhouyw_2020@163.com

[†]These authors have contributed equally to this work

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

Received: *21 September 2020*
Accepted: *08 February 2021*
Published: *26 February 2021*

Citation:
He H, Yu F, Shen W, Chen K, Zhang L, Lou S, Zhang Q, Chen S, Yuan X, Jia X and Zhou Y (2021) The Novel Key Genes of Non-obstructive Azoospermia Affect Spermatogenesis: Transcriptomic Analysis Based on RNA-Seq and scRNA-Seq Data. Front. Genet. 12:608629. doi: 10.3389/fgene.2021.608629

INTRODUCTION

About 10–15% of people of childbearing age are infertile in the world, of which male infertility accounts for about 50% (Gifford, 2015). Male infertility is closely related to sexual dysfunction, varicocele, reproductive system infection, endocrine, obstructive azoospermia (OA), non-obstructive azoospermia (NOA), etc. (Arafat et al., 2017). The incidence of NOA in men is about 1%, accounting for 10–15% of infertile men, and it is one of the most important causes of male infertility (Wosnitzer et al., 2014).

Non-obstructive azoospermia is a type of male infertility caused by spermatogenic dysfunction of testicular tissue. Patients with NOA cannot produce sperm or can only produce a very small amount of sperm. In patients with NOA, the structure of the seminiferous tubules in the testis is disordered, while the maturation of spermatogenic cells is blocked, and the meiosis of spermatogenic cells is arrested (Kohn and Pastuszak, 2018). In recent years, some studies (Vij et al., 2018) have shown that there are focal and heterogeneous tissues in the spermatogenesis disorder. Even if sperm are not found in most of the seminiferous tubules in the testicular tissue, it cannot be completely denied that there may be a very small amount of sperm in some seminiferous tubules. Studies have confirmed that some sperm found from the testis or epididymis in patients whose spermatogenic cells stop during meiosis can also be conceived by intracytoplasmic sperm injection (ICSI) diagnosis and treatment (Silber et al., 1996). There are also studies reporting that the use of round or long sperm can also enable patients to gain fertility (Sofikitis et al., 1998; Ghazzawi et al., 1999). Reproductive technology has greatly reduced requirements for sperm quantity and sperm maturity compared with natural conception conditions, so that it can become a reality for NOA patients to obtain genetic offspring (Corona et al., 2019; Song et al., 2020). Because reproductive technology bypasses the natural selection mechanism in the process of sperm formation, the risk of genetic defects being passed to the next generation is also significantly increased. Therefore, it is particularly important to grasp the key factors that affect the process of sperm development.

At present, the research on NOA mainly focuses on including chromosomal abnormalities, Y chromosome microdeletion, and epigenetics. However, the spermatogenesis barriers and pathogenesis related to gene expression levels are still to be explored (Fang et al., 2020). In this study, we mainly used the gene expression profile microarray transcriptome data and single cell sequencing data in the NCBI Gene Expression Comprehensive Database (NCBI-GEO) database. By analyzing transcriptome data, we identified differentially expressed genes (DEGs) between patients with and without spermatogenesis disorders. Subsequently, we performed Gene Ontology (GO) function enrichment analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment pathway analysis, protein-protein interaction (PPI) network analysis, and identification of hub genes. By analyzing the sperm single-cell sequencing data, we annotated the cell clusters, revealing the expression of the hub gene in the cell clusters, and finally performed a pseudo-chronological analysis of the hub gene to reveal the role of the hub gene in sperm development.

MATERIALS AND METHODS

Data Source

Gene Expression Omnibus (GEO) is a public genome database that provides gene expression data, microarray, and single cell sequencing data.[1] The transcriptome data of this study were from two datasets: GSE108886 and GSE145467. People without spermatogenesis disorders were included in the control group, including normal people and patients with OA, and patients with NOA were included in the disease group. The population of the GSE108886 data set consists of one normal person, three OA patients and eight NOA patients. The GSE145467 data set consists of 10 OA patients and 10 NOA patients. The single-cell sequencing data were from the GSE109037 data set, which contained 11 samples from four sub-data sets (spermatogenesis, spermatocytes, spermatogonia, and spermatids) during sperm development.

Identification of Differentially Expressed Genes

The GEO2R online analysis software on GEO database was used to identify DEGs,[2] and the screening threshold was set to $|\log2\ FC| > 2$ and $p < 0.01$.

Functional Enrichment Analysis of the DEGs

The DAVID online analysis tool was used for GO and KEGG enrichment analyses of DEGs.[3] We then use the imageGP tool for visual analysis.[4] $p < 0.05$ was considered statistically significant.

Construction of PPI Network and Identification of Hub Genes

We use the String database to build the PPI network of DEGs,[5] and set the interaction with combined score to be greater than or equal to 0.4. Subsequently, the data after String analysis were imported into Cytoscape software (version 3.6.1) to construct the PPI network again. The five calculation methods (Degree, DMNC, EPC, MCC, and MNC) in the CytoHubba plug-in were used to identify hub genes in the PPI network. A total of 50 hub genes (TOP10 in each algorithm) were selected to be involved in the Veen Diagram, and the core hub genes were then identified for subsequent analysis.

Cell Clustering and Annotation of Single Cell Sequencing Data

We used the Seurat package (version 3.1.5.9915) of the R software (version 4.02) to analyze the four sub-sets of GSE109037. After a series of data filtering, quality control, and normalization, we performed principal component analysis (PCA) and non-linear dimensionality reduction UMAP method for cell clustering.

[1]http://www.ncbi.nlm.nih.gov/geo
[2]http://www.ncbi.nlm.nih.gov/geo/geo2r/
[3]https://david.ncifcrf.gov/
[4]http://www.ehbio.com/ImageGP/index.php/Home
[5]http://string-db.org

Through the FindAllMarkers function, we obtained the marker genes of the cell clustering, and then we queried the Human Cell Landscape (HPL) database for cell annotation.[6] The DotPlot and VlnPlot functions were used to draw bubble charts and violin charts to visualize the expression of hub genes in different cell clusters.

Constructing Trajectories of Hub Genes in Single Cell

The monocle3 package (version 0.2.3.0) was used to perform pseudotime time analysis to construct the developmental trajectory of sperm cells and hub genes. The main functions used were learn_graph, label_leaves, label_branch_points, plot_cells, order_cells, and root_pr_node.

RESULTS

Identification of DEGs

In the GSE108886 data set, a total of 520 DEGs were obtained, including 511 downregulated genes and nine upregulated genes (**Figure 1A**), while in the GSE145467 data set, a total of 1,622 DEGs were obtained, including 1,566 downregulated genes and 56 upregulated genes (**Figure 1B**). There were 430 DEGs overlapping the two data sets, and all of them were downregulated genes (**Figures 1C,D**).

GO Functional Enrichment Analysis and KEGG Pathway Analysis

The results of GO functional enrichment analysis showed that the DEGs were mainly enriched in biological processes (BP) and cell composition (CC). In the BP category, they were mainly enriched in spermatogenesis, multicellular organism development, cell differentiation, spermatid development, and sperm motility (**Figure 2A**). In the CC category, they were mainly enriched in modile cilium, microtubule, acrosomal vesicle, and nucleus (**Figure 2B**). In KEGG pathway analysis, we found that these candidate genes might affect spermatogenesis signaling pathways including glycolysis/gluconeogenesis, protein processing in endoplasmic reticulum, carbon metabolism, and cell cycle (**Figure 2C**).

Construction of PPI Network of DEGs and Identification of Hub Genes

We mapped these DEGs to the STRING database to construct a PPI network, and then re-imported these interactive network data into Cytoscape software (version 3.6.0), and finally 304 nodes and 933 edges constituted the PPI network (**Figure 3A**). We used five algorithms in cytoHubba to obtain the top 10 hub genes of each algorithm (**Figures 3B–F**). Then, these hub genes were integrated, and 15 of them overlapped in the five algorithms (ACTRT2, ADAM32, AKAP4, ALS2CR11, C22orf23, CAPZA3, CRISP2, FAM71F1, GKAP1, ODF1, PGK2, PRM2, TNP1, TSACC, and TTC25). The 15 hub genes were considered to be key genes (**Figure 3G**), indicating that they might play important roles in spermatogenesis. After consulting relevant literature (Choi et al., 2003; Malcher et al., 2013; Liu et al., 2015; Javadian-Elyaderani et al., 2016; Malla and Bhandari, 2017; Park et al., 2018; Lim et al., 2019; Tapia Contreras and Hoyer-Fender, 2019; Wang et al., 2019a; Oud et al., 2020; Zhang et al., 2020), among those 15 genes, we screened out three hub genes (C22orf23, TSACC, and TTC25) that have not been reported to be related to spermatogenesis.

Expression of the Hub Genes in Sperm Cells

To clarify the expression of these three hub genes in different types of sperm cells, we analyzed the single-cell sequencing data covering the whole process of sperm development. Through the Seurat package nonlinear dimensionality reduction method (UMAP), a total of 13 cell clusters were found (**Figure 4A**). Next, we use the FindAllMarkers function to identify and visualize the marker genes of cell clusters (**Figure 4B**). By consulting the HPL database through the marker gene, 12 types of sperm cells were identified, and the relationship between them was shown in a heat map (**Figure 4C**). The analysis results showed that the three hub genes were expressed in late primary spermatocyte, round spermatid, elongated spermatid, sperm1, and sperm2, while testis-specific serine kinase 6 activating co-chaperone (TSACC) had the highest expression level among the three hub genes. In early primary spermatocyte and differentiating spermatogonia cells, TTC25 was almost not expressed, while C22orf23 was slightly expressed, and TSACC expression was higher (**Figures 4D,E**). This result indicated that there were differences in the expression of hub genes in the development of sperm cells.

Pseudotime Time Analysis of Hub Genes in Sperm Cells

The developmental trajectory of sperm cells was constructed using monocle 3 software. The results showed that during sperm development, the initial cell of sperm development was identified as spermatogonial stem cell, and the only branch appeared in the elongated spermatid (**Figures 5A,B**). We constructed the trajectory of C22orf23, TSACC, and TTC25 in sperm development, and found that TTC25 has been maintained at a low level, while C22orf23 began to rise from the differentiating spermatogonia and the first peak appeared in the late primary spermatocyte, and then gradually decreased, there was an upward trend in the process from elongated spermatid to sperm formation. During the whole process from spermatogonial stem cell to the sperm formation, the expression of TSACC showed an upward trend (**Figure 5C**).

DISCUSSION

The process of spermatogenesis is divided into three stages: mitosis stage, meiosis stage, and sperm cell forming stage. Subsequently, mature sperm are freely released into the lumen

[6]http://bis.zju.edu.cn/HCL/

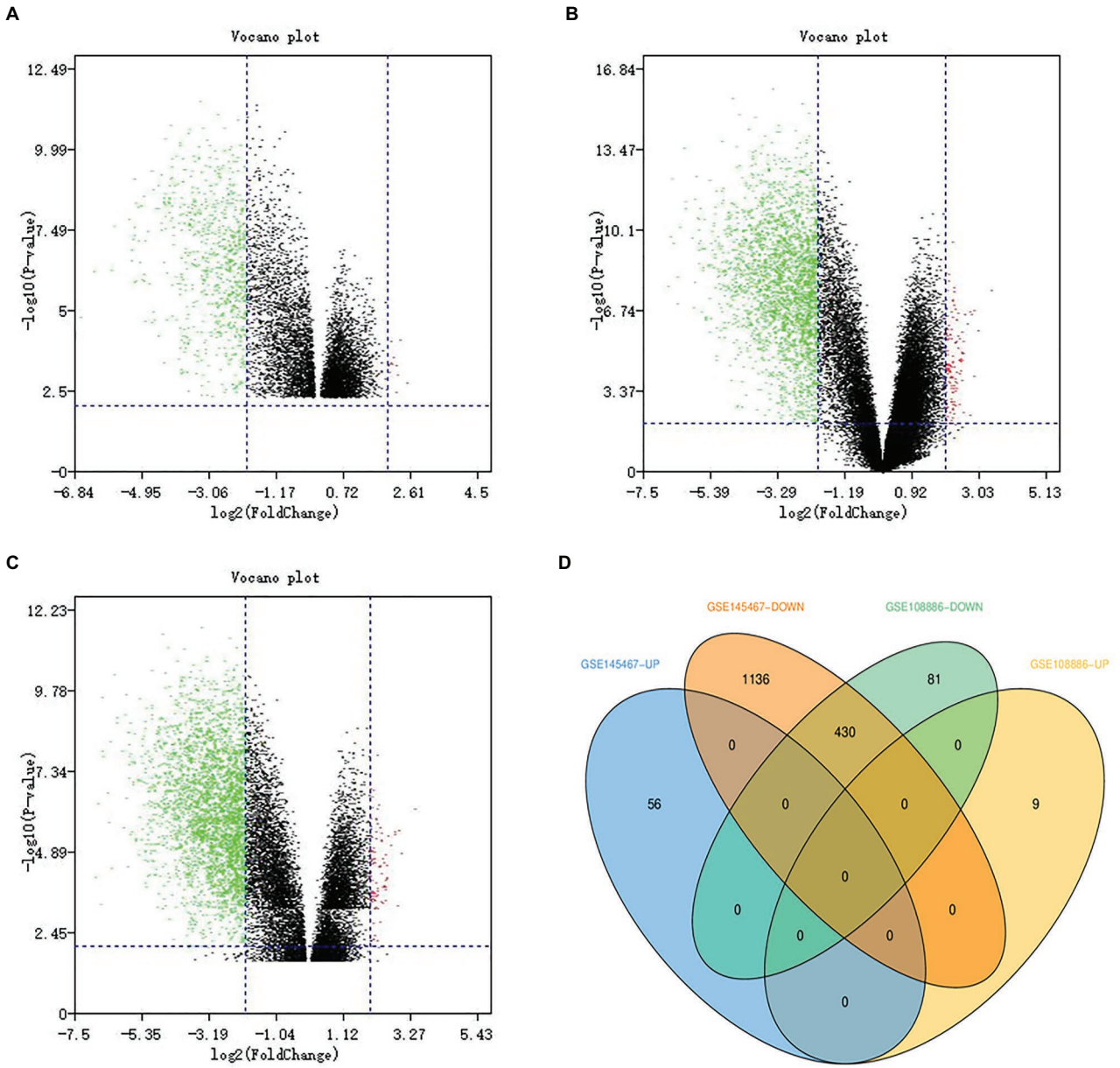


FIGURE 1 | Identification of differentially expressed genes (DEGs). Green dots represent downregulated genes, and red dots represent upregulated genes. **(A)** The volcano map showing DEGs in the GSE108886 data set. **(B)** The volcano map showing DEGs in the GSE145467 data set. **(C)** The volcano map showing DEGs after data set integration. **(D)** The Venn diagram showing that there are 430 DEGs overlapping these two data sets.

of seminiferous tubules, and migrate to the epididymal tissue for energy and storage (Neto et al., 2016). In the whole process of spermatogenesis, any abnormality in any link can lead to NOA, which can lead to male infertility (La and Hobbs, 2019). The process of spermatogenesis is a complex process involving a variety of regulatory mechanisms, among which the regulation of gene transcription level has been a hotspot of research in recent years (Singh et al., 2019).

A number of studies have shown (Jan et al., 2017; Law et al., 2019; Trigg et al., 2019) that changes in transcriptome genes can affect the development of different types of sperm cells and ultimately lead to the occurrence of NOA. Our analysis of GSE108886 and GSE145467 found that 430 DEGs were identified in both data sets at the same time and all were downregulated. Subsequent GO analysis showed that these genes were mainly enriched in BP category including spermatogenesis, multicellular organism development, cell differentiation, spermatid development, and sperm motility, as well as in CC category including motile cilium, microtubule, acrosomal vesicle and nucleus. These are also the hotspots of research on male

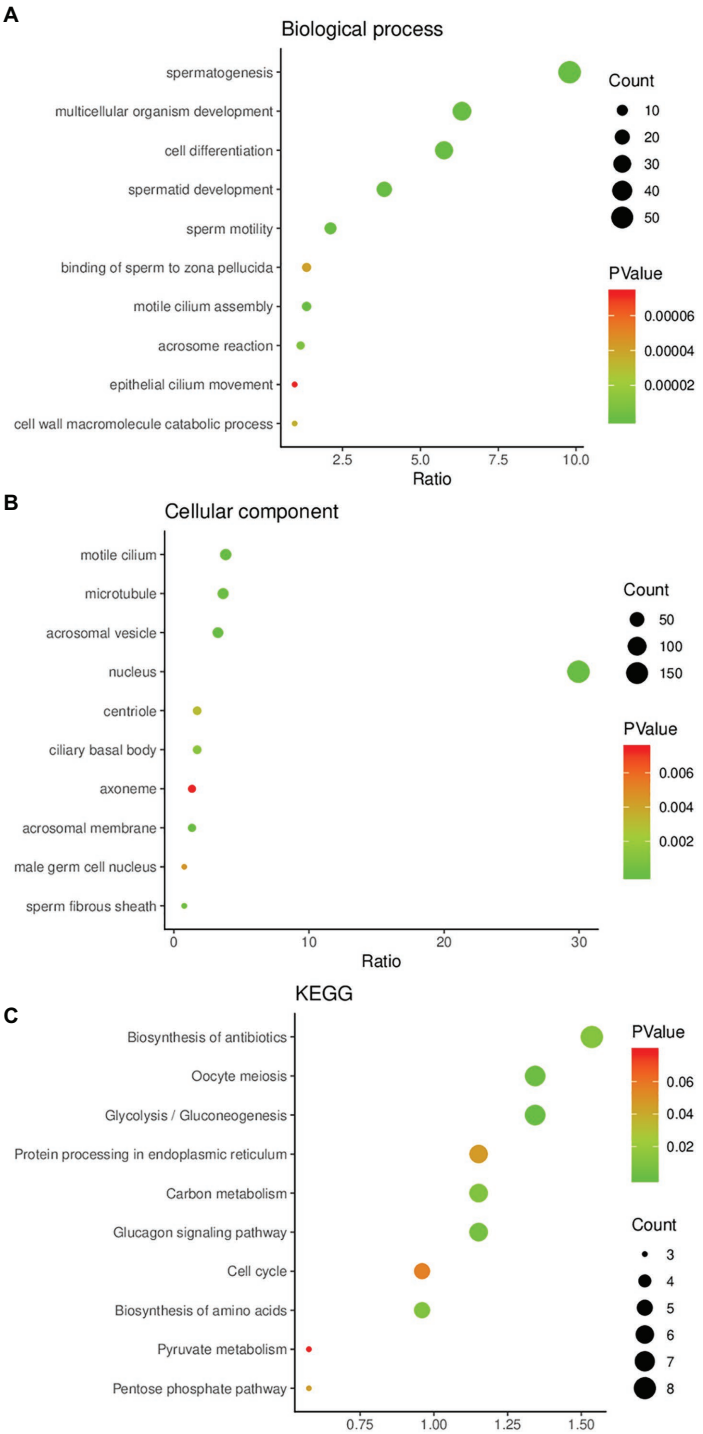


FIGURE 2 | Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of DEGs. **(A)** The result of GO enrichment analysis (biological process). **(B)** The result of GO enrichment analysis (cellular component). **(C)** The result of KEGG pathway enrichment analysis.

infertility from the transcriptome level in recent years (Moye et al., 2019; Wang et al., 2019b; Yuan et al., 2019; Goto, et al., 2020). KEGG pathway analysis results indicated that the pathways that DEGs might affect sperm development were mainly enriched in glycolysis/gluconeogenesis, protein processing in endoplasmic reticulum, carbon metabolism and cell cycle. Studies have shown that One-carbon metabolism may affect spermatogenesis and lead to male infertility (Singh and Jaiswal, 2013). The protein

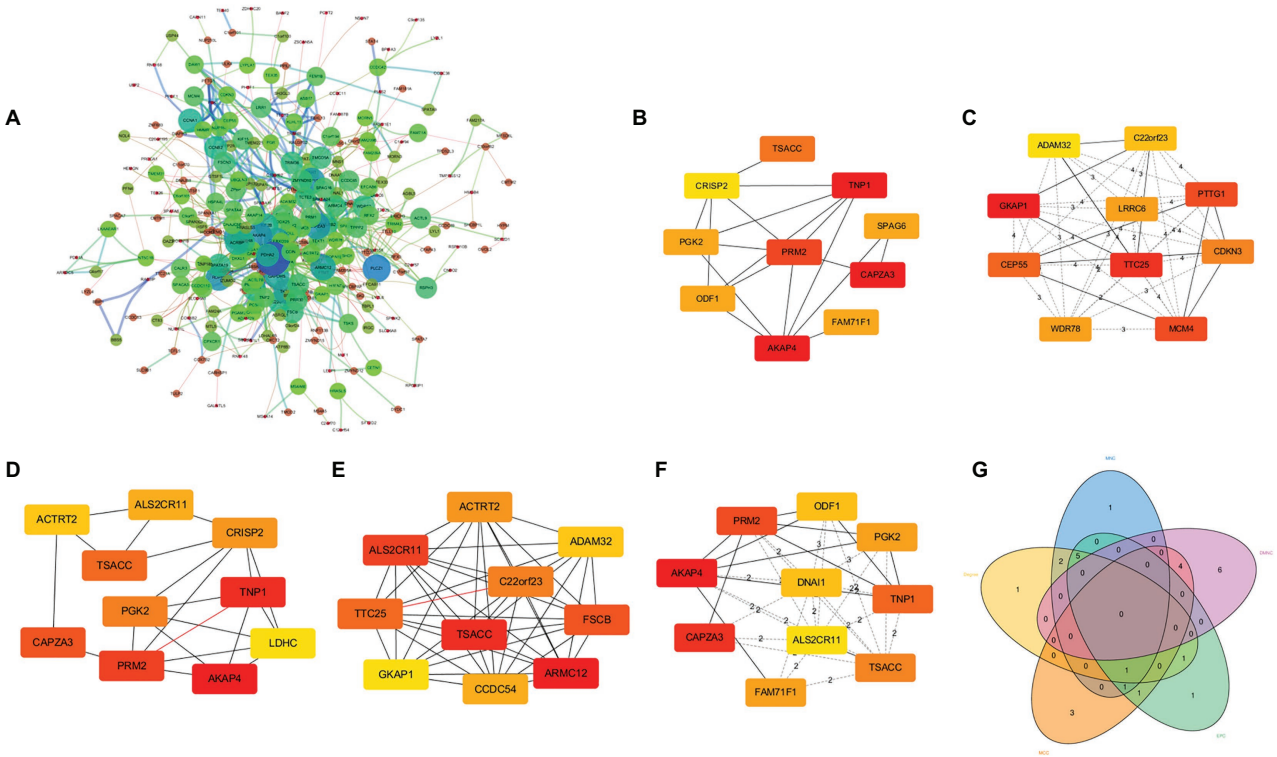


FIGURE 3 | Protein-protein interaction (PPI) network of DEGs and identification of hub genes. **(A)** PPI network composed of 304 nodes and 933 edges. **(B)** Hub genes obtained by Degree algorithm. **(C)** Hub genes obtained by Density of Maximum Neighborhood Component algorithm. **(D)** Hub genes obtained by Edge Percolated component algorithm. **(E)** Hub genes obtained by Maximal Clique Centrality algorithm. **(F)** Hub genes obtained by Maximum Neighborhood Component algorithm. **(G)** Venn diagram showing 15 overlapped hub genes.

processing in endoplasmic reticulum signaling pathway affects histone ubiquitination and acetylation during sperm formation, which is crucial for sperm development (Kavarthapu et al., 2020). In proteomic studies, it was found that the glycolysis/gluconeogenesis pathway is different between normal people and asthenospermia patients (Saraswat et al., 2017), but till now, there is no report on the association between glycolysis/gluconeogenesis and NOA patients.

In a total of 50 hub genes obtained by five different algorithms in the cytoHubba plug-in, we obtained 15 genes that overlap in multiple algorithms. After consulting database documents such as NCBI, for the first time, we identified three genes that might be closely related to spermatogenesis: C22orf23, TSACC, and TTC253.

In order to further study the expression of C22orf23, TSACC, and TTC25 in sperm cells, we used the seurat package of the R software to integrate and analyze the corresponding single-cell sequencing data sets. After comparing the data of the HCL database (Han et al., 2020), we identified 12 kinds of sperm cells. Our research results showed that during spermatogenesis, hub genes were expressed in late primary spermatocyte, round spermatid, elongated spermatid, and sperm. The expression of TSACC was of the highest level. In differentiating spermatogonia and early primary spermatocyte cells, TTC25 was almost not expressed, while C22orf23 was lowly expressed, and expression of TSACC is higher than the former two. The above research showed that the expression of the three genes was different in the process of sperm development, and there might be differences in their mechanisms that affect spermatogenesis. The potential mechanisms of spermatogenesis mainly include: (1) the functional structure and cell composition of spermatogenesis, such as blood testis barrier, seminiferous tubules, spermatogenic cells, and testicular somatic cells, etc.; (2) the whole process of spermatogenesis and sperm kinetics of occurrence; (3) endocrine regulation of spermatogenesis; (4) testicular local regulation of spermatogenesis, etc. (Caroppo and Colpi, 2021; Park and Pang, 2021; Rodriguez-Casuriaga and Geisinger, 2021). Differentiating spermatogonia and early primary spermatocyte belong to the representative cells of spermatogenesis during meiosis, so the difference in the expression of these three genes is mainly reflected in the meiosis stage.

Sperm cells have clear developmental trajectories. We used the monocle3 package to successfully construct the developmental trajectory of sperm cells in the GSE109037 dataset. The spematogonial stem cell was recognized as the starting point of spermatogenesis, and a branch was found in the elongated spermatid. The trajectory of C22orf23, TSACC, and TTC25

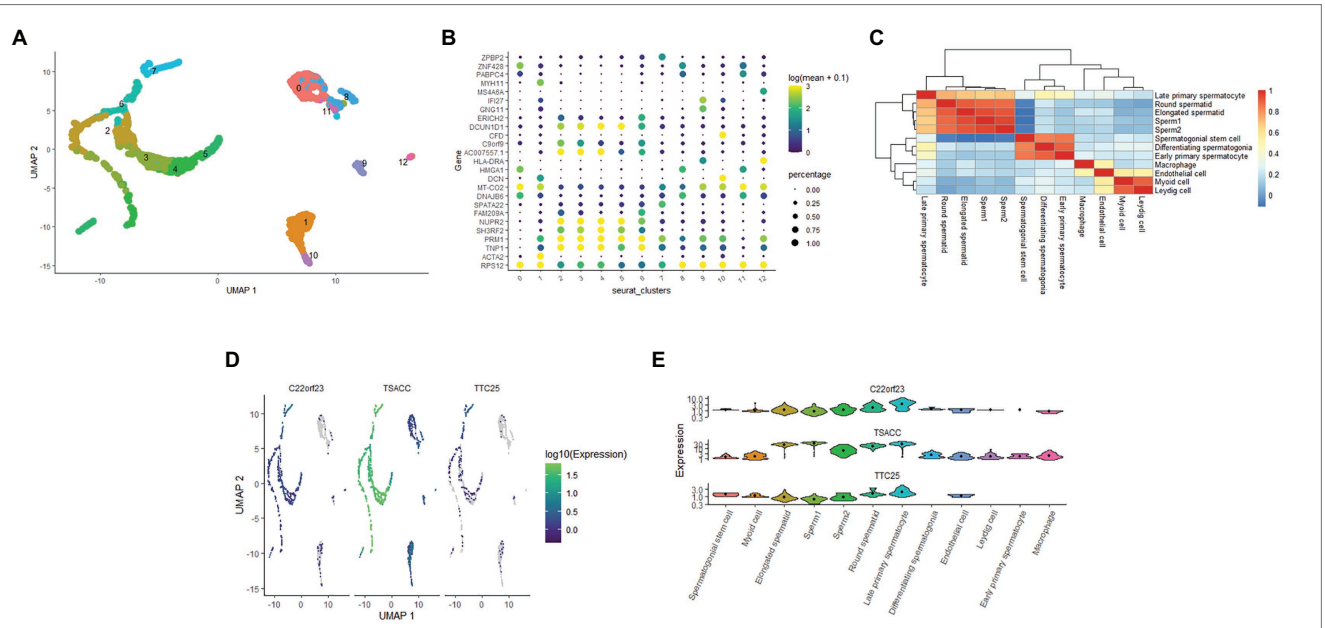


FIGURE 4 | The expression of hub genes in sperm cells. **(A)** Thirteen cell clusters (0–12) of the GSE109037 single-cell sequencing data set. **(B)** The bubble chart showing the marker gene expression of each cell clusters. **(C)** The heat map showing the correlation between each cell clusters. **(D)** Distribution of three hub genes in cell clusters. **(E)** The violin diagram showing the expression of three hub genes in various sperm cells.

genes in sperm development was constructed, and it was found that the expression of TTC25 gene was always maintained at a low level. The C22orf23 gene showed two uplifts during the process from differentiating spermatogonia to sperm production. The TSACC gene showed a slow upward trend during the whole process from spermatogonial stem cell to sperm formation. At present, there are few literatures on the correlation between sperm development and C22orf23 gene. However, the TTC25 gene encodes a tetratricopeptide repeat domain-containing protein, which is located in the ciliary axons and plays a role in the docking of the outer actin arm with the cilia. Mutations in this gene can cause primary ciliary dyskinesia (Wallmeier et al., 2016; Emiralioğlu et al., 2020), but whether it is related to sperm development remains to be further studied. TSACC gene is also called testis-specific serine kinase 6-Activating Co-Chaperone Protein (TSSK6-Activating Co-Chaperone), and TSSK6 belongs to a member of the testis-specific serine/threonine kinase family. Studies have confirmed that TSSK family is expressed after meiosis in male germ cells and mature mammalian sperm. When TSSK family expression is restricted after meiosis, it can affect sperm development through phosphorylation signal transduction (Wang et al., 2015, 2016; Salicioni et al., 2020). Animal experiments have confirmed (Spiridonov et al., 2005) that male mice knocked out of the TSSK6 gene can cause spermatogenesis disorders, including decreased sperm count, decreased motility and survival rate, and increased number of abnormal sperm. Su et al. (2010) studied the relationship between TSSK6 gene mutation and human spermatogenesis and found that TSSK6 gene polymorphism is closely related to male infertility. TSACC is a Co-Chaperone that activates TSSK6, and is closely related to the expression of TSSK6. Based on our analysis results, we speculate that TSACC plays an important role in germ cell differentiation and/or sperm function. TSACC is a Co-Chaperone that specifically activates TSSK6 (Spiridonov et al., 2005). Kula and other scholars (Jha et al., 2010) have shown that TSACC is specifically expressed in the testis, and experiments have confirmed that when sperm cells undergo reorganization and chromatin condensation, TSACC and TSSK6 are in sperm co-localization in the cytoplasm of cells, combined with our analysis results, speculate that TSACC plays an important role in germ cell differentiation.

Non-obstructive azoospermia is one of the most common and important causes of male infertility. Our research used existing public databases to integrate gene expression profile microarray and single-cell sequencing data for bioinformatics analysis. In this way, more reliable and accurate results could be obtained. However, to confirm the relationship between these genes and spermatogenesis, molecular biology experiments are needed.

CONCLUSION

This study identified 430 DEGs in NOA patients, as well as their GO functional annotations and signaling pathways, and further identified three hub genes that might play important roles in sperm development, which deserved further in-depth study. These new findings may provide important enlightenment for revealing the pathogenesis of NOA. These hub genes may be biomarkers suggesting abnormal spermatogenesis, providing novel options for precise diagnosis and precise treatment of NOA.

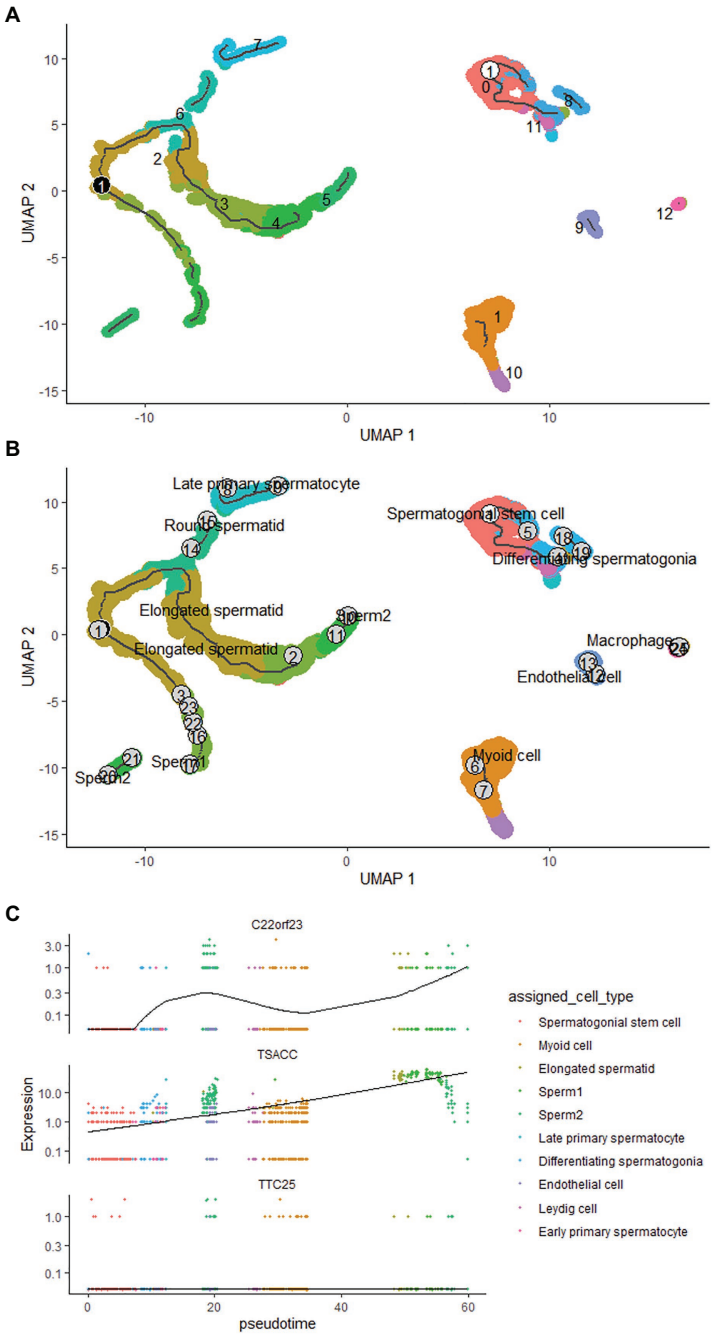


FIGURE 5 | The pseudotime time analysis of sperm development. **(A)** The developmental trajectory of sperm cells. The ① on the white background indicates that the development of sperm starts from cluster 0 cells, and the ① on the black background indicates branches that appear in the sperm development trajectory. **(B)** The sperm cells of each developmental stage are marked, the numbers indicate the sequence of each sperm cell's development trajectory, and words indicate the result of annotating the clustered cells. **(C)** The results of trajectory analysis of the three hub genes in sperm development, which covers the whole process from spermatogonial stem cell to sperm formation.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

HH, YZ, XJ, and FY launched the study. HH and FY performed the data analysis. WS, KC, LZ, QZ, SL, XY, and SC participated in reference collecting. HH, FY, and WS

completed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The study received grants from Research Project in Health Care of Bao'an, Shenzhen, Guangdong (no. 2016CX300). The study is supported by Shenzhen Hospital of Southern Medical University, Research Promotion Funds for the Key Discipline Construction Program (no. ZDXKKYTS006).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.608629/full#supplementary-material

REFERENCES

Arafat, M., Har-Vardi, I., Harlev, A., Levitas, E., Zeadna, A., Abofoul-Azab, M., et al. (2017). Mutation in TDRD9 causes non-obstructive azoospermia in infertile men. *J. Med. Genet.* 54, 633–639. doi: 10.1136/jmedgenet-2017-104514

Caroppo, E., and Colpi, G. M. (2021). Hormonal treatment of men with nonobstructive azoospermia: what does the evidence suggest? *J. Clin. Med.* 10:387. doi: 10.3390/jcm10030387

Choi, I., Woo, J. M., Hong, S., Jung, Y. K., and Cho, C. (2003). Identification and characterization of ADAM32 with testis-predominant gene expression. *Gene* 304, 151–162. doi: 10.1016/S0378-1119(02)01202-7

Corona, G., Minhas, S., Giwercman, A., Bettocchi, C., Dinkelman-Smit, M., Dohle, G., et al. (2019). Sperm recovery and ICSI outcomes in men with non-obstructive azoospermia: a systematic review and meta-analysis. *Hum. Reprod. Update* 25, 733–757. doi: 10.1093/humupd/dmz028

Emiralioğlu, N., Taşkıran, E. Z., Koşukcu, C., Bilgiç, E., Atilla, P., Kaya, B., et al. (2020). Genotype and phenotype evaluation of patients with primary ciliary dyskinesia: first results from Turkey. *Pediatr. Pulmonol.* 55, 383–393. doi: 10.1002/ppul.24583

Fang, F., Li, Z., Zhao, Q., Ye, Z., Gu, X., Pan, F., et al. (2020). Induced pluripotent stem cells derived from two idiopathic azoospermia patients display compromised differentiation potential for primordial germ cell fate. *Front. Cell Dev. Biol.* 8:432. doi: 10.3389/fcell.2020.602574

Ghazzawi, I. M., Alhasani, S., Taher, M., and Souso, S. (1999). Reproductive capacity of round spermatids compared with mature spermatozoa in a population of azoospermic men. *Hum. Reprod.* 14, 736–740. doi: 10.1093/humrep/14.3.736

Gifford, J. A. (2015). The role of WNT signaling in adult ovarian folliculogenesis. *Reproduction* 150, R137–R148. doi: 10.1530/REP-14-0685

Goto, C., Tamura, K., Nishimaki, S., Maruyama, D., and Hara-Nishimura, I. (2020). The nuclear envelope protein KAKU4 determines the migration order of the vegetative nucleus and sperm cells in pollen tubes. *J. Exp. Bot.* 71, 6273–6281. doi: 10.1093/jxb/eraa367

Han, X., Zhou, Z., Fei, L., Sun, H., Wang, R., Chen, Y., et al. (2020). Construction of a human cell landscape at single-cell level. *Nature* 581, 303–309. doi: 10.1038/s41586-020-2157-4

Jan, S. Z., Vormer, T. L., Jongejan, A., Röling, M. D., Silber, S. J., de Rooij, D. G., et al. (2017). Unraveling transcriptome dynamics in human spermatogenesis. *Development* 144, 3659–3673. doi: 10.1242/dev.152413

Javadian-Elyaderani, S., Ghaedi, K., Tavalaee, M., Rabiee, F., Deemeh, M. R., and Nasr-Esfahani, M. H. (2016). Diagnosis of genetic defects through parallel assessment of PLCzeta and CAPZA3 in infertile men with history of failed oocyte activation. *Iran. J. Basic Med. Sci.* 19, 281–289.

Jha, K. N., Wong, L., Zerfas, P. M., De Silva, R. S., Fan, Y. X., Spiridonov, N. A., et al. (2010). Identification of a novel HSP70-binding cochaperone critical to HSP90-mediated activation of small serine/threonine kinase. *J. Biol. Chem.* 285, 35180–35187. doi: 10.1074/jbc.M110.134767

Kavarthapu, R., Anbazhagan, R., Sharma, A. K., Shiloach, J., and Dufau, M. L. (2020). Linking phospho-gonadotropin regulated testicular RNA helicase (GRTH/DDX25) to histone ubiquitination and acetylation essential for spermatid development during spermiogenesis. *Front. Cell Dev. Biol.* 8:310. doi: 10.3389/fcell.2020.00310

Kohn, T. P., and Pastuszak, A. W. (2018). Non-obstructive azoospermia and shortened leukocyte telomere length: further evidence linking poor health and infertility. *Fertil. Steril.* 110, 629–630. doi: 10.1016/j.fertnstert.2018.06.013

La, H. M., and Hobbs, R. M. (2019). Mechanisms regulating mammalian spermatogenesis and fertility recovery following germ cell depletion. *Cell. Mol. Life Sci.* 76, 4071–4102. doi: 10.1007/s00018-019-03201-6

Law, N. C., Oatley, M. J., and Oatley, J. M. (2019). Developmental kinetics and transcriptome dynamics of stem cell specification in the spermatogenic lineage. *Nat. Commun.* 10:2787. doi: 10.1038/s41467-019-10596-0

Lim, S., Kierzek, M., O'Connor, A. E., Brenker, C., Merriner, D. J., Okuda, H., et al. (2019). CRISP2 is a regulator of multiple aspects of sperm function and male fertility. *Endocrinology* 160, 915–924. doi: 10.1210/en.2018-01076

Liu, Y., Guo, Y., Song, N., Fan, Y., Li, K., Teng, X., et al. (2015). Proteomic pattern changes associated with obesity-induced asthenozoospermia. *Andrology* 3, 247–259. doi: 10.1111/andr.289

Malcher, A., Rozwadowska, N., Stokowy, T., Kolanowski, T., Jedrzejczak, P., Zietkowiak, W., et al. (2013). Potential biomarkers of nonobstructive azoospermia identified in microarray gene expression analysis. *Fertil. Steril.* 100, 1686–1694. doi: 10.1016/j.fertnstert.2013.07.1999

Malla, A. B., and Bhandari, R. (2017). IP6K1 is essential for chromatoid body formation and temporal regulation of Tnp2 and Prm2 expression in mouse spermatids. *J. Cell Sci.* 130, 2854–2866. doi: 10.1242/jcs.204966

Moye, A. R., Bedoni, N., Cunningham, J. G., Sanzhaeva, U., Tucker, E. S., Mathers, P., et al. (2019). Mutations in ARL2BP, a protein required for ciliary microtubule structure, cause syndromic male infertility in humans and mice. *PLoS Genet.* 15:e1008315. doi: 10.1371/journal.pgen.1008315

Neto, F. T., Bach, P. V., Najari, B. B., Li, P. S., and Goldstein, M. (2016). Spermatogenesis in humans and its affecting factors. *Semin. Cell Dev. Biol.* 59, 10–26. doi: 10.1016/j.semcdb.2016.04.009

Oud, M. S., Okutman, Ö., Hendricks, L. A., de Vries, P. F., Houston, B. J., Vissers, L. E., et al. (2020). Exome sequencing reveals novel causes as well as new candidate genes for human globozoospermia. *Hum. Reprod.* 35, 240–252. doi: 10.1093/humrep/dez246

Park, H. J., Lee, W. Y., Park, C., Hong, K. H., Kim, J. H., and Song, H. (2018). Species-specific expression of phosphoglycerate kinase 2 (PGK2) in the developing porcine testis. *Theriogenology* 110, 158–167. doi: 10.1016/j.theriogenology.2018.01.007

Park, Y. J., and Pang, M. G. (2021). Mitochondrial functionality in male fertility: from spermatogenesis to fertilization. *Antioxidants* 10:98. doi: 10.3390/antiox10010098

Rodriguez-Casuriaga, R., and Geisinger, A. (2021). Contributions of flow cytometry to the molecular study of spermatogenesis in mammals. *Int. J. Mol. Sci.* 22:1151. doi: 10.3390/ijms22031151

Salicioni, A. M., Gervasi, M. G., Sosnik, J., Tourzani, D. A., Nayyab, S., Caraballo, D. A., et al. (2020). Testis-specific serine kinase protein family in male fertility and as targets for non-hormonal male contraceptiondagger. *Biol. Reprod.* 103, 264–274. doi: 10.1093/biolre/ioaa064

Saraswat, M., Joenväärä, S., Jain, T., Tomar, A. K., Sinha, A., Singh, S., et al. (2017). Human spermatozoa quantitative proteomic signature classifies normo- and asthenozoospermia. *Mol. Cell. Proteomics* 16, 57–72. doi: 10.1074/mcp.M116.061028

Silber, S. J., Van Steirteghem, A., Nagy, Z., Liu, J., Tournaye, H., and Devroey, P. (1996). Normal pregnancies resulting from testicular sperm extraction and intracytoplasmic sperm injection for azoospermia due to maturation arrest. *Fertil. Steril.* 66, 110–117. doi: 10.1016/S0015-0282(16)58396-4

Singh, K., and Jaiswal, D. (2013). One-carbon metabolism, spermatogenesis, and male infertility. *Reprod. Sci.* 20, 622–630. doi: 10.1177/1933719112459232

Singh, V., Jaiswal, D., Singh, K., Trivedi, S., Agrawal, N. K., Gupta, G., et al. (2019). Azoospermic infertility is associated with altered expression of DNA repair genes. *DNA Repair* 75, 39–47. doi: 10.1016/j.dnarep.2019.01.006

Sofikitis, N. V., Yamamoto, Y., Miyagawa, I., Mekras, G., Mio, Y., Toda, T., et al. (1998). Ooplasmic injection of elongating spermatids for the treatment of non-obstructive azoospermia. *Hum. Reprod.* 13, 709–714. doi: 10.1093/humrep/13.3.709

Song, J., Gu, L., Ren, X., Liu, Y., Qian, K., Lan, R., et al. (2020). Prediction model for clinical pregnancy for ICSI after surgical sperm retrieval in different types of azoospermia. *Hum. Reprod.* 35, 1972–1982. doi: 10.1093/humrep/deaa163

Spiridonov, N. A., Wong, L., Zerfas, P. M., Starost, M. F., Pack, S. D., Paweletz, C. P., et al. (2005). Identification and characterization of SSTK, a serine/threonine protein kinase essential for male fertility. *Mol. Cell. Biol.* 25, 4250–4261. doi: 10.1128/MCB.25.10.4250-4261.2005

Su, D., Zhang, W., Yang, Y., Zhang, H., Liu, Y. Q., Bai, G., et al. (2010). c.822+126T>G/C: a novel triallelic polymorphism of the TSSK6 gene associated with spermatogenic impairment in a Chinese population. *Asian J. Androl.* 12, 234–239. doi: 10.1038/aja.2009.80

Tapia Contreras, C., and Hoyer-Fender, S. (2019). CCDC42 localizes to manchette, HTCA and tail and interacts with ODF1 and ODF2 in the formation of the male germ cell cytoskeleton. *Front. Cell Dev. Biol.* 7:151. doi: 10.3389/fcell.2019.00151

Trigg, N. A., Eamens, A. L., and Nixon, B. (2019). The contribution of epididymosomes to the sperm small RNA profile. *Reproduction* 157, R209–R223. doi: 10.1530/REP-18-0480

Vij, S. C., Sabanegh, E. Jr., and Agarwal, A. (2018). Biological therapy for non-obstructive azoospermia. *Expert Opin. Biol. Ther.* 18, 19–23. doi: 10.1080/14712598.2018.1380622

Wallmeier, J., Shiratori, H., Dougherty, G. W., Edelbusch, C., Hjeij, R., Loges, N. T., et al. (2016). TTC25 deficiency results in defects of the outer dynein arm docking machinery and primary ciliary dyskinesia with left-right body asymmetry randomization. *Am. J. Hum. Genet.* 99, 460–469. doi: 10.1016/j.ajhg.2016.06.014

Wang, X., Li, H., Fu, G., Wang, Y., Du, S., Yu, L., et al. (2016). Testis-specific serine/threonine protein kinase 4 (Tssk4) phosphorylates Odf2 at Ser-76. *Sci. Rep.* 6:22861. doi: 10.1038/srep22861

Wang, W., Tu, C., Nie, H., Meng, L., Li, Y., Yuan, S., et al. (2019b). Biallelic mutations in CFAP65 lead to severe asthenoteratospermia due to acrosome hypoplasia and flagellum malformations. *J. Med. Genet.* 56, 750–757. doi: 10.1136/jmedgenet-2019-106031

Wang, H., Wang, G., Dai, Y., Li, Z., Zhu, Y., and Sun, F. (2019a). Functional role of GKAP1 in the regulation of male germ cell spontaneous apoptosis and sperm number. *Mol. Reprod. Dev.* 86, 1199–1209. doi: 10.1002/mrd.23236

Wang, X. L., Wei, Y. H., Fu, G. L., and Yu, L. (2015). Testis specific serine/threonine protein kinase 4 (TSSK4) leads to cell apoptosis relying on its kinase activity. *J. Huazhong Univ. Sci. Technolog. Med. Sci.* 35, 235–240. doi: 10.1007/s11596-015-1417-2

Wosnitzer, M., Goldstein, M., and Hardy, M. P. (2014). Review of azoospermia. *Spermatogenesis* 4:e28218. doi: 10.4161/spmg.28218

Yuan, S., Liu, Y., Peng, H., Tang, C., Hennig, G. W., Wang, Z., et al. (2019). Motile cilia of the male reproductive system require miR-34/miR-449 for development and function to generate luminal turbulence. *Proc. Natl. Acad. Sci. U. S. A.* 116, 3584–3593. doi: 10.1073/pnas.1817018116

Zhang, A. L., Tang, S. F., Yang, Y., Li, C. Z., Ding, X. J., Zhao, H., et al. (2020). Histone demethylase JHDM2A regulates H3K9 dimethylation in response to arsenic-induced DNA damage and repair in normal human liver cells. *J. Appl. Toxicol.* 40, 1661–1672. doi: 10.1002/jat.4026

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

frontiers in Genetics

CASE REPORT
published: 05 March 2021
doi: 10.3389/fgene.2020.589899

Case Report: Identification of a *de novo* Missense Mutation in the *F8* Gene, p.(Phe690Leu)/c.2070C > A, Causing Hemophilia A: A Case Report

*Haiyan Bai, Xia Xue, Li Tian, Xi Tong Liu and Qian Li**

Assisted Reproductive Center, Women's & Children's Hospital of Northwest, Xi'an, China

Hemophilia A is an X-linked recessive bleeding disorder caused by various types of pathological defects in the factor VIII gene (*F8*/FVIII). Preimplantation genetic testing for monogenic disease (PGT-M) is a powerful tool to tackle the transmission of monogenic inherited disorders from generation to generation. In our case, a mutation in *F8* had passed through female carriers in a hemophilia A family and resulted in two male patients with hemophilia A. To identify the etiological genetic variants of *F8*, next-generation sequencing (NGS) was used for chromosome copy number variation detection, Sanger sequencing to verify mutation sites, single nucleotide polymorphism (SNP) for site amplification, and sequencing to validate the genetic linkage. Finally, a novel missense mutation, p. (Phe690Leu)/c.2070C > A, occurring in exon 13 of *F8*, was screened out as a pathogenic mutation. Following this, an *F8* normal euploid blastocyst was transferred. At the 18th week, the pregnant mother underwent amniocentesis, NGS, Sanger sequencing, and SNP typing that further confirmed that the fetus had a healthy genotype. After delivery, a neonatal blood sample was sent for FVIII concentration detection, and the result established that the FVIII protein was rescued to a nearly average level. We first identified a new type of pathogenic mutation in *F8*, which has not been previously reported, selected a genetically healthy progeny for an affected family, and provided valuable knowledge of the diagnosis and treatment of hemophilia A.

Keywords: hemophilia A, preimplantation genetic testing for monogenic disease, p.(Phe690Leu)/c.2070C > A, next-generation sequencing, chromosome copy number variation, single nucleotide polymorphism

OPEN ACCESS

Edited by:
Hu Hao,
The Sixth Affiliated Hospital of Sun Yat-sen University, China

Reviewed by:
Lu-Hong Xu,
Sun Yat-sen Memorial Hospital, China
Liangzhong Sun,
Southern Medical University, China

***Correspondence:**
Qian Li
liqianlinqing@163.com

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

Received: *19 October 2020*
Accepted: *23 November 2020*
Published: *05 March 2021*

Citation:
Bai H, Xue X, Tian L, Liu XT and Li Q (2021) Case Report: Identification of a de novo Missense Mutation in the F8 Gene, p.(Phe690Leu)/c.2070C > A, Causing Hemophilia A: A Case Report. Front. Genet. 11:589899. doi: 10.3389/fgene.2020.589899

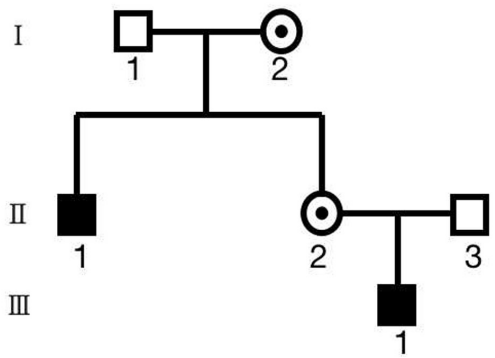
INTRODUCTION

Hemophilia A (OMIM 306700) is an inherited X-linked recessive bleeding disorder. It is caused by the deficiency of blood coagulant activities of factor VIII (FVIII), due to abnormalities in the *F8* coding gene (Keeney et al., 2005). The father with a homozygous mutation will certainly pass the mutation to a female child who would phenotypically be a normal female hemizygous carrier, whereas none of the male children born will be affected. A mother with homozygous mutation will give birth to male children with homozygous mutation and a has a 50% risk of giving birth

to a hemizygous carrier female child. Both male and female descendants could be affected by the disease if both parents are homozygous for the abnormal gene. Therefore, hemophilia A rarely occurs in female individuals, as it always appears in a heterozygote form. Conversely, the incidence is estimated to be 1:5,000–10,000 in male individuals (Hallden et al., 2012). Activity and circulating plasma levels of FVIII protein are used to evaluate the severity of hemophilia A, which can be classified as severe, moderate, and mild, corresponding to FVIII protein levels of ≤1, 2–5, and 5–30%, respectively (Srivastava et al., 2013).

The *F8* gene (MIM + 300841) is located at the distal end of the long arm of the X chromosome (Xq28), which

A



B Blood coagulation state testing of Ⅲ-1

Item code	Item	Result	Reference	Unit
XII%	Factor XII coagulant activity	25.1↓	71.7-113.1	%
XII%	Factor XI coagulant activity	45.8↓	81.6-118.4	%
IX%	Factor IX coagulant activity	17.9↓	67.7-128.5	%
VIII%	Factor VIII coagulant activity	9.5↓	77.3-128.7	%

C

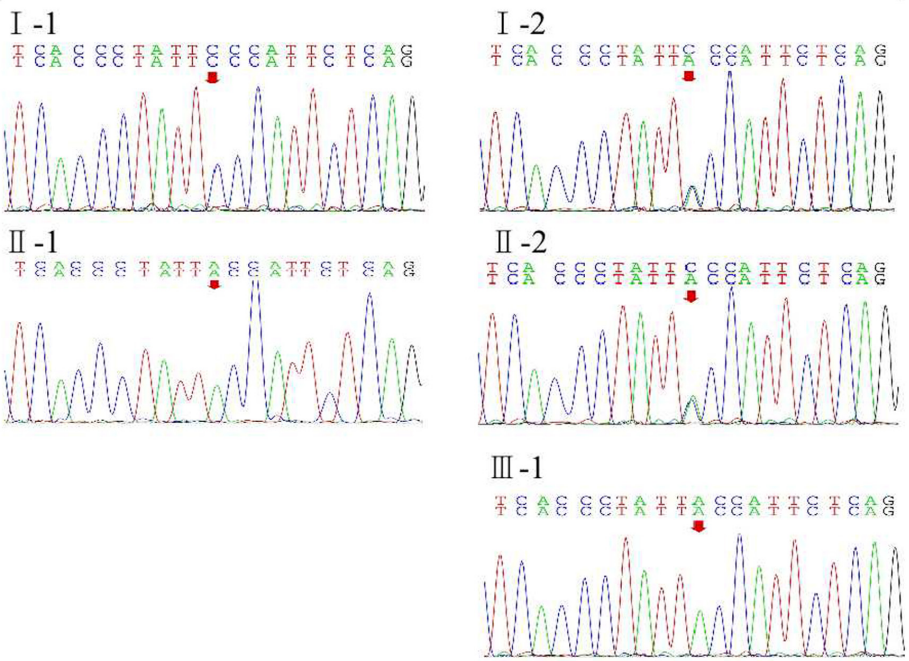


FIGURE 1 | (A) Pedigree of the family. The white square represents healthy male individuals; the black square indicates affected male individuals; the black circle with a dot in the middle indicates carrier female individuals. **(B)** The inspection report of III-1. **(C)** Mutation screening of the principle of family members using Sanger sequencing. The red arrow indicates the mutation site of (Phe690Leu)/c.2070C > A of *F8* at an individual person.

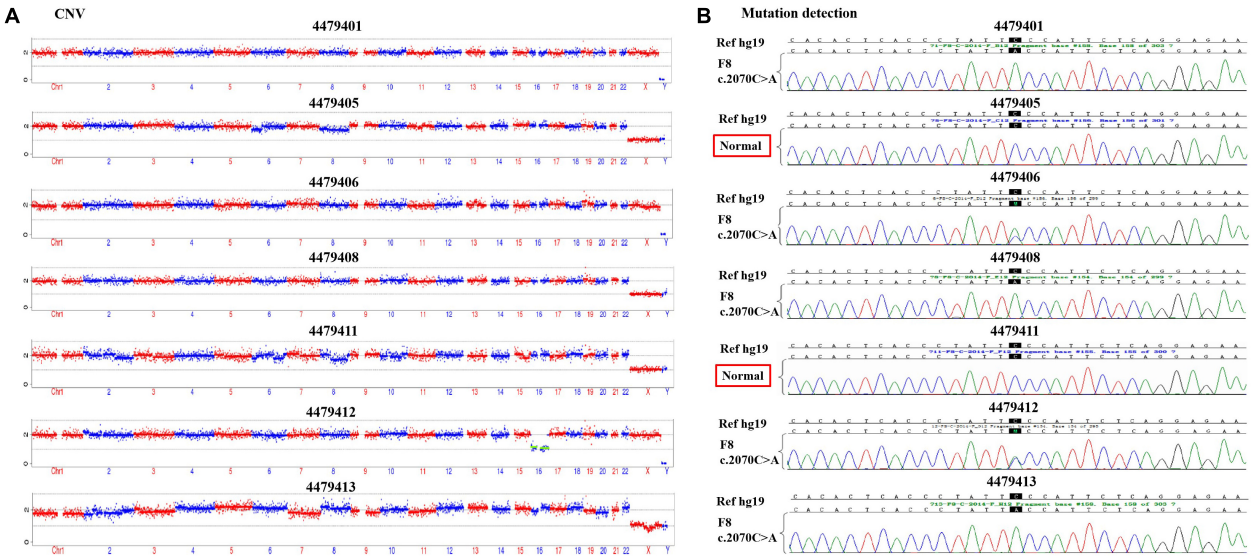


FIGURE 2 | (A) The copy number variation (CNV) detection and **(B)** the mutation screening of the seven blastocysts.

is 186 kb in size [hg19: chrX:154064064-154250998; UCSC genome browser[1]], comprising 26 exons and 25 introns. Since the first publication of the *F8* sequence in 1984, more than 2,000 gene mutations corresponding to 5,472 individual case reports of hemophilia A have been reported. These mutations are described in the Human Gene Mutation Database (HGMD[2]) and Factor VIII Variant Database[3]. Currently, different heterogeneous genetic mutations in the *F8* gene cause hemophilia A, including inversion in intron 22 and intron 1, point mutations, small deletions/insertions/duplications, and large deletions/duplications. Point mutations caused by missense mutations usually result in 57 mild or moderate illness. Other types of variants always cause severe outcomes (Peyvandi et al., 2013; Lannoy and Hermans, 2016). Therefore, the identification of *F8* mutations can be crucial in genetic counseling and prenatal diagnosis.

In this study, with the use of next-generation sequencing (NGS), Sanger sequencing, and single nucleotide polymorphism (SNP) typing, we reported the identification and characteristics of a novel missense mutation (p.(Phe690Leu)/c.2070C > A) occurring in exon 13 of *F8* in a hemophilia A family, which caused the disease. This mutation has not been previously reported in HGMD and PubMed.

CASE DESCRIPTION

Assessment of Family F8 Mutation

A woman (II-2) presented with no apparent clinical phenotype of hemophilia A came to our assisted reproduction center for PGT-M counseling because her son (III-1) and her brother (II-1) were both affected by hemophilia A (**Figure 1A**). The activated partial thromboplastin time (APTT) in her son was 121.5 s, and FVIII coagulant activity was 9.50%, corresponding to mild hemophilia A. The hemophilia A patients suffered from recurrent hemarthrosis and continuous bleeding after an injury (**Figure 1B** and **Supplementary Figure S1**).

To recognize the potential etiological genetic variants, we screened the *F8* gene in her family. The first stage was the pedigree pretest. Genomic DNA was isolated from blood samples based on the principle of family members. Sanger sequencing was used to screen for *F8* mutations in these family members. The result is shown in **Figure 1C**; a variation, c.2070C > A in exon 13 of *F8*, was found in the mother (II-2). The results also revealed that the mutated nucleotide A was acquired from the II-2's mother (I-2) and therefore passed to the brother (II-1), the mother herself (II-2), and the mother's son (III-1). The father and husband of II-2 (I-1 and II-3) had a normal *F8* gene. II-2 and her mother were heterozygous for the mutation without disease phenotype. However, her brother and her son were homozygotes for the mutation, and both were affected by hemophilia A. The mutation results in an amino acid change of phenylalanine to leucine (p. Phe690Leu). The current knowledge of this kind of mutation suggests that it may have an impact on the configuration and conformation of FVIII protein, leading to the loss of its function. This is likely to be a pathogenic cause of hemophilia A (Gouw et al., 2012; Chen et al., 2016; Stephensen et al., 2019).

PGT-M Procedure

According to the mother-II-2's wish, we performed intracytoplasmic sperm injection with PGT-M for her assisted reproduction trial. Finally, seven blastocysts (4479401, 4479405, 4479406, 4479408, 4479411, 4479412, and 4479413) were obtained, and trophectoderm biopsy was performed and sent

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

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

[3]http://www.factorviii-db.org/

TABLE 1 | Single nucleotide polymorphism (SNP) typing of the family and preimplantation genetic testing for monogenic disease (PGT-M) embryos.

SNP Number	I-1	I-2	II-1	II-2 normal	II-2 carry	4479401		4479405	4479406		4479408	4479411	4479412		4479413
	Haplotype	Haplotype	Haplotype	Haplotype	Haplotype	Maternal haplotype	Maternal haplotype	Maternal haplotype	Paternal haplotype	Maternal haplotype	Maternal haplotype	Maternal haplotype	Paternal haplotype	Maternal haplotype	Maternal haplotype
YK-F8-SNP01	T	T/C	C	T	C	C	C	T	C	C	C	T	C	C	C
YK-F8-SNP02	C	C/A	A	C	A	A	A	C	A	A	A	C	A	A	A
YK-F8-SNP07	C	A/A	A	C	A	A	A	C	A	A	A	C	A	A	A
YK-F8-SNP13	C	T/T	C	C	T	T	T	C	C	T	T	C	C	T	T
YK-F8-SNP14	T	C/C	T	T	C	C	C	T	T	C	C	T	T	C	C
YK-F8-SNP15	C	G/C	C	C	G	G	G	C	C	G	G	C	C	G	G
YK-F8-SNP16	T	C/C	T	T	C	C	C	T	T	C	C	T	T	C	C
YK-F8-SNP17	T	G/G	G	T	G	G	G	T	G	G	G	T	G	G	G
YK-F8-SNP18	T	C/C	C	T	C	C	C	T	C	C	C	T	C	C	C
YK-F8-SNP20	T	A/A	T	T	A	A	A	T	T	A	A	T	T	A	A
YK-F8-SNP22	C	C/A	C	C	A	A	A	C	C	A	A	C	C	A	A
YK-F8-SNP23	C	T/T	C	C	T	T	T	C	C	T	T	C	C	T	T
YK-F8-SNP25	C	G/G	C	C	G	G	G	C	C	G	G	C	C	G	G
YK-F8-SNP27	A	G/A	A	A	G	G	G	A	A	G	G	A	A	G	G
YK-F8-SNP32	T	C/T	C	T	C	C	C	T	C	C	C	T	C	C	C
YK-F8-SNP33	T	C/T	C	T	C	C	C	T	C	C	C	T	C	C	C
YK-F8-SNP34	G	G/T	T	G	T	T	T	G	T	T	T	G	T	T	T
YK-F8-SNP35	C	C/T	T	C	T	T	T	C	T	T	T	C	T	T	T
YK-F8-SNP37	G	G/A	A	G	A	A	A	G	A	A	A	G	A	A	A
YK-F8-SNP38	A	A/C	C	A	C	C	C	A	C	C	C	A	C	C	C
YK-F8-SNP39	C	T/C	T	C	T	T	T	C	T	T	T	C	T	T	T
YK-F8-SNP40	T	C/C	T	T	C	C	C	T	T	C	C	T	T	C	C
YK-F8-SNP42	G	A/A	A	G	A	A	A	G	A	A	A	G	A	A	A
YK-F8-SNP44	A	G/G	A	A	G	G	G	A	A	G	G	A	A	G	G
YK-F8-SNP45	C	G/C	C	C	G	G	G	C	C	G	G	C	C	G	G
YK-F8-SNP46	T	C/T	T	T	C	C	C	T	T	C	C	T	T	C	C
YK-F8-SNP47	T	T/A	T	T	A	A	A	T	T	A	A	T	T	A	A
YK-F8-SNP48	T	T/C	T	T	C	C	C	T	T	C	C	T	T	C	C
YK-F8-SNP49	C	G/G	C	C	G	G	G	C	C	G	G	C	C	G	G
YK-F8-SNP50	G	A/A	G	G	A	A	A	G	G	A	A	G	G	A	A
YK-F8-SNP51	G	A/A	G	G	A	A	A	G	G	A	A	G	G	A	A
YK-F8-SNP58	A	A/G	G	A	G	G	G	A	G	G	G	A	G	G	G
YK-F8-SNP61	A	A/G	G	A	G	G	G	A	G	G	G	A	G	G	G
YK-F8-SNP62	C	C/T	T	C	T	T	T	C	T	T	T	C	T	T	T
YK-F8-SNP63	A	A/G	G	A	G	G	G	A	G	G	G	A	G	G	G
YK-F8-SNP65	C	C/T	T	C	T	T	T	C	T	T	T	C	T	T	T
YK-F8-SNP66	A	A/G	G	A	G	G	G	A	G	G	G	A	G	G	G

(Continued)

TABLE 1 | Continued

SNP Number	I-1	I-2	II-1	II-2 normal	II-2 carry	4479401		4479405	4479406		4479408	4479411	4479412		4479413	
	Haplotype	Haplotype	Haplotype	Haplotype	Haplotype	Maternal haplotype	Maternal haplotype	Maternal haplotype	Paternal haplotype	Maternal haplotype	Maternal haplotype	Maternal haplotype	Paternal haplotype	Maternal haplotype	Maternal haplotype	Maternal haplotype
YK-F8-SNP67	A	G/A	G	A	G	G	G	A	G	G	G	A	G	G	G	G
YK-F8-SNP68	C	C/T	T	C	T	T	T	C	T	T	T	C	T	T	T	T
YK-F8-SNP72	A	A/G	G	A	G	G	G	A	G	G	G	A	G	G	G	G
YK-F8-SNP73	G	A/G	A	G	A	A	A	G	A	A	A	G	A	A	A	A
YK-F8-SNP75	C	G/C	G	C	G	G	G	C	G	G	G	C	G	G	G	G
YK-F8-SNP79	G	C/C	G	G	C	C	C	G	G	C	C	G	G	C	C	C
YK-F8-SNP80	T	C/C	T	T	C	C	C	T	T	C	C	T	T	C	C	C
YK-F8-SNP81	T	C/T	T	T	C	C	C	T	T	C	C	T	T	C	C	C

for copy number variation (CNV) testing by NGS, mutation detection through Sanger sequencing, and SNP site amplification and sequencing. First, all blastocysts detected without CNV at 4 M resolution level except for embryo 4479412 [45, XN, -16(× 1)] lost chromosome 16 (**Figure 2A**). Second, after Sanger sequencing, only two out of the seven blastocysts were detected as normal as c.2070C > C in exon 13 of *F8* (**Figure 2B**). Furthermore, SNP typing validated that the two euploid and non-mutation embryos detected in the first two steps inherited the normal X chromosome from their mother (**Table 1**). The two embryos were numbered 4479405 and 4479411 (highlighted by the red rectangular box or red color), and both were male. Embryos 4479401 and 4479406 were female, and embryo 4479408 and 4479413 were male. These four embryos inherited SNP abnormal chromosome X from the mother (II-2), consistent with the Sanger result and thus were excluded.

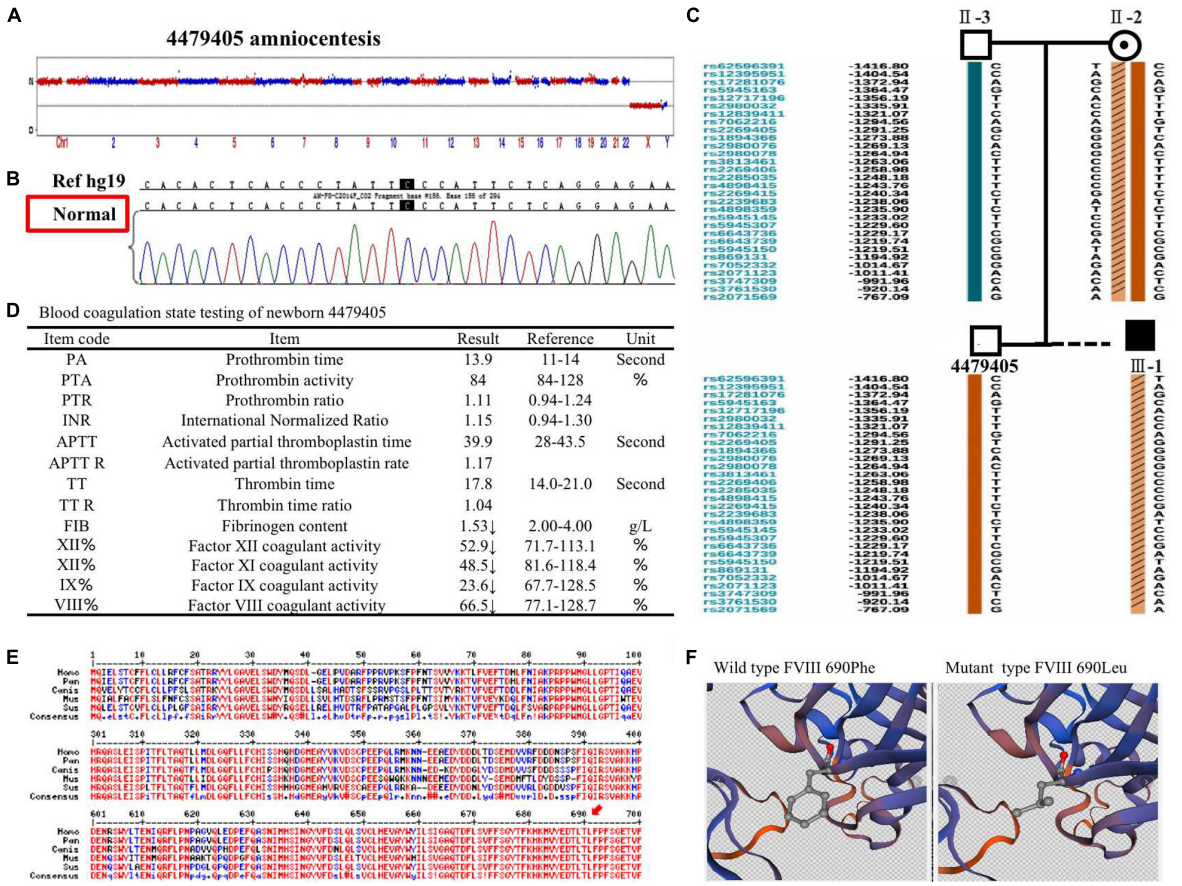
Amniocentesis Testing and Detection of Newborn FVIII Activity

The 4479405 embryo was selected for blastocyst transfer. At the 18th week, amniocentesis was applied to the mother-II-2 to further confirm the genetic health of the fetus. The result was the same as that in the blastocyst examination. There was no significant abnormal CNV detection at the 4 M resolution level as well as no genetic mutation (**Figures 3A,B**). Regarding the SNP typing, as illustrated in **Figure 3C**, the fetus 4479405 inherited the SNP normal X chromosome from the mother, while the previous III-1, who was affected by hemophilia A, inherited the X chromosome with the pathogenic SNP.

After delivery, the FVIII concentration of newborn 4479405 was analyzed after 1 month. The result is shown in **Figure 4A**. The FVIII activity was rescued to 66.5% compared with the 9.50% of the first child (III-1) with hemophilia A of the mother-II-2. According to the doctor's opinion, FVIII activity became almost normal (**Figure 3D** and **Supplementary Figure S2**). These results showed that the mutation site p. (Phe690Leu)/c.2070C > A was exactly the etiological site responsible for hemophilia A. The protein conservation of Phe690 of F8 showed high consensus among *Homo sapiens*, *Pan troglodytes*, *Canis lupus familiaris*, *Mus musculus*, and *Sus scrofa*; high consensus was labeled as red color (**Figure 3E**, arrow). The possible damaging effect of the mutation to FVIII protein 3D structure and software-predicted damaging parameters are illustrated in **Figure 3F** and listed in **Supplementary Figure S3**. Compared with wild-type FVIII, the mutant 690 Leu FVIII lacks the benzene ring of the original phenylalanine; the bioinformatic parameters all indicate that such change has harmful effect on FVIII protein.

DISCUSSION

Due to the X-linked recessive mode of inheritance, hemophilia A usually affects male individuals, while the female family members commonly present as deficient gene carriers who may pass the mutation gene to their progeny (Lacroix-Desmazes et al., 2020; Ling and Tuddenham, 2020).



Item code	Item	Result	Reference	Unit
PA	Prothrombin time	13.9	11-14	Second
PTA	Prothrombin activity	84	84-128	%
PTR	Prothrombin ratio	1.11	0.94-1.24	
INR	International Normalized Ratio	1.15	0.94-1.30	
APTT	Activated partial thromboplastin time	39.9	28-43.5	Second
APTT R	Activated partial thromboplastin rate	1.17		
TT	Thrombin time	17.8	14.0-21.0	Second
TT R	Thrombin time ratio	1.04		
FIB	Fibrinogen content	1.53↓	2.00-4.00	g/L
XII%	Factor XII coagulant activity	52.9↓	71.7-113.1	%
XII%	Factor XI coagulant activity	48.5↓	81.6-118.4	%
IX%	Factor IX coagulant activity	23.6↓	67.7-128.5	%
VIII%	Factor VIII coagulant activity	66.5↓	77.1-128.7	%

FIGURE 3 | (A) The copy number variation (CNV), **(B)** mutation, and **(C)** single nucleotide polymorphism (SNP) detection of amniotic fluid at the 18th pregnancy week after transfer blastocyst 4479405. **(D)** The inspection report of newborn 4479405. **(E)** The consensus of 690 Phe of *F8* among different specials. **(F)** The 3D structure of wild-type FVIII and mutant FVIII. The red arrow represents the conservation of the site Phe690 of *F8* among different species.

Based on the present report, we detected a novel missense mutation occurring in exon 13 of the *F8* gene, p. (Phe690Leu)/c.2070C > A. In the inspected family, this novel mutation had been passed on by the phenotypically normal heterozygote female family members (I-2 and II-2). Male descendants were all affected by inheriting their mother's deficient gene (II-1 and III-1). In the PGT-M treatment cycle of mother-II-2, two out of the seven embryos were confirmed as euploid as well as possessing a non-mutated *F8* gene through CNV testing, mutation detection, and SNP typing. The genetic selection of embryo transfer led to an *F8* normal fetus as seen at the 18th week. Finally, the FVIII concentration was rescued to nearly normal levels after the baby was born. Through PGT-M diagnosis and embryo selection, we found a novel *F8* mutation and ensured that a normal *F8* gene and a healthy CNV were present. The outcome provided us with the confidence that the mutation we found was pathogenic, which is relevant to hemophilia A knowledge.

One of the major genetic challenges is to distinguish causal mutations from polymorphisms. The mutation c.2070C > A in exon 13 was not registered in the SNP database (dbSNP) and has not been reported in HGMD and PubMed. Thus, the mutation was recognized as a novel *F8* mutation. This type of mutation may be induced during the normal process of DNA replication, termed a replication error, which occurs in a remarkably low frequency, or may be caused by cell metabolism through depurination/depyrimidination and deamination, called a spontaneous lesion. It was estimated that it occurs in 10,000–1,000,000 nucleotides per cell per day (de Ligt et al., 2013; Lannoy and Hermans, 2016).

The effect of p. (Phe690Leu)/c.2070C > A mutation on FVIII protein was confirmed by bioinformatics analysis (**Figures 3E,F** and **Supplementary Figure S3**). It is generally assumed that the more conservative the locus is, the more likely it is to cause harmful results when it is mutant. The Phe690 site of FVIII protein was highly conserved among different specials, which means that it has a great possibility to affect protein function when changed. In a previous study, the mutation of Phe to Leu in allene oxide synthase (AOS) changes its catalysis activity to epoxyalcohol synthase (EAS), which is another fatty acid hydroperoxide metabolizing enzymes of the CYP74 family (Toporkova et al., 2020). In another study, the mutation of Phe to Leu in the 3C-like protease (3CLpro) of severe acute respiratory syndrome coronavirus (SARS-CoV) leads to conformational

change and autoinhibition of the enzyme (Muramatsu et al., 2016). In our report, all bioinformatic software parameters indicate that the mutation we found has detrimental effect on protein function.

From the blood coagulation state testing of mother II-2's first children affected with hemophilia A and 4479405 (**Figures 1B**, **3D**), we can see that not only factor VIII decreased but also factor IX apparently declined, which is responsible for hemophilia B. FIX is a serine protease that plays a vital role in the coagulation cascade. Its cofactor, FVIII, is critical for FIX enzymatic activity. They cooperate in coagulation. Therefore, mutations in either protein may lead to lack of coagulation activity in the other, resulting in frequent spontaneous bleeds in patients (Nienhuis et al., 2017).

The strength of this report is to find out a confirmed *de novo* F8 gene mutation that is related to hemophilia A. The limitation is that the mutation we found is only responsible for mild hemophilia A; this depends on the patient we met with. However, whether the mutation also contributes to severe hemophilia A, that we do not know.

CONCLUSION

In conclusion, in this case, the F8 gene mutation-causing hemophilia A has affected two male patients in a family, which was found out through CNV detection by NGS, mutation sites verification by Sanger sequencing, and genetic linkage analysis by SNP. Finally, a novel missense mutation, p. (Phe690Leu)/c.2070C > A, occurring in exon 13 of F8, was screened out as a pathogenic mutation. Following blastocyst transfer of the non-mutation euploid embryo, at the 18th week, the pregnant mother underwent amniocentesis; genetic testing further confirmed that the fetus had a healthy genotype. After delivery, blood testing proved that the FVIII protein has rescued to a nearly average level. This case reveals a new F8 mutation, which provides valuable knowledge for genetic counseling and treatment decisions.

DATA AVAILABILITY STATEMENT

All data and material generated or analyzed during this study are included in this published article.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Northwest Women's and Children's Hospital. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the authors and participated patients for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

HB was in charge of genetic consultation, diagnostic evaluation, and clinical management of the couple and wrote parts of the manuscript. XX participated in the laboratory procedures. LT and XL provided clinical assistance. QL was responsible for the literature review, writing, and coordinating the submission of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the State Natural Science Fund Projects 81771657.

ACKNOWLEDGMENTS

We would like to present our sincere acknowledgment to Yikon Gene Company, who did all the sequencing and bio-information analysis for us.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.589899/full#supplementary-material

REFERENCES

Chen, M., Chang, S.-P., Ma, G.-C., Lin, W.-H., Chen, H.-F., Chen, S.-U., et al. (2016). Preimplantation genetic diagnosis of hemophilia A. *Thromb. J.* 14:33. doi: 10.1186/s12959-016-0098-9

de Ligt, J., Veltman, J. A., and Vissers, L. E. (2013). Point mutations as a source of de novo genetic disease. *Curr. Opin. Genet. Dev.* 23, 257–263. doi: 10.1016/j.gde.2013.01.007

Gouw, S. C., van der Bom, J. G., Oldenburg, J., Astermark, J., de Groot, P. G., Margaglione, M., et al. (2012). F8 gene mutation type and inhibitor development in patients with severe Hemophilia a: systematic review and meta-analysis. *Blood* 119, 2922–2934. doi: 10.1182/blood-2011-09-379453

Hallden, C. L., Knobe, K. E., Sjorin, E., Nilsson, D., and Ljung, R. (2012). Investigation of disease-associated factors in haemophilia A patients without detectable mutations. *Haemophilia* 18, e132–e137. doi: 10.1111/j.1365-2516.2011.02737.x

Keeney, S., Mitchell, M., Goodeve, A., and UK Haemophilia Center Doctors' Organization Haemophilia Genetics Laboratory Network (2005). The molecular analysis of haemophilia A: a guideline from the UK haemophilia centre doctors'organization haemophilia genetics laboratory network. *Haemophilia* 11, 387–397. doi: 10.1111/j.1365-2516.2005.01111.x

Lacroix-Desmazes, S., Voorberg, J., Lillicrap, D., Scott, D. W., and Pratt, K. P. (2020). Tolerating factor VIII:Recent progress. *Front. Immunol.* 10:2991. doi: 10.3389/fimmu.2019.02991

Lannoy, N., and Hermans, C. (2016). Principles of genetic variations and molecular diseases: applications in Hemophilia A. *Crit. Rev. Oncol. Hematol.* 104, 1–8. doi: 10.1016/j.critrevonc.2016.04.005

Ling, G., and Tuddenham, E. G. D. (2020). Factor VIII: the protein, cloning its gene, synthetic factor and now-35 years alter-gene therapy; what happened in between. *Br. J. Haematol.* 189, 400–407. doi: 10.1111/bjh.16311

Muramatsu, T., Takemoto, C., Kim, Y. T., Wang, H., Nishii, W., Terada, T., et al. (2016). SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel subsite cooperativity. *Proc. Natl. Acad. Sci. U.S.A.* 113, 12997–13002. doi: 10.1073/pnas.1601327113

Nienhuis, A. W., Nathwani, A. C., and Davidoff, A. M. (2017). Gene therapy for hemophilia. *Mol. Ther.* 25, 1163–1167. doi: 10.1016/j.ymthe.2017.03.033

Peyvandi, F., Kunicki, T., and Lillicrap, D. (2013). Genetic sequence analysis of inherited bleeding diseases. *Blood* 122, 3423–3431. doi: 10.1182/blood-2013-05-505511

Srivastava, A., Brewer, A. K., Mauser-Bunschoten, E. P., Key, N. S., Kitchen, S., Llinas, A., et al. (2013). Guidelines for the management of hemophilia. *Haemophilia* 19, e1–e47. doi: 10.1111/j.1365-2516.2012.02909.x

Stephensen, D., de Kleijn, P., Matlary, R. E. D., Katzerova, M., McLaughlin, P., Ryan, A., et al. (2019). Scope of practice of haemophilia physiotherapists: A European survey. *Haemophilia* 25, 514–520. doi: 10.1111/hae.13727

Toporkova, Y. Y., Smirnova, E. O., Mukhtarova, L. S., Gorina, S. S., and Grechkin, A. N. (2020). Catalysis by allene oxide synthases (CYP74A and CYP74C): alterations by the Phe/Leu mutation at the SRS-1 region. *Phytochemistry* 169:112152. doi: 10.1016/j.phytochem.2019.112152

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

frontiers in Genetics

CASE REPORT
published: 08 April 2021
doi: 10.3389/fgene.2021.651878

Case Report: Compound Heterozygous Variants in *MOCS3* Identified in a Chinese Infant With Molybdenum Cofactor Deficiency

Qi Tian[1†], *Yang Cao*[2†], *Li Shu*[1,3,4], *Yongjun Chen*[5], *Ying Peng*[1], *Yaqin Wang*[6], *Yuanyuan Chen*[7], *Hua Wang*[1,3*] and *Xiao Mao*[1,3*]

[1] Department of Medical Genetics, Maternal and Child Health Hospital of Hunan Province, Changsha, China, [2] Department of Radiology, Chenzhou First People's Hospital, Chenzhou, China, [3] National Health Commission Key Laboratory of Birth Defects Research, Prevention and Treatment, Hunan Provincial Maternal and Child Health Care Hospital, Changsha, China, [4] Department of School of Life Sciences, Central South University, Changsha, China, [5] Department of Neurology, Nanhua Affiliated Hospital, University of South China, Hengyang, China, [6] Health Management Center, The Third Xiangya Hospital, Central South University, Changsha, China, [7] Reproductive Center of Maternal and Child Health Hospital of Hunan Province, Changsha, China

OPEN ACCESS

Edited by:
Can Liao,
Prenatal Diagnostic Center, Guangzhou Women and Children's Medical Center, China

Reviewed by:
James A. Poulter, University of Leeds, United Kingdom
Shabeesh Balan, RIKEN Center for Brain Science (CBS), Japan

***Correspondence:**
Xiao Mao
gbtechies@outlook.com
Hua Wang
wanghua_213@hotmail.com

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

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

Received: 11 January 2021
Accepted: 12 March 2021
Published: 08 April 2021

Citation:
Tian Q, Cao Y, Shu L, Chen Y, Peng Y, Wang Y, Chen Y, Wang H and Mao X (2021) Case Report: Compound Heterozygous Variants in MOCS3 Identified in a Chinese Infant With Molybdenum Cofactor Deficiency. Front. Genet. 12:651878. doi: 10.3389/fgene.2021.651878

Background: The molybdenum cofactor (Moco) deficiency in humans results in the inactivity of molybdenum-dependent enzymes and is caused by pathogenic variants in *MOCS1* (*Molybdenum cofactor synthesis 1*), *MOCS2* (*Molybdenum cofactor synthesis 2*), and *GPHN* (*Gephyrin*). These genes along with *MOCS3* (*Molybdenum cofactor synthesis 3*) are involved in Moco biosynthesis and providing cofactors to Moco-dependent enzymes. Until now, there was no study to confirm that *MOCS3* is a causative gene of Moco deficiency.

Methods: Detailed clinical information was collected in the pedigree. The Whole-exome sequencing (WES) accompanied with Sanger sequencing validation were performed.

Results: We described the clinical presentations of an infant, born to a non-consanguineous healthy family, diagnosed as having *MOCS3* variants caused Moco deficiency and showing typical features of Moco deficiency including severe neurologic symptoms and cystic encephalomalacia in the brain MRI, resulting in neonatal death. Compound heterozygous variants in the *MOCS3* gene were identified by WES. Positive sulfite and decreased levels of uric acid in plasma and urine were detected.

Conclusion: To our knowledge, this is the first case of *MOCS3* variants causing Moco deficiency. Our study may contribute to genetic diagnosis of Moco deficiency and future genetic counseling.

Keywords: *MOCS3*, neurodevelopmental outcome, sulfite oxidase, whole exome sequencing, molybdenum cofactor deficiency

INTRODUCTION

The molybdenum cofactor (Moco) deficiency in humans results in the inactivity of molybdenum-dependent enzymes including sulfite oxidase, xanthine oxidoreductase, and aldehyde oxidase (Mayr et al., 2021). The inactivation of these enzymes in neurons was related to severe neurological symptoms and residual

diseases (Zaki et al., 2016). The etiology of Moco deficiency was caused by pathogenic variants in *MOCS1* (*Molybdenum cofactor synthesis 1*), *MOCS2* (*Molybdenum cofactor synthesis 2*), and *GPHN* (*Gephyrin*). The typical clinical features of Moco deficiency patients with these variants are characterized by neonatal-onset intractable seizures, followed by feeding difficulty, developmental delay, and neonatal deaths (Mechler et al., 2015).

The signs and symptoms of patients with Moco deficiency are attributed to the pathogenesis of inactivating the molybdenum-dependent enzymes (Atwal and Scaglia, 2016). The enzyme inactivation subsequently leads to the accumulation of metabolites including sulfite, xanthine, hypoxanthine, etc. It is believed that the pathogenesis is caused by accumulated metabolites, and it may exert toxic effects on neurons and lead to seizures, encephalopathy, and other neurological symptoms (Mechler et al., 2015).

Until now, there have been no cases reported with the *MOCS3 (Molybdenum cofactor synthesis 3)* variation that presented with the typical features of Moco deficiency. The only reported case related to *MOCS3* showed a slight change in sulfite metabolites, and as a result the patient only presented mild neurologic disorder symptoms such as high muscular tone and limited speech ability. Except for the thin splenium of corpus callosum, the neurologic imaging was also unremarkable (Huijmans et al., 2017). Therefore, there was not enough evidence to support the pathogenesis of *MOCS3* variation and studies illustrating the phenotypic spectrum of *MOCS3*-induced Moco deficiency are needed.

Here we described a case of Moco deficiency identified with compound heterozygous variations in the *MOCS3* gene. The infant presented aggressive phenotypes of Moco deficiency including early onset of neonatal seizures, hypertonia, feeding difficulty, and severe developmental delay. The MRI indicated global cystic encephalomalacia and cortical necrosis. The patient showed poor reaction to treatment and resulted in infant death. To the best of our knowledge, our report is probably the first study to prove that *MOCS3* is a pathogenic gene of Moco deficiency. Our study may contribute to genetic diagnosis of Moco deficiency and future genetic counseling.

CASE DESCRIPTION

Our patient is the second child born to the non-consanguineous Chinese parents. The first male child presented neonatal seizures, hypotonia, and developmental delay. He died at the age of 8 months and the suspected diagnosis was hypoxic-ischemic encephalopathy, which was made by Magnetic resonance imaging (MRI) presentation.

The second female child was born at full term (40 weeks) with no abnormality in pregnancy. Without any inducement, generialized tonic-clonic seizures were observed within 24 h after birth and seizures were not controlled after multiple treatments. Her after birth physical examination indicated microcephaly [occipitofrontal head circumference 30.0 cm, < -3 Standard deviation (SD)], hypotonia, and hyper-reflexia. At the age of 5 months the patient was admitted to hospital for the reoccurrence of the seizures. She did not develop the ability to lift her head and trace. MRI showed frontotemporal, parietal, and occipital lobes cystic encephalomalacia and bilateral basal ganglia atrophy (**Figure 1**). The patient showed severe neurodevelopmental delay. She was still unable to trace, lift her head, roll over, and sit at 10 months. The patient had feeding difficulties resulting in low breast milk intake and died at age of 12 months because of lung infection.

DIAGNOSTIC ASSESSMENT

Whole-exome sequencing (WES) accompanied with Sanger sequencing validation was performed in the proband and her parents (**Figure 2**). The variants were firstly filtered according to HGMD and ACMG guidelines Disease-causing mutations (DM) and probable/possible pathological mutation (DM) in the HGMD (Professional version 2019.1) database, and pathogenic (P) and likely pathogenic (LP) variants interpreted by ACMG guidelines for interpretation of genetic variants were included. Secondly, variants were filtered according to allele frequency, variant type, and inheritance mode. Variants with minor allele frequencies (MAF) $<0.1\%$, variant depth of coverage ≥ 20 and alteration base depth of coverage ≥ 4 were taken on for further analyses. Filtering was conducted on the remaining variants according to variant type and inheritance model of the associated disease. The proband was identified with compound heterozygous variants in NM_014484.49 (*MOCS3*): c.1375C>T; p.Gln459Ter (chr20:49576754-49576754) and NM_014484.49 (*MOCS3*): c.325C>G; p.Leu109Val (chr20:49575704-49575704). Her father was a heterozygous carrier of the *MOCS3* c.1375C>T variant and the mother carried another variant in the *MOCS3* c.325C>G. All the segregated variants of this case were listed in the **Supplementary Table 1** and the annotation of each variant is given.

The variants are evaluated by MutationTaster (v2.0), CADD (v1.6), FATHMM (v2.3). The scores of each variant of *MOCS3* from the bioinformatic prediction tools are as follows: for c.325C>G the scores are 1.000, 22.8, and 1.19 and for c.1375C>T (premature termination mutation) the scores are 0.958, 36, NA, respectively. The allele frequency of c.325C>G was 6.181e−05 in gnomAD database, 5.182e−05 in the ExAC and 0 in the 1,000 genomes database. The allele frequency of c.1375C>T was 0 in gnomAD the ExAC and the 1,000 genomes databases.

Moco deficiency-related laboratory tests were conducted. The urine and plasma uric acid levels were detected by Beckman Coulter AU680 and enzymatic colorimetric assay on the Cobas 8000 system. The urine sulfite was tested by a urine sulfite test strip. The urine uric acid level was 32 mmol/mol creatinine (normal range 350–2,500 mmol/mol creatinine) and plasma uric acid level was 0.01 mmol/l (normal range 0.08–0.37 mmol/l). The urine sulfite test strip was positive.

DISCUSSION

Moco deficiency is characterized by severe neonatal neuropathologic symptoms, typical neuroimaging findings

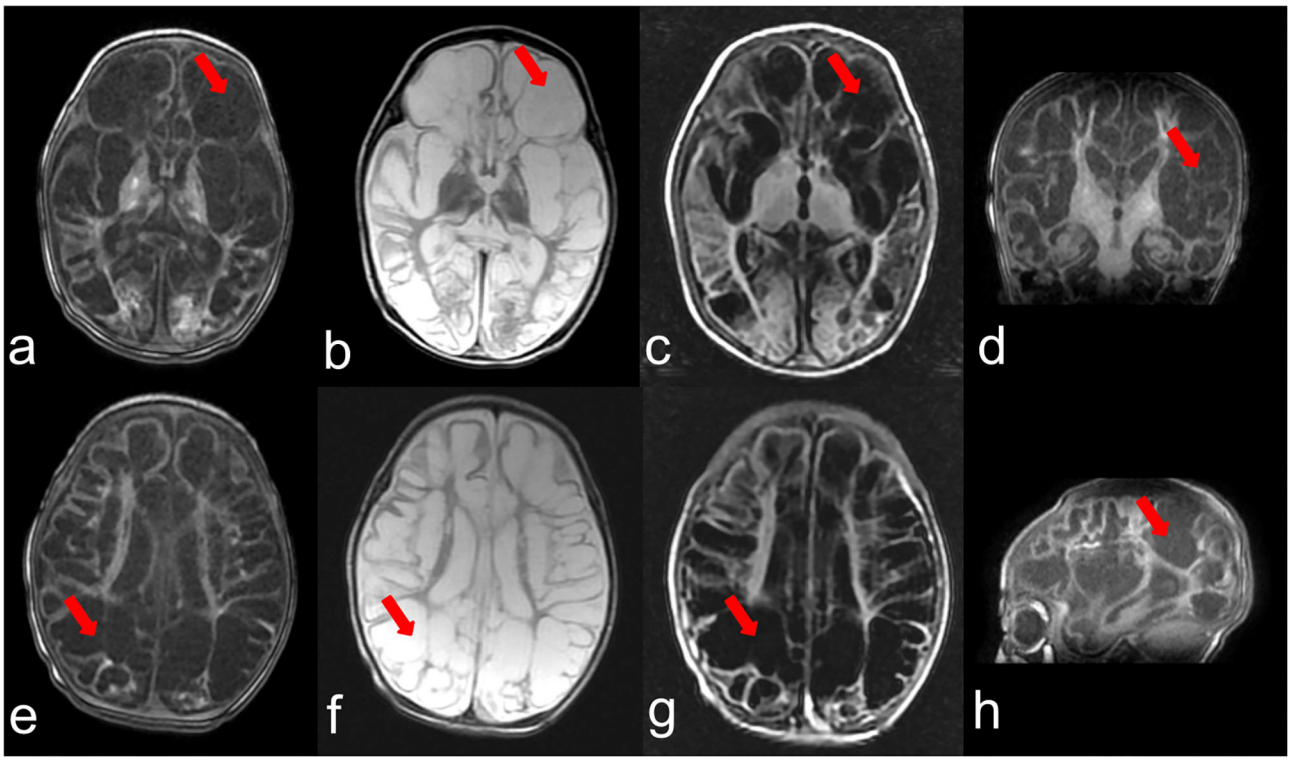


FIGURE 1 | Brain MRI at the age of 3 months. **(a–c)** represent T1,T2,T2-FLAIR weighted images of one section and **(e–g)** indicate T1,T2,T2-FLAIR weighted images of another section, respectively. **(d,h)** show T1 weighted images of the coronal plane and sagittal plane. The arrows point out the representing lesions in each picture.

and variants in causative genes *MOCS1*, *MOCS2*, and *GPHN* (Atwal and Scaglia, 2016). These genes along with *MOCS3* are involved in Moco biosynthesis and providing cofactors to Moco dependent enzymes. In reported Moco deficiency cases, two thirds are caused by *MOCS1* variants, followed by *MOCS2* and *GPHN*. Most of these cases presented typical symptoms of Moco deficiency (Reiss and Johnson, 2003; Zaki et al., 2016; Mayr et al., 2021). Until now there was no evidence that *MOCS3* was one of the causative genes of Moco deficiency. In 2017 there was one case with the *MOCS3* variant who showed a slight disorder of sulfite metabolites and near to normal neurologic symptoms (Huijmans et al., 2017). In our study, we identified a Moco deficiency patient with variants in *MOCS3* presented typical features of Moco deficiency including early onset severe neurological symptoms accompanied with typical MRI manifestations and biochemical changes. The published variant in 2017 was a missense variant and the variants in our study included a premature termination variation, and it could be a possible reason for the severity of our case. To our knowledge, our study is the first study to show that *MOCS3* is a pathogenic gene of Moco deficiency.

The clinical diagnosis of Moco deficiency is often supported by typical MRI readings (brain edema, cystic encephalomalacia, atrophy of cortex and white matter, focal or bilateral changes of globus pallidus, thalamus, etc.), and it has to be confirmed by genetic and biochemical testing (Durmaz and Ozbakir, 2018; Arican et al., 2019). The primary features of the disease were reported to be seizures (72%), feeding difficulties (26%), hypotonia (11%), and developmental delay (9%). The median onset age was the first day of life and the median survival was 36 months (Mechler et al., 2015). In our study, the proband presented similar symptoms as the first infant, and the inherited diseases were considered and further tests were performed. Although the first infant was suspected as hypoxic-ischemic encephalopathy, in some cases the neurological manifestations of Moco deficiency may overlap with severe hypoxic-ischemic encephalopathy, and shows seizures, developmental delay, and cystic encephalomalacia in MRI etc. (Zaki et al., 2016; Yoganathan et al., 2018). However, Moco deficiency usually presents more aggressive neurological symptoms. The positive urine test for sulfite with decreased plasma and urine uric acid would assist diagnosis (Yoganathan et al., 2018; Bender et al., 2019).

MOCS3 might involve in the pathogenesis of neurodevelopmental diseases by impacting the Moco biosynthesis pathway, affecting neuronal receptors and causing neuronal death. *MOCS3* knocked out cells were reported to cause combined deficiency of sulfite oxidase, xanthine oxidoreductase, and aldehyde oxidase (Chowdhury et al., 2012). This resulted in accumulation of many metabolites, especially sulfite, a neurotoxin proved to react with cystin to form the S-sulfocysteine and thiosulfate (Neukranz et al., 2019). The structure of S-sulfocysteine highly resembled the excitatory

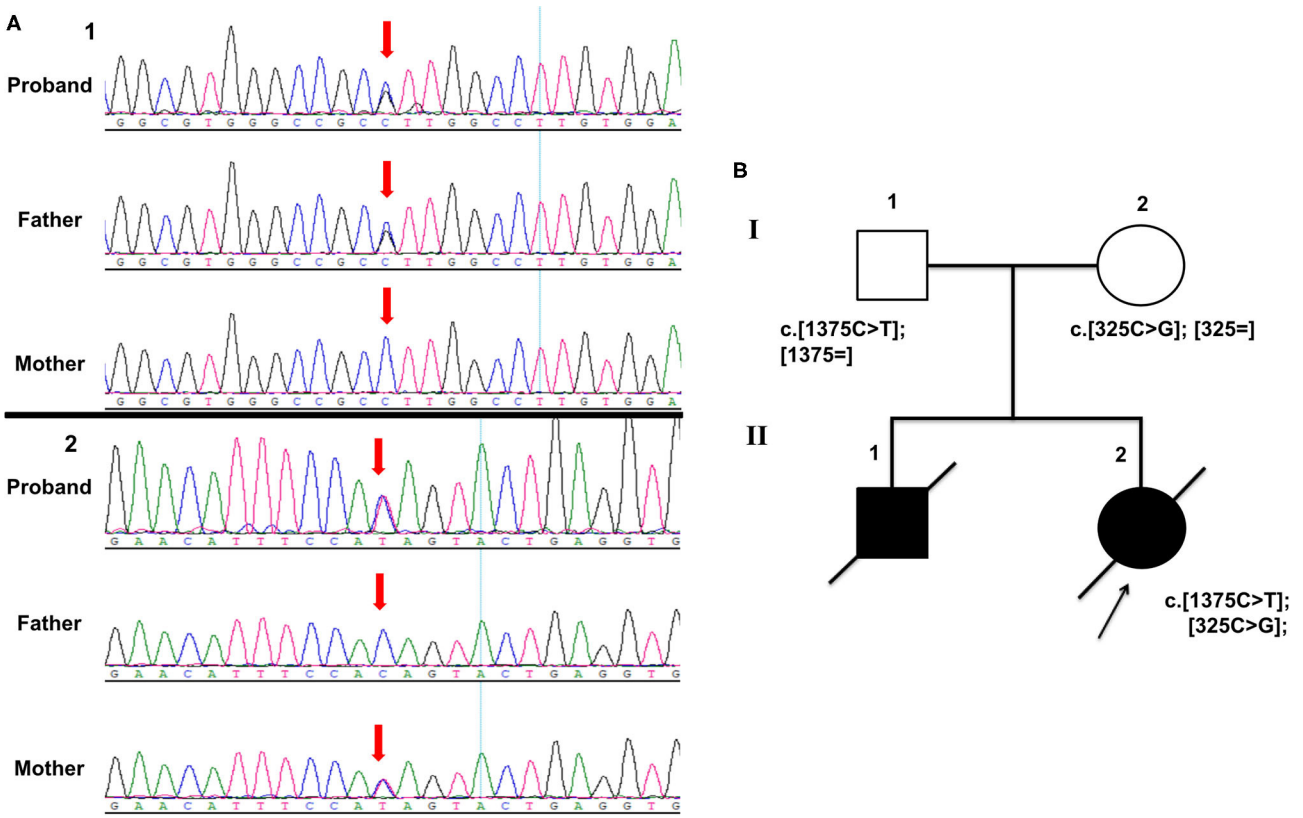


FIGURE 2 | (A) Sanger sequencing electropherograms showing compound heterozygous variants in *MOCS3* (**A** showing c.1375C>T and **B** showing c.325C>G). **(B)** Pedigree chart of the family.

neurotransmitter glutamate that could bind and stimulate glutamatergic receptors. It has been proved in a zebrafish model that S-sulfocysteine could stimulate glutamatergic receptors, induce seizure-like movements, and increase cell death in the central nervous system (Zaki et al., 2016; Plate et al., 2019). Accumulation of S-sulphocysteine and thiosulfate would activate the N-methyl-D-aspartate receptor and increase intracellular magnesium and calcium. This could enhance the permeability of the mitochondria, compromise the mitochondrial energy supply, and reduce the cell viability in cerebral cortex of rat brain (Grings et al., 2014; Marelja et al., 2018; Bender et al., 2019; Plate et al., 2019). All these processes may be the possible mechanisms underlying the neurological dysfunction observed in our patient.

In summary, we reported a novel causative gene *MOCS3* in a Moco deficiency patient. Our study may contribute to genetic diagnosis of Moco deficiency and future genetic counseling.

MATERIALS AND METHODS

Leukocyte DNA was extracted from peripheral blood using the phenol-chloroform method. DNA was then sheared to ∼200 bp by the Biorupter UCD-200 (Diagenode). The DNA fragments were then repaired at the end, and one A base was added to the 3′ end. The DNA fragments were connected with sequencing adaptors, and fragments of ∼320 bp were collected by XP beads. After PCR amplification, the DNA fragments were hybridized and captured by IDT's xGen Exome Research Panel (Integrated DNA Technologies, San Diego, USA) according to the manufacturer's protocol. The hybrid products were eluted and collected. Then, DNA was PCR amplified and purified. The libraries were tested for enrichment by qPCR, and size distribution and concentration were determined using an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). WES was performed on the HiSeq 2500 system (Illumina) with an average coverage depth of 100× of the variants for sequencing the genomic DNA of the family. Raw sequence reads were mapped to the hg19/GRCh37 using Elandv2e (CASAVA1.8.2; Illumina) to remove duplicated reads and CASAVA1.8.2 was further used to call single-nucleotide variants and short insertions/deletions (Markus et al., 2020; Tessarech et al., 2020). The variants were annotated using ANNOVAR (Wang et al., 2010).

A public database (1,000 Genomes Project, ExAC, gnomAD) was used to detect variants frequencies. The pathogenicity of variants was predicted using following software programs:

Combined Annotation Dependent Depletion [CADD] [https://cadd.gs.washington.edu/] (Kircher et al., 2014; Rentzsch et al., 2019).

MutationTaster [http://www.mutationtaster.org/] (Schwarz et al., 2014).

FATHMM [http://fathmm.biocompute.org.uk/fathmmMKL.htm] (Shihab et al., 2013a,b).

Sanger sequencing was performed on the DNA of the proband's parents to validate the variants found in WES using standard methods on ABI 3730 automated sequencer with BigDye™ Terminator v3.1 Cycle Sequencing Kit, as described previously (Stockley et al., 2013). The primers are as follows: forward primer (5′-CTCTGTCCCGAGATGAGATTCT-3′) and reverse primer (5′-CGTTGTCCGAGCAGTCAGC-3'), forward primer (5′-AAGAAGCAATCTGGGAAGAGAAG) and reverse primer (5′-ATGACCCTCTTTCTGAATAATTAAA).

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 study was approved by the Ethics Committee of the Maternal and Child Health Hospital of Hunan Province (2020-S003). Informed consent was obtained from the legal guardian of the participant for the publication of this case report (including all data and images).

AUTHOR CONTRIBUTIONS

XM and HW designed the research. QT and YaC interpreted the data and wrote the manuscript. HW, LS, YoC, YW, and YuC did the follow-up study and collected, evaluated the clinical, and genetic evidence. QT and HW revised the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (81801136), Major Scientific and Technological Projects for Collaborative Prevention and Control of Birth Defects in Hunan Province (2019SK1010, 2019SK1012, 2019SK1014), the National Key R&D Program of China (No. 2019YFC1005100), the China Postdoctoral Science Foundation (2019M662804), and the Changsha Municipal Natural Science Foundation (kq2007048).

ACKNOWLEDGMENTS

We thank the families and clinical staff for participation in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.651878/full#supplementary-material

REFERENCES

Arican, P., Gencpinar, P., Kirbiyik, O., Bozkaya Yilmaz, S., Ersen, A., Oztekin, O., et al. (2019). The clinical and molecular characteristics of molybdenum cofactor deficiency due to MOCS2 mutations. *Pediatr. Neurol.* 99, 55–59. doi: 10.1016/j.pediatrneurol.2019.04.021

Atwal, P. S., and Scaglia, F. (2016). Molybdenum cofactor deficiency. *Mol. Genet. Metab.* 117, 1–4. doi: 10.1016/j.ymgme.2015.11.010

Bender, D., Kaczmarek, A. T., Santamaria-Araujo, J. A., Stueve, B., Waltz, S., Bartsch, D., et al. (2019). Impaired mitochondrial maturation of sulfite oxidase in a patient with severe sulfite oxidase deficiency. *Hum. Mol. Genet.* 28, 2885–2899. doi: 10.1093/hmg/ddz109

Chowdhury, M. M., Dosche, C., Lohmannsroben, H. G., and Leimkuhler, S. (2012). Dual role of the molybdenum cofactor biosynthesis protein MOCS3 in tRNA thiolation and molybdenum cofactor biosynthesis in humans. *J. Biol. Chem.* 287, 17297–17307. doi: 10.1074/jbc.M112.351429

Durmaz, M. S., and Ozbakir, B. (2018). Molybdenum cofactor deficiency: neuroimaging findings. *Radiol. Case Rep.* 13, 592–595. doi: 10.1016/j.radcr.2018.02.025

Grings, M., Moura, A.P., Amaral, A. U., Parmeggiani, B., Gasparotto, J., Moreira, J. C, et al. (2014). Sulfite disrupts brain mitochondrial energy homeostasis and induces mitochondrial permeability transition pore opening via thiol group modification. *Biochim. Biophys. Acta* 1842, 1413–1422. doi: 10.1016/j.bbadis.2014.04.022

Huijmans, J. G. M., Schot, R., de Klerk, J. B. C., Williams, M., de Coo, R. F. M., Duran, M., et al. (2017). Molybdenum cofactor deficiency: identification of a patient with homozygote mutation in the MOCS3 gene. *Am. J. Med. Genet. A* 173, 1601–1606. doi: 10.1002/ajmg.a.38240

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

Marelja, Z., Leimkuhler, S., and Missirlis, F. (2018). Iron sulfur and molybdenum cofactor enzymes regulate the drosophila life cycle by controlling cell metabolism. *Front. Physiol.* 9:50. doi: 10.3389/fphys.2018.00050

Markus, F., Angelini, C., Trimouille, A., Rudolf, G., Lesca, G., Goizet, C., et al. (2020). Rare variants in the GABAA receptor subunit epsilon identified in patients with a wide spectrum of epileptic phenotypes. *Mol. Genet. Genomic Med.* 8:e1388. doi: 10.1002/mgg3.1388

Mayr, S. J., Mendel, R. R., and Schwarz, G. (2021). Molybdenum cofactor biology, evolution and deficiency. *Biochim. Biophys. Acta Mol. Cell. Res.* 1868:118883. doi: 10.1016/j.bbamcr.2020.118883

Mechler, K., Mountford, W. K., Hoffmann, G. F., and Ries, M. (2015). Ultra-orphan diseases: a quantitative analysis of the natural history of molybdenum cofactor deficiency. *Genet. Med.* 17, 965–970. doi: 10.1038/gim.2015.12

Neukranz, Y., Kotter, A., Beilschmidt, L., Marelja, Z., Helm, M., Graf, R., et al. (2019). Analysis of the cellular roles of MOCS3 identifies a MOCS3-independent localization of NFS1 at the tips of the centrosome. *Biochemistry* 58, 1786–1798. doi: 10.1021/acs.biochem.8b01160

Plate, J., Sassen, W. A., Hassan, A. H., Lehne, F., Koster, R. W., and Kruse, T. (2019). S-sulfocysteine induces seizure-like behaviors in zebrafish. *Front. Pharmacol.* 10:122. doi: 10.3389/fphar.2019.00122

Reiss, J., and Johnson, L. J. (2003). Mutations in the molybdenum cofactor biosynthetic genes MOCS1, MOCS2, and GEPH. *Hum. Mutat.* 21, 569–576. doi: 10.1002/humu.10223

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., and Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* 47, D886–D894. doi: 10.1093/nar/gky1016

Schwarz, J. M., Cooper, D. N., Schuelke, M., and Seelow, D. (2014). MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* 11, 361–362. doi: 10.1038/nmeth.2890

Shihab, H. A., Gough, J., Cooper, D. N., Day, I. N., and Gaunt, R. T. (2013b). Predicting the functional consequences of cancer-associated amino acid substitutions. *Bioinformatics* 29, 1504–1510. doi: 10.1093/bioinformatics/btt182

Shihab, H. A., Gough, J., Cooper, D. N., Stenson, P. D., Barker, G. L., Edwards, K. J., et al. (2013a). Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Hum. Mutat.* 34, 57–65. doi: 10.1002/humu.22225

Stockley, J., Morgan, N. V., Bem, D., Lowe, G. C., Lordkipanidze, M., Dawood, B., et al. (2013). Enrichment of FLI1 and RUNX1 mutations in families with excessive bleeding and platelet dense granule secretion defects. *Blood* 122, 4090–4093. doi: 10.1182/blood-2013-06-506873

Tessarech, M., Gorce, M., Boussion, F., Bault, J. P., Triau, S., Charif, M., et al. (2020). Second report of RING finger protein 113A (RNF113A) involvement in a Mendelian disorder. *Am. J. Med. Genet. A* 182, 565–569. doi: 10.1002/ajmg.a.61384

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

Yoganathan, S., Sudhakar, S., Thomas, M., Kumar Dutta, A., Danda, S., and Chandran, M. (2018). Novel Imaging finding and novel mutation in an infant with molybdenum cofactor deficiency, a mimicker of hypoxic-ischaemic encephalopathy. *Iran. J. Child Neurol.* 12, 107–112. doi: 10.22037/ijcn.v12i2.12671

Zaki, M. S., Selim, L., El-Bassyouni, H. T., Issa, M. Y., Mahmoud, I., Ismail, S., et al. (2016). Molybdenum cofactor and isolated sulphite oxidase deficiencies: Clinical and molecular spectrum among Egyptian patients. *Eur. J. Paediatr. Neurol.* 20, 714–722. doi: 10.1016/j.ejpn.2016.05.011

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 May 2021
doi: 10.3389/fgene.2021.642849

Renal and Skeletal Anomalies in a Cohort of Individuals With Clinically Presumed Hereditary Nephropathy Analyzed by Molecular Genetic Testing

Michaela Stippel[1†], Korbinian M. Riedhammer[1,2†], Bärbel Lange-Sperandio[3], Michaela Geßner[4], Matthias C. Braunisch[1,2], Roman Günthner[1,2], Martin Bald[5], Miriam Schmidts[6], Peter Strotmann[7], Velibor Tasic[8], Christoph Schmaderer[2], Lutz Renders[2], Uwe Heemann[2] and Julia Hoefele[1]*

[1] Institute of Human Genetics, Klinikum rechts der Isar, School of Medicine, Technical University of Munich, Munich, Germany, [2] Department of Nephrology, Klinikum rechts der Isar, School of Medicine, Technical University of Munich, Munich, Germany, [3] Division of Pediatric Nephrology, Dr. v. Hauner Children's Hospital, Ludwig-Maximilians University, Munich, Germany, [4] Pediatric Nephrology, University Children's Hospital, Tübingen, Germany, [5] Pediatric Nephrology, Olgahospital, Klinikum Stuttgart, Stuttgart, Germany, [6] Center for Pediatrics and Adolescent Medicine, University Hospital Freiburg, Freiburg University Faculty of Medicine, Freiburg, Germany, [7] Pediatric Nephrology, Children's Hospital, München-Klinik Schwabing, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany, [8] Medical Faculty of Skopje, University Children's Hospital, Skopje, Macedonia

OPEN ACCESS

Edited by:
Yanling Yang,
Peking University First Hospital, China

Reviewed by:
Gavin Arno,
University College London,
United Kingdom
Reuben J. Pengelly,
University of Southampton,
United Kingdom

***Correspondence:**
Julia Hoefele
julia.hoefele@tum.de

[†] These authors have contributed equally to this work

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

Received: 16 December 2020
Accepted: 11 March 2021
Published: 26 May 2021

Citation:
Stippel M, Riedhammer KM, Lange-Sperandio B, Geßner M, Braunisch MC, Günthner R, Bald M, Schmidts M, Strotmann P, Tasic V, Schmaderer C, Renders L, Heemann U and Hoefele J (2021) Renal and Skeletal Anomalies in a Cohort of Individuals With Clinically Presumed Hereditary Nephropathy Analyzed by Molecular Genetic Testing. Front. Genet. 12:642849. doi: 10.3389/fgene.2021.642849

Background: Chronic kidney disease (CKD) in childhood and adolescence occurs with a median incidence of 9 per million of the age-related population. Over 70% of CKD cases under the age of 25 years can be attributed to a hereditary kidney disease. Among these are hereditary podocytopathies, ciliopathies and (monogenic) congenital anomalies of the kidney and urinary tract (CAKUT). These disease entities can present with a vast variety of extrarenal manifestations. So far, skeletal anomalies (SA) have been infrequently described as extrarenal manifestation in these entities. The aim of this study was to retrospectively investigate a cohort of individuals with hereditary podocytopathies, ciliopathies or CAKUT, in which molecular genetic testing had been performed, for the extrarenal manifestation of SA.

Material and Methods: A cohort of 65 unrelated individuals with a clinically presumed hereditary podocytopathy (focal segmental glomerulosclerosis, steroid resistant nephrotic syndrome), ciliopathy (nephronophthisis, Bardet-Biedl syndrome, autosomal recessive/dominant polycystic kidney disease), or CAKUT was screened for SA. Data was acquired using a standardized questionnaire and medical reports. 57/65 (88%) of the index cases were analyzed using exome sequencing (ES).

Results: 8/65 (12%) index individuals presented with a hereditary podocytopathy, ciliopathy, or CAKUT and an additional skeletal phenotype. In 5/8 families (63%), pathogenic variants in known disease-associated genes (1x *BBS1*, 1x *MAFB*, 2x *PBX1*, 1x *SIX2*) could be identified.

Conclusions: This study highlights the genetic heterogeneity and clinical variability of hereditary nephropathies in respect of skeletal anomalies as extrarenal manifestation.

Keywords: hereditary nephropathy, CAKUT, podocytopathy, FSGS, SRNS, ciliopathy, skeletal anomaly

INTRODUCTION

Chronic kidney disease (CKD) in childhood and adolescence occurs with a median incidence of 9 per million of the age-related population (Harambat et al., 2012). A hereditary (monogenic) nephropathy accounts for over 70% of CKD cases with onset under 25 years of age (Vivante and Hildebrandt, 2016). Hereditary nephropathies comprise a group of clinically and genetically heterogeneous kidney disorders representing a significant risk for the development of ESRD (end-stage renal disease) (Stavljenic-Rukavina, 2009). About 6% of children up to the age of 14 years on renal replacement therapy (RRT) have a diagnosed hereditary nephropathy. Additionally, congenital anomalies of the kidney and urinary tract (CAKUT) and cystic kidney diseases, in which various monogenic disorders have been identified, are present in more than 50% of children up to the age of 14 years on RRT (Chesnaye et al., 2014). Hence, hereditary nephropathies, which can occur in isolated or syndromic forms, pose a substantial burden on mortality and morbidity.

More than 200 clinically distinct syndromes have been described to involve renal/urinary tract malformations (van der Ven et al., 2018). However, skeletal anomalies (SA) as extrarenal feature of hereditary nephropathies are infrequently described in the literature, but can be part of syndromic forms of hereditary podocytopathies (e.g., focal segmental glomerulosclerosis, steroid resistant nephrotic syndrome), ciliopathies (e.g., nephronophthisis), or CAKUT (Solomon, 2011; Romani et al., 2013; Priya et al., 2016; Riedhammer et al., 2017; Titieni and König, 2018; Haffner and Petersen, 2019; König et al., 2019). The frequency of the phenotypic combination of renal and skeletal anomalies depends on the underlying disease/disease entity and can be observed, for example, in individuals with CAKUT in up to 50% (Stoll et al., 2014). This is not surprising, as both the skeleton and kidney originate from the mesoderm and their development share important pathways, such as Wnt- and hedgehog signaling (Barker et al., 2014; Krause et al., 2015).

This study highlights eight individuals with extrarenal features of SA in a cohort of individuals with clinical suspicion of a hereditary podocytopathy, ciliopathy, or CAKUT in which molecular genetic testing had been performed.

MATERIALS AND METHODS

The study was approved by the local Ethics Committee of the Technical University of Munich and performed according to the standard of the Helsinki Declaration of 2013. Written informed consent was obtained by all participants or their legal guardians. All individuals were clinically examined by pediatric or adult nephrologists.

Study Cohort

All individuals were referred to the Institute of Human Genetics of the Klinikum rechts der Isar of the Technical University of Munich, a tertiary care center. Clinical data was acquired using a standardized questionnaire and medical reports. The study cohort encompassed 65 unrelated index individuals (index cases; 42/65 [65%] of Non-Finnish European descent) with a clinically presumed hereditary podocytopathy, ciliopathy, or CAKUT and was screened for skeletal anomalies. The included individuals are part of a larger cohort of ∼300 families with clinically presumed hereditary kidney disease ("NephroGen") collected between 2015 and 2018. The NephroGen cohort additionally contains families with clinically presumed hereditary kidney diseases like type-IV-collagen-related nephropathy (Alport syndrome, thin basement membrane nephropathy), tubulopathies, and autosomal dominant tubulointerstitial kidney disease (ADTKD), which were not included in this study.

Molecular-genetic analysis was performed primarily by exome sequencing (ES). However, in selected cases—depending on the clinical phenotype—single-gene Sanger sequencing, single nucleotide polymorphism (SNP) array or multiplex ligation-dependent probe amplification (MLPA) was performed (see Results).

Exome Sequencing

DNA was extracted from peripheral blood using either the Gentra Puregene Blood Kit (Qiagen, Hilden, Germany) or the Chemagic 360 (Perkin Elmer, Waltham) according to the manufacturer's instructions. ES was performed using a Sure Select Human All Exon 60 Mb V6 Kit (Agilent) and a HiSeq4000 (Illumina) or a Sure Select Human All Exon 50 Mb V5 Kit (Agilent) and a HiSeq2500 (Illumina) (Kremer et al., 2017). Mitochondrial DNA was derived from off-target exome reads (Griffin et al., 2014). Reads were aligned to the human reference genome (UCSC Genome Browser build hg19) using Burrows-Wheeler Aligner (v.0.7.5a). Detection of single-nucleotide variants and small insertions and deletions (indels) was performed with SAMtools (version 0.1.19). ExomeDepth was used for the detection of copy number variants (CNVs). A noise threshold of 2.5 was accepted for diagnostic analysis (Plagnol et al., 2012). Called CNVs were visualized by the Integrative Genomics Viewer (IGV, https://software.broadinstitute.org/software/igv/) to check for sufficient coverage of the inspected region and plausibility of the CNV. CNVs were compared with publicly available control databases like the Genome Aggregation Database (gnomAD, https://gnomad.broadinstitute.org/about), the Database of Genomic Variants (DGV, http://dgv.tcag.ca/dgv/app/home) and databases for pathogenic CNVs like DECIPHER (https://decipher.sanger.ac.uk/) and ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/). Rating was done according to American College of Medical

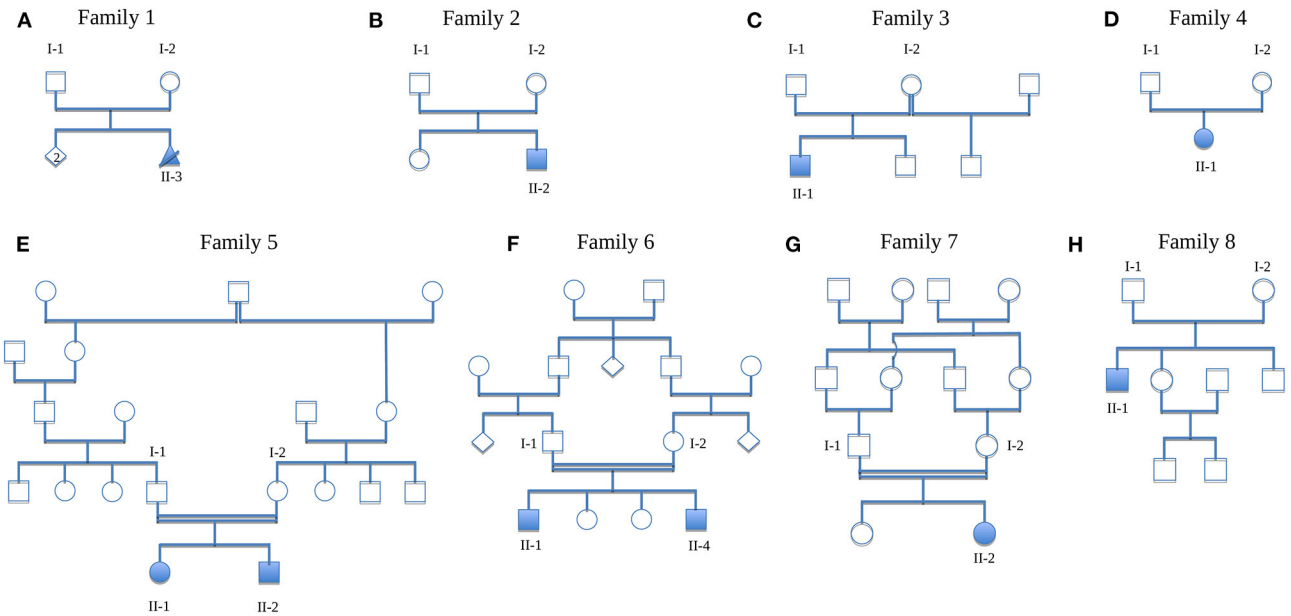


FIGURE 1 | Pedigrees of the families. **(A)** Family 1 (HN-F18), **(B)** Family 2 (HN-F151), **(C)** Family 3 (HN-F75), **(D)** Family 4 (HN-F80), **(E)** Family 5 (HN-F191), **(F)** Family 6 (HN-F198), **(G)** Family 7 (HN-F241), **(H)** Family 8 (HN-F305). Solid symbols, affected individuals; circles, females; squares, males; crossed triangle, miscarriage; diamond, unknown gender; double lines, consanguinity.

Genetics (ACMG) guidelines and the recommendation of the ACGS (Richards et al., 2015; Ellard et al., 2019; Riggs et al., 2020). For the analysis of *de novo,* autosomal dominant and mitochondrial variants, only variants with a minor allele frequency (MAF) of <0.1% in an in-house database ("Munich Exome Server") containing over 21,000 exomes were considered. For the analysis of autosomal recessive and X-linked variants (homozygous, hemizygous or putative compound heterozygous) variants with a MAF of <1.0% were considered. As there are pathogenic alleles in hereditary nephropathies (HNs) with a MAF of more than 1.0% like the *NPHS2* p.Arg229Gln allele (also known as p.R229Q), in unsolved cases, an additional search for recessive and X-linked variants a MAF up to 3% was accepted. The p.Arg229Gln allele is known to cause steroid resistant nephrotic syndrome when *in trans* with specific 3′ *NPHS2* variants (Tory et al., 2014; Miko et al., 2018). Identified variants were compared with publicly available databases for pathogenic variants like ClinVar, the Human Gene Mutation Database (HGMD®, http://www.hgmd.cf.ac.uk) and the Leiden Open Variation Database (LOVD, https://www.lovd.nl). Only variants rated as "likely pathogenic" or "pathogenic" according to the classification of the ACMG and with a genotype in agreement with the mode of inheritance led to the designation "solved case" (Kearney et al., 2011; Richards et al., 2015; Riggs et al., 2020).

Sanger Sequencing, SNP Array, and MLPA

Molecular genetic testing of selected cases or segregation analysis of identified variants in parents and affected relatives was performed by Sanger sequencing, SNP array or MLPA as appropriate. Oligonucleotide primer sequences used for Sanger sequencing are available upon request. DNA alignment and sequence variant analysis were carried out using the Sequence PilotCE software (JSI Medical Systems GmbH, Kippenheim, Germany) and compared to EMBL (European Molecular Biology Laboratory) and GenBank databases as well as our in-house database. For SNP array, Affymetrix® CytoScanTM 750 K Array (Affymetrix® Inc., Santa Clara, CA, USA) with an average space between two oligonucleotides of 4 kilobases (kb) was used. Scanning was performed by the Affymetrix® GeneChip Scanner 3000 7G (resolution 0.51–2.5 μm). The data analysis was conducted using the Affymetrix® Chromosome Analysis Suite Software (ChAS), version 3.0, hg19. For MLPA, the SALSA MLPA kit P241-D2 Kit (maturity-onset diabetes of the young, renal cysts and diabetes syndrome, *HNF1B*) was purchased from MRC-Holland (Amsterdam, the Netherlands) and performed according to the manufacturer's instructions. Fragment analysis was conducted using GeneMarker software (www.softgenetics.com, Soft Genetics, State College, PA, USA).

Molecular genetic results were communicated to the families either via the referring clinician or as part of genetic counseling.

RESULTS

The total cohort of 65 index individuals comprised 18 (28%) individuals with the clinical picture of a hereditary podocytopathy (one exception of an index case [HN-F203-II-2] who was initially classified as having a type-IV-collagen-related nephropathy based on the clinical phenotype and later reclassified as hereditary podocytopathy on genetic diagnosis; see **Supplementary Table 1**), 20 (31%) with a ciliopathy, and 27 (41%) with CAKUT. A total of eight individuals (12%)

displayed one or more skeletal anomalies additionally to their renal phenotype (**Figure 1**, **Table 1**, family description below).

In 57/65 (88%) index cases, ES was performed. 8/65 (12%) index cases were analyzed either by phenotype-guided panel diagnostics (4/8; 50%), single-gene Sanger sequencing (1/8; 13%), SNP array (2/8; 25%), or MLPA (1/8; 13%) (**Figure 2**). 22/65 (34%) index cases could be genetically solved (i.e., [likely] pathogenic variant with a fitting genotype identified).

The eight index cases with an additional skeletal phenotype are presented as follows. The remaining 57 index cases are not specifically featured in this paper. Phenotypic and genotypic data of these cases can be found in the **Supplementary Table 1**.

Family 1 (HN-F18; CAKUT)

In Family 1, the third, unborn male child (II-3; Non-Finnish European descent) was pre-natally diagnosed with bilateral renal agenesis. The pregnancy was therefore pre-maturely terminated. The fetus presented with dysmorphological features such as knuckle pads at the proximal interphalangeal joint of both hands, short and pointed fingers and pes equinovarus. Other conspicuous facial features were hypoplastic nasal wings, a flat philtrum and micrognathia. There were no known affected relatives, parents and siblings of the index individual were healthy.

A previously conducted chromosome analysis revealed no pathological findings. Afterwards, trio ES was performed on the index and the parents, which also was inconspicuous.

Family 2 (HN-F51; Hereditary Podocytopathy)

The son (F2-II-2; Non-Finnish European descent) of healthy parents presented with proteinuria at the age of 2 years. He additionally had shortened ulna and fingers, which were broad at the basis and tapered to the end. The base of the thumb was thickened and blue, the hands were doughy and toenails and toes hypoplastic. Further signs of dysmorphia were a broad root and a spherical tip of the nose, discreet epicanthus, a long philtrum, and a narrow upper lip. There were no affected relatives.

Proband-only ES identified a heterozygous pathogenic missense variant NM_005461.4:c.184A>G, p.(Thr62Ala) in *MAFB*. Targeted Sanger sequencing of the parents confirmed its *de novo* status. *MAFB* has been associated with multicentric carpotarsal osteolysis syndrome (MTCO, MIM 166300). It encodes a transcription factor which impacts the differentiation and activation of osteoclasts as well as renal development. The disease occurs very rarely and it features aggressive osteolysis especially in the area of carpal and tarsal bones. It is frequently associated with a progressive kidney failure (Zankl et al., 2012).

Family 3 (HN-F75; CAKUT)

This case has already been published (Riedhammer et al., 2017). The renal phenotype of the male index case in this family (II-1; Non-Finnish European descent) consisted of bilateral renal dysplasia which was first seen at the age of 3 months. At the age of 4 years, he had already developed renal insufficiency. Furthermore, different skeletal anomalies like hypoplastic claviculae, cleidocranial dysplasia, clinodactyly and brachydactyly of the little finger as well as a narrow thorax and shoulders were present. Moreover, he had intellectual disability. There were no other affected family members.

First, SNP array analysis was performed but did not reveal any causative microdeletions or microduplications. Afterwards, proband-only ES was performed in the index case identifying the pathogenic frameshift variant NM_002585.3:c.413_419del, p.(Gly138Valfs*40) in *PBX1*. *PBX1* is a known CAKUT disease-associated gene (MIM 617641) (Heidet et al., 2017). The parents were examined using Sanger sequencing, which confirmed the *de novo* status of this variant.

Family 4 (HN-F80; CAKUT)

This female individual (II-1; Non-Finnish European descent) of Family 4 had developed ESRD already in childhood as a result of renal hypodysplasia. Extrarenal manifestations were decreased vision in one eye, frontonasal dysplasia with retracted nasal root, and short stature due to skeletal dysplasia. Her parents were healthy as well as other family members.

Trio ES was conducted and identified a *de novo* heterozygous deletion of at least 2.9 kb including the complete exonic regions of *SIX2* (approx. chr2:g.45233309-45236248). While *SIX2* missense variants have been associated with renal hypodysplasia, *SIX2* haploinsufficiency has not been linked to a renal phenotype so far but to frontonasal dysplasia, ptosis, and hearing loss (Weber et al., 2008; Guan et al., 2016; Henn et al., 2018).

Family 5 (HN-F191; Ciliopathy)

Family 5 featured two siblings with a syndromal phenotype, fitting best into the spectrum of ciliopathies. Their parents were healthy, of Turkish origin and consanguineous. The female index case (II-1) as well as her younger brother (II-2) showed various congenital anomalies. The girl presented with cardiovascular anomalies such as mitral valve insufficiency, atrial septal defect, persistent ductus arteriosus and pulmonary hypertension. Furthermore, she suffered from bronchial hyperreactivity, which lead to recurrent pneumonia. Noticeable skeletal anomalies were disproportionate short stature and bilateral genu vara. The renal phenotype was designated as nephropathy of unclear origin. A biopsy revealed glomerular and tubulo-interstitial scarring as well as minimal tubular ectasia of the cortex. Her clinical phenotype was comparable with a Bardet-Biedl-like syndrome. Her brother showed similar anomalies. He presented with pulmonary symptoms such as recurrent, partly obstructive bronchitis and short stature. Furthermore, hepatomegaly was observed in him.

Using ES in the female index individual, no causative variants in known disease-associated genes could be identified.

Family 6 (HN-F198; Ciliopathy)

In Family 6, two individuals had the clinical tentative diagnosis of Bardet-Biedl syndrome. The male index individual (II-1) as well as his younger brother (II-4) were affected. Their parents were healthy, of Pakistani origin and consanguineous. Other relatives did not show any signs of kidney disorder. The index individual had already developed ESRD at the age of 4 years as a consequence of recurrent pyelonephritis

TABLE 1 | Overview of the index cases and also affected siblings including a detailed description of phenotype and genotype.

Individuals and sex (index, further affected individuals)	Genetically solved	Gene (transcript)	Inheritance	Variant with cDNA and protein position and inheritance (in parenthesis)	Chromosomal position (hg19)	Zygosity	ACMG rating	Applied ACMG criteria for variant interpretation/ ACMG CNV score[#]	ClinVar[##]	Clinical tentative diagnosis or genetic diagnosis with MIM number in parenthesis (if case was solved genetically)	Age at first presentation	Renal phenotype	Skeletal phenotype	Additional phenotype
Family 1 (HN-F18) II-3 male	No	-	-	-	-	-	-	-	-	CAKUT	24th week of gestation	Bilateral renal agenesis	Micrognathia	Pes equinovarus, hypoplastic nasal wings, flat philtrum, short and pointed fingers, knuckle pads
Family 2 (HN-F51) II-2 male	Yes	*MAFB* (NM_005461.4)	AD	c.184A>G, p.(Thr62Ala) (*de novo*)	chr20:g.39317307T>C	Het	Pathogenic	PS2 PM2 PM5 PP3 PP4	n.l.	Multicentric carpotarsal osteolysis syndrome (166300)	2 years	Small right kidney, proteinuria	Shortened ulna and fingers	Broad root and spherical tip of the nose, discreet epicanthus, long philtrum, narrow upper lip, thick and blue thumb base, the doughy hands, hypoplastic toes
Family 3 (HN-F75) II-1 male	Yes	*PBX1* (NM_002585.3)	AD	c.413_419del p.(Gly138Valfs*40) (*de novo*)	chr1:g.164761878_164761884del	Het	Pathogenic	PVS1 PS2 PM2	n.l.	CAKUTHED (617641)	At birth	Bilateral renal dysplasia	Hypoplastic claviculae, cleidocranial dysplasia, clinodactyly and brachydactyly of the little finger	Intellectual disability
Family 4 (HN-F80) II-1 female	No	*SIX2* (NM_016932.5)	AD	Deletion - at least 2.9 kb involving the complete exonic regions of *SIX2** (*de novo*)	Approx. chr2:g.45233309-45236248	Het	Pathogenic[###]	1,00	n.l.	Unknown syndromic kidney disorder	6 months	Bilateral renal hypodysplasia	Frontonasal dysplasia with retracted nasal root, disproportionate short stature	Decreased vision in one eye
Family 5 (HN-F191) II-1 female	No	-	-	-	-	-	-	-	-	Bardet-Biedl-like phenotype	At birth	Chronic kidney disease, tubule-interstitial scarring, renal tubular ectasia	Disproportionate short stature, bilateral genu vara	Mitral valve insufficiency, atrial septal defect, persistent ductus arteriosus, pulmonary hypertension, bronchial hyperreactivity
Family 5 (HN-F191) II-2 male	No	-	-	-	-	-	-	-	-	Bardet-Biedl-like phenotype	At birth	Chronic kidney disease	Short stature	Recurrent obstructive bronchitis, hepatomegaly

(Continued)

TABLE 1 | Continued

Individuals and sex (index, further affected individuals)	Genetically solved	Gene (transcript)	Inheritance	Variant with cDNA and protein position and inheritance (in parenthesis)	Chromosomal position (hg19)	Zygosity	ACMG rating	Applied ACMG criteria for variant interpretation/ ACMG CNV score[#]	ClinVar[##]	Clinical tentative diagnosis or genetic diagnosis with MIM number in parenthesis (if case was solved genetically)	Age at first presentation	Renal phenotype	Skeletal phenotype	Additional phenotype
Family 6 (HN-F198) II-1 male	Yes	*BBS1* (NM_024649.4)	AR	Complex allele c.[831-3C>G; 1285C>T];[831-3C>G;1285C>T], p.[?;Arg429*];[?;Arg429*] (maternal, paternal)	chr11:g.66290924C>G and chr11:g.66294224C>T	Hom	Pathogenic	PS4_supporting PM2 PM3 PP3 and PVS1 PS4_supporting PM2	c.831-3C>G: pathogenic and c.1285C>T: likely pathogenic	Bardet-Biedl syndrome 1 (209900)	1 year	Right vesicoureteral reflux	Polydactyly of the feet, oligodactyly right hand, left split hand	Microcephaly, developmental delay, anal atresia, severely impaired vision and oculocutaneous albinism
Family 6 (HN-F198) II-4 male	Yes	*BBS1* (NM_024649.4)	AR	Complex allele c.[831-3C>G; 1285C>T];[831-3C>G;1285C>T], p.[?;Arg429*];[?;Arg429*] (maternal, paternal)	chr11:g.66290924C>G and chr11:g.66294224C>T	Hom	Pathogenic	PS4_supporting PM2 PM3 PP3 and PVS1 PS4_supporting PM2	c.831-3C>G: pathogenic and c.1285C>T: likely pathogenic	Bardet-Biedl syndrome 1 (209900)	At birth	Hypoplastic horseshoe kidney	Hexadactyly of both hands and feet, narrow thorax	Liver fibrosis, anal atresia
Family 7 (HN-F241) II-2 female	No	-	-	-		-	-	-	-	Dysostosis multiplex	2 years	Right sided double kidney	Short stature	Coarse facial features
Family 8 (HN-F305) II-1 male	yes	*PBX1* (NM_002585.3)	AD	c.661G>T, p.(Glu221*) (*de novo*)	chr1:g.164769086G>T	Het	Pathogenic	PVS1 PS2 PM2	n.l.	CAKUTHED (617641)	1 year	Right renal agenesis, left renal hypoplasia	Short stature	Choanal atresia, facial dysmorphic signs

[#]*See (Richards et al., 2015; Ellard et al., 2019; Riggs et al., 2020). In general, PVS1 indicates a very strong level of pathogenicity, PS designations indicate strong evidence of pathogenicity, PM designations indicate a moderate level of pathogenicity, and PP designations indicate supporting evidence of pathogenicity.*

[##]*Based on https://www.ncbi.nlm.nih.gov/clinvar/ (ClinVar ratings solely by our institute are not listed).*

[###]*This is a pathogenic variant. However, haploinsufficiency of SIX2 has been associated with frontonasal dysplasia with ptosis and hearing loss but is not known to be associated with renal hypodysplasia so far, hence no exact genetic diagnosis (see Discussion).*

AD, autosomal dominant; AR, autosomal recessive; CAKUT, congenital anomalies of the kidney and urinary tract; CAKUTHED, congenital anomalies of kidney and urinary tract syndrome with or without hearing loss, abnormal ears, or developmental delay; het, heterozygous; hom, homozygous; n.l., not listed.

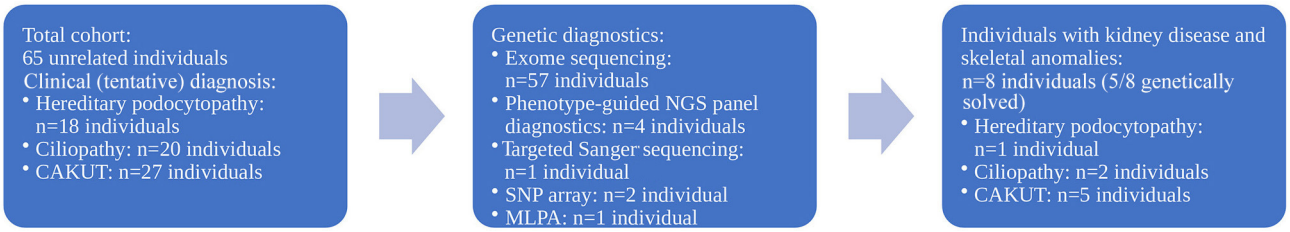


FIGURE 2 | Flowchart of the cohort. CAKUT, congenital anomalies of the kidney and urinary tract; MLPA, multiplex ligation-dependent probe amplification; NGS, next-generation sequencing; SNP, single nucleotide polymorphism.

with abscess formation due to vesicoureteral reflux on the right side. Kidney transplantation was performed at the age of 4 years. In addition to the renal phenotype, the individual had oculocutaneous albinism with serious vision problems (no signs for retinal dystrophy), severe intellectual disability, epilepsy, multiple intra-abdominal abscesses and anal atresia. He also had Factor V Leiden thrombophilia. Concerning the skeletal appearance, he had bilateral fibular polydactyly, oligodactyly of the right hand and a cleft hand on the left side. His younger brother had a hypoplastic-cystic horseshoe kidney. He additionally presented with liver cirrhosis and anal atresia. Similar to the index individual, he had skeletal anomalies such as hexadactyly of both hands and feet. Furthermore, a slim thorax and lateralized mammilla were observed.

Molecular genetic diagnostics using duo ES was performed. Both affected individuals had a pathogenic homozygous complex allele NM_024649.4:c.[831-3C>G;1285C>T];[831-3C>G;1285C>T], p.[?;Arg429*];[?;Arg429*] in *BBS1*. The parents were heterozygous carriers for the complex allele. *BBS1* is associated with autosomal recessive Bardet-Biedl syndrome (MIM 209900) (Mykytyn et al., 2002).

Family 7 (HN-F241; CAKUT)

In Family 7, the healthy parents were of Egyptian origin and consanguineous. One of the daughters (II-2) presented with double kidney on the right side as well as dysostosis multiplex. Coarse facial features and short stature were striking in appearance. Both parents and the sister of the affected individual as well as other relatives were healthy.

Proband-only ES was performed and was inconspicuous.

Family 8 (HN-F305; CAKUT)

The son (II-1; Non-Finnish European descent) of healthy parents of Family 8 showed dysplasia of the right and hypoplasia of the left kidney. Extrarenal manifestations were choanal atresia, short stature, and facial dysmorphic signs. At the age of 6 years he underwent the first kidney transplantation, a second kidney was transplanted at the age of 11 years. Consanguinity or further affected relatives were not reported.

Trio ES was conducted and revealed the pathogenic heterozygous *de novo* variant NM_002585.3:c.661G>T, p.(Glu221*) in *PBX1*. Further information on *PBX1* are already mentioned in Family 3.

Summary

Of the total 65 index cases, SA were present in 1/18 (6%) individuals with a hereditary podocytopathy, 2/20 (10%) individuals with a ciliopathy, and 5/27 (19%) individuals with CAKUT.

In terms of diagnostic yield in individuals with an additional extrarenal phenotype of SA, 5/8 (63%) cases could be genetically solved: 1/5 (20%) with a hereditary podocytopathy, 1/5 (20%) with a ciliopathy, and 3/5 (60%) with CAKUT (**Figure 2**). Of note, concerning Family 4, haploinsufficiency of *SIX2* has not been associated with a renal phenotype so far (see Discussion).

DISCUSSION

In this study, 65 unrelated index cases affected with a suspected hereditary podocytopathy, ciliopathy or CAKUT, in which molecular genetic diagnostics had been performed, were examined for a skeletal phenotype. All three disease entities comprise a broad monogenic spectrum with variable clinical phenotypes. In this cohort, eight index cases (12%) were affected by a combination of both renal and skeletal anomalies.

Almost every fifth individual with CAKUT in this study presented with an additional skeletal phenotype. Stoll et al. investigated extrarenal manifestations in children born with CAKUT over a period of 26 years and could observe skeletal anomalies in 97/1678 (6%) individuals (Stoll et al., 2014).

Furthermore, skeletal anomalies like limb defects in individuals with urinary tract malformations have been observed with an incidence of 1:20.000 live births (Natarajan et al., 2013).

Ciliopathies include a wide range of syndromal diseases and can be accompanied by various extrarenal manifestations such as retinal degeneration (Senior-Løken syndrome), vermis hypoplasia, ataxia, and intellectual disability (Joubert syndrome), or occipital meningoencephalocele (Meckel's syndrome) (Hildebrandt et al., 2011). SA are well-known as an extrarenal manifestation, e.g., with Jeune syndrome, Sensenbrenner syndrome, and short-rib polydactyly syndrome (Bredrup et al., 2011). In the present study, the index individual of Family 6 affected by a ciliopathy and SA was genetically diagnosed with a Bardet-Biedl syndrome. Individuals with Bardet-Biedl syndrome show renal anomalies in 53–85% and polydactyly in 77% (Rudling et al., 1996; Forsythe et al., 2018).

The combination of a renal and skeletal phenotype in individuals with an underlying hereditary podocytopathy has mainly been described within Schimke immunoosseous dysplasia (MIM 242900) (Haffner and Petersen, 2019). ES of the index in this study cohort (Family 2; HN-F51) revealed a likely pathogenic variant in *MAFB*. This finding lead to the very rare diagnosis of autosomal dominant MTCO. Zankl et al. published a study comprising 11 index cases with MTCO. Every individual had skeletal anomalies, eight also had an additional renal phenotype consisting of non-specific glomerulosclerosis and severe tubulointerstitial fibrosis (Zankl et al., 2012).

The percentage of individuals with a skeletal and renal phenotype in this study cohort seems rather high (8/65, 12%). However, given the low absolute number of cases ($n = 8$) with a skeletal and renal phenotype, this finding has to be interpreted with caution. In order to achieve a clear statement on the prevalence of SA in hereditary podocytopathies, ciliopathies and (monogenic) CAKUT, the study population for each disease entity would have to be extended. Nevertheless, it is worth mentioning that, at least in this small cohort, almost every fifth individual (5/27, 19%) with CAKUT presented with an additional skeletal anomaly, which is much higher compared to the other subgroups and should encourage investigation of SA as extrarenal manifestation in CAKUT.

It has to be mentioned that the present study has several limitations: (i) the study cohort was small, therefore no assumptions on frequency of renal and skeletal anomalies in CAKUT, hereditary podocytopathies, and ciliopathies can be drawn (see above); (ii) although comprehensive in most cases, the molecular genetic testing strategy is inconsistent as only in 88% of the individuals ES was performed. Nonetheless, all individuals with an additional skeletal phenotype were examined with ES. Therefore, no bias concerning the diagnostic yield should be seen in individuals with skeletal anomalies; and (iii) haploinsufficiency of *SIX2* is not known to be associated with a renal phenotype so far but SIX2 is described to be involved as a key factor within the kidney mesenchyme (Kobayashi et al., 2008; Henn et al., 2018). Therefore, even though the heterozygous deletion of the complete exonic regions of *SIX2* in Family 4 can be classified as pathogenic, its contribution to the renal phenotype is unknown to date.

In conclusion, this study should highlight the association of renal and skeletal phenotypes and their genetic heterogeneity and prompt further systematic and large-scale investigation of the combination of these phenotypic features.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the Technical University of Munich, Munich, Germany. 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

MSt, KR, and JH collected all data, performed mutational analysis, and were responsible for writing and revision of the manuscript. BL-S, MG, MBr, RG, MBa, MSc, PS, VT, CS, LR, and UH cared for the patients and provided the clinical data. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the German Research Foundation (DFG) and the Technical University of Munich (TUM) in the framework of the Open Access Publishing Program.

ACKNOWLEDGMENTS

We sincerely thank the probands and the physicians for participating in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.642849/full#supplementary-material

REFERENCES

Barker, A. R., Thomas, R., and Dawe, H. R. (2014). Meckel-Gruber syndrome and the role of primary cilia in kidney, skeleton, and central nervous system development. *Organogenesis* 10, 96–107. doi: 10.4161/org.27375

Bredrup, C., Saunier, S., Oud, M. M., Fiskerstrand, T., Hoischen, A., Brackman, D., et al. (2011). Ciliopathies with skeletal anomalies and renal insufficiency due to mutations in the IFT-A gene WDR19. *Am. J. Hum. Genet.* 89, 634–643. doi: 10.1016/j.ajhg.2011.10.001

Chesnaye, N., Bonthuis, M., Schaefer, F., Groothoff, J. W., Verrina, E., Heaf, J. G., et al. (2014). Demographics of paediatric renal replacement therapy in Europe: a report of the ESPN/ERA–EDTA registry. *Pediatr. Nephrol.* 29, 2403–2410. doi: 10.1007/s00467-014-2884-6

Ellard, S., Baple, E. L., Berry, I., Forrester, N., Turnbull, C., Owens, M., et al. (2019). *ACGS Best Practice Guidelines for Variant Classification 2019*. Available online at: https://www.acgs.uk.com/media/11285/uk-practice-guidelines-for-variant-classification-2019-v1-0-3.pdf (accessed April 22, 2021).

Forsythe, E., Kenny, J., Bacchelli, C., and Beales, P. L. (2018). Managing bardet–biedl syndrome—now and in the future. *Front. Pediatr.* 6:23. doi: 10.3389/fped.2018.00023

Griffin, H. R., Pyle, A., Blakely, E. L., Alston, C. L., Duff, J., Hudson, G., et al. (2014). Accurate mitochondrial DNA sequencing using off-target reads provides a single test to identify pathogenic point mutations. *Genet. Med*. 16, 962–971. doi: 10.1038/gim.2014.66

Guan, J., Wang, D., Cao, W., Zhao, Y., Du, R., Yuan, H., et al. (2016). SIX2 haploinsufficiency causes conductive hearing loss with ptosis in humans. *J. Hum. Genet*. 61, 917–922. doi: 10.1038/jhg.2016.86

Haffner, D., and Petersen, C. (2019). "Niere und Harnwege interdisziplinär," in *Pädiatrie,* eds C.P. Speer, M. Gahr, and J. Dötsch (Berlin, Heidelberg: Springer), 607–642. doi: 10.1007/978-3-662-57295-5_25

Harambat, J., van Stralen, K. J., Kim, J. J., and Tizard, E. J. (2012). Epidemiology of chronic kidney disease in children. *Pediatr. Nephrol*. 27, 363–373. doi: 10.1007/s00467-011-1939-1

Heidet, L., Moriniere, V., Henry, C., De Tomasi, L., Reilly, M. L., Humbert, C., et al. (2017). Targeted exome sequencing identifies PBX1 as involved in monogenic congenital anomalies of the kidney and urinary tract. *J. Am. Soc. Nephrol*. 28, 2901–2914. doi: 10.1681/ASN.2017010043

Henn, A., Weng, H., Novak, S., Rettenberger, G., Gerhardinger, A., Rossier, E., et al. (2018). SIX2 gene haploinsufficiency leads to a recognizable phenotype with ptosis, frontonasal dysplasia, and conductive hearing loss. *Clin. Dysmorphol*. 27, 27–30. doi: 10.1097/MCD.0000000000000213

Hildebrandt, F., Benzing, T., and Katsanis, N. (2011). Ciliopathies. *N. Engl. J. Med*. 364, 1533–1543. doi: 10.1056/NEJMra1010172

Kearney, H. M., Thorland, E. C., Brown, K. K., Quintero-Rivera, F., South, S. T., and Working Group of the American College of Medical Genetics Laboratory Quality Assurance, C. (2011). American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet. Med*. 13, 680–685. doi: 10.1097/GIM.0b013e3182217a3a

Kobayashi, A., Valerius, M. T., Mugford, J. W., Carroll, T. J., Self, M., Oliver, G., et al. (2008). Six2 defines and regulates a multipotent self-renewing nephron progenitor population throughout mammalian kidney development. *Cell Stem Cell* 3, 169–181. doi: 10.1016/j.stem.2008.05.020

König, J., Habbig, S., and Liebau, M. C. (2019). Management von Ziliopathien im Kindes- und Jugendalter. *Der Nephrologe*. 14, 192–198. doi: 10.1007/s11560-019-0312-4

Krause, M., Rak-Raszewska, A., Pietila, I., Quaggin, S. E., and Vainio, S. (2015). Signaling during kidney development. *Cells* 4, 112–132. doi: 10.3390/cells4020112

Kremer, L. S., Bader, D. M., Mertes, C., Kopajtich, R., Pichler, G., Iuso, A., et al. (2017). Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nat. Commun*. 8:15824. doi: 10.1038/ncomms15824

Miko, A. D. K. M., Kaposi, A., Antignac, C., and Tory, K. (2018). The mutation-dependent pathogenicity of NPHS2 p.R229Q: a guide for clinical assessment. *Hum. Mutat*. 39, 1854–1860. doi: 10.1002/humu.23660

Mykytyn, K., Nishimura, D. Y., Searby, C. C., Shastri, M., Yen, H.-j., Beck, J. S., et al. (2002). Identification of the gene (BBS1) most commonly involved in Bardet-Biedl syndrome, a complex human obesity syndrome. *Nat. Genet*. 31, 435–438. doi: 10.1038/ng935

Natarajan, G., Jeyachandran, D., Subramaniyan, B., Thanigachalam, D., and Rajagopalan, A. (2013). Congenital anomalies of kidney and hand: a review. *Clin. Kidney J*. 6, 144–149. doi: 10.1093/ckj/sfs186

Plagnol, V., Curtis, J., Epstein, M., Mok, K. Y., Stebbings, E., Grigoriadou, S., et al. (2012). A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics* 28, 2747–2754. doi: 10.1093/bioinformatics/bts526

Priya, S., Nampoothiri, S., Sen, P., and Sripriya, S. (2016). Bardet-Biedl syndrome: genetics, molecular pathophysiology, and disease management. *Indian J. Ophthalmol*. 64, 620–627. doi: 10.4103/0301-4738.194328

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

Riedhammer, K. M., Siegel, C., Alhaddad, B., Montoya, C., Kovacs-Nagy, R., Wagner, M., et al. (2017). Identification of a novel heterozygous de novo 7-bp frameshift deletion in PBX1 by whole-exome sequencing causing a multi-organ syndrome including bilateral dysplastic kidneys and hypoplastic clavicles. *Front. Pediatr*. 5:251. doi: 10.3389/fped.2017.00251

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

Romani, M., Micalizzi, A., and Valente, E. M. (2013). Joubert syndrome: congenital cerebellar ataxia with the molar tooth. *Lancet Neurol*. 12, 894–905. doi: 10.1016/S1474-4422(13)70136-4

Rudling, O., Riise, R., Tornqvist, K., and Jonsson, K. (1996). Skeletal abnormalities of hands and feet in Laurence-Moon-Bardet-Biedl (LMBB) syndrome: a radiographic study. *Skeletal Radiol*. 25, 655–660. doi: 10.1007/s002560050153

Solomon, B. D. (2011). VACTERL/VATER Association. *Orphanet J. Rare Dis*. 6:56. doi: 10.1186/1750-1172-6-56

Stavljenic-Rukavina, A. (2009). Hereditary kidney disorders. *EJIFCC* 20, 33–40.

Stoll, C., Dott, B., Alembik, Y., and Roth, M. P. (2014). Associated nonurinary congenital anomalies among infants with congenital anomalies of kidney and urinary tract (CAKUT). *Eur. J. Med. Genet*. 57, 322–328. doi: 10.1016/j.ejmg.2014.04.014

Titieni, A., and König, J. (2018). Nephronophthise und assoziierte Ziliopathien. *medizinische genetik* 30, 461–468. doi: 10.1007/s11825-018-0213-3

Tory, K., Menyhard, D. K., Woerner, S., Nevo, F., Gribouval, O., Kerti, A., et al. (2014). Mutation-dependent recessive inheritance of NPHS2-associated steroid-resistant nephrotic syndrome. *Nat. Genet*. 46, 299–304. doi: 10.1038/ng.2898

van der Ven, A. T., Vivante, A., and Hildebrandt, F. (2018). Novel insights into the pathogenesis of monogenic congenital anomalies of the kidney and urinary tract. *J. Am. Soc. Nephrol*. 29, 36–50. doi: 10.1681/ASN.2017050561

Vivante, A., and Hildebrandt, F. (2016). Exploring the genetic basis of early-onset chronic kidney disease. *Nat. Rev. Nephrol*. 12, 133–146. doi: 10.1038/nrneph.2015.205

Weber, S., Taylor, J. C., Winyard, P., Baker, K. F., Sullivan-Brown, J., Schild, R., et al. (2008). SIX2 and BMP4 mutations associate with anomalous kidney development. *J. Am. Soc. Nephrol*. 19, 891–903. doi: 10.1681/ASN.2006111282

Zankl, A., Duncan, E. L., Leo, P. J., Clark, G. R., Glazov, E. A., Addor, M.-C., et al. (2012). Multicentric carpotarsal osteolysis is caused by mutations clustering in the amino-terminal transcriptional activation domain of MAFB. *Am. J. Hum. Genet.* 90, 494–501. doi: 10.1016/j.ajhg.2012.01.003

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

BRIEF RESEARCH REPORT
published: 25 June 2021
doi: 10.3389/fgene.2021.682565

Challenging Disease Ontology by Instances of Atypical *PKHD1* and *PKD1* Genetics

*Jonathan de Fallois[1], Ria Schönauer[1], Johannes Münch[1], Mato Nagel[2], Bernt Popp[3] and Jan Halbritter[1]**

[1] Department of Endocrinology, Nephrology and Rheumatology, University of Leipzig Medical Center, Leipzig, Germany, [2] Center for Nephrology and Metabolic Disorders, Weißwasser, Germany, [3] Institute of Human Genetics, University of Leipzig Medical Center, Leipzig, Germany

Background: Autosomal polycystic kidney disease is distinguished into dominant (ADPKD) and recessive (ARPKD) inheritance usually caused by either monoallelic (*PKD1/PKD2*) or biallelic (*PKHD1*) germline variation. Clinical presentations are genotype-dependent ranging from fetal demise to mild chronic kidney disease (CKD) in adults. Additionally, exemptions from dominant and recessive inheritance have been reported in both disorders resulting in respective phenocopies. Here, we comparatively report three young adults with microcystic-hyperechogenic kidney morphology based on unexpected genetic alterations beyond typical inheritance.

Methods: Next-generation sequencing (NGS)-based gene panel analysis and multiplex ligation-dependent probe amplification (MLPA) of PKD-associated genes, familial segregation analysis, and reverse phenotyping.

Results: Three unrelated individuals presented in late adolescence for differential diagnosis of incidental microcystic-hyperechogenic kidneys with preserved kidney and liver function. Upon genetic analysis, we identified a homozygous hypomorphic *PKHD1* missense variant causing pseudodominant inheritance in a family, a large monoallelic *PKDH1*-deletion with atypical transmission, and biallelic *PKD1* missense hypomorphs with recessive inheritance.

Conclusion: By this report, we illustrate clinical presentations associated with atypical PKD-gene alterations beyond traditional modes of inheritance. Large monoallelic *PKHD1*-alterations as well as biallelic hypomorphs of both *PKD1* and *PKHD1* may lead to mild CKD in the absence of prominent macrocyst formation and functional liver impairment. The long-term renal prognosis throughout life, however, remains undetermined. Increased detection of atypical inheritance challenges our current thinking of disease ontology not only in PKD but also in Mendelian disorders in general.

Keywords: PKD1, PKHD1, ADPKD, ARPKD, cystic kidneys, chronic kidney disease

OPEN ACCESS

Edited by:
Xiu-An Yang,
Chengde Medical College, China

Reviewed by:
Gloria Pérez-Rubio,
Instituto Nacional de Enfermedades Respiratorias, Mexico
Espiridión Ramos-Martínez,
Universidad Nacional Autónoma de México, Mexico

***Correspondence:**
Jan Halbritter
jan.halbritter@medizin.uni-leipzig.de

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

Received: *18 March 2021*
Accepted: *05 May 2021*
Published: *25 June 2021*

Citation:
de Fallois J, Schönauer R, Münch J, Nagel M, Popp B and Halbritter J (2021) Challenging Disease Ontology by Instances of Atypical PKHD1 and PKD1 Genetics. Front. Genet. 12:682565. doi: 10.3389/fgene.2021.682565

INTRODUCTION

Polycystic kidney disease (PKD) is a genetically and clinically heterogeneous condition. In recent years, an increasing number of genetic alterations were associated with PKD (McConnachie et al., 2020). Apart from ongoing detection of new disease-causing alleles in known PKD genes, novel disease genes, such as *GANAB* (Porath et al., 2016), *ALG9* (Besse et al., 2019), *ALG8* (Besse et al., 2017), *DNAJB11* (Cornec-Le Gall et al., 2018a), and *DZIP1L* (Lu et al., 2017), have been identified. By mode of inheritance, PKD is categorized into autosomal dominant (ADPKD, MIM#173900/MIM#613095) and autosomal recessive (ARPKD, MIM#263200). While ADPKD is a quite frequent adult-onset condition due to monoallelic *PKD1* or *PKD2* variation, ARPKD is a rare childhood-onset disorder.

With an estimated prevalence of about 1:1,000, ADPKD is indeed the leading genetic cause of end-stage kidney disease (ESKD) among adults (Cornec-Le Gall et al., 2018b; Lanktree et al., 2018). Patients with protein truncating *PKD1* variants are at higher risk of developing ESKD early in life than patients with non-truncating *PKD1* variants, and *PKD2* carriers express an even milder form of ADPKD often without requirement of renal replacement therapies throughout life (Hwang et al., 2016; Bergmann et al., 2018). However, even in patients with truncating *PKD1* variants, exceptionally mild renal phenotypes can be found similar to patients with *PKD2* disease (Lanktree et al., 2021).

On the contrary, ARPKD affects about 1:20,000 live births and results from biallelic variants in *PKHD1*, encoding the ciliary protein fibrocystin. Although perinatal manifestation is most common, genetic studies in adults demonstrated that ARPKD is also diagnosed later in life with an almost ADPKD-like phenotype (Cornec-Le Gall et al., 2018b; Schönauer et al., 2020). Similar to ADPKD, genotype–phenotype correlations have been established in ARPKD. While biallelic truncating variants are mostly associated with fetal demise (Bergmann et al., 2005; Erger et al., 2017), combinations of truncating and missense variants can be found in children surviving the first year of life, and biallelic missense variants are commonly associated with milder forms of disease compatible with ESKD in adulthood (Bergmann et al., 2018; Schönauer et al., 2020). Nevertheless, exceptional cases with biallelic truncation surviving the perinatal period (Ebner et al., 2017; Burgmaier et al., 2018) and biallelic missense variants linked to severe courses (Furu et al., 2003; Bergmann et al., 2005) were reported. Other studies suggested an impact of the type of affected PKHD1-protein domain on the clinical phenotype (Furu et al., 2003; Bergmann et al., 2005).

Clinical presentations of ADPKD and ARPKD are hence genotype-dependent ranging from fetal demise to mild chronic kidney disease (CKD) in adults (Bergmann et al., 2018). In addition to CKD, hepatic involvement presents differently in ADPKD and ARPKD. While in ADPKD, cystic liver enlargement overlapping with autosomal dominant polycystic liver disease (ADPLD) is a characteristic hallmark, ARPKD often leads to rather microcystic liver fibrosis and more often results in consecutive liver transplantation (Cornec-Le Gall et al., 2018b).

Besides classic ADPKD or ARPKD, other rare cystic kidney disorders include nephronophthisis-related ciliopathies, based on biallelic variants in multiple genes encoding ciliary and centrosomal proteins (Bergmann et al., 2018; Adamiok-Ostrowska and Piekiełko-Witkowska, 2020).

Before the advent of next-generation sequencing (NGS) techniques, establishing the correct diagnosis was mainly based on clinical presentation and kidney imaging. NGS has since increased diagnostic accuracy, also in PKD, and led to an improved understanding of the underlying disease mechanisms (Obeidova et al., 2020).

While exemptions from dominant and recessive inheritance in ADPKD and ARPKD have been reported before (Mantovani et al., 2020; Obeidova et al., 2020), we aim at illustrating the complexity of PKD genetics through peculiar genetic alterations of both *PKD1* and *PKHD1* in three young adult females and one male infant. The presented examples demonstrate unusual modes of inheritance and exceptional clinical presentations, challenging the given disease ontology in PKD.

MATERIALS AND METHODS

Patients

Index patients and their families were recruited from the outpatient clinic for *Hereditary Kidney Disorders* at the University of Leipzig Medical Center. All families were included upon written informed consent approved by the local Institutional Review Board (IRB) at the University of Leipzig, Germany (IRB00001750; #402/16-ek).

Next-Generation Sequencing and Variant Analysis

All study participants underwent genetic analysis conducted from blood-derived DNA samples by NGS-based gene-panel analysis and multiplex ligation-dependent probe amplification (MLPA). Gene panels included all known PKD-associated genes displayed in **Supplementary Table 1** (*PKD1, PKD2,* and *PKHD1*, among others). Further segregation analysis was done by Sanger-sequencing of respective gene variants in available family members. MLPA was used to detect copy number variations (CNVs) and to validate the NGS data set in complex gene regions, ensuring that no structural alteration was overlooked. *In silico* prediction with combined annotation-dependent depletion score (CADD-PHRED v1.6: 23.80) (Rentzsch et al., 2019) was used to categorize newly detected variants in terms of pathogenicity. Variant analysis and interpretation was carried out in accordance with the criteria established by the *American College of Medical Genetics and Genomics* (ACMG) (Richards et al., 2015).

Reverse Phenotyping

Reverse phenotyping was done in all family members available. Laboratory tests (e.g., serum-creatinine, urea, urine albumin/creatinine ratio) and kidney imaging [renal ultrasound, magnetic resonance imaging (MRI)] were ascertained to derive genotype–phenotype correlations.

RESULTS

Family 1: *PKHD1* Disease With Pseudodominant Inheritance

Clinical Description

The asymptomatic index patient (ID 1, female, 18 years) of Family 1 was from non-consanguineous Romany origin without any former medical conditions. She was seeking medical advice during her first pregnancy at 17 years of age. Pregnancy had been complicated by oligohydramnios and fetal hyperechogenic kidneys upon prenatal ultrasound. After birth, the newborn (ID 1.1) demonstrated small cystic alterations of both kidneys in accordance with infantile ARPKD. Extrarenally, bilateral pneumothoraces but no hepatic involvement were diagnosed in the newborn. Upon closer examination, kidney ultrasound in the mother (ID 1) revealed hyperechogenic parenchyma and microcystic alterations at the age of 18 years (**Figure 1**). Laboratory analysis showed normal serum creatinine with an increased estimated glomerular filtration rate (eGFR CKD-EPI: 122 ml/min/1.73 m^2), mild proteinuria (albumin/creatinine ratio: 174 mg/g creatinine), and normal liver function (**Table 1**).

Molecular Genetics

With affected individuals in two subsequent generations, the pedigree suggested dominant inheritance (**Figure 1**). Genetic analyses of the index patient (ID 1), however, revealed a known pathogenic *PKHD1* missense variant [NM_138694.3: c.664A > G, p.(Ile222Val)] in the homozygous state. This variant affects an evolutionary conserved amino acid residue located near the N-terminal extracellular tail of fibrocystin within the second IPT/TIG (Ig-like, plexins, and transcription factors) domain[1] (Gunay-Aygun et al., 2010). Allele frequency in Europeans (non-Finish) amounts to 0.0056% with no homozygous cases reported in population databases (gnomAD). *In silico* prediction using CADD yielded a score of 16.14 (CADD-PHRED v1.6) (Rentzsch et al., 2019), compatible with pathogenicity in a recessive disease model.

Surprisingly, the newborn, who was diagnosed with congenital kidney disease *in utero*, was found to carry compound heterozygous *PKDH1* variants: the aforementioned maternal missense variant [c.664A > G, p.(Ile222Val)] in exon 9 plus another pathogenic *PKHD1* nonsense variant in exon 50 out of 67 [NM_138694.3: c.7916C > A, p.(Ser2639*)] on the paternal allele (CADD-PHRED v1.6: 40.00) (Denamur et al., 2010; Rentzsch et al., 2019). For the latter variant, population allele frequency is estimated at 0.0039% (gnomAD). Its deduced impact on the protein level leads to the introduction of a stop codon after the sixth PbH1 (parallel beta-helix) repeat resulting in the loss of the critical second G8 domain containing eight conserved glycine residues likely important for protein stability (see text foot note 1) (He et al., 2006).

Family 2: *PKHD1* Disease With Atypical Transmission of Monoallelic CNV

Clinical Description

In family 2 from central Europe, the index patient (ID 2; female, 27 years) came to obtain medical advice after previous genetic testing. At the age of 18 years, ID 2 was hospitalized due to repeated urinary tract infections and

[1]https://www.uniprot.org/uniprot/P08F94

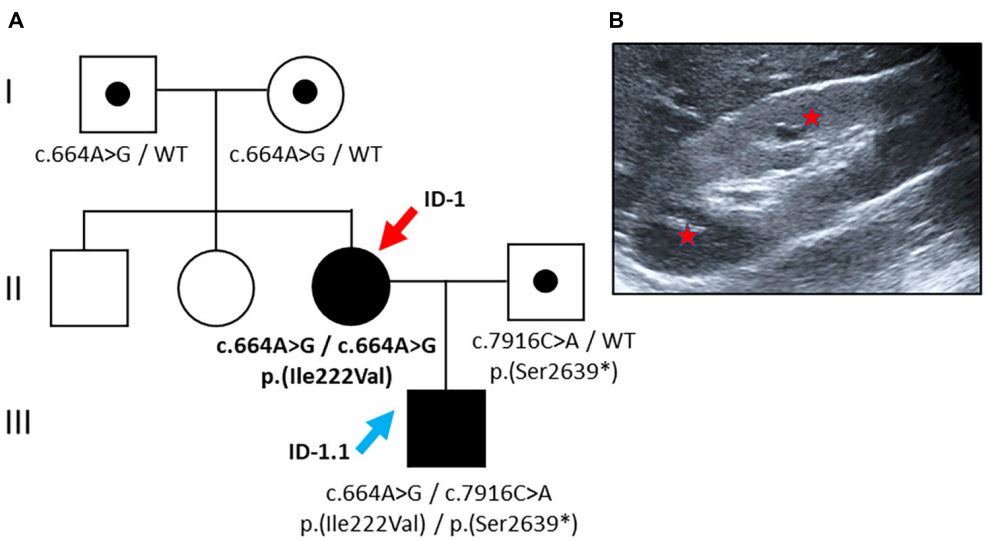


FIGURE 1 | Pedigree and renal imaging (Family 1). **(A)** Family pedigree: Individuals suffering from PKD are indicated as black symbols, and variant carriers are denoted with a black dot. The female index patient (red arrow) carries a homozygous missense variant [NM_138694.3: c.664A > G, p.(Ile222Val)]. Her son (blue arrow) carries compound heterozygous *PKDH1*-variants: maternal missense variant plus a putative paternally transmitted pathogenic nonsense variant [NM_138694.3: c.7916C > A, p.(Ser2639*)]. WT: wild-type. **(B)** Kidney ultrasound of ID-1 at the age of 18 years displaying mild renal phenotype with hyperechogenic parenchyma and microcystic alterations (red stars highlight small cysts).

TABLE 1 | Phenotypic and genotypic characteristics of patients included in this case series.

Patient	FAM 1 – ID 1	FAM 1 – ID 1.1	FAM 2 – ID 2	FAM 3 – ID 3
Gene	*PKHD1*	*PKHD1*	*PKHD1*	*PKD1*
Variant	hom.	comp.	het.	comp. het.
1st allele – c. position	c.664A > G	het. c.664A > G	c.(?_5909)_(12225_?)del	c.11723T > C
1st allele – p. position	p.(Ile222Val)	p.(Ile222Val)	Ex37_Ex67del	p.(Leu3908Pro)
allele frequency (gnomAD)	0.0056%	0.0056%	Novel	Novel
2nd allele – c. position	c.664A > G	c.7916C > A	WT	c.4709C > T
2nd allele – p. position	p.(Ile222Val)	p.(Ser2639*)		p.(Thr1570Met)
allele frequency (gnomAD)	0.0056%	0.0039%		0.0031%
Sex	Female	Male	Female	Female
Origin	Romany	Romany	German	German
Parental consanguinity	No	Not reported	No	No
Age at inclusion	18 years	1.5 years	27 years	19 years
Age at 1st manifestation	18 years	Prenatal oligohydramnion	18 years	Prenatal renal hyper-echogenicity
Chronic kidney disease (stage)[#]	G1	G1-2	G1	G1
Proteinuria: albumin/creatinine ratio (mg/g creatinine)	174	31	6	19.5
Increased echogenicity	+	+	+	+
Renal cysts	+ Bilateral small cysts	++ bilateral multiple cysts	+++ bilateral multiple cysts, enlarged kidneys	+++ bilateral multiple, tiny cysts "Salt and pepper pattern"
Kidney stones	–	–	+	–
Extrarenal manifestation	–	Bilateral pneumothoraces	–	–
Arterial hypertension	–	+	–	+
Liver phenotype	–	–	–	–

The variants reported refer to RefSeq NM_138694.3 (PKDH1), NM_001009944.2 (PKD1).
Abbreviations: comp. het., compound heterozygous; hom., homozygous; het, heterozygous; WT, wild-type; y, years; [#]Chronic kidney disease refers to the classification of the "Kidney Disease: Improving Global Outcomes" (KDIGO) initiative. gnomAD v2.1.1 – population specific allele frequency (http://gnomad.broadinstitute.org/).

conspicuous renal morphology with increased parenchymal echogenicity. Abdominal MRI showed symmetric enlargement of both kidneys [diameter (*length* × *width*): right: 130 × 60 mm; left: 146 × 72 mm] with small cystic alterations but regular hepatic morphology (**Figure 2**). At the time of presentation in our clinic, the kidney function was preserved (eGFR CKD-EPI: 130 ml/min/1.73 m^2), and no proteinuria was detected (albumin/creatinine ratio: 6 mg/g) (**Table 1**). The siblings and parents of the patient showed no renal or liver phenotype.

Molecular Genetics

Genetic analysis revealed a large heterozygous deletion of *PKHD1* c.(?_5909)_(12225_?)del (designated as Ex37_Ex67del) encompassing the extracellular domains of PbH1-9 and one of the critical G8 repeats of fibrocystin close to the cell membrane (see text foot note 1). Familial segregation analysis of both parents showed no genetic alteration of *PKHD1*, suggesting the *de novo* nature of this finding. However, further segregation analysis including the younger siblings of ID 2 who were non-identical twins yielded duplication of the same region c.(?_5909)_(12225_?)dup (designated as Ex37_Ex67dup) in one of the twin sisters without clinical manifestation.

Family 3: *PKD1* Disease With Recessive Inheritance

Clinical Description

This index patient (ID 3, female, 19 years) of central European descent had been congenitally diagnosed with infantile PKD and suspected ARPKD because of early-onset cystic kidneys. The kidney function, however, was completely preserved (eGFR CKD-EPI 133 ml/min/1.73 m^2) with merely mild proteinuria (albumin/creatinine ratio: 29 mg/g creatinine) (**Table 1**). While liver imaging presented normal, kidney ultrasound showed bilaterally enlarged kidneys [diameter (*length* × *width*): right: 117 × 52 mm; left: 120 × 57 mm] with increased parenchymal echogenicity and abnormal corticomedullary differentiation (the so-called "salt and pepper" pattern) (Iorio et al., 2020) (**Figure 3**).

Molecular Genetics

Gene-panel analysis of PKD-associated genes in the index patient yielded two rare heterozygous *PKD1* missense variants. Diagnostic variants in *PKHD1* or other known kidney cyst genes were not detected. Both *PKD1* variants were of uncertain significance by ACMG classification [NM_001009944.2: c.11723T > C, p.(Leu3908Pro) and NM_001009944.2: c.4709C > T, p.(Thr1570Met)]

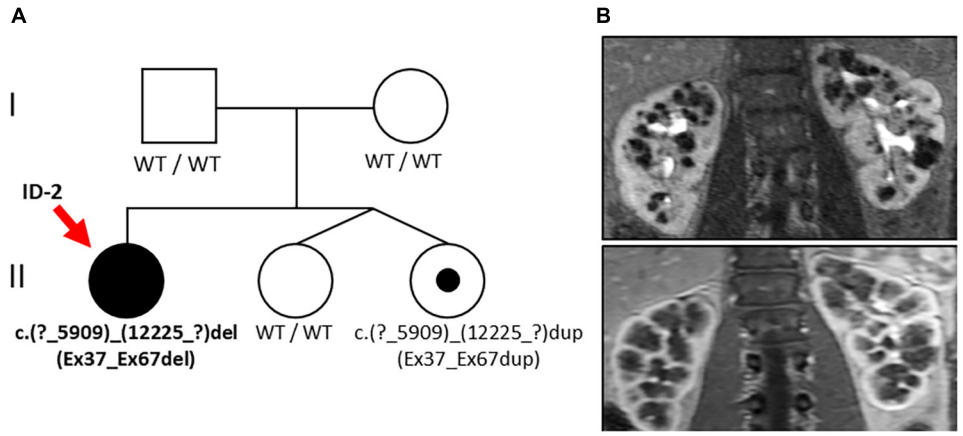


FIGURE 2 | Pedigree and renal imaging (Family 2). **(A)** Family pedigree: Individuals suffering from PKD are indicated as black symbols and variant carriers with a black dot. Genetic analysis in the index patient (ID-2, red arrow) revealed the large heterozygous deletion NM_138694.3: c.(?_5909)_(12225_?)del of *PKHD1* (Ex37_Ex67del) and familial segregation analysis in one younger sibling yielded duplication of the same alteration c.(?_5909)_(12225_?)dup (Ex37_Ex67dup). **(B)** Abdominal MRI of ID 2 at the age of 18 years showing bilateral cystic kidney enlargement.

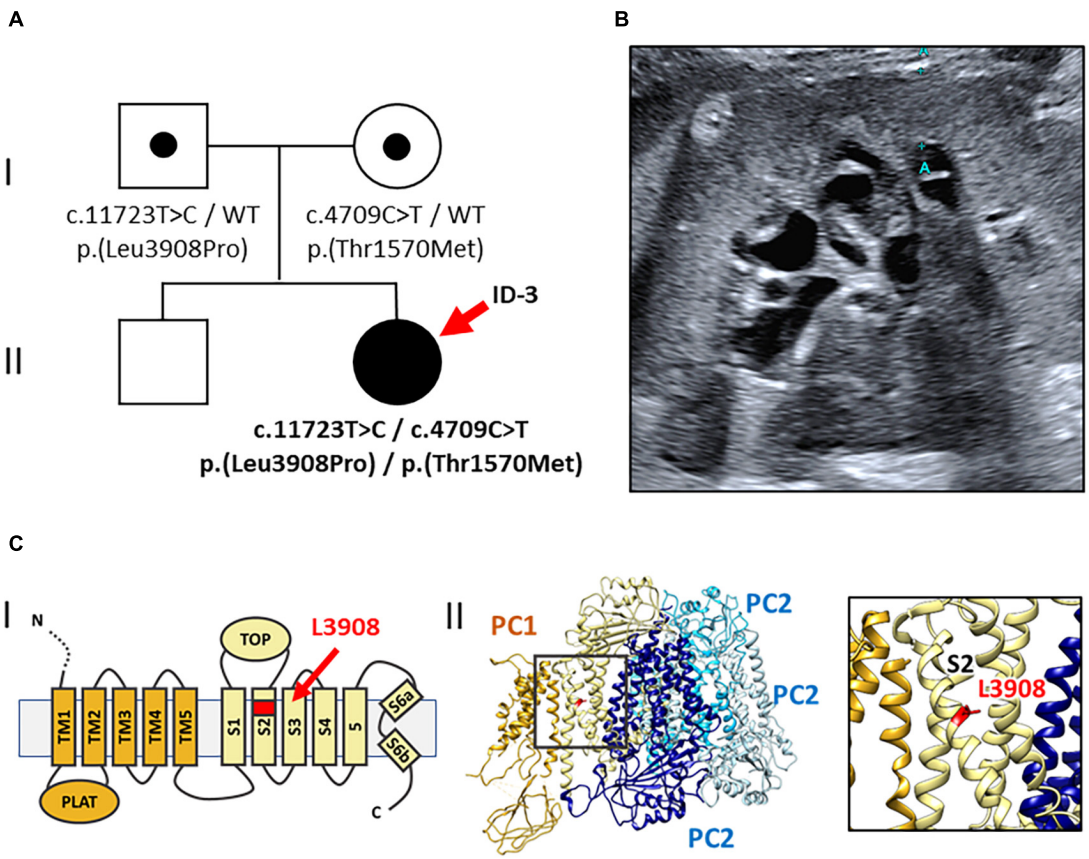


FIGURE 3 | Pedigree, renal imaging, and 3D-PC1-protein structure (Family 3). **(A)** Family pedigree: ID-3 (red arrow) with compound heterozygous rare *PKD1* missense variants both of uncertain significance [NM_001009944.2: c.11723T > C, p.(Leu3908Pro) and c.4709C > T, p.(Thr1570Met)]. WT: wild-type. **(B)** Kidney ultrasound of ID-3 at the age of 19 years shows enlarged, microcystic-hyperechogenic kidneys. **(C)** Localization of the novel variant p.(Leu3908Pro) in the PC1 protein structure. **(I)** Schematic illustration of the PC1 transmembrane region; L3908 is located in transmembrane helix S2 (red arrow). **(II)** Cryo-EM structure of human PC1/PC2 complex (pdb 6A70) (left) and zoom-in view (right) showing the position of L3908 in helix S2 (red) (Su et al., 2018).

(Kane et al., 2006; Richards et al., 2015). Parental segregation analysis confirmed compound heterozygosity in the index patient with c.11723T > C [p.(Leu3908Pro)] being paternally inherited and c.4709C > T [p.(Thr1570Met)] being maternally transmitted. Of note, clinical examination of both parents was unremarkable, including absence of kidney and liver cysts at the age of 32 years. The first *PKD1* variant is novel, neither reported in population databases (gnomAD) nor disease databases [HGMD Professional Version 2020.4/Mayo PKD-database[2] (Stenson et al., 2020)]. The respective amino acid residue Leu3908 is evolutionary conserved in vertebrates belonging to the transmembrane helix S2 domain of polycystin 1 (**Figure 3**)(Su et al., 2018). Introduction of a proline residue into the alpha-helix is likely deleterious, leading to conformational impairments. Hence, *in silico* prediction using CADD also suggested pathogenicity (CADD-PHRED v1.6: 23.80)(Rentzsch et al., 2019).

The second *PKD1* variant [c.4709C > T, p.(Thr1570Met)] was previously reported in the literature (Vujic et al., 2010), where it was found to be associated with *in utero* onset disease in the homozygous state but incomplete penetrance in heterozygous carriers. Thr1570 constitutes a highly conserved residue in vertebrates and localizes to the PKD11 domain (CADD-PHRED v1.6: 26.30)(Rentzsch et al., 2019). Its allele frequency in Europeans is 0.0031% without homozygous findings in population databases (gnomAD).

DISCUSSION

Genetic variant interpretation and its clinical correlation has become one of the major challenges of our times. Clinical geneticists and treating physicians are increasingly confronted with new rare variants that require meaningful translation into clinical practice. Furthermore, genetic results often question our traditional categories and force us to take a new perspective on disease classification and terminology. This dilemma is exemplified by the presented cases of three young females with mild cystic kidney alterations illustrating the difficulties and chances of adequate genetic counseling and disease terminology in the era of broadly accessible genetic testing. All index patients presented clinically asymptomatic with preserved kidney function as young adults.

The common denominator is hampered disease prediction through atypical and novel genetic findings with little clinical experience derivable from previously published cases. When it comes to disease ontology, neither single heterozygous *PKHD1* germline variants nor biallelic *PKD1* hypomorphs do match the given terminology of autosomal-recessive and autosomal-dominant PKD but mimic each other's renal manifestation. Therefore, common risk stratification tools like the Mayo imaging classification (Irazabal et al., 2015) or the PROPKD-score (Cornec-Le Gall et al., 2016) are not applicable, complicating genetic counseling and decision-making in respective families.

The pedigree of family 1 revealed affected family members in two subsequent generations suggesting dominant inheritance. Genetic testing, however, showed *PKHD1* variants defining ARPKD and thereby led to correction of previous assumptions. While pseudodominant inheritance was reported in other Mendelian kidney diseases, e.g., Gitelman's disease (de La Faille et al., 2011), this is the first report of pseudodominant inheritance in ARPKD. The two affected family members, the index patient (Fam 1 – ID 1) and her newborn son (ID 1.1), impressively illustrate the broad clinical spectrum of ARPKD ranging from severe congenital manifestation to mild CKD in adults depending on the nature and combination of respective *PKHD1* alleles (Burgmaier et al., 2019). The homozygous pathogenic *PKHD1* missense variant [p.(Ile222Val)] in the index patient leads to a mild course of disease potentially even compatible with normal life expectancy. The very missense variant has been reported frequently in compound heterozygous state in patients with ARPKD before (Bergmann et al., 2005; Denamur et al., 2010; Gunay-Aygun et al., 2010; Brinkert et al., 2013; Obeidova et al., 2020), but has not yet been reported homozygously to the best of our knowledge. Therefore, long-term prognoses remain speculative. Asymptomatic presentation with normal renal and liver function at age 18 suggests that biallelic occurrence of the Ile222Val missense variant is linked to a mild renal phenotype and that only additional second/third hits result in more severe courses of ARPKD. This theory is supported by the clinical presentation of her newborn carrying an additional known *PKHD1* null allele (p.Ser2639*), which is obviously associated with aggravated disease severity (Denamur et al., 2010; Melchionda et al., 2016; Szabó et al., 2018).

The second family represents another example of mild ARPKD in adults. This time, the index patient (Fam 2 – ID 2) was found to carry a large heterozygous deletion in *PKHD1* (Ex37_Ex67del). As clinically healthy parents showed *PKHD1* wild-type sequences, *de novo* nature of the deletion was assumed until further segregation analysis in the sister yielded duplication of the very same region without clinical manifestation. This unexpected finding leads to verification of our initial assumption and established yet another mechanism of inheritance in terms of recurrent unequal meiotic crossovers or an unbalanced cryptic chromosomal translocation. This case demonstrates sporadic manifestation of ARPKD based on single heterozygous alterations of *PKHD1*, an observation that was increasingly reported recently (Besse et al., 2017). Additionally, data in mice suggest that heterozygous *PKHD1* carriers may develop ARPKD-associated liver cysts only later in life after the occurrence of a second hit affecting the *PKHD1* wild-type allele in somatic tissues (Shan et al., 2019; Besse et al., 2020). However, given the large size and complex exon structure of *PKHD1*, we cannot exclude a deep intronic splice variant *in trans* with the herein described deletion as an alternative explanation (Michel-Calemard et al., 2009; Chen et al., 2019). *PKHD1* is not sufficiently expressed in blood to be easily accessible to RNA sequencing approaches, but future analyses from cultured urinary renal epithelial cells might enable screening for these rare variants in atypical cases with single *PKHD1* variants.

[2]http://pkdb.pkdcure.org

Lastly, the third case is an example for biallelic hypomorphic *PKD1* variants mimicking ARPKD. First, the index patient presented in childhood with enlarged kidneys and morphological signs of ARPKD upon kidney ultrasound (Iorio et al., 2020); second, the pedigree of the family indicated recessive inheritance. Surprisingly, genetic analysis in the index patient yielded no diagnostic findings in *PKHD1*, but compound heterozygosity for two rare *PKD1* missense variants, both of uncertain significance. One of them [c.4709C > T, p.(Thr1570Met)], however, was previously associated with early-onset ADPKD in the homozygous state (Vujic et al., 2010). In this case, two *PKD1* variants *in trans* manifested as phenocopy of ARPKD. One of these hypomorphic alleles alone likely results in a normal phenotype, but two alleles *in trans* cause early-onset cystic kidney disease. As cyst initiation is assumed to depend on the dosage of functional Polycystin 1 (Rossetti et al., 2009), the combination of two hypomorphic variants could additively reduce the *PKD1* function below a critical threshold to result in clinically apparent disease. Coincidence of biallelic hypomorphic *PKD1* variants represents an example for PKD without an apparent family history (Iliuta et al., 2017).

Given these examples, we suggest revisiting and adapting the classification of genetic PKD to enable the description of the apparent continuum of AR/ADPKD genotypes and phenotypes. In a subset of cases, clinical presentation and associated genetic findings do not fit into the current ontology. As a first step, a more flexible disease terminology including the respective gene name was suggested in 2018 without being widely adopted to date (Cornec-Le Gall et al., 2018b). In other renal conditions, however, such as autosomal-dominant tubulointerstitial kidney disease (ADTKD), a new disease classification was successfully established, combining the mode of inheritance, major clinical manifestation, and underlying genetic defect in one term (ADTKD-*UMOD*; ADTKD-*REN*, ADTKD-*MUC1*, and ADTKD-*HNF1B*) (Bleyer et al., 2017).

In summary, this report demonstrates the genetic complexity of PKD. Attributing the correct diagnosis and providing adequate genetic counseling to affected patients and their families is not self-evident. Atypical presentations should prompt immediate genetic testing to increase diagnostic accuracy by verification or correction of clinically assumed diagnoses.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: requests to access these datasets should be directed to JH, jan.halbritter@medizin.uni-leipzig.de.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by local Institutional Review Board (IRB) at the Universities of Leipzig, Germany (IRB00001750; #402/16-ek). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

JF and JH contributed to the conception and design of the study, and helped in investigation. JF wrote the first draft of the manuscript. RS, JM, and BP wrote the sections of the manuscript. JF, RS, MN, BP, and JH contributed to the analysis. JF and RS helped in visualization. All authors contributed to manuscript revision, read, and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.682565/full#supplementary-material

REFERENCES

Adamiok-Ostrowska, A., and Piekiełko-Witkowska, A. (2020). Ciliary genes in renal cystic diseases. *Cells* 9:907. doi: 10.3390/cells9040907

Bergmann, C., Guay-Woodford, L. M., Harris, P. C., Horie, S., Peters, D. J. M., and Torres, V. E. (2018). Polycystic kidney disease. *Nat. Rev. Dis. Primers* 4:50. doi: 10.1038/s41572-018-0047-y

Bergmann, C., Senderek, J., Windelen, E., Küpper, F., Middeldorf, I., Schneider, F., et al. (2005). Clinical consequences of PKHD1 mutations in 164 patients with autosomal-recessive polycystic kidney disease (ARPKD). *Kidney Int.* 67, 829–848. doi: 10.1111/j.1523-1755.2005.00148.x

Besse, W., Chang, A. R., Luo, J. Z., Triffo, W. J., Moore, B. S., Gulati, A., et al. (2019). ALG9 mutation carriers develop kidney and liver cysts. *J. Am. Soc. Nephrol.* 30, 2091–2102. doi: 10.1681/ASN.2019030298

Besse, W., Dong, K., Choi, J., Punia, S., Fedeles, S. V., Choi, M., et al. (2017). Isolated polycystic liver disease genes define effectors of polycystin-1 function. *J. Clin. Invest.* 127, 1772–1785. doi: 10.1172/JCI90129

Besse, W., Roosendaal, C., Tuccillo, L., Roy, S. G., Gallagher, A.-R., and Somlo, S. (2020). Adult inactivation of the recessive polycystic kidney disease gene causes polycystic liver disease. *Kidney360* 1, 1068–1076. doi: 10.34067/kid.0002522020

Bleyer, A. J., Kidd, K., Živná, M., and Kmoch, S. (2017). Autosomal dominant tubulointerstitial kidney disease. *Adv. Chronic Kidney Dis.* 24, 86–93. doi: 10.1053/j.ackd.2016.11.012

Brinkert, F., Lehnhardt, A., Montoya, C., Helmke, K., Schaefer, H., Fischer, L., et al. (2013). Combined liver-kidney transplantation for children with autosomal recessive polycystic kidney disease (ARPKD): indication and outcome. *Transpl. Int.* 26, 640–650. doi: 10.1111/tri.12098

Burgmaier, K., Kilian, S., Bammens, B., Benzing, T., Billing, H., Büscher, A., et al. (2019). Clinical courses and complications of young adults with autosomal recessive polycystic kidney disease (ARPKD). *Sci. Rep.* 9:7919. doi: 10.1038/s41598-019-43488-w

Burgmaier, K., Kunzmann, K., Ariceta, G., Bergmann, C., Buescher, A. K., Burgmaier, M., et al. (2018). Risk factors for early dialysis dependency in autosomal recessive polycystic kidney disease. *J. Pediatr.* 199, 22–28.e6. doi: 10.1016/j.jpeds.2018.03.052

Chen, J., Ma, N., Zhao, X., Li, W., Zhang, Q., Yuan, S., et al. (2019). A rare deep intronic mutation of PKHD1 gene, c.8798-459 C > A, causes autosomal recessive polycystic kidney disease by pseudoexon activation. *J. Hum. Genet.* 64, 207–214. doi: 10.1038/s10038-018-0550-8

Cornec-Le Gall, E., Audrézet, M.-P., Rousseau, A., Hourmant, M., Renaudineau, E., Charasse, C., et al. (2016). The PROPKD score: a new algorithm to predict renal

survival in autosomal dominant polycystic kidney disease. *J. Am. Soc. Nephrol.* 27, 942–951. doi: 10.1681/ASN.2015010016

Cornec-Le Gall, E., Olson, R. J., Besse, W., Heyer, C. M., Gainullin, V. G., Smith, J. M., et al. (2018a). Monoallelic mutations to DNAJB11 cause atypical autosomal-dominant polycystic kidney disease. *Am. J. Hum. Genet.* 102, 832–844. doi: 10.1016/j.ajhg.2018.03.013

Cornec-Le Gall, E., Torres, V. E., and Harris, P. C. (2018b). Genetic complexity of autosomal dominant polycystic kidney and liver diseases. *J. Am. Soc. Nephrol.* 29, 13–23. doi: 10.1681/ASN.2017050483

de La Faille, R., Vallet, M., Venisse, A., Nau, V., Collet-Gaudillat, C., Houillier, P., et al. (2011). A pseudo-dominant form of Gitelman's syndrome. *Clin. Kidney J.* 4, 386–389. doi: 10.1093/ndtplus/sfr094

Denamur, E., Delezoide, A. L., Alberti, C., Bourillon, A., Gubler, M. C., Bouvier, R., et al. (2010). Genotype-phenotype correlations in fetuses and neonates with autosomal recessive polycystic kidney disease. *Kidney Int.* 77, 350–358. doi: 10.1038/ki.2009.440

Ebner, K., Dafinger, C., Ortiz-Bruechle, N., Koerber, F., Schermer, B., Benzing, T., et al. (2017). Challenges in establishing genotype–phenotype correlations in ARPKD: case report on a toddler with two severe PKHD1 mutations. *Pediatr. Nephrol.* 32, 1269–1273. doi: 10.1007/s00467-017-3648-x

Erger, F., Brüchle, N. O., Gembruch, U., and Zerres, K. (2017). Prenatal ultrasound, genotype, and outcome in a large cohort of prenatally affected patients with autosomal-recessive polycystic kidney disease and other hereditary cystic kidney diseases. *Arch. Gynecol. Obstet.* 295, 897–906. doi: 10.1007/s00404-017-4336-6

Furu, L., Onuchic, L. F., Gharavi, A., Hou, X., Esquivel, E. L., Nagasawa, Y., et al. (2003). Milder presentation of recessive polycystic kidney disease requires presence of amino acid substitution mutations. *J. Am. Soc. Nephrol.* 14, 2004–2014. doi: 10.1097/01.asn.0000078805.87038.05

Gunay-Aygun, M., Tuchman, M., Font-Montgomery, E., Lukose, L., Edwards, H., Garcia, A., et al. (2010). PKHD1 sequence variations in 78 children and adults with autosomal recessive polycystic kidney disease and congenital hepatic fibrosis. *Mol. Genet. Metab.* 99, 160–173. doi: 10.1016/j.ymgme.2009.10.010

He, Q., Liu, X., Li, Q., Studholme, D. J., Li, X., and Liang, S. (2006). G8: a novel domain associated with polycystic kidney disease and non-syndromic hearing loss. *Bioinformatics* 22, 2189–2191. doi: 10.1093/bioinformatics/btl123

Hwang, Y. H., Conklin, J., Chan, W., Roslin, N. M., Liu, J., He, N., et al. (2016). Refining genotype-phenotype correlation in autosomal dominant polycystic kidney disease. *J. Am. Soc. Nephrol.* 27, 1861–1868. doi: 10.1681/ASN.2015060648

Iliuta, I. A., Kalatharan, V., Wang, K., Cornec-Le Gall, E., Conklin, J., Pourafkari, M., et al. (2017). Polycystic kidney disease without an apparent family history. *J. Am. Soc. Nephrol.* 28, 2768–2776. doi: 10.1681/ASN.2016090938

Iorio, P., Heidet, L., Rutten, C., Garcelon, N., Audrézet, M.-P., Morinière, V., et al. (2020). The "salt and pepper" pattern on renal ultrasound in a group of children with molecular-proven diagnosis of ciliopathy-related renal diseases. *Pediatr. Nephrol.* 35, 1033–1040. doi: 10.1007/s00467-020-04480-z

Irazabal, M. V., Rangel, L. J., Bergstralh, E. J., Osborn, S. L., Harmon, A. J., Sundsbak, J. L., et al. (2015). Imaging classification of autosomal dominant polycystic kidney disease: a simple model for selecting patients for clinical trials. *J. Am. Soc. Nephrol.* 26, 160–172. doi: 10.1681/ASN.2013101138

Kane, I., Lindley, R., Lewis, S., and Sandercock, P. (2006). Impact of stroke syndrome and stroke severity on the process of consent in the third international stroke trial. *Cerebrovasc. Dis.* 21, 348–352. doi: 10.1159/000091541

Lanktree, M. B., Guiard, E., Akbari, P., Pourafkari, M., Iliuta, I.-A., Ahmed, S., et al. (2021). Patients with protein-truncating PKD1 mutations and mild ADPKD. *Clin. J. Am. Soc. Nephrol.* 16, 374–383. doi: 10.2215/CJN.11100720

Lanktree, M. B., Haghighi, A., Guiard, E., Iliuta, I. A., Song, X., Harris, P. C., et al. (2018). Prevalence estimates of polycystic kidney and liver disease by population sequencing. *J. Am. Soc. Nephrol.* 29, 2593–2600. doi: 10.1681/ASN.2018050493

Lu, H., Galeano, M. C. R., Ott, E., Kaeslin, G., Kausalya, P. J., Kramer, C., et al. (2017). Mutations in DZIP1L, which encodes a ciliary-transition-zone protein, cause autosomal recessive polycystic kidney disease. *Nat. Genet.* 49, 1025–1034. doi: 10.1038/ng.3871

Mantovani, V., Bin, S., Graziano, C., Capelli, I., Minardi, R., Aiello, V., et al. (2020). Gene panel analysis in a large cohort of patients with autosomal dominant polycystic kidney disease allows the identification of 80 potentially causative novel variants and the characterization of a complex genetic architecture in a subset of families. *Front. Genet.* 11:464. doi: 10.3389/fgene.2020.00464

McConnachie, D. J., Stow, J. L., and Mallett, A. J. (2020). Ciliopathies and the kidney: a review. *Am. J. Kidney Dis.* 66, 50–62. doi: 10.1053/j.ajkd.2020.08.012

Melchionda, S., Palladino, T., Castellana, S., Giordano, M., Benetti, E., De Bonis, P., et al. (2016). Expanding the mutation spectrum in 130 probands with ARPKD: identification of 62 novel PKHD1 mutations by sanger sequencing and MLPA analysis. *J. Hum. Genet.* 61, 811–821. doi: 10.1038/jhg.2016.58

Michel-Calemard, L., Dijoud, F., Till, M., Lambert, J. C., Vercherat, M., Tardy, V., et al. (2009). Pseudoexon activation in the PKHD1 gene: a French founder intronic mutation IVS46+653A>G causing severe autosomal recessive polycystic kidney disease. *Clin. Genet.* 75, 203–206. doi: 10.1111/j.1399-0004.2008.01106.x

Obeidova, L., Seeman, T., Fencl, F., Blahova, K., Hojny, J., Elisakova, V., et al. (2020). Results of targeted next-generation sequencing in children with cystic kidney diseases often change the clinical diagnosis. *PLoS One* 15:e0235071. doi: 10.1371/journal.pone.0235071

Porath, B., Gainullin, V. G., Cornec-Le Gall, E., Dillinger, E. K., Heyer, C. M., Hopp, K., et al. (2016). Mutations in GANAB, encoding the glucosidase IIα subunit, cause autosomal-dominant polycystic kidney and liver disease. *Am. J. Hum. Genet.* 98, 1193–1207. doi: 10.1016/j.ajhg.2016.05.004

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., and Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* 47, D886–D894. doi: 10.1093/nar/gky1016

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* 17, 405–424. doi: 10.1038/gim.2015.30

Rossetti, S., Kubly, V. J., Consugar, M. B., Hopp, K., Roy, S., Horsley, S. W., et al. (2009). Incompletely penetrant PKD1 alleles suggest a role for gene dosage in cyst initiation in polycystic kidney disease. *Kidney Int.* 75, 848–855. doi: 10.1038/ki.2008.686

Schönauer, R., Baatz, S., Nemitz-Kliemchen, M., Frank, V., Petzold, F., Sewerin, S., et al. (2020). Matching clinical and genetic diagnoses in autosomal dominant polycystic kidney disease reveals novel phenocopies and potential candidate genes. *Genet. Med.* 22, 1374–1383. doi: 10.1038/s41436-020-0816-3

Shan, D., Rezonzew, G., Mullen, S., Roye, R., Zhou, J., Chumley, P., et al. (2019). Heterozygous Pkhd1C642* mice develop cystic liver disease and proximal tubule ectasia that mimics radiographic signs of medullary sponge kidney. *Am. J. Physiol. Renal Physiol.* 316, F463–F472. doi: 10.1152/ajprenal.00181.2018

Stenson, P. D., Mort, M., Ball, E. V., Chapman, M., Evans, K., Azevedo, L., et al. (2020). The human gene mutation database (HGMD®): optimizing its use in a clinical diagnostic or research setting. *Hum. Genet.* 139, 1197–1207. doi: 10.1007/s00439-020-02199-3

Su, Q., Hu, F., Ge, X., Lei, J., Yu, S., Wang, T., et al. (2018). Structure of the human PKD1-PKD2 complex. *Science* 361:eaat9819. doi: 10.1126/science.aat9819

Szabó, T., Orosz, P., Balogh, E., Jávorszky, E., Máttyus, I., Bereczki, C., et al. (2018). Comprehensive genetic testing in children with a clinical diagnosis of ARPKD identifies phenocopies. *Pediatr. Nephrol.* 33, 1713–1721. doi: 10.1007/s00467-018-3992-5

Vujic, M., Heyer, C. M., Ars, E., Hopp, K., Markoff, A., Örndal, C., et al. (2010). Incompletely penetrant PKD1 alleles mimic the renal manifestations of ARPKD. *J. Am. Soc. Nephrol.* 21, 1097–1102. doi: 10.1681/ASN.2009101070

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: 06 July 2021
doi: 10.3389/fgene.2021.697085

Spectrum of Mutations in Pediatric Non-glomerular Chronic Kidney Disease Stages 2–5

Xiaoyuan Wang, Huijie Xiao, Yong Yao, Ke Xu, Xiaoyu Liu, Baige Su, Hongwen Zhang, Na Guan, Xuhui Zhong, Yanqin Zhang, Jie Ding and Fang Wang**

Department of Pediatrics, Peking University First Hospital, Beijing, China

Renal hypodysplasia and cystic kidney diseases, the common non-glomerular causes of pediatric chronic kidney disease (CKD), are usually diagnosed by their clinical and imaging characteristics. The high degree of phenotypic heterogeneity, in both conditions, makes the correct final diagnosis dependent on genetic testing. It is not clear, however, whether the frequencies of damaged alleles vary among different ethnicities in children with non-glomerular CKD, and this will influence the strategy used for genetic testing. In this study, 69 unrelated children (40 boys, 29 girls) of predominantly Han Chinese ethnicity with stage 2–5 non-glomerular CKD caused by suspected renal hypodysplasia or cystic kidney diseases were enrolled and assessed by molecular analysis using proband-only targeted exome sequencing and array-comparative genomic hybridization. Targeted exome sequencing discovered genetic etiologies in 33 patients (47.8%) covering 10 distinct genetic disorders. The clinical diagnoses in 13/48 patients (27.1%) with suspected renal hypodysplasia were confirmed, and two patients were reclassified carrying mutations in nephronophthisis (*NPHP*) genes. The clinical diagnoses in 16/20 patients (80%) with suspected cystic kidney diseases were confirmed, and one patient was reclassified as carrying a deletion in the hepatocyte nuclear factor-1-beta gene (*HNF1B*). The diagnosis of one patient with unknown non-glomerular disease was elucidated. No copy number variations were identified in the 20 patients with negative targeted exome sequencing results. *NPHP* genes were the most common disease-causing genes in the patients with disease onsets above 6 years of age (14/45, 31.1%). The children with stage 2 and 3 CKD at onset were found to carry causative mutations in paired box gene 2 (*PAX2*) and *HNF1B* gene (11/24, 45.8%), whereas those with stage 4 and 5 CKD mostly carried causative mutations in *NPHP* genes (19/45, 42.2%). The causative genes were not suspected by the kidney imaging patterns at disease onset. Thus, our data show that in Chinese children with non-glomerular renal dysfunction caused by renal hypodysplasia and cystic kidney diseases, the common causative genes vary with age and CKD stage at disease onset. These findings have the potential to improve management and genetic counseling of these diseases in clinical practice.

Keywords: non-glomerular, chronic kidney disease, renal hypodysplasia, cystic kidney diseases, targeted exome sequencing, genetic diagnosis

OPEN ACCESS

Edited by:
Xiu-An Yang,
Chengde Medical College, China

Reviewed by:
Nancy Monroy-Jaramillo,
National Institute of Neurology and Neurosurgery, Mexico
Desheng Liang,
Central South University, China

****Correspondence:***
Fang Wang
wangfangped@163.com
Jie Ding
djnc_5855@126.com

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

Received: *18 April 2021*
Accepted: *07 June 2021*
Published: *06 July 2021*

Citation:
Wang X, Xiao H, Yao Y, Xu K, Liu X, Su B, Zhang H, Guan N, Zhong X, Zhang Y, Ding J and Wang F (2021) Spectrum of Mutations in Pediatric Non-glomerular Chronic Kidney Disease Stages 2–5. Front. Genet. 12:697085. doi: 10.3389/fgene.2021.697085

INTRODUCTION

The presence of structural or functional abnormalities in the kidney over a 3-month period is defined as chronic kidney disease (CKD), and is classified into five stages based on the glomerular filtration rate (Andrassy, 2013). End-stage renal disease (ESRD), which is the most serious CKD stage, requires the use of renal replacement therapy. Pediatric CKDs are less common than in adults, but affected children are at increased risk of early mortality and disabling physical comorbidities, which highlights the need for appropriate management of the affected children.

Congenital anomalies of the kidney and urinary tract (CAKUT) are the most common non-glomerular presentations of pediatric ESRD followed by cystic kidney diseases (Smith et al., 2007; Wuhl et al., 2014). In fact, the most prevalent malformation is reported to be renal hypodysplasia, which includes renal aplasia, hypoplasia, and dysplasia (Smith et al., 2007). Renal ultrasound provides essential diagnostic information about renal hypodysplasia and cystic kidney diseases (Sanna-Cherchi et al., 2007; Vester et al., 2010; Gimpel et al., 2019). For example, hypodysplastic kidney is defined by renal ultrasonography findings as a reduced renal size of greater than two standard deviations from the mean size in terms of age and loss of corticomedullary differentiation, and the sonographic signs of parenchymal hyperechogenicity and renal enlargement in a child are highly suggestive of polycystic kidney disease. However, a clinically definitive diagnosis of hypodysplastic kidney disease or cystic kidney disease remains challenging to arrive at because the sonographic appearance of these two conditions is observed in a variety of renal diseases.

With advances in genomic DNA sequencing technologies, the genetic mechanisms leading to renal hypodysplasia and cystic kidney diseases have been more readily assessed (Vivante and Hildebrandt, 2016; Sanna-Cherchi et al., 2018; Armstrong and Thomas, 2019; Devlin and Sayer, 2019; Nigam et al., 2019), and this improves the diagnostic accuracy of genetic testing. Our previous study reported that the genetic test results for pediatric steroid-resistant nephrotic syndrome vary by ethnicity (Wang et al., 2017b). It is not clear, however, whether a similar phenomenon exists with the pediatric chronic renal dysfunction caused by renal hypodysplasia and cystic kidney diseases, which provides the impetus for the reasonable selection of genetic testing approaches. To address this question, index-only targeted exome sequencing and array-comparative genomic hybridization (CGH) were performed in a cohort of 69 unrelated children with non-glomerular stage 2–5 CKD who were clinically suspected of having renal hypodysplasia or cystic kidney diseases.

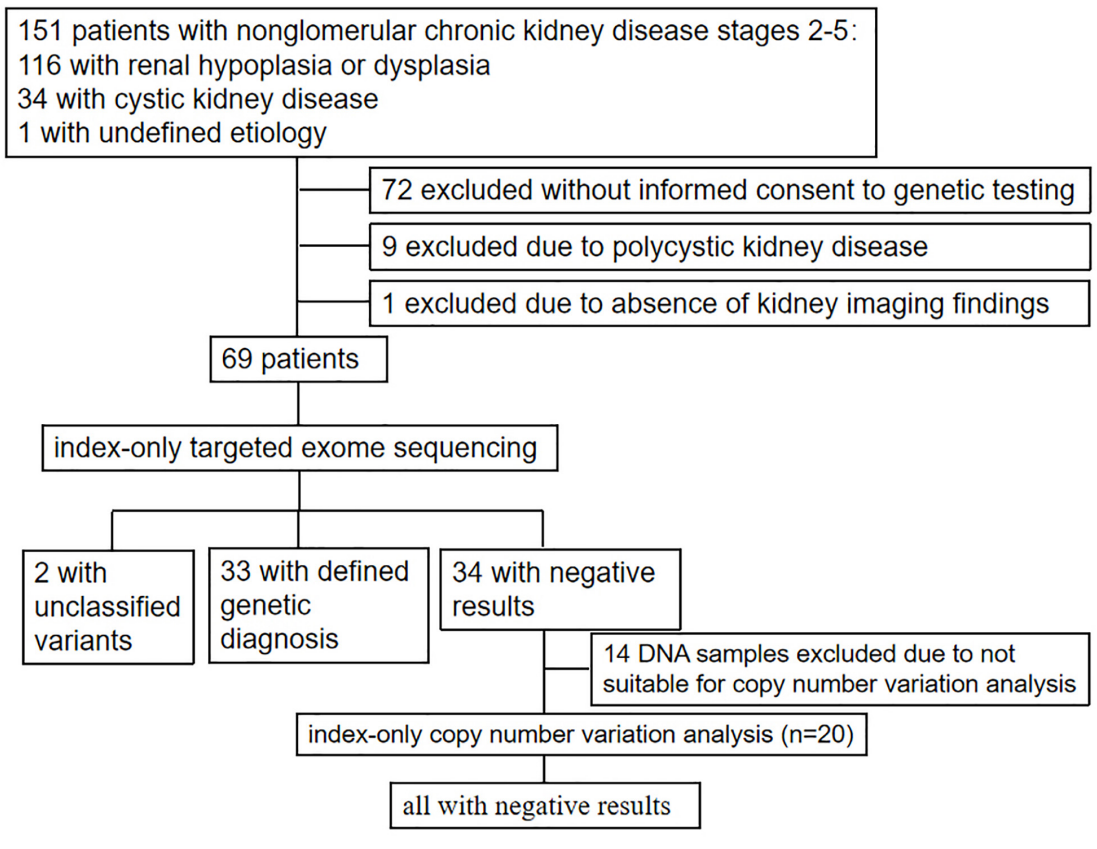


FIGURE 1 | Flow diagram for patient selection and genetic testing.

MATERIALS AND METHODS

Patients

Patients were enrolled in the study between January 2011 and September 2018 by a group of trained pediatric nephrologists from the Department of Pediatrics, Peking University First Hospital based on fulfillment of the following criteria: (i) the presence of stage 2–5 CKD below the age of 18 years; (ii) clinical diagnosis or suspicion of renal hypoplasia/dysplasia, cystic kidney diseases, or unknown non-glomerular diseases. Patients with polycystic kidney disease, incomplete clinical data (especially the absence of kidney imaging results), and an unwillingness to participate in the study were excluded. The study was approved by the Ethics Committee of Peking University First Hospital and was performed in accordance with the Declaration of Helsinki.

Comprehensive clinical data [including age of onset, age of renal failure, urinalysis, examination of urinary protein, renal imaging, estimated glomerular filtration rate using 24-h endogenous creatinine clearance or the Schwartz formula (Schwartz et al., 1987), extrarenal manifestations, renal biopsy, information from the last follow-up, and family history] and demographics were extracted from the Chinese Registry Database of Hereditary Kidney Diseases and then analyzed. Sonographic measurements of the longitudinal sections of both kidneys in each patient were compared with those of age-matched controls (Loftus et al., 1998).

After receiving informed consent from the patients or their parents/legal guardians, blood samples and comprehensive clinical data were collected and analyzed.

Genetic Examination

Genomic DNA was extracted from peripheral white blood cells using the QIAamp DNA Blood Mini Kit (A1120, Qiagen, Germany). DNA quantity and quality were determined by NanoDrop (Thermo Fisher Scientific, United States). When available, DNA samples from the participants' relatives were obtained.

Because targeted exome sequencing is a cost-effective diagnostic strategy for identifying the genetic causes of kidney disorders (Vivante and Hildebrandt, 2016; Groopman et al., 2018), we used it to simultaneously examine 30 genes that are known to be associated with renal hypodysplasia and 118 genes associated with cystic kidney diseases (**Supplementary Table 1**). These genes were selected from the relevant literature (Devuyst et al., 2014; Mann et al., 2019; Nigam et al., 2019). DNA library preparation, capture, enrichment, next-generation sequencing, and data analysis were performed at BGI-Shenzhen, China, as described previously (Wang et al., 2017b). Variants with minor allele frequencies <0.01 were selected based on the control database such as NCBI dbSNP (snp137), 1000 Genomes Project (phase I), Exome sequencing project (ESP6500), Exome Aggregation Consortium (ExAC), Genome Aggregation Database (gnomAD), and the BGI in-house database. The Human Gene Mutation Database (HGMD) and ClinVar were used to detect previously reported pathogenic variants. The prioritized variants were classified according to the American College of Medical Genetics and Genomics (ACMG) guidelines (Richards et al., 2015).

To detect copy number variations (CNVs), array CGH was performed using the Agilent SurePrint G3 Human 8 × 60 K CGH Microarray (Agilent Technologies, Technologies, Santa Clara, CA, United States). DNA labeling, array hybridization, scanning, and data analysis were conducted at the Department of Central Laboratory, Peking University First Hospital, Beijing, China, as described previously (Yi et al., 2016). Public CNV databases including DGV, NCBI, DECIPHER, ClinGen, OMIM, and ISCA were used to detect known CNVs. The prioritized CNVs were classified according to the ACMG guidelines (Riggs et al., 2020).

Validation of all candidate pathogenic or likely pathogenic variants was performed using Sanger sequencing or quantitative PCR (qPCR) on the genomic DNAs of the probands. Hepatocyte nuclear factor-1-beta gene (*HNF1B*) and the nephronophthisis type 1 (*NPHP1*) gene were used to normalize the gene dosage in

TABLE 1 | Clinical features of the 69 patients from the present study.

Parameter	Patients with molecular diagnosis* (*n* = 35)	Patients without molecular diagnosis (*n* = 34)
Gender (M, F)	15, 20	25, 9
Age of onset, years	9.4 (0–16.7)	8.1 (0–15.0)
Follow-up time, months	27 (7–120)	17 (3–120)
Age of genetic test, years	10.0 (1.0–16.7)	8.8 (0.2–16.0)
Clinical diagnosis		
Renal hypoplasia/dysplasia	25	23
Cystic kidney disease	9	11
Unknown non-glomerular disease	1	0
CKD stage at disease onset		
(CKD stage 2, 3, 4, 5, *n*)	1, 0, 1, 1	0, 1, 1, 2
<1 year	0, 2, 2, 1	0, 1, 2, 1
1−3 years	1, 2, 0, 2	0, 1, 1, 2
3−6 years	1, 2, 5, 7	2, 6, 1, 6
6−12 years	1, 1, 3, 2	1, 1, 0, 5
12−18 years		
Extrarenal manifestations	14	9
Renal histopathologic diagnosis		
Chronic tubulointerstitial nephropathy with or without glomerular lesions	10	2
Oligomeganephronia with atypical membranous nephropathy	1	0
Oligomeganephronia?	0	1
Focal proliferative sclerosing purpura nephritis with glomerular hypertrophy	1	0
Focal segmental glomerular sclerosis	1	1
Family history	5	5

**Including two cases carrying putatively pathogenic NPHP3 variants that required further functional verification.*

TABLE 2 | Pathogenic or likely pathogenic variants detected by targeted exome sequencing*.

Patient ID	Gender	Age at onset	Clinical diagnosis	Renal ultrasound findings	Renal biopsy (at age)	Extrarenal manifestations	Follow-up (at age)	Gene	Nucleotide alteration	Genomic position and SNP	Amino acid changes	Location (zygosity, segregation)	ACMG classify sequence variants	ACMG interpretation	Ref.
19	Female	11M	Unknown non-glomerular diseases, CKD4	Normal size kidneys without cyst	ND	NO	Loss to follow-up	*ACE*	c.793C >T	g.61557835C >T (rs138873311)	p.Arg265*	EX5 (het, mother)	PVS1 PM2 PP3 PP4	Pathogenic	Gribouval et al., 2012
								ACE	c.1028G >A	g.61559009G >A (rs11466112)	p.Trp343*	EX7 (het, paternal)	PVS1 PM2 PP3 PP4	Pathogenic	Gribouval et al., 2012
36	Male	3Y	Renal hypoplasia/ dysplasia, CKD3	Small size kidneys without cyst	ND	NO	Loss to follow-up	*HNF1B*	EX1-9del	–	–	The whole gene (het, *de novo*)	PVS1 PS2 PM2 PP4 PP3	Pathogenic	Weber et al., 2006
4	Female	11Y	Cystic kidney diseases, CKD4	Normal size kidneys with a cyst	ND	NO	Transplant (15Y)	*NPHP1*	EX1-20 del	–	–	The whole gene (hom, ?)	PVS1 PM2 PP3 PP4	Pathogenic	Kang et al., 2016
11	Female	2Y	Renal hypoplasia/ dysplasia, CKD4	Small size kidneys without cyst	Chronic tubulointerstitial nephropathy	NO	PD (17Y)	*NPHP1*	EX1-20 del	–	–	The whole gene (hom, ?)	PVS1 PM2 PP3 PP4	Pathogenic	Kang et al., 2016
14	Female	6Y4M	Renal hypoplasia/ dysplasia, CKD5	Small size kidneys without cyst	ND	NO	CKD5 (10Y)	*NPHP1*	EX1-20 del	–	–	The whole gene (hom, maternal)	PVS1 PM2 PP3 PP4	Pathogenic	Kang et al., 2016
20	Female	11Y6M	Renal hypoplasia/ dysplasia, CKD5	Small size kidneys with a cyst	ND	NO	Transplant (12Y7M)	*NPHP1*	EX1-20 del	–	–	The whole gene (hom, paternal, maternal)	PVS1 PM2 PP3 PP4	Pathogenic	Kang et al., 2016
27	Male	13Y6M	Renal hypoplasia/ dysplasia, CKD4	Small size kidneys without cyst	ND	Short stature	Loss to follow-up	*NPHP1*	EX1-20 del	–	–	The whole gene (hom, maternal)	PVS1 PM2 PP3 PP4	Pathogenic	Kang et al., 2016
33	Male	9Y	Cystic kidney diseases, CKD5	Normal size kidneys with cysts	ND	Astigmatism, strabismus	CKD5 (10Y)	*NPHP1*	EX1-20 del	–	–	The whole gene (hom, ?)	PVS1 PM2 PP3 PP4	Pathogenic	Kang et al., 2016
37	Male	13Y2M	Cystic kidney diseases, CKD5	Normal size kidneys with a cyst	Chronic tubulointerstitial nephropathy	NO	PD (13Y9M)	*NPHP1*	c.1122+4 delA	g.110919176delT	–	IVS10 (het, maternal)	PVS1 PM2 PP3 PP4	Pathogenic	This report
								NPHP1	EX1-20 del	–	–	The whole gene (het, paternal)	PVS1 PM2 PP3 PP4	Pathogenic	Kang et al., 2016
52	Female	16Y8M	Cystic kidney diseases, CKD4	Normal size kidneys without cyst	ND	NO	Loss to follow-up	*NPHP1*	EX1-20 del	–	–	The whole gene (hom, paternal, maternal)	PVS1 PM2 PP3 PP4	Pathogenic	Kang et al., 2016
9	Female	3Y9M	Cystic kidney diseases, CKD5	Normal size kidneys with cysts	ND	NO	Transplant (6Y)	*INVS (NPHP2)*	c.2782C >T	g.103055321C > T(rs376879175)	p.Arg928*	EX14 (het, paternal)	PVS1 PM2 PP3 PP4	Pathogenic	Halbritter et al., 2013
								INVS (NPHP2)	c.2666_2667 delTG	g.103055205_10 3055206delTG	p.Val889Glufs*3	EX14 (het, maternal)	PVS1 PM2 PP3 PP4	Pathogenic	This report
26	Male	11M	Renal hypoplasia/ dysplasia, CKD5	Small size kidneys without cyst	Chronic tubulointerstitial nephropathy with glomerular lesions	NO	Died (2Y)	*INVS (NPHP2)*	c.2701C >T (het)	g.1030552 40C >T	p.Gln901*	EX14 (het, maternal)	PVS1 PM2 PP3 PP4	Pathogenic	This report
								INVS (NPHP2)	c.2786+2T >C	g.103055327T > C(rs1322951938)	-	IVS14 (het, paternal)	PVS1 PM2 PP3 PP4	Pathogenic	Otto et al., 2003
60	Male	6Y	Renal hypoplasia/ dysplasia, CKD5	Small size kidneys without cyst	ND	NO	CKD5 (7Y)	*NPHP3*	c.909C >A	g.132433977G >T	p.Tyr303*	EX5 (het, paternal)	PVS1 PM2 PP3 PP4	Pathogenic	This report
								NPHP3	c.3202-2A >G	g.1324052 33T >C		IVS22 (het, maternal)	PVS1 PM2 PP3 PP4	Pathogenic	This report

(Continued)

TABLE 2 | Continued

Patient ID	Gender	Age at onset	Clinical diagnosis	Renal ultrasound findings	Renal biopsy (at age)	Extrarenal manifestations	Follow-up (at age)	Gene	Nucleotide alteration	Genomic position and SNP	Amino acid changes	Location (zygosity, segregation)	ACMG classify sequence variants	ACMG interpretation	Ref.
2	Male	9Y9M	Renal hypoplasia/ dysplasia, CKD5	Small size kidneys with cysts	ND	Cryptorchidism	Died (12Y)	*NPHP3*	c.1082C > G (het)	g.132432006G >C (rs146250226)	p.Ser361Cys	EX6 (het, maternal)	PM1 PM2 PP3 PP4	Likely pathogenic	ClinVar
								NPHP3	c.1986-2A >G	g.132416208T >C	–	IVS13 (het, paternal)	PVS1 PM2 PP3 PP4	Pathogenic	This report
61	Female 11Y cystic kidney diseases, CKD4 normal size kidneys with cysts MPGN, tubulointerstitial histopathology growth retardation CKD5(13y)	11Y	Renal hypoplasia/ dysplasia, CKD	Normal size kidneys with cysts	MPGN, tubulointerstitial histopathology	Growth retardation	CKD5 (13Y)	*NPHP4*	c.992+1G >A	g.6008129C >T	–	IVS8 (het, paternal)	PVS1 PM2 PP3 PP4	Pathogenic	This report
								NPHP4	c.2260G >A	g.5950972C >T (rs373962831)	p.Gly754Arg	EX17 (het, maternal)	PM1 PM2 PP3 PP4	Likely pathogenic	Otto et al., 2002
1	Female	14Y	Renal hypoplasia/ dysplasia, CKD5	Small size kidneys with a cyst	Chronic tubulointerstitial nephropathy	Nystagmus, hypermetropia, astigmatism	Loss to follow-up	*IQCB1 (NPHP5)*	c.1090C >T	g.121508959G >A (rs727503968)	p.Arg364*	EX11 (het, paternal)	PVS1 PM2 PP3 PP4	Pathogenic	Khanna et al., 2009
								IQCB1 (NPHP5)	c.1333C >T	g.121500667G >A	p.Arg445*	EX13 (het, maternal)	PVS1 PM2 PP3 PP4	Pathogenic	Halbritter et al., 2012
63	Female	8Y6M	Renal hypoplasia/ dysplasia, CKD5	Small size kidneys without cyst		NO	Loss to follow-up	*WDR19*	c.641T >A	g.39206811T >A (rs751290509)	p.Leu214*	EX8 (het, maternal)	PVS1 PM2 PP3 PP4	Pathogenic	VKGL-NL_AMC[$]
								WDR19	c.904G >T	g.39216234G >T	p.Asp302Tyr	EX10 (het, *de novo*)	PM1 PM2 PM3 PP3 PP4	Pathogenic	This report
16	Female	6Y9M	Cystic kidney diseases, CKD4	Normal size kidneys with cysts	ND	NO	CKD5 (10Y9M)	*UMOD*	c.178G >T	g.20360445C >A	p.Gly60Cys	EX3 (het, *de novo*)	PS2 PM1 PM2 PP3	Likely pathogenic	This report

**We assessed the Human Gene Mutation Database and the Leiden Open Variation Database (LOVD) to check novel variants in April 2021. Nucleotide positions are numbered in accordance with the reference sequences (NM_000789.3 for ACE, NM_000458.4 for HNF1B, NM_003990.3 for PAX2, NM_000272.3 for NPHP1, NM_014425.3 for INVS, NM_153240.4 for NPHP3, NM_015102.3 for NPHP4, NM_001023570.2 for IQCB1, NM_025132.3 for WDR19, and NM_003361.2 for UMOD) using the first coding ATG of exon 1 as the initiation codon.*

Patients 8, 10, 13, 18, 21, 24, 30, 32, 34, and 38 have already been reported (Deng et al., 2019). Patients 6, 28, and 29 have also been reported (Wang et al., 2017a). Patients 23 and 31 have also been reported (Deng et al., 2020).

[$]LOVD.

ACMG, the American College of Medical Genetics and Genomics; CKD, chronic kidney disease; ND, not done; PD, peritoneal dialysis; IVS, intron; EX, exon; Hom, homozygous; Het, heterozygous; M, months; Y, years; PVS, pathogenic very strong; PS, pathogenic strong; PM, pathogenic moderate; PP, pathogenic supporting.

qPCR, and they were analyzed in triplicate. Segregation analyses were performed for all the available first-degree relatives.

RESULTS

Clinical Features

As shown in **Figure 1** and **Table 1**, 69 unrelated patients (40 boys, 29 girls) were enrolled in this study. They were from 18 provinces, municipalities, and autonomous Chinese regions and were predominantly the Han Chinese ethnicity (61/69). Renal dysfunction was detected in most of these patients (36/69), either accidentally or for other reasons at disease onset, whereas complaints of fatigue or a sallow complexion were observed in 19 patients, edema in eight patients, short stature in four patients, polydipsia and polyuria in one patient, and enuresis in another patient. Their median age at disease onset was 8.5 years (range, 0 day–16.7 years). Renal hypoplasia/dysplasia, cystic kidney disease, and unknown non-glomerular disease were diagnosed or suspected in 48, 20, and 1 patient, respectively. There were seven patients with stage 2 CKD, 17 patients with stage 3 CKD, 16 patients with stage 4 CKD, and 29 patients with stage 5 CKD. Of the 17 patients undergoing renal biopsy, chronic tubulointerstitial nephropathy was the most common histopathological diagnosis. The patients' extrarenal manifestations included short stature, ocular abnormalities (including ametropia, strabismus, microphthalmia, retinopathy, vitreous opacity, and nystagmus), auricle malformation, preauricular fistula, spina bifida, cryptorchidism, skeletal deformities (including polydactylism, tetradactylism, straw sandal-like feet, strephenopodia, and fourth metatarsal microsomia), elevated liver enzymes, ovarian teratoma, microcephaly, ventricular septal defect, and patent arterial duct in 23 patients. Parental consanguinity was reported in only one patient, whereas eight patients had positive family histories of ESRD, one patient had a positive family history of proteinuria, and one patient had a positive family history of renal cystic disease. Eight patients had received a renal transplant (median age, 12.9 years; range, 6–18 years), and no disease recurrence in their allografts was documented. Six patients died at a median age of 5.9 (2–12) years.

Genetic Study

Twenty-two pathogenic variants and six likely pathogenic variants in 10/148 targeted genes, including nine non-sense, seven missense, six splice sites, three small deletions, two whole gene deletion, and one small insertion, were detected in 33/69 patients (47.8%), and these variants encompassed 10 distinct genetic disorders (**Table 2** and **Figure 2**). Of these variants, the 14 (50.0%) novel ones included three variants that we reported on previously (Deng et al., 2019), whereas the remaining 14 variants were previously reported. Of the 18 patients harboring diagnostic variants in recessive genes, compound heterozygous variants were found in 10 patients and homozygous variants were found in eight patients. Of the 48 patients with suspected renal hypodysplasia, the targeted exome sequencing confirmed the clinical diagnoses of 13 patients (27.1%), and reclassified the clinical diagnoses of two patients carrying mutations in nephronophthisis (NPHP) genes (*INVS* and *WDR19*). Of the 20 patients with suspected cystic kidney diseases, the clinical diagnoses for 16 patients (80%) were confirmed, and that of one patient with a deletion in *HNF1B* was reclassified. The diagnosis of the remaining patient (patient 19) with renal dysfunction (serum creatinine, 122 μmol/L; evaluated glomerular filtration rate, 22.9 ml/min/1.73 m^2), moderate anemia (70 g/L), short stature (height, 71 cm), and normal-sized, non-cystic kidneys combined with parenchymal hyperechogenicity and poor corticomedullary differentiation on renal ultrasonography (11 months of age) was classified by the compound heterozygous non-sense mutations present in the gene encoding the angiotensin-converting enzyme (*ACE*).

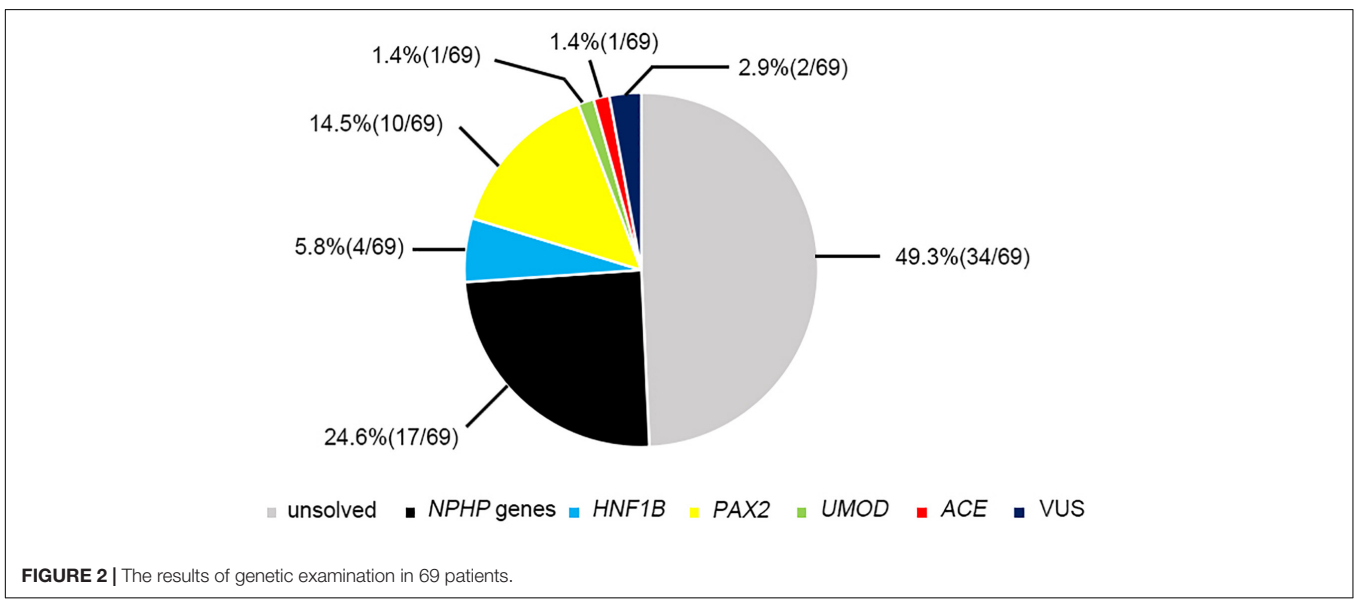


FIGURE 2 | The results of genetic examination in 69 patients.

Patients 22 and 69 were strongly suspected of having NPHP based on the combination of ESRD before the age of 7 years without proteinuria or hematuria, normal-sized kidneys with hyperechogenicity and the absence of corticomedullary differentiation, and chronic tubulointerstitial nephropathy (in patient 22) or elevated liver enzymes of unknown cause (in patient 69), whereas targeted exome sequencing revealed three rare and predicted deleterious variants in *NPHP3* that were classified as having unknown significance using ACMG criteria (**Table 3**). We assumed that these variants are pathogenic, although functional analyses on them are required.

No CNVs were identified in the 20 patients with negative targeted exome sequencing results.

The likelihood of establishing an accurate molecular diagnosis of non-glomerular CKD did not improve with increasing age and remained roughly the same (at about 50%) (**Figure 3A**). We detected diagnostic *PAX2* and *NPHP* gene variants in all four age groups, and the *NPHP* genes were the most common disease-causing ones in the patients whose disease onset was above 6 years of age (14/45, 31.1%).

The molecular diagnostic performance and common mutated genes differed in line with the increased CKD stage at disease onset (**Figure 3B**). The genetic diagnostic yield was highest in the patients with stage 4 CKD at onset (11/16, 68.7%). The children whose CKD onset was stage 2 and 3 carried mutations in *PAX2* and *HNF1B* genes (11/24, 45.8%), whereas those whose CKD onset was stage 4 and 5 mostly carried mutations in *NPHP* genes (19/45, 42.2%).

Because renal ultrasonography is used in the first instance to diagnose renal hypodysplasia and cystic nephropathies, we analyzed the renal imaging patterns at disease onset and the mutation detection rates in the patients (**Figure 3C**). An etiological diagnosis was found in 52.1% of the children with small kidneys (25/48), 47.6% with normal-sized kidneys (10/21), 48.9% without cysts (23/47), 54.5% with a single cyst or multiple cysts (12/22), 66.7% without corticomedullary differentiation (12/18), and 45.1% with distinct corticomedullary differentiation (23/51). Renal parenchymal hyperechogenicity was observed in all 69 patients.

The most prevalent genetic diagnosis in our study was NPHP. Chaki et al. (2011) reported that 100% of 440 patients with NPHP-related ciliopathies carried biallelic pathogenic variants in *NPHP* genes. We therefore analyzed the genetic test results from the patients who met at least one of the four criteria for NPHP used by Chaki et al. (2011). Hence, we were able to clinically diagnose 52 patients as having NPHP and found that 26 patients (50%) had pathogenic or likely pathogenic variants in the causative genes. Of these patients, 17 had *NPHP*, whereas we identified mutations in *PAX2*, *HNF1B*, *ACE*, and *UMOD* genes in the remaining patients (**Figure 4**).

DISCUSSION

In the present study, we used targeted exome sequencing and array CGH to depict the genetic features of 69 unrelated children with non-glomerular stage 2–5 CKD caused by suspected renal

TABLE 3 | Variants of unknown significance that were detected by targeted exome sequencing.

Patient ID	Gender	Age at onset	Clinical diagnosis	Renal ultrasound findings	Renal biopsy (at age)	Extrarenal manifestations	Follow-up (at age)	Gene	Nucleotide alteration	Genomic position and SNP	Deduced amino acid changes	Location (zygosity, segregation)	ACMG classify sequence variants	ACMG interpretation	Ref.
22	Female	2Y2M	Cystic kidney diseases, CKD5	Normal size kidneys without cysts	Chronic tubulointerstitial nephropathy	Short stature	CKD5 (6Y6M)	*NPHP3*	c.3813-3A >G	g.132400937T >C	–	IVS26 (hom, mother)	PM2 PP3 PP4	Unknown significance	This report
69	Male	3Y	Cystic kidney diseases, CKD5	Normal size kidneys with cysts	FSGS	Abnormal liver function	PD (3Y8M)	*NPHP3*	c.3813-3A >G	g.132400937T >C	–	IVS26 (het, father)	PM2 PP3 PP4	Unknown significance	This report
								NPHP3	c.1135T >C	g.132427085A >G	p.Cys379Arg	EX7 (het, mother)	PM2 PP3 PP4	Unknown significance	This report

Accession no.: NPHP3, NM_153240.4. ACMG, the American College of Medical Genetics and Genomics; CKD, chronic kidney disease; FSGS, focal segmental glomerulosclerosis; PD, peritoneal dialysis; IVS, intron; EX, exon; Hom, homozygous; Het, heterozygous; PM, pathogenic moderate; PP, pathogenic supporting.

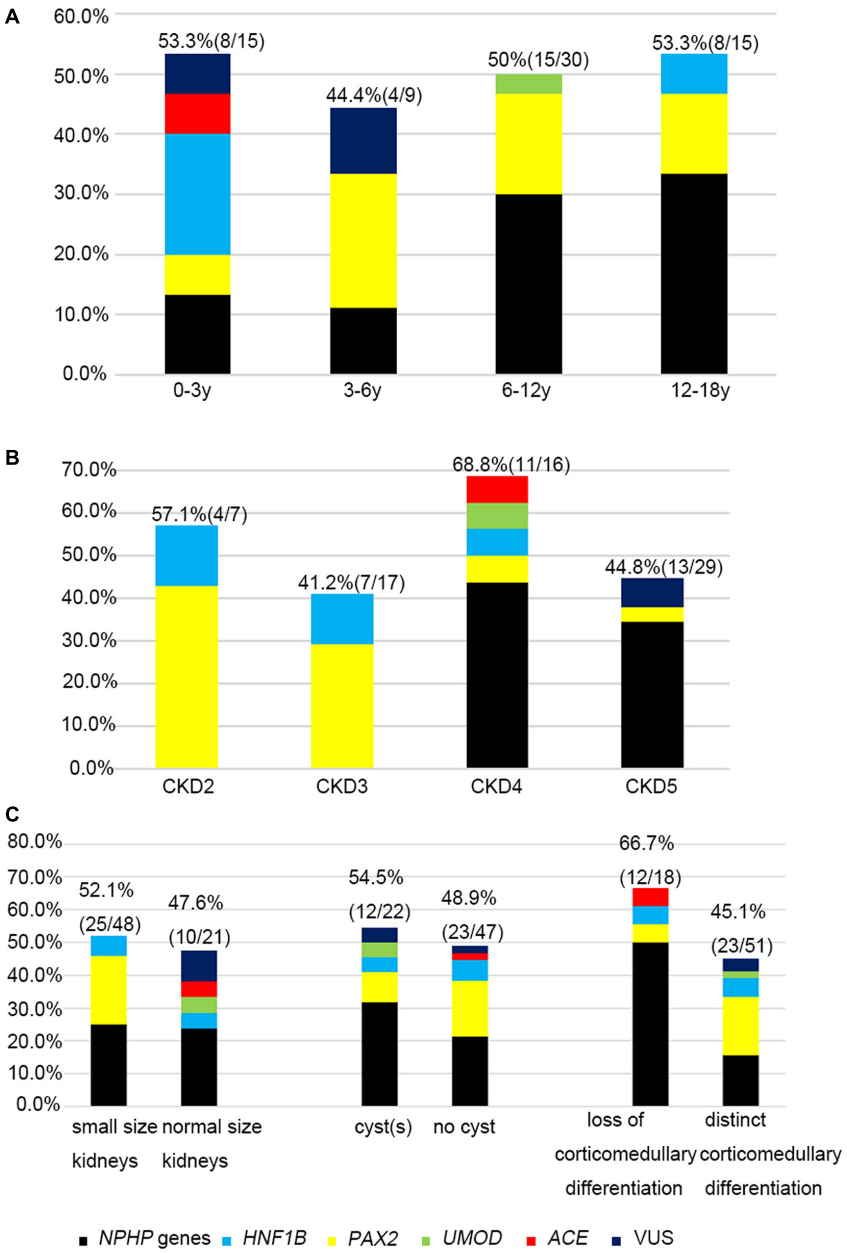


FIGURE 3 | The relationship between causative genes and clinical features. Histograms indicate fractions (in percentage) of patients with disease gene detected per group. **(A)** Genetic diagnosis and the age at onset for non-glomerular CKD. **(B)** Genetic diagnosis and CKD stages. **(C)** Genetic diagnosis and renal ultrasound findings.

hypodysplasia or cystic kidney diseases. First, 27.1% of our patients with suspected renal hypodysplasia obtained a molecular diagnosis, and *PAX2* was the most common mutated gene (found in ten patients). In contrast, in one cohort of 159 Chinese CAKUT children (Rao et al., 2019), only four carried *PAX2* mutations. Ishiwa et al. (2019) performed genetic analysis on 66 Japanese patients with CAKUT (the number of patients with renal hypodysplasia is not available) and identified the etiologies in 14 individuals with renal hypodysplasia (21.2%). Ahn et al. (2020) identified the causative genes responsible for renal hypodysplasia in 12/76 Korean children (15.8%), and predominance of *HNF1B* mutations was seen in these patients. The mutation detection rates range from 7 to 17% in patients from Europe and the United States with renal hypodysplasia combined with or without renal failure (Weber et al., 2006; Thomas et al., 2011; Hwang et al., 2014; Nicolaou et al., 2016). The *SALL1* (spalt-like transcription factor 1) gene was detected more frequently in one cohort of patients (Hwang et al., 2014), whereas *PAX2* and *HNF1B* were detected in another cohort (Weber et al., 2006; Thomas et al., 2011; Nicolaou et al., 2016). One possible

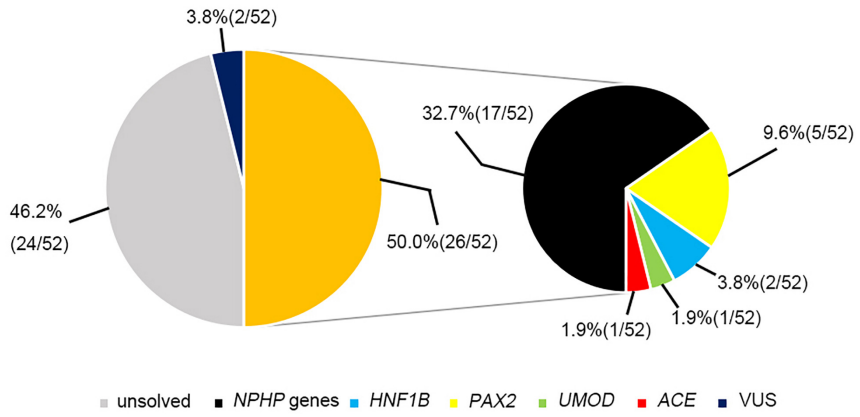


FIGURE 4 | The genetic examination results for 52 patients who were highly suspected of having nephronophthisis.

explanation for this discrepancy relates to the criteria used for selecting patients: our cohort contained patients with bilateral renal lesions, whereas other studies have contained patients with bilateral and unilateral renal hypodysplasia. Another possible explanation is the high genetic heterogeneity in this condition. Second, causative genes were identified in 85% of our patients with suspected cystic kidney diseases, and 10 patients carried *NPHP1* mutations, making it the most prevalent mutated gene. However, in other patient cohorts, about 70% of the children with cystic kidney diseases had monogenic disease, and the most frequent molecular diagnosis was autosomal recessive polycystic kidney disease or polycystic kidney disease (Bullich et al., 2018; Rao et al., 2019; Obeidova et al., 2020). Excluding polycystic kidney disease from our study may in part explain this discrepancy. Finally, a genetic diagnosis was obtained in three children where phenotypic overlapping caused the initial disease to be clinically misdiagnosed, one case of which had undiagnosed stage 4 CKD, which stresses the importance of genetic testing as one of the diagnostic workups in the pediatric CKD population.

It is worth noting the difference we observed for the common causative genes in relation to the age and CKD stage at disease onset. *NPHP* genes were the most frequently mutated genes in the patients whose onset exceeded 6 years of age with stage 4–5 CKD, whereas mutations in *HNF1B* and *PAX2* together were more prevalent in patients whose onset was less than 6 years of age and had become stage 2–3 CKD. These findings suggest that performing genetic testing in accordance with the age and CKD stage at disease onset may be an efficient strategy for the molecular diagnosis of children with non-glomerular CKD. In contrast, Weber's study showed that *HNF1B* and *PAX2* mutations caused CKDs with an age of onset between 10 and 23 years (Weber et al., 2006), and autosomal recessive polycystic kidney disease was reported to be the most prevalent etiology in neonatal-onset cystic kidney diseases (Obeidova et al., 2020). The difference is likely impacted by the use of small populations of patients and the patients' ethnic origins. Because early stage renal hypodysplasia and cystic kidney diseases are often clinically silent, patients with renal insufficiency who are usually detected accidentally may be referred for clinical diagnosis. Renal ultrasonography is currently the diagnostic mainstay. The presence of small-sized kidneys in a child always leads clinicians to make a diagnosis of renal hypodysplasia, and the presence of renal cysts support the diagnosis of cystic kidney disease. However, the phenotypic and genetic variability of these two conditions makes establishing the final clinical diagnosis challenging. As our study has shown, small- to normal-sized kidneys with or without cyst formation or changes in corticomedullary differentiation can be caused by mutant *NPHP* and *HNF1B* genes (Chaki et al., 2011; Avni et al., 2015), whereas mutations in *PAX2* lead to small-sized kidneys that often show distinct corticomedullary differentiation and no cysts (Bower et al., 2012).

NPHP is one of the most common inherited diseases leading to pediatric ESRD, and the phenotypes and genotypes in Chinese children with *NPHP* have been described (Tang et al., 2020; Yue et al., 2020). However, our finding shows that non-*NPHP* genes can also cause NPHP-like phenotypes, which emphasizes the difficulty in diagnosing NPHP in clinical settings. Similar phenomena have been reported elsewhere (Bullich et al., 2018; Mann et al., 2019).

Pathogenic CNVs, recognizably important etiological factors underlying renal hypodysplasia (Sanna-Cherchi et al., 2012; Verbitsky et al., 2019), are recommended to be detected by array-based technologies. However, consistent with prior reports on the use of targeted exome sequencing as a tool for identifying CNVs (Roberts et al., 2017; Ahn et al., 2020), our targeted exome sequencing and qPCR, we detected the deletion of the whole *HNF1B* and *NPHP1* genes in 4 and 10 patients, respectively, but no additional CNVs were identified using array CGH.

Obtaining a definite molecular diagnosis is very important for patients and their families and for facilitating genetic counseling. For example, *HNF1B* mutations are associated with diabetes mellitus (Clissold et al., 2015), and *NPHP1* genetic variants may cause multisystemic diseases and Joubert syndrome, among others (Soliman et al., 2012). Early discovery of related hidden symptoms and timely treatments are very important for patients.

We are conscious of some limitations in our study, which include the comparatively small patient cohort, the stringent

clinical criteria for selecting patients, the absence of whole-exome sequencing in patients lacking a genetic diagnosis, and the lack of functional verification of novel unknown significance variants. False-negative results from targeted exome sequencing were possible in the patients with no detectable variants. Nonetheless, our study was performed in one of the largest referral centers on mainland China, and the patients were from 18 out of 34 provincial administrative Chinese regions, indicating that our study is somewhat representative of Chinese children with renal dysfunction caused by renal hypodysplasia and cystic kidney diseases, and our findings provide the genotypic features seen in them. To the best of our knowledge, this is the first cohort study to provide evidence about the association between causative mutations and the stage of CKD onset.

In summary, the Chinese children with non-glomerular renal dysfunction caused by renal hypodysplasia and cystic kidney diseases in this study that carried the common causative genes varied in the age and CKD stage at disease onset. This new knowledge should help with improving the management and genetic counseling of the abovementioned diseases in clinical practice.

DATA AVAILABILITY STATEMENT

The datasets for this article are not publicly available due to concerns regarding participant/patient anonymity. Requests to access the datasets should be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of Peking University First Hospital approved the procedures in this study. 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

FW and JD: conceptualization, formal analysis, writing – review and editing, visualization, supervision, project Administration, and funding acquisition. XW: methodology, software, investigation, and resources. HX, YY, KX, HZ, XL, BS, NG, XZ, and YZ: data curation. XW and FW: writing – original draft preparation. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from the National Key Research and Development Program of China (No. 2016YFC0901505) (the registry study of rare diseases in children) and the Beijing Key Laboratory of Molecular Diagnosis and Study on Pediatric Genetic Diseases (BZ0317).

ACKNOWLEDGMENTS

We thank the patients, families, and physicians who contributed to this project. We are very grateful to the NCBI, LOVD, and UCSC databases for their free use.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.697085/full#supplementary-material

Supplementary Table 1 | List of genes list associated with renal hypodysplasia and cystic kidney diseases.

REFERENCES

Ahn, Y. H., Lee, C., Kim, N. K. D., Park, E., Kang, H. G., Ha, I. S., et al. (2020). Targeted exome sequencing provided comprehensive genetic diagnosis of congenital anomalies of the kidney and urinary tract. *J. Clin. Med.* 9:751. doi: 10.3390/jcm9030751

Andrassy, K. M. (2013). Comments on 'KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease'. *Kidney Int.* 84, 622–623. doi: 10.1038/ki.2013.243

Armstrong, M. E., and Thomas, C. P. (2019). Diagnosis of monogenic chronic kidney diseases. *Curr. Opin. Nephrol. Hypertens* 28, 183–194. doi: 10.1097/MNH.0000000000000486

Avni, F. E., Lahoche, A., Langlois, C., Garel, C., Hall, M., and Vivier, P. H. (2015). Renal involvement in children with HNF1beta mutation: early sonographic appearances and long-term follow-up. *Eur. Radiol.* 25, 1479–1486. doi: 10.1007/s00330-014-3550-x

Bower, M., Salomon, R., Allanson, J., Antignac, C., Benedicenti, F., Benetti, E., et al. (2012). Update of PAX2 mutations in renal coloboma syndrome and establishment of a locus-specific database. *Hum. Mutat.* 33, 457–466. doi: 10.1002/humu.22020

Bullich, G., Domingo-Gallego, A., Vargas, I., Ruiz, P., Lorente-Grandoso, L., Furlano, M., et al. (2018). A kidney-disease gene panel allows a comprehensive genetic diagnosis of cystic and glomerular inherited kidney diseases. *Kidney Int.* 94, 363–371. doi: 10.1016/j.kint.2018.02.027

Chaki, M., Hoefele, J., Allen, S. J., Ramaswami, G., Janssen, S., Bergmann, C., et al. (2011). Genotype-phenotype correlation in 440 patients with NPHP-related ciliopathies. *Kidney Int.* 80, 1239–1245. doi: 10.1038/ki.2011.284

Clissold, R. L., Hamilton, A. J., Hattersley, A. T., Ellard, S., and Bingham, C. (2015). HNF1B-associated renal and extra-renal disease-an expanding clinical spectrum. *Nat. Rev. Nephrol.* 11, 102–112. doi: 10.1038/nrneph.2014.232

Deng, H., Zhang, Y., Xiao, H., Yao, Y., Liu, X., Su, B., et al. (2019). Diverse phenotypes in children with PAX2-related disorder. *Mol. Genet. Genomic Med.* 7:e701. doi: 10.1002/mgg3.701

Deng, H., Zhang, Y., Yao, Y., Xiao, H., Su, B., Xu, K., et al. (2020). Interpretation of autosomal recessive kidney diseases with "presumed homozygous" pathogenic variants should consider technical pitfalls. *Front. Pediatr.* 8:165. doi: 10.3389/fped.2020.00165

Devlin, L. A., and Sayer, J. A. (2019). Renal ciliopathies. *Curr. Opin. Genet. Dev.* 56, 49–60. doi: 10.1016/j.gde.2019.07.005

Devuyst, O., Knoers, N. V., Remuzzi, G., Schaefer, F., Board of the Working Group for Inherited Kidney Diseases of the European Renal, A., European, D., et al. (2014). Rare inherited kidney diseases: challenges, opportunities, and perspectives. *Lancet* 383, 1844–1859. doi: 10.1016/S0140-6736(14)60659-0

Gimpel, C., Avni, E. F., Breysem, L., Burgmaier, K., Caroli, A., Cetiner, M., et al. (2019). Imaging of kidney cysts and cystic kidney diseases in children: an international working group consensus statement. *Radiology* 290, 769–782. doi: 10.1148/radiol.2018181243

Gribouval, O., Moriniere, V., Pawtowski, A., Arrondel, C., Sallinen, S. L., Saloranta, C., et al. (2012). Spectrum of mutations in the renin-angiotensin system genes in autosomal recessive renal tubular dysgenesis. *Hum. Mutat.* 33, 316–326. doi: 10.1002/humu.21661

Groopman, E. E., Rasouly, H. M., and Gharavi, A. G. (2018). Genomic medicine for kidney disease. *Nat. Rev. Nephrol.* 14, 83–104. doi: 10.1038/nrneph.2017.167

Halbritter, J., Diaz, K., Chaki, M., Porath, J. D., Tarrier, B., Fu, C., et al. (2012). High-throughput mutation analysis in patients with a nephronophthisis-associated ciliopathy applying multiplexed barcoded array-based PCR amplification and next-generation sequencing. *J. Med. Genet.* 49, 756–767. doi: 10.1136/jmedgenet-2012-100973

Halbritter, J., Porath, J. D., Diaz, K. A., Braun, D. A., Kohl, S., Chaki, M., et al. (2013). Identification of 99 novel mutations in a worldwide cohort of 1,056 patients with a nephronophthisis-related ciliopathy. *Hum. Genet.* 132, 865–884. doi: 10.1007/s00439-013-1297-0

Hwang, D. Y., Dworschak, G. C., Kohl, S., Saisawat, P., Vivante, A., Hilger, A. C., et al. (2014). Mutations in 12 known dominant disease-causing genes clarify many congenital anomalies of the kidney and urinary tract. *Kidney Int.* 85, 1429–1433. doi: 10.1038/ki.2013.508

Ishiwa, S., Sato, M., Morisada, N., Nishi, K., Kanamori, T., Okutsu, M., et al. (2019). Association between the clinical presentation of congenital anomalies of the kidney and urinary tract (CAKUT) and gene mutations: an analysis of 66 patients at a single institution. *Pediatr. Nephrol.* 34, 1457–1464. doi: 10.1007/s00467-019-04230-w

Kang, H. G., Lee, H. K., Ahn, Y. H., Joung, J. G., Nam, J., Kim, N. K., et al. (2016). Targeted exome sequencing resolves allelic and the genetic heterogeneity in the genetic diagnosis of nephronophthisis-related ciliopathy. *Exp. Mol. Med.* 48:e251. doi: 10.1038/emm.2016.63

Khanna, H., Davis, E. E., Murga-Zamalloa, C. A., Estrada-Cuzcano, A., Lopez, I., den Hollander, A. I., et al. (2009). A common allele in RPGRIP1L is a modifier of retinal degeneration in ciliopathies. *Nat. Genet.* 41, 739–745. doi: 10.1038/ng.366

Loftus, W. K., Gent, R. J., LeQuesne, G. W., and Metreweli, C. (1998). Renal length in Chinese children: sonographic measurement and comparison with western data. *J. Clin. Ultrasound* 26, 349–352. doi: 10.1002/(sici)1097-0096(199809)26:7<349::aid-jcu4<3.0.co;2-9

Mann, N., Braun, D. A., Amann, K., Tan, W., Shril, S., Connaughton, D. M., et al. (2019). Whole-Exome sequencing enables a precision medicine approach for kidney transplant recipients. *J. Am. Soc. Nephrol.* 30, 201–215. doi: 10.1681/ASN.2018060575

Nicolaou, N., Pulit, S. L., Nijman, I. J., Monroe, G. R., Feitz, W. F., Schreuder, M. F., et al. (2016). Prioritization and burden analysis of rare variants in 208 candidate genes suggest they do not play a major role in CAKUT. *Kidney Int.* 89, 476–486. doi: 10.1038/ki.2015.319

Nigam, A., Knoers, N., and Renkema, K. Y. (2019). Impact of next generation sequencing on our understanding of CAKUT. *Semin. Cell Dev. Biol.* 91, 104–110. doi: 10.1016/j.semcdb.2018.08.013

Obeidova, L., Seeman, T., Fencl, F., Blahova, K., Hojny, J., Elisakova, V., et al. (2020). Results of targeted next-generation sequencing in children with cystic kidney diseases often change the clinical diagnosis. *PLoS One* 15:e0235071. doi: 10.1371/journal.pone.0235071

Otto, E., Hoefele, J., Ruf, R., Mueller, A. M., Hiller, K. S., Wolf, M. T., et al. (2002). A gene mutated in nephronophthisis and retinitis pigmentosa encodes a novel protein, nephroretinin, conserved in evolution. *Am. J. Hum. Genet.* 71, 1161–1167. doi: 10.1086/344395

Otto, E. A., Schermer, B., Obara, T., O'Toole, J. F., Hiller, K. S., Mueller, A. M., et al. (2003). Mutations in INVS encoding inversin cause nephronophthisis type 2, linking renal cystic disease to the function of primary cilia and left-right axis determination. *Nat. Genet.* 34, 413–420. doi: 10.1038/ng1217

Rao, J., Liu, X., Mao, J., Tang, X., Shen, Q., Li, G., et al. (2019). Genetic spectrum of renal disease for 1001 Chinese children based on a multicentre registration system. *Clin. Genet.* 96, 402–410. doi: 10.1111/cge.13606

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

Roberts, S. C. M., Fuentes, L., Kriz, R., Williams, V., and Upadhyay, U. D. (2017). Corrigendum to "Implications for women of Louisiana's law requiring abortion providers to have hospital admitting privileges" [Contraception (2015) 91:368-372. doi: 10.1016/j.contraception.2015.02.001] *Contraception* 95, 221–222. doi: 10.1016/j.contraception.2016.08.015

Sanna-Cherchi, S., Caridi, G., Weng, P. L., Scolari, F., Perfumo, F., Gharavi, A. G., et al. (2007). Genetic approaches to human renal agenesis/hypoplasia and dysplasia. *Pediatr. Nephrol.* 22, 1675–1684. doi: 10.1007/s00467-007-0479-1

Sanna-Cherchi, S., Kiryluk, K., Burgess, K. E., Bodria, M., Sampson, M. G., Hadley, D., et al. (2012). Copy-number disorders are a common cause of congenital kidney malformations. *Am. J. Hum. Genet.* 91, 987–997. doi: 10.1016/j.ajhg.2012.10.007

Sanna-Cherchi, S., Westland, R., Ghiggeri, G. M., and Gharavi, A. G. (2018). Genetic basis of human congenital anomalies of the kidney and urinary tract. *J. Clin. Invest.* 128, 4–15. doi: 10.1172/JCI95300

Schwartz, G. J., Brion, L. P., and Spitzer, A. (1987). The use of plasma creatinine concentration for estimating glomerular filtration rate in infants, children, and adolescents. *Pediatr. Clin. North Am.* 34, 571–590. doi: 10.1016/s0031-3955(16)36251-4

Smith, J. M., Stablein, D. M., Munoz, R., Hebert, D., and McDonald, R. A. (2007). Contributions of the transplant registry: the 2006 annual report of the North American Pediatric renal trials and collaborative studies (NAPRTCS). *Pediatr. Transpl.* 11, 366–373. doi: 10.1111/j.1399-3046.2007.00704.x

Soliman, N. A., Hildebrandt, F., Otto, E. A., Nabhan, M. M., Allen, S. J., Badr, A. M., et al. (2012). Clinical characterization and NPHP1 mutations in nephronophthisis and associated ciliopathies: a single center experience. *Saudi J. Kidney Dis. Transpl.* 23, 1090–1098. doi: 10.4103/1319-2442.100968

Tang, X., Liu, C., Liu, X., Chen, J., Fan, X., Liu, J., et al. (2020). Phenotype and genotype spectra of a Chinese cohort with nephronophthisis-related ciliopathy. *J. Med. Genet* doi: 10.1136/jmedgenet-2020-107184 [Epub ahead of print],

Thomas, R., Sanna-Cherchi, S., Warady, B. A., Furth, S. L., Kaskel, F. J., and Gharavi, A. G. (2011). HNF1B and PAX2 mutations are a common cause of renal hypodysplasia in the CKiD cohort. *Pediatr. Nephrol.* 26, 897–903. doi: 10.1007/s00467-011-1826-9

Verbitsky, M., Westland, R., Perez, A., Kiryluk, K., Liu, Q., Krithivasan, P., et al. (2019). The copy number variation landscape of congenital anomalies of the kidney and urinary tract. *Nat. Genet.* 51, 117–127. doi: 10.1038/s41588-018-0281-y

Vester, U., Kranz, B., and Hoyer, P. F. (2010). The diagnostic value of ultrasound in cystic kidney diseases. *Pediatr. Nephrol.* 25, 231–240. doi: 10.1007/s00467-008-0981-0

Vivante, A., and Hildebrandt, F. (2016). Exploring the genetic basis of early-onset chronic kidney disease. *Nat. Rev. Nephrol.* 12, 133–146. doi: 10.1038/nrneph.2015.205

Wang, F., Yao, Y., Yang, H. X., Shi, C. Y., Zhang, X. X., Xiao, H. J., et al. (2017a). [Clinical phenotypes of hepatocyte nuclear factor 1 homeobox b-associated disease]. *Zhonghua Er Ke Za Zhi* 55, 658–662. doi: 10.3760/cma.j.issn.0578-1310.2017.09.006

Wang, F., Zhang, Y., Mao, J., Yu, Z., Yi, Z., Yu, L., et al. (2017b). Spectrum of mutations in Chinese children with steroid-resistant nephrotic syndrome. *Pediatr. Nephrol.* 32, 1181–1192. doi: 10.1007/s00467-017-3590-y

Weber, S., Moriniere, V., Knuppel, T., Charbit, M., Dusek, J., Ghiggeri, G. M., et al. (2006). Prevalence of mutations in renal developmental genes in children

with renal hypodysplasia: results of the ESCAPE study. *J. Am. Soc. Nephrol.* 17, 2864–2870. doi: 10.1681/ASN.2006030277

Wuhl, E., van Stralen, K. J., Wanner, C., Ariceta, G., Heaf, J. G., Bjerre, A. K., et al. (2014). Renal replacement therapy for rare diseases affecting the kidney: an analysis of the ERA-EDTA Registry. *Nephrol. Dial. Transplant* 29 Suppl 4, iv1–iv8. doi: 10.1093/ndt/gfu030

Yi, Z., Pan, H., Li, L., Wu, H., Wang, S., Ma, Y., et al. (2016). Chromosome Xq28 duplication encompassing MECP2: clinical and molecular analysis of 16 new patients from 10 families in China. *Eur. J. Med. Genet.* 59, 347–353. doi: 10.1016/j.ejmg.2016.05.004

Yue, Z., Lin, H., Li, M., Wang, H., Liu, T., Hu, M., et al. (2020). Clinical and pathological features and varied mutational spectra of pathogenic genes in 55 Chinese patients with nephronophthisis. *Clin. Chim. Acta* 506, 136–144. doi: 10.1016/j.cca.2020.03.015

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 07 July 2021
doi: 10.3389/fgene.2021.688241

Investigation of a Novel *LRP6* Variant Causing Autosomal-Dominant Tooth Agenesis

Yan-xia Huang[1], Chun-yan Gao[1], Chun-yan Zheng[1], Xu Chen[1], You-sheng Yan[2], Yong-qing Sun[2], Xing-yue Dong[1], Kai Yang[2] and Dong-liang Zhang[1]**

[1] Department of Orthodontics, School of Stomatology, Beijing Stomatological Hospital, Capital Medical University, Beijing, China, [2] Prenatal Diagnosis Center, Beijing Obstetrics and Gynecology Hospital, Capital Medical University, Beijing, China

Background: The low-density lipoprotein receptor-related protein 6 (*LRP6*) gene is a recently defined gene that is associated with the autosomal-dominant inherited tooth agenesis (TA). In the present study, a family of four generations having TA was recruited and subjected to a series of clinical, genetic, *in silico*, and *in vitro* investigations.

Methods: After routine clinical evaluation, the proband was subjected to whole-exome sequencing (WES) to detect the diagnostic variant. Next, *in silico* structural and molecular dynamics (MD) analysis was conducted on the identified novel missense variant for predicting its intramolecular impact. Subsequently, an *in vitro* study was performed to further explore the effect of this variant on protein maturation and phosphorylation.

Results: WES identified a novel variant, designated as *LRP6*: c.2570G > A (p.R857H), harbored by six members of the concerned family, four of whom exhibited varied TA symptoms. The *in silico* analysis suggested that this novel variant could probably damage the Wnt bonding function of the LRP6 protein. The experimental study demonstrated that although this novel variant did not affect the *LRP6* gene transcription, it caused a impairment in the maturation and phosphorylation of LRP6 protein, suggesting the possibility of the disruption of the Wnt signaling.

Conclusion: The present study expanded the mutation spectrum of human TA in the *LRP6* gene. The findings of the present study are insightful and conducive to understanding the functional significance of specific *LRP6* variants.

Keywords: tooth agenesis, *LRP6* gene, whole-exome sequencing, molecular dynamics analysis, experimental study

OPEN ACCESS

Edited by:
Yanling Yang,
Peking University First Hospital, China

Reviewed by:
Lucia Taja-Chayeb,
National Institute of Cancer (INCAN), Mexico
Teresa Villarreal-Molina,
Instituto Nacional de Medicina Genómica (INMEGEN), Mexico

****Correspondence:***
Dong-liang Zhang
zhangdongliang@mail.ccmu.edu.cn
Kai Yang
yk19830919@ccmu.edu.cn

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

Received: *30 March 2021*
Accepted: *16 June 2021*
Published: *07 July 2021*

Citation:
Huang Y-x, Gao C-y, Zheng C-y, Chen X, Yan Y-s, Sun Y-q, Dong X-y, Yang K and Zhang D-l (2021) Investigation of a Novel LRP6 Variant Causing Autosomal-Dominant Tooth Agenesis. Front. Genet. 12:688241. doi: 10.3389/fgene.2021.688241

INTRODUCTION

Tooth agenesis (TA) is one of the most prevalent congenital craniofacial malformations occurring in humans, which may lead to masticatory dysfunction, speech alteration, and malocclusion, besides aesthetic problems (Matalova et al., 2008). TA occurs in over 300 syndromic or non-syndromic conditions, with a remarkable phenotypic heterogeneity[1]. Among the familial cases

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

of TA, autosomal dominant inheritance is the most frequent pattern observed (Matalova et al., 2008). On the basis of the number of missing teeth, TA is classified into hypodontia (<6 teeth) and oligodontia (≥6 teeth). Hypodontia is quite common, with a prevalence of 3–10% depending on the population, while oligodontia is rare, with a prevalence of <1% (Dhamo et al., 2018).

Genetic variations contribute greatly to the pathogenesis of congenital TA, and may also provide insights into tooth development (Yin and Bian, 2015). To date, 15 genes in the WNT/β-catenin, TGF-β/BMP, and Eda/Edar/NF-κB pathways, namely, *WNT10A* (MIM *606268), *WNT10B* (MIM *601906), *LRP6* (MIM *603507), *DKK1* (MIM *605189), *KREMEN1* (MIM *609898), *AXIN2* (MIM *604025), *PAX9* (MIM *167416), *MSX1* (MIM *142983), *GREM2* (MIM *608832), *BMP4* (MIM *112262), *LTBP3* (MIM *602090), *EDA* (MIM *300451), *EDAR* (MIM *604095), *EDARADD* (MIM *606603), and *SMOC2* (MIM *607223), have been reported to be responsible for non-syndromic TA (see text footnote 1).

The WNT/β-catenin pathway plays a pivotal role in cell differentiation, proliferation, and migration involved in the formation and homeostasis of bone and teeth (Duan and Bonewald, 2016). The *LRP6* gene encodes the low-density lipoprotein receptor-related protein 6 (LRP6), which is a single-pass transmembrane receptor of Wnts in the WNT/β-catenin pathway (Tamal et al., 2000; Kokubu et al., 2004). The mutations in the *LRP6* gene were initially reported to be associated with a broad spectrum of anomalies in human and animals, such as neural tube defects (Carter et al., 2005), early coronary disease (Mani et al., 2007), and metabolic syndromes (Singh et al., 2013). Massink et al. (2015) reported that the variations in the *LRP6* gene could cause autosomal dominant TA. Later, three studies on TA identified other variants of *LRP6*, which corroborated this causality (Ockeloen et al., 2016; Dinckan et al., 2018; Yu et al., 2020). In the most recent study, Yu et al. (2020) revealed the spatial/temporal expression pattern of Lrp6 in mouse molar development, which provided further insight into the dynamic function of the WNT/β-catenin pathway in tooth formation.

In the present study, a Chinese family with four generations having autosomal dominant TA was recruited and subjected to a comprehensive genetic investigation which revealed a novel missense variant in *LRP6*. In order to confirm the pathogenicity of the identified novel variant and understand its impact on the protein structure and function, western blotting (WB) and *in silico* molecular dynamics (MD) simulation analyses were performed.

MATERIALS AND METHODS

Subjects and Clinical Evaluation

The present study was designed as a prospective review and was approved by the Ethics Committee of the Capital Medical University Affiliated Beijing Stomatological Hospital. Informed consent was provided by all the participants included in the study. All procedures performed in the present study were in accordance with the Declaration of Helsinki 1964 and its later amendments or comparable ethical standards.

A family with four generations having hereditary TA was recruited at the Beijing Stomatological Hospital in December 2019. Twelve members of this family, including seven females and five males, participated in the study. The medical histories of all participants were thoroughly surveyed, and a pedigree diagram was drawn. Panoramic radiography was conducted for the members affected by TA.

Genetic Studies

Genomic DNA was extracted from the peripheral blood sample from each of the 12 subjects using QIAamp DNA Blood Mini kit (Qiagen, Germany).

Whole-exome sequencing was employed to detect the sequence variants in the sample from the proband, as described in a previous study (Yang et al., 2019). Briefly, the target-region sequences were enriched using the Agilent Sure Select Human Exon Sequence Capture Kit, V5 + UTR (Agilent, United States). The DNA libraries were screened using quantitative PCR, the size, distribution, and concentration of which were determined using Agilent Bioanalyzer 2100 (Agilent, United States). The NovaSeq6000 platform (Illumina, Inc.) and ∼150 bp pair-end reads were employed to sequence the DNA (∼300 pM per sample) using the NovaSeq Reagent kit. The sequencing raw reads (quality level Q30 > 90%)[2] were aligned to the human reference genome (accession no: hg19/GRCh37) in Burrows–Wheeler Aligner tool, and the PCR duplicates were removed using Picard v1.57. Variant calling was performed using the Verita Trekker® Variants Detection system (v2.0; Berry Genomics, China) and the Genome Analysis Tool Kit[3] (detailed variant filtering criteria included in **Supplementary Material**). Subsequently, the annotation tools, ANNOVAR v2.0 (Wang et al., 2010) and Enliven® Variants Annotation Interpretation systems (Berry Genomics), were used to provide information for the establishment of the criteria of the common guidelines provided by the American College of Medical Genetics and Genomics (ACMG) (Richards et al., 2015). In order to assist in the interpretation of pathogenicity, three frequency databases (1000G_2015aug_eas, ExAC_EAS, and gnomAD_exome_EAS)[4][5][6] and HGMD pro v2019 (Human Gene Mutation Database) were referred to. The Revel score (for pathogenicity prediction) (Ioannidis et al., 2016) and the pLI score (representing the tolerance for truncating variants) were also utilized.

As a confirmatory method, Sanger sequencing was performed using the 3500DX Genetic Analyzer (Applied Biosystems, United States). The details regarding the sequencing PCR primers, reaction conditions, and reagents are provided in

[2]https://www.illumina.com/science/technology/next-generation-sequencing/plan-experiments/quality-scores.html

[3]https://software.broadinstitute.org/gatk/

[4]https://www.internationalgenome.org

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

[6]http://gnomad.broadinstitute.org

Supplementary Material. The evolutionary conservatism of the amino acid (AA) residue affected by the identified novel missense variant was analyzed using MEGA7[7] with the default parameters.

Structural and Molecular Dynamics (MD) Analysis

The SWISS-MODEL program was applied for modeling the LRP6 PE3/4 domains containing the mutation site, using the crystal structure PDB:3S2K[8] with X-resolution of 2.8 Å as the template (Ahn et al., 2011).

Next, the molecular dynamics (MD) prediction analysis was performed for the wild type (WT) LRP6 and R857H-LRP6 models generated by Modeler 9v17 (Šali et al., 1995). The CHARMM22 program was applied to add hydrogen atoms and N-terminal and C-terminal patches to the models (MacKerell et al., 1998). The generated models were solvated and neutralized with TIP3P water within a box at a minimum distance of 13 Å between the model and the wall of the box. All simulations were run using NAMD 2.9 and by applying periodic boundary conditions (PBC) (Phillips et al., 2020). The temperature was set at 300 K, the pressure was set at 1 atm, and the time step was set to 2 fs. The particle mesh Ewald method was used for modeling the electrostatics, and the threshold for van der Waals interactions was set at 12 Å. Both the models followed a three-step pre-equilibration totaling 600 ps, with the last snapshots selected as the beginning structures for 60-ns productive simulations without constraints.

Plasmids Construction, Cell Transfection, and Wnt3a Treatment

In order to construct the expression plasmid vectors containing the coding sequence of the wild-type (WT) or mutant *LRP6*, the cDNA sequences of both WT and mutant *LRP6* were obtained using RT-PCR from the mRNA sample extracted from the proband. The obtained cDNA sequences were subcloned into the pET vectors and verified using Sanger sequencing. Next, the cDNA sequences were inserted into the pcDNA3.1(+) vectors, and the resultants were designated as LRP6-WT and LRP6-Mut, which represented the WT and mutant ones, respectively.

Commercial HEK (human embryonic kidney) 293T cells were purchased and cultured in 24-well plates. Subsequently, the cells were transfected with LRP6-WT or LRP6-Mut using Lipofectamine 3000 (Invitrogen, United States) at a suitable confluence.

This process lasted for 48 h, after which the cells were (±) treated with recombinant human Wnt3a protein (Cat No. ab153563, Abcam, United States) via Lipofectamine 3000.

Western Blotting (WB)

In order to analyze the expression levels and the phosphorylation levels of the LRP6 protein in the HEK 293T cells, WB was performed using the monoclonal antibodies against LRP6 (Cat No. 2560, CST, United States) and phospho-LRP6 (Ser1490) (Cat No. 2568, CST, United States), respectively. Detailed methods were described in **Supplementary Material**.

RNA Expression Analysis With Quantitative Fluorescent RT-PCR

At 48 h after transfection, cells were harvested and total RNA was extracted with an RNeasy Mini Kit (QIAGEN). Reverse transcription was performed with the Prime Script RT Reagent Kit with the gDNA Eraser (Takara). The expression level of *LRP6* was assessed by quantitative fluorescent RT-PCR using SYBR Premix Ex Taq II (Perfect Real Time) (Takara) with ABI 7500 system (see details in **Supplementary Material**).

Statistical Analysis

Statistical calculations were performed using the SPSS v22.0 software. Student's *t*-test was used for determining the statistical significance and $p \leq 0.05$ was considered significant.

RESULTS

Clinical Manifestations

The constructed pedigree diagram is depicted in **Figure 1A**. Four members of the recruited family, designated as II-2 (the proband), II-3, III-2, and III-3, exhibited apparent permanent tooth loss, with a remarkable phenotypic heterogeneity in the number of missing teeth. In addition, both II-2 and III-3 had sparse hair (data not presented as requested by the participants). The phenotype of generation IV was not conclusive as all three subjects in this generation (IV-1, IV-2, and IV-3) were at the age of deciduous teeth. Detailed clinical information regarding the four affected members is provided in **Table 1** and the panoramic radiographs and the corresponding schematics are presented in **Figure 1B**. The images of the four patients in this study were included in **Supplementary Material**.

Genetic Findings

Whole-exome sequencing identified a novel heterozygous missense variant, which was named *LRP6*: NM: 002336.3: c.2570G > A (p.R857H). This identified variant was later verified using Sanger sequencing (**Figures 2A,B**). In addition, Sanger sequencing revealed that six members in this family, namely, II-2, II-3, III-2, III-3, IV-2, and IV-3, carried the identified novel variant, while the other six members carried the wild-type variant (**Figure 2B**). As depicted in the linear structural patterns of the genes and proteins, the mutation was located in the PE3 (P, the YWTD β-propeller domain; E, the epidermal growth factor-like domain) domain (**Figure 2C**). The AA residue that was affected by this novel mutant variant, namely R857, remained evolutionarily conserved among species (**Figure 2D**). None of the three frequency databases searched had this variant included, and both REVEL and pLI predictions indicated it as being "deleterious" to gene function (see **Supplementary Material**). Accordingly, this variant fulfills the evidence level PM2, PP1, PP2, PP3, and PP4.

[7]http://www.megasoftware.net/previousVersions.php

[8]http://www1.rcsb.org/

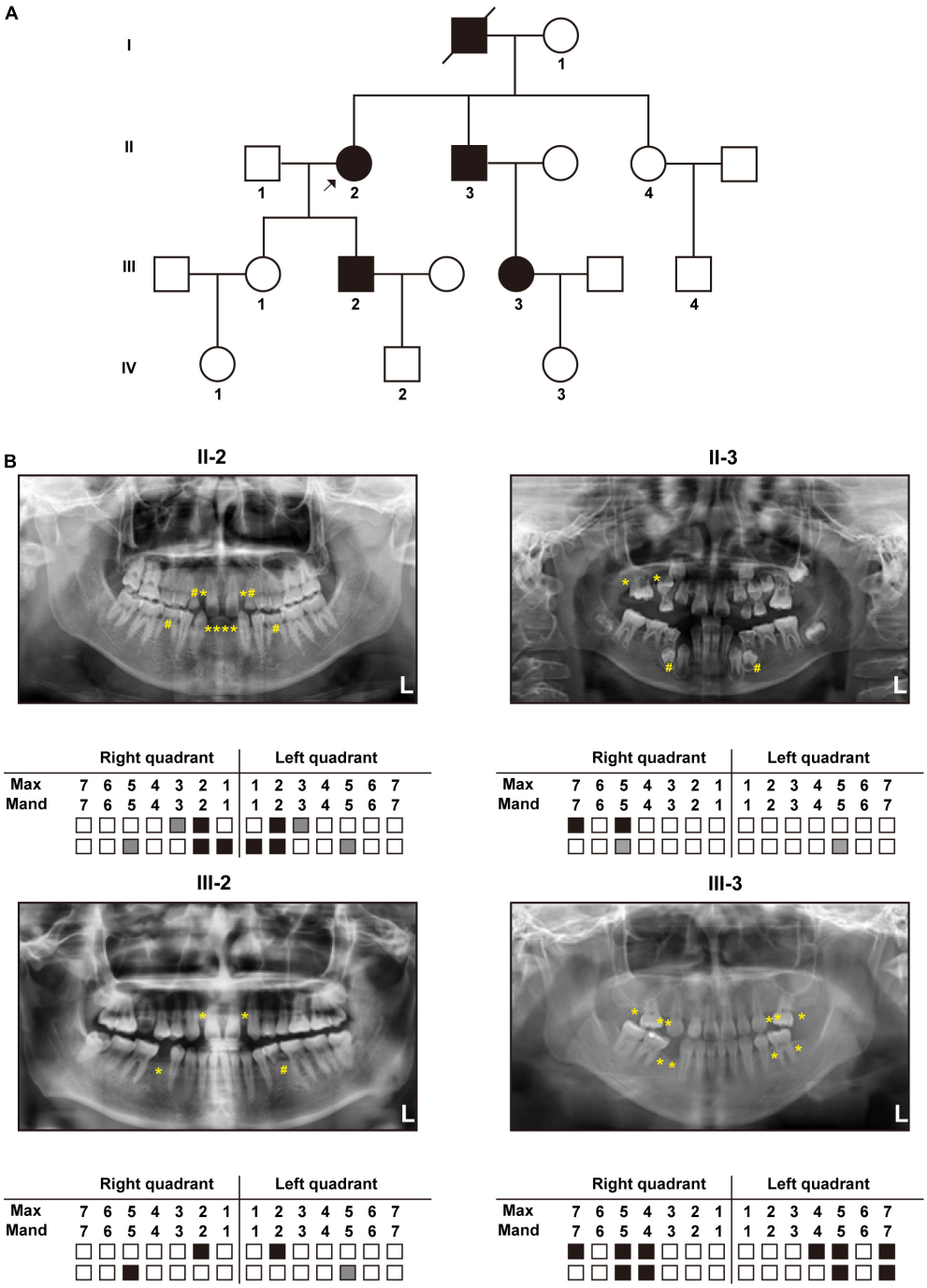


FIGURE 1 | Clinical manifestations: **(A)** Pedigree diagram of the recruited family with autosomal dominant tooth agenesis (TA) (The arrow indicates the proband). **(B)** Panoramic radiographs and schematics of the four patients in the recruited family (White blocks represent the normal eruption of permanent teeth; Gray blocks and "#" marks represent residual deciduous teeth; Black blocks and "*" marks represent permanent tooth agenesis).

The *LRP6*^{R857H} Variant Impact the Protein Stability and Secondary Structure

The final and converged models out of structure prediction are depicted in **Figure 3A**. It could be inferred from the trajectory of Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuation (RMSF) that the R857H model was extremely flexible compared to the WT model (**Figures 3B–D**). Unsurprisingly, amino acid residue R857 formed a greater number of hydrogen bonds with the other residues compared to the variant residue R857H (**Figure 3E**). Furthermore, R857H

variant affected the corresponding secondary structures of the LRP6 protein (**Figure 3F**).

LRP6^{R857H} Impacts the Protein Maturation and Phosphorylation

As depicted in **Figure 4A**, both WT and Mut LRP6 vectors could be expressed exogenously, although the LRP6-Mut vector generated only a single lower molecular weight (MW) protein band while the LRP6-WT vector generated two adjacent bands. Moreover, after treatment with Wnt3a, a significantly higher amount of the phosphorylated LRP6 protein was produced by WT compared to the Mut.

Figure 4B presents the comparison of the transcription levels of *LRP6*, with no significant difference indicated.

DISCUSSION

The *LRP6* gene has been cloned and studied for several decades now (Brown et al., 1998; Pinson et al., 2000). Nonetheless, the comprehensive function of this gene in the physiological process of humans remains to be elucidated. The possibility of the pathogenic variants in *LRP6* leading to autosomal dominant TA was revealed quite recently (Massink et al., 2015). So far, to the best of our knowledge, only 15 TA-related pathogenic variants in *LRP6* are reported, which includes seven truncating variants, four missense variants, and four splicing variants (**Figure 2C** and **Table 2**; Massink et al., 2015; Ockeloen et al., 2016; Dinckan et al., 2018; Yu et al., 2020). The present study is the fifth one on TA that reports the identification of a novel *LRP6* variant. Most mutations (11 of 16) are loss of function mutations (frameshift or splice site), although a dominant negative effect of mutations has been suggested by other authors (Yu et al., 2020). Despite the current number of variants and their distribution on various domains, and the evidences provided by functional experiments, the genotype-phenotype correlations of LRP6 gene still cannot be well established.

In the present study, a hereditary variant, which was designated as *LRP6*: c.2570G > A (p.R857H), was detected in six members of the affected family, including two children in the deciduous phase of tooth development. The four adult patients exhibited remarkable intrafamilial phenotypic variation in the number of missing teeth and the presence of other ectodermal dysplasia-like phenotypes. Yu et al. (2020) reported that this phenotype difference exists between affected individuals; yet, in patients from one family, this is the first time we observed this phenotypic variation. We suppose that this phenotypic difference may be related to the individual genetic background or the non-penetrance of this variant, yet both expanded sample size of patients and specific mechanism has to be studied to clearly elucidate this. Greater attention should be paid to the two children who have not yet developed TA during their future development. On the basis of the common standards for the interpretation of genetic variations (Richards et al., 2015), the novel variant identified in the present study was interpreted as "likely pathogenic," as evidenced by "PM2, PP1, PP2, PP3, and PP4." Further functional studies, such as the dimerization of LRP6 and the interaction between LRP6 and Wnts, need to be implemented to demonstrate whether the variant meets the PS3 evidence.

TABLE 1 | Clinical features of the four patients in this family.

Subject ID	Gender	Age	Number of a genetic teeth (excluding third molar)	Other manifestation
II-2	F	61 yr	10	Sparse hair
II-3	M	58 yr	4	NA
III-2	M	29 yr	4	NA
III-3	F	28 yr	10	Sparse hair

F, female; M, Male; yr, years old; NA, not applicable.

The low-density lipoprotein (LDL)-related receptors (LRPs), such as LRP4, LRP5, and LRP6, represent a group of evolutionary-conserved receptors that are involved in the regulation of a wide range of cellular processes through the modulation of several pathways, including the canonical WNT signaling pathway (Huybrechts et al., 2020). The extracellular region of the majority of the LRP receptors contains one ligand-binding domain comprising cysteine-rich ligand-binding-type repeats and one epidermal growth factor (EGF)-precursor homology domain comprising a YWTD/β-propeller (P) domain and EGF repeats (E) (Huybrechts et al., 2020). Presumably, different types of *LRP6* variations lead to increased or decreased WNT signaling activity, thereby inducing different phenotypes (Huybrechts et al., 2020). Previously reported TA-related variations in the coding regions of *LRP6* were more concentrated in the extracellular PE functional domains (**Figure 2C**), which could affect the binding of LRP6 to Wnts. In order to determine the effect of the novel variant on protein function, MD predictive simulation was conducted, the results of which indicated that the amino acid residue R857 located in the β-strand formed a hydrogen bond with residue T867. Then, in the mutation-resultant R857H, which replaced the strongly basic arginine with a less basic amino acid, the hydrogen bonds formed by the side chain of R857 residue were expectantly broken, thereby changing its potential distribution. In addition, the effect of this variation on the secondary structure of LRP6 (**Figure 3F**) could significantly perturb the dimerization of this protein or its binding to Wnt, which is necessary for the formation of the Wnt-Fzd-LRP5-LRP6 complex for the activation of Wnt signaling (Boudin et al., 2013). So far, the evidence suggested the loss of LRP6 protein function, yet it should be demonstrated by further study.

Furthermore, in the *in vitro* study, it was demonstrated that the transcription of *LRP6* was not disrupted by the identified novel variant (**Figure 4B**). However, according to the WB results, the glycosylated mature form of LRP6 was missing in the Mut samples as only a low MW band appeared ("LRP6" line in **Figure 4A**; Massink et al., 2015). In this case, the treatment of the cells with Wnt3a could not mediate the proper LRP6 phosphorylation ("p-LRP6" line in **Figure 4A**; Pan et al., 2008). These results strongly suggested that the p.R857H variant caused loss of function in LRP6. However, the underlying reason for the heterogeneity in the disease phenotype with the same variant

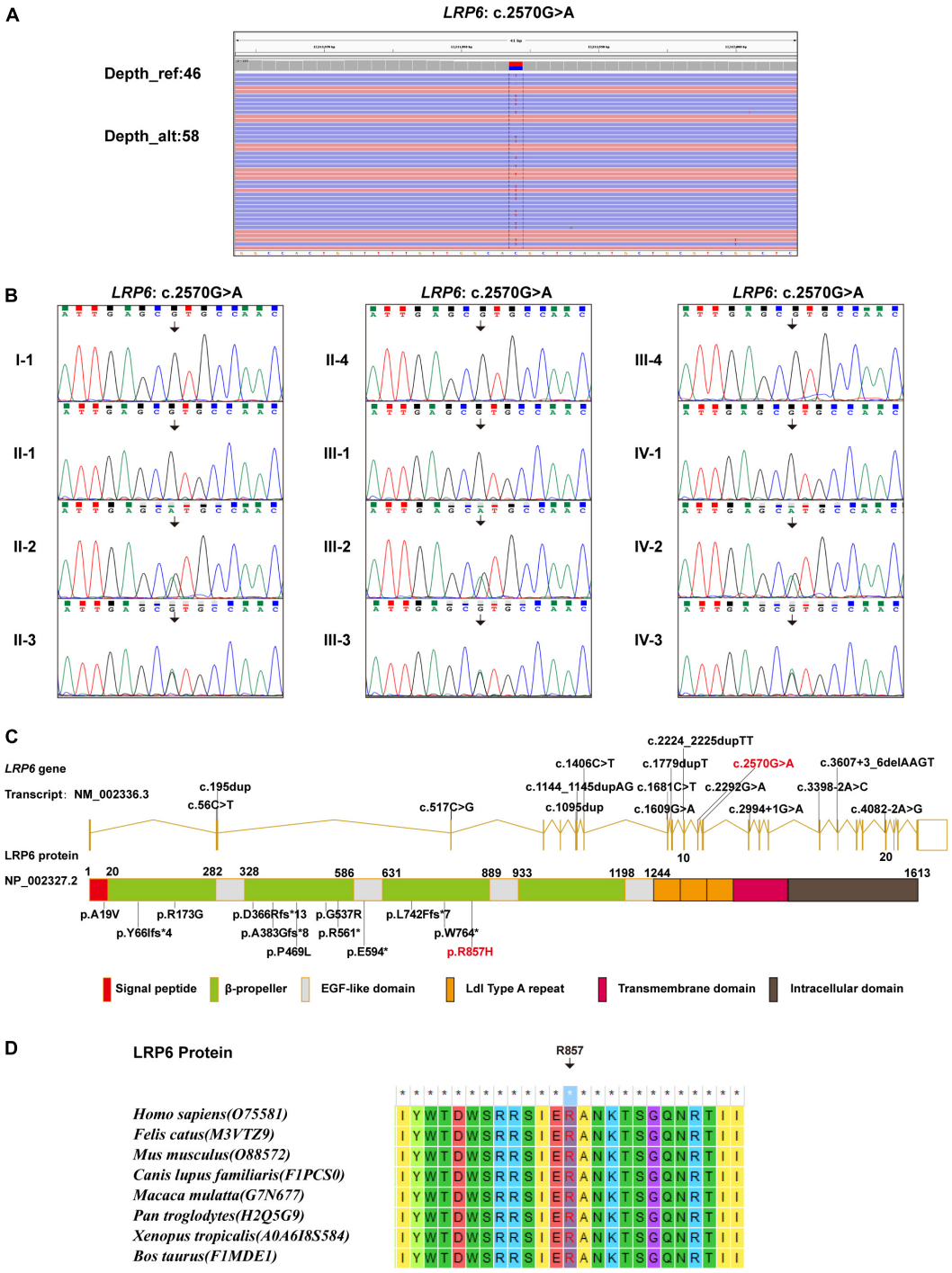


FIGURE 2 | Genetic findings: **(A)** WES data demonstrating the *LRP6*: c.2570G > A variant. **(B)** Sanger sequencing results demonstrating the carrying status of the *LRP6*: c.2570G > A variant in the 12 members of the recruited family. **(C)** All *LRP6* variants associated with TA phenotype reported in literature, illustrated in gene and protein schematics (Red fonts represent the variant in the present study). **(D)** The conservatism of the amino acid (R857) affected by c.2570G > A variant across species.

and the effect of this variant on the Wnt signaling cascade remain to be explored.

One of the limitations of our study is that there is no *in vivo* validation of the variant for the time being, which will be complemented by further studies. Moreover, limited by funding, we have not tested the novel variant with sufficient interaction experiments between Wnts and LRP6, which will be fully supplemented when we are funded.

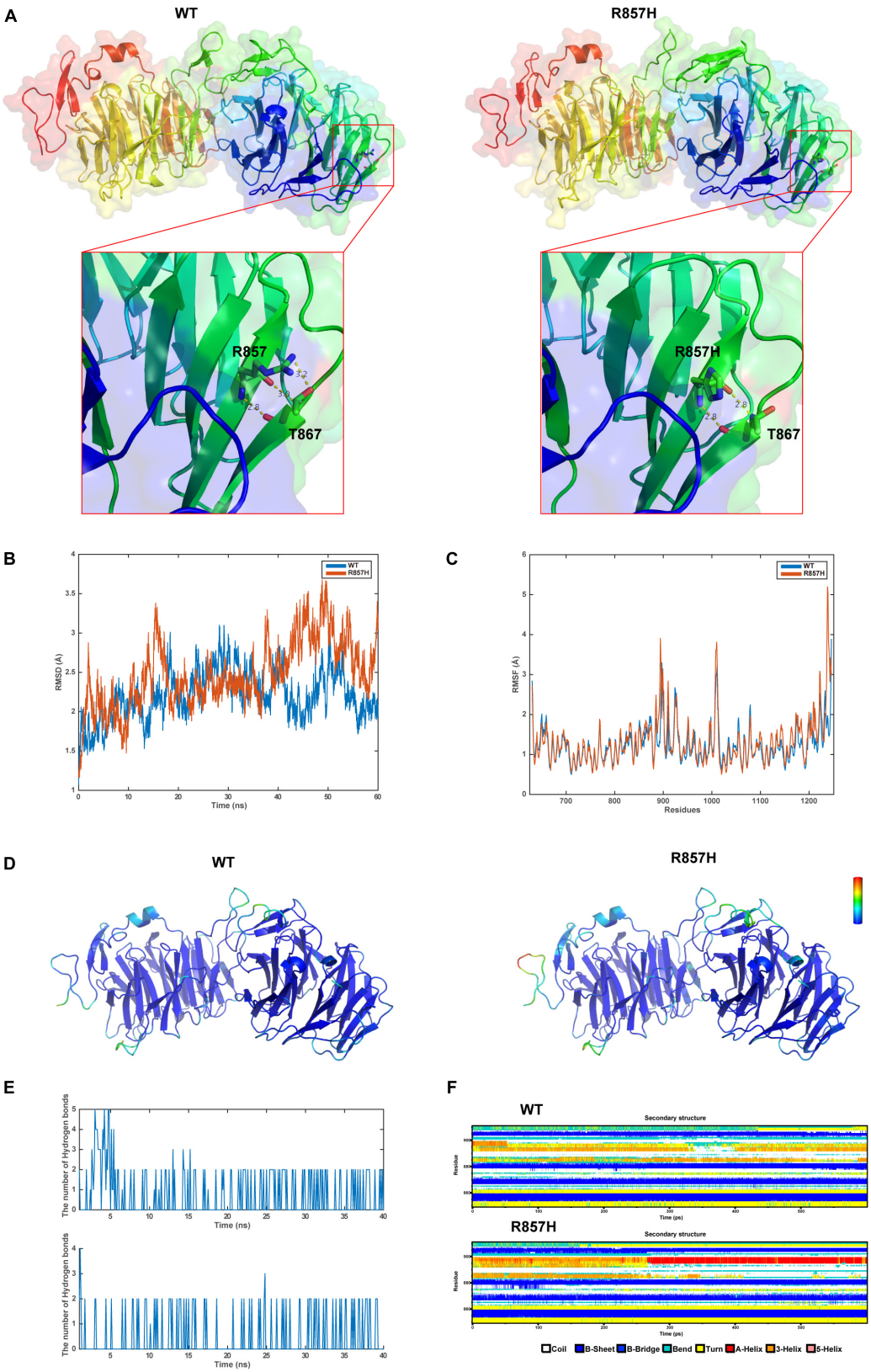


FIGURE 3 | Continued

FIGURE 3 | Structural and MD analysis results: **(A)** The structures of the domain containing the WT and R857H models (Residues forming hydrogen bonds with the residue R857 or R857H are depicted in stick representation; Dotted yellow lines represent the hydrogen bonds). **(B)** The trajectory of RMSD (Cα) for the two proteins, which compared every structure in the trajectory to the reference/initial frame (0 ns) by computing the root mean square deviation (RMSD). RMSD is a numerical measurement representing the difference between two structures. In molecular dynamics, we are interested in how structures and parts of structures change over time as compared to the starting point, so the trajectory of RMSD can be used to identify large changes in protein structure as compared to the starting point. **(C)** RMSF of the two proteins calculated from each simulation, which computed the root mean square fluctuation (RMSF) of atomic positions in the trajectory after fitting to the reference/initial frame (0 ns). RMSF is a numerical measurement similar to RMSD, but instead of indicating positional differences between entire structures over time, RMSF is a calculation of individual residue flexibility, or how much a particular residue moves (fluctuates) during a simulation. **(D)** The two models (WT and R857H) colored according to RMSF. **(E)** The number of hydrogen bonds formed between the residue R857 (up) or R857H (blew) and the other residues for each structure in the trajectory. Although the hydrogen bond is much weaker than a covalent bond, the large number of imide and carbonyl groups in peptide chains results in the formation of numerous hydrogen bonds, and these are important for structures to stabilize the folding of the peptide backbone and facilitate molecular interactions. **(F)** Secondary structural components of the corresponding region as a function of time. Secondary structures, refer to local folded structures that form within a polypeptide due to interactions between atoms of the backbone, linked topologically to form 3D structures. The secondary structures are changed after mutation, especially for residues 890–900.

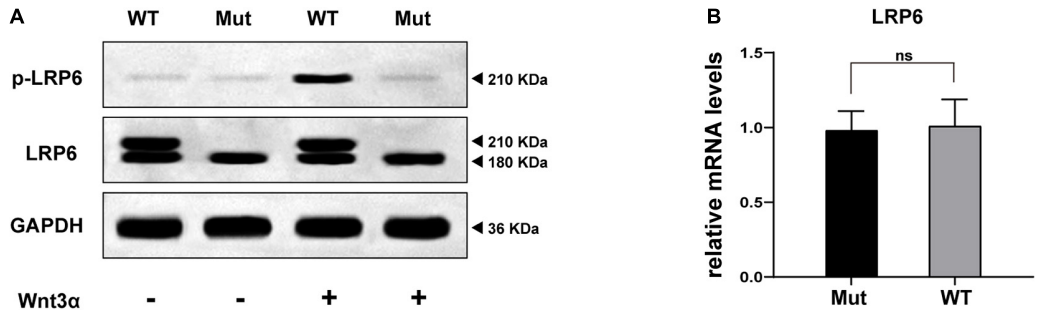


FIGURE 4 | *In vitro* study results: **(A)** WB bands representing the expression, maturation, and phosphorylation levels of LRP6 protein, compared between the two groups of cells (transfected by WT or Mutant LRP6 cDNA, respectively). **(B)** The result of real-time fluorescence quantitative PCR demonstrating the transcriptional level of *LRP6*, compared between the two groups of cells.

TABLE 2 | Tooth agenesis-related variants in the *LRP6* gene (cited from HGMD database, in chronological order of reports).

No.	Genomic coordinates	Reference base (s)	Variant base (s)	HGVS description (NM_002336.3)	Protein alteration	References (PMID)
1	12:12315180-12315180	C	CAA	c.2224_2225dupTT	p.L742Ffs*7	PMID: 26387593
2	12:12397589-12397589	G	A	c.56C > T	p.A19V	PMID: 26387593
3	12:12317479-12317479	C	CA	c.1779dupT	p.E594*	PMID: 26387593
4	12:12334204-12334204	C	CCT	c.1144_1145dupAG	p.A383Gfs*8	PMID: 26387593
5	12:12318166-12318166	C	T	c.1609G > A	p.G537R	PMID: 26963285
6	12:12291470-12291470	T	G	c.3398-2A > C	Splice site	PMID: 26963285
7	12:12303769-12303769	C	T	c.2994 + 1G > A	Splice site	PMID: 26963285
8	12:12356267-12356267	G	C	c.517C > G	p.R173G	PMID: 26963285
9	12:12332883-12332883	G	A	c.1406C > T	p.P469L	PMID: 26963285
10	12:12279857-12279857	T	C	c.4082-2A > G	Splice site	PMID: 26963285
11	12:12291252-12291256	CACTT	C	c.3607 + 3_6delAAGT	Splice site	PMID: 28813618
12	12:12312886-12312886	C	T	c.2292G > A	p.W764*	PMID: 33164649
13	12:12397449-12397450	T	TT	c.195dup	p.Y66Ifs*4	PMID: 33164649
14	12:12334254_12334255	T	TT	c.1095dup	p.D366Rfs*13	PMID: 33164649
15	12:12318094-12318094	G	A	c.1681C > T	p.R561*	PMID: 33164649
16	12:12311984-12311984	C	T	c.2570G > A	p.R857H	This study

HGMD, The Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php).
HGVS, Human Genome Variation Society (http://www.hgvs.org/).
PMID, PubMed paper ID (https://pubmed.ncbi.nlm.nih.gov/).

CONCLUSION

In summary, the present study investigated a large family with TA and identified a novel diagnostic variant in the *LRP6* gene, thereby expanding the mutation spectrum of human tooth agenesis. Moreover, it was confirmed through *in vitro* experiments and intramolecular effects that the identified variant could lead to loss of function in the LRP6 protein, which

provided a novel perspective for the verification of the impact of missense variations.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://figshare.com/articles/online_resource/LRP6_zip/14595420, 1.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Capital Medical University Affiliated Beijing Stomatological Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

D-LZ and KY planned and designed the whole study. Y-XH, C-YG, C-YZ, and XC recruited the family, performed the clinical evaluation, and analyzed the data. Y-SY, Y-QS, and KY conducted genetic detections and corresponding data analysis. Y-XH and X-YD performed structural and MD analysis. C-YG, C-YZ, and KY performed *in vitro* experimental study. All authors wrote, reviewed, and corrected the manuscript.

ACKNOWLEDGMENTS

We thank the patients and their families for their participation in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.688241/full#supplementary-material

REFERENCES

Ahn, V. E., Chu, M. L., Choi, H. J., Tran, D., Abo, A., and Weis, W. I. (2011). Structural Basis of Wnt Signaling Inhibition by Dickkopf Binding to LRP5/6. *Dev. Cell* 21, 862–873. doi: 10.1016/j.devcel.2011.09.003

Boudin, E., Fijalkowski, I., Piters, E., and Van Hul, W. (2013). The role of extracellular modulators of canonical Wnt signaling in bone metabolism and diseases. *Semin. Arthritis Rheum.* 43, 220–240. doi: 10.1016/j.semarthrit.2013.01.004

Brown, S. D., Twells, R. C. J., Hey, P. J., Cox, R. D., Levy, E. R., Soderman, A. R., et al. (1998). Isolation and Characterization of LRP6, a Novel Member of the Low Density Lipoprotein Receptor Gene Family. *Biochem. Bioph. Res. Commun.* 248, 879–888.

Carter, M., Chen, X., Slowinska, B., Minnerath, S., Glickstein, S., Shi, L., et al. (2005). Crooked tail (Cd) model of human folate-responsive neural tube defects is mutated in Wnt coreceptor lipoprotein receptor-related protein 6. *Proc. Natl. Acad. Sci. U. S. A.* 102, 12843–12848. doi: 10.1073/pnas.0501963102

Dhamo, B., Kuijpers, M. A. R., Balk-Leurs, I., Boxum, C., Wolvius, E. B., and Ongkosuwito, E. M. (2018). Disturbances of dental development distinguish patients with oligodontia-ectodermal dysplasia from isolated oligodontia. *Orthod. Craniofac. Res.* 21, 48–56. doi: 10.1111/ocr.12214

Dinckan, N., Du, R., Petty, L. E., Coban-Akdemir, Z., Jhangiani, S. N., Paine, I., et al. (2018). Whole-Exome Sequencing Identifies Novel Variants for Tooth Agenesis. *J. Dent. Res.* 97, 49–59. doi: 10.1177/0022034517724149

Duan, P., and Bonewald, L. F. (2016). The role of the wnt/beta-catenin signaling pathway in formation and maintenance of bone and teeth. *Int. J. Biochem. Cell Biol.* 77, 23–29. doi: 10.1016/j.biocel.2016.05.015

Huybrechts, Y., Mortier, G., Boudin, E., and Van Hul, W. (2020). WNT Signaling and Bone: lessons From Skeletal Dysplasias and Disorders. *Front. Endocrinol.* 11:165. doi: 10.3389/fendo.2020.00165

Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., et al. (2016). REVEL: an Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am. J. Hum. Genet.* 99, 877–885. doi: 10.1016/j.ajhg.2016.08.016

Kokubu, C., Heinzmann, U., Kokubu, T., Sakai, N., Kubota, T., Kawai, M., et al. (2004). Skeletal defects in ringelschwanz mutant mice reveal that Lrp6 is required for proper somitogenesis and osteogenesis. *Development* 131, 5469–5480. doi: 10.1242/dev.01405

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., et al. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* 102, 3586–3616. doi: 10.1021/jp973084f

Mani, A., Radhakrishnan, J., Wang, H., Mani, A., Mani, M. A., Nelson-Williams, C., et al. (2007). LRP6 mutation in a family with early coronary disease and metabolic risk factors. *Science* 315, 1278–1282. doi: 10.1126/science.1136370

Massink, M. P., Creton, M. A., Spanevello, F., Fennis, W. M., Cune, M. S., Savelberg, S. M., et al. (2015). Loss-of-Function Mutations in the WNT Co-receptor LRP6 Cause Autosomal-Dominant Oligodontia. *Am. J. Hum. Genet.* 97, 621–626. doi: 10.1016/j.ajhg.2015.08.014

Matalova, E., Fleischmannova, J., Sharpe, P. T., and Tucker, A. S. (2008). Tooth Agenesis: from Molecular Genetics to Molecular Dentistry. *J. Dent. Res.* 87, 617–623. doi: 10.1177/154405910808700715

Ockeloen, C. W., Khandelwal, K. D., Dreesen, K., Ludwig, K. U., Sullivan, R., van Rooij, I., et al. (2016). Novel mutations in LRP6 highlight the role of WNT signaling in tooth agenesis. *Genet. Med.* 18, 1158–1162. doi: 10.1038/gim.2016.10

Pan, W., Choi, S.-C., Wang, H., Qin, Y., Volpicelli-Daley, L., Swan, L., et al. (2008). Wnt3a-Mediated Formation of Phosphatidylinositol 4,5-Bisphosphate Regulates LRP6 Phosphorylation. *Science* 321, 1350–1353.

Phillips, J. C., Hardy, D. J., Maia, J. D. C., Stone, J. E., Ribeiro, J. V., Bernardi, R. C., et al. (2020). Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* 153:044130. doi: 10.1063/5.0014475

Pinson, K. I., Brennan, J., Monkley, S., Avery, B. J., and Skarnes, W. C. (2000). An LDL-receptor-related protein mediates Wnt signaling in mice. *Nature* 407, 535–538.

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

Šali, A., Potterton, L., Yuan, F., van Vlijmen, H., and Karplus, M. (1995). Evaluation of comparative protein modeling by MODELLER. *Proteins* 23, 318–326. doi: 10.1002/prot.340230306

Singh, R., Smith, E., Fathzadeh, M., Liu, W., Go, G. W., Subrahmanyan, L., et al. (2013). Rare nonconservative LRP6 mutations are associated with metabolic syndrome. *Hum. Mutat.* 34, 1221–1225. doi: 10.1002/humu.22360

Tamal, K., Semenov, M., Kato, Y., Spokony, R., Liu, C., Katsuyama, Y., et al. (2000). LDL-receptor-related protein in Wnt signal transduction. *Nature* 407, 530–535.

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from next-generation sequencing data. *Nucleic Acids Res.* 38:e164.

Yang, K., Shen, M., Yan, Y., Tan, Y., Zhang, J., Wu, J., et al. (2019). Genetic Analysis in Fetal Skeletal Dysplasias by Trio Whole-Exome Sequencing. *BioMed. Res. Int.* 2019, 1–8. doi: 10.1155/2019/2492590

Yin, W., and Bian, Z. (2015). The Gene Network Underlying Hypodontia. *J. Dent. Res.* 94, 878–885. doi: 10.1177/0022034515583999

Yu, M., Fan, Z., Wong, S. W., Sun, K., Zhang, L., Liu, H., et al. (2020). Lrp6 Dynamic Expression in Tooth Development and Mutations in Oligodontia. *J. Dent. Res.* 100, 415–422. doi: 10.1177/0022034520970459

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: 19 July 2021
doi: 10.3389/fgene.2021.682050

Cohort Analysis of 67 Charcot-Marie-Tooth Italian Patients: Identification of New Mutations and Broadening of Phenotype Expression Produced by Rare Variants

*Rosangela Ferese[1], Rosa Campopiano[1], Simona Scala[1], Carmelo D'Alessio[1], Marianna Storto[1], Fabio Buttari[1], Diego Centonze[1,2], Giancarlo Logroscino[3,4], Chiara Zecca[3], Stefania Zampatti[1,5], Francesco Fornai[1,6], Vittoria Cianci[7], Elisabetta Manfroi[8], Emiliano Giardina[5,9], Mauro Magnani[10], Antonio Suppa[1,11], Giuseppe Novelli[1,9] and Stefano Gambardella[1,10]**

[1]*IRCCS Neuromed, Pozzilli, Italy,* [2]*Laboratory of Synaptic Immunopathology, Department of Systems Medicine, Tor Vergata University, Rome, Italy,* [3]*Center for Neurodegenerative Diseases and the Aging Brain, Department of Clinical Research in Neurology, The University of Bari "Aldo Moro," "Pia Fondazione Card G. Panico" Hospital Tricase, Lecce, Italy,* [4]*Department of Basic Medicine Neuroscience and Sense Organs, University "Aldo Moro" Bari, Bari, Italy,* [5]*Genomic Medicine Laboratory, IRCCS Fondazione Santa Lucia, Rome, Italy,* [6]*Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy,* [7]*Regional Epilepsy Centre, Great Metropolitan Hospital Bianchi-Melacrino-Morelli, Reggio Calabria, Italy,* [8]*Department of Neuroscience- Neurogenetics, Santa Maria Hospital, Terni, Italy,* [9]*Department of Biomedicine and Prevention, University of Rome "Tor Vergata," Rome, Italy,* [10]*Department of Biomolecular Sciences, University of Urbino "Carlo Bo," Urbino, Italy,* [11]*Department of Human Neurosciences, Sapienza University of Rome, Rome, Italy*

OPEN ACCESS

Edited by:
Xiu-An Yang,
Chengde Medical College, China

Reviewed by:
Rincic Martina,
University of Zagreb, Croatia
Raji Grewal,
Saint Francis Medical Center, United States

***Correspondence:**
Stefano Gambardella
stefano.gambardella@uniurb.it

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

Received: *17 March 2021*
Accepted: *17 June 2021*
Published: *19 July 2021*

Citation:
Ferese R, Campopiano R, Scala S, D'Alessio C, Storto M, Buttari F, Centonze D, Logroscino G, Zecca C, Zampatti S, Fornai F, Cianci V, Manfroi E, Giardina E, Magnani M, Suppa A, Novelli G and Gambardella S (2021) Cohort Analysis of 67 Charcot-Marie-Tooth Italian Patients: Identification of New Mutations and Broadening of Phenotype Expression Produced by Rare Variants. Front. Genet. 12:682050. doi: 10.3389/fgene.2021.682050

Charcot-Marie-Tooth (CMT) disease is the most prevalent inherited motor sensory neuropathy, which clusters a clinically and genetically heterogeneous group of disorders with more than 90 genes associated with different phenotypes. The goal of this study is to identify the genetic features in the recruited cohort of patients, highlighting the role of rare variants in the genotype-phenotype correlation. We enrolled 67 patients and applied a diagnostic protocol including multiple ligation-dependent probe amplification for copy number variation (CNV) detection of *PMP22* locus, and next-generation sequencing (NGS) for sequencing of 47 genes known to be associated with CMT and routinely screened in medical genetics. This approach allowed the identification of 26 patients carrying a whole gene CNV of *PMP22*. In the remaining 41 patients, NGS identified the causative variants in eight patients in the genes *HSPB1*, *MFN2*, *KIF1A*, *GDAP1*, *MTMR2*, *SH3TC2*, *KIF5A*, and *MPZ* (five new vs. three previously reported variants; three sporadic vs. five familial variants). Familial segregation analysis allowed to correctly interpret two variants, initially reported as "variants of uncertain significance" but re-classified as pathological. In this cohort is reported a patient carrying a novel familial mutation in the tail domain of *KIF5A* [a protein domain previously associated with familial amyotrophic lateral sclerosis (ALS)], and a CMT patient carrying a *HSPB1* mutation, previously reported in ALS. These data indicate that combined tools for gene association in medical genetics allow dissecting

unexpected phenotypes associated with previously known or unknown genotypes, thus broadening the phenotype expression produced by either pathogenic or undefined variants.

Clinical trial registration: ClinicalTrials.gov (NCT03084224).

Keywords: neurogenetics, Charcot-Marie-Tooth disease, multiple ligation dependent probe amplification, next-generation sequencing, diagnosis

INTRODUCTION

Charcot-Marie-Tooth (CMT) disease, also known as hereditary motor and sensory neuropathy, is a common, clinically heterogeneous group of inherited peripheral neuropathies with an estimated prevalence of one in 2,500 individuals (Wiszniewski et al., 2013; Di Vincenzo et al., 2014).

CMT is clinically, neurophysiologically, and genetically heterogeneous. It is most commonly characterized by sensory loss that starts in the lower limbs and progresses slowly in a length-dependent manner. This produces progressive distal muscle atrophy, weakness, distal sensory loss, foot deformities, and depressed tendon reflexes (Shy et al., 2005; Rossor et al., 2013).

The clinical classification is based on age at onset, distribution of muscle weakness, sensory loss, walking difficulties, and foot deformities (Rossor et al., 2013). Neurophysiology allows subdividing the disease into a demyelinating (CMT1) and axonal (CMT2) forms depending on whether the median motor nerve conduction velocity (NCV) is below or above 38 m/s, respectively. A third form, intermediate CMT, has both demyelinating and axonal features and NCV between 25 and 45 m/s (Reilly and Shy, 2009; Berciano et al., 2012).

The duplication of *PMP22* is the most common cause of CMT, with a prevalence up to 40% in some populations (Reilly and Shy, 2009; Abe et al., 2011; Saporta et al., 2011; Murphy et al., 2012; Østern et al., 2013; Sivera et al., 2013). Then, approximately 100 different genes have been linked to CMT-like phenotypes which are associated with related conditions involved in axonal transport, myelin structure, and membrane metabolism that have been found in multiple unrelated families or confirmed by functional studies (Saifi et al., 2003; Saporta et al., 2011; Azzedine et al., 2012; Vallat et al., 2013; IPNMDB, 2014; OMIM, 2014; Washington University, 2014).

Therefore, the large spectrum of genetically identifiable disease alleles complicates the molecular diagnosis. Genetic heterogeneity is associated with a wide spectrum of phenotypes, complicated by the fact that mutations in the same gene cause different phenotypes (Reilly and Shy, 2009; Azzedine et al., 2012; OMIM, 2014). Furthermore, sporadic cases of CMT are not uncommon due to autosomal recessive inheritance, reduced penetrance, late-onset, small family size, and *de novo* mutations (Braathen et al., 2011; Saporta et al., 2011; Vallat et al., 2013).

In this study, we enrolled 67 patients and applied a diagnostic protocol including multiple ligation-dependent probe amplification (MLPA) for copy number variation (CNV) detection of *PMP22* locus, and next-generation sequencing (NGS) for sequencing of 47 genes known to be associated with CMT and routinely screened in medical genetics.

The goal of this study is to identify the genetic features in the recruited cohort of patients, highlighting the role of rare variants in the genotype-phenotype correlation.

MATERIALS AND METHODS

Study Population

We collected blood or DNA samples from 67 unrelated patients with a clinical diagnosis of CMT from January 2013 to December 2019 at IRCCS Neuromed Institute (Italy). Genomic DNA was isolated from peripheral blood leukocytes according to standard procedures (QIAamp DNA Blood Mini Kit – QIAGEN).

Clinical Data

Diagnosis of CMT includes the presence of slowly progressive neuropathy with or without family history and after exclusion of other common causes of acquired neuropathy. CMT subtype was classified as CMT if both motor and sensory nerves were similarly affected, and dHMN or HSN if the neuropathy showed exclusive or predominant involvement of motor or sensory nerves, respectively. CMT patients were further subdivided into demyelinating CMT if conduction velocity of the nondominant median or ulnar nerve was $\leq$38 m/s and axonal or intermediate CMT if >38 m/s.

Literature Review

A systematic review of the literature was conducted to identify the detection rate of genetic variants and the clinical phenotype of CMT patients. Pubmed, Medline, and Embase database identified 23 cohort analysis studies consisting of Italian and European CMT patients in the period between 1997 and 2020 (**Supplementary Table S1**).

Multiple Ligation-Dependent Probe Amplification

The commercially available kit P405 (MRC-Holland, Amsterdam, Netherlands) was used for the multiplex dosage. This SALSA MLPA Probemix contains 42 MLPA probes with amplification products between 130 and 445 nt: 15 probes located in the 17p12 region (*PMP22* gene), two flanking probes, seven probes in the *MPZ* gene, five probes in the *GJB1* gene, 10 reference probes detecting autosomal chromosomes, and three probes on the X-chromosome. The MLPA was performed on DNA from patients, and four normal subjects were used as internal controls.

Next-Generation Sequencing Panel

The NGS analysis was performed using the Seq Cap EZ Choice Enrichment Kits (Hoffmann-La Roche, Basel) on an Illumina MiSeq (San Diego, CA). A full list of genes sequenced is provided in **Table 1**. All coding exons of the RefSeq transcripts of the genes and 15 base pairs of the flanking introns were targeted, except for *GJB1*, for which the target region is extended 860 bases upstream of the ATG start codon to include the nerve-specific promoter region. 99% of the coding exons were sequenced with a minimal read depth of 30X.

GenomeUp software[1] was used for data analysis. It provides automated annotation (Best Practices workflows of GATK v4.1 for germline variant calling), alignment of sequence reads to the reference genome GRCh37/hg19, and selection of potentially pathogenic variants. Direct evaluation of data sequence was performed by the Integrative Genomics Viewer v.2.3. Mutation re-sequencing and segregation analysis were performed by the Sanger sequencing ABI 3130xl Genetic Analyzer (Applied Biosystems).

Data Analysis and Variant Interpretation

Variants were classified as pathogenic (class 5), likely pathogenic (class 4), and variants of uncertain significance (VoUS; class 3) according to American College of Medical Genetics Guideline for germline variant classification (Li et al., 2017). To this aim, public databases were used (VarSome https://varsome.com; GnomAD https://gnomad.broadinstitute.org). In silico analyses of variants were performed using SIFT,[2] PolyPhen2,[3] PROVEAN,[4] and Mutation Assessor.[5] The new mutations identified have been submitted in ClinVar database.[6]

RESULTS

Multiple Ligation-Dependent Probe Amplification Analysis

The variation of the whole *PMP22* gene CNV was assessed through MLPA, confirming the clinical diagnosis in 26 out of 67 patients (38.8%). In detail, 20/67 (30%) were carriers of the heterozygote whole gene duplication, thus confirming the clinical diagnosis of CMT1A; 5/67 (7.4%) were carriers of the heterozygote whole gene deletion, confirming HNPP clinical diagnosis; and one was a carrier of a rare whole gene mosaic duplication of 1.5-Mb in heterozygous in *PMP22* gene (MIM 601097), considered as pathogenic (Rautenstrauss et al., 1998) and thus responsible for CMT disease 1A, autosomal dominant (MIM 118220). (Family ID 564, II:1; **Figure 1A**), classified as sporadic on anamnesis.

[1]https://platform.genomeup.com
[2]http://sift.jcvi.org
[3]http://genetics.bwh.harvard.edu/pph2
[4]http://provean.jcvi.org/index.php
[5]http://mutationassessor.org
[6]https://www.ncbi.nlm.nih.gov/clinvar/

Next-Generation Sequencing Analysis

The remaining 41 patients were tested by NGS using a target panel that considers 49 genes associated with CMT (**Table 1**). This approach identified the causative variants (pathogenic or likely pathogenic), in 8/41 patients (19.5%; **Figure 2A**). Thus, NGS improved the detection rate to 50.8% (38.8% MLPA + 12% NGS). The onset of all patients with a genetic diagnosis ranges from 7 to 57 years old, and the clinical features, consistent with phenotypes reported in OMIM database, are summarized in **Table 2**.

These variants fall in eight different genes: *MFN2*, *MPZ*, *GDAP1*, *SH3TC2*, *HSPB1*, *KIF5A*, *MTMR2*, *KIF1A*, responsible for demyelinating CMT (5/64, 8%), and axonal or intermediate CMT (3/64, 4.7%; **Table 2**).

Of these variants, three have been previously reported as causative of CMT (Families ID 896, 125, 402; Warner et al., 1996; Capponi et al., 2016; Piscosquito et al., 2016), five are new variants (**Figure 1**).

In six out of eight probands, familiar members were available for segregation analysis. Of these, four variants were already classified as pathological (Families ID 882, 1141, 721, 961), while two patients were carriers of variants classified as VoUS, but reclassified as pathological through data obtained from familiar segregation (Families ID 580, 184; **Figure 1**). The variants identified have been classified as sporadic (3/8, 37.5%) or familiar (5/8, 62.5%).

Causative Variants Identified as Sporadic

Family ID 882

The proband, a 45 years old female (II:1; **Figure 1B**), onset at 7 years old, presented with a moderate hyposthenia, axonal neuropathy, and normal conduction velocity (>45 m/s). She is a carrier of the new heterozygous stop mutation: NM_001127660.1:c.[2258dupT], NP_001121132.1:p.(Gln754AlafsTer9; rs773371488) in mitofusin-2 (*MFN2*; MIM 608507) considered as pathogenic and thus responsible for CMT disease 2A2A and autosomal dominant (MIM 609260). The variant was considered as sporadic since it is not present in the mother (I:2), and her father (I:1) is reported as neurological healthy.

Family ID 896

The proband, a 45 years old male (II:1; **Figure 1C**), onset at 20 years old, presented with axonal and demyelinating neuropathy with a very slow conduction velocity (>15 m/s). Molecular analysis identified the known heterozygous stop mutation: NM_000530.8: c.[306delA], NP_000521: p.(Asp104ThrfsTer14), (rs281865125) in *MPZ* (MIM 159440) considered as pathogenic (Warner et al., 1996) and thus responsible for CMT disease-dominant intermediate D (MIM 607791). The variant was considered sporadic. Family members were not available for testing, but both his mother I:2 and father I:1 were reported as neurological healthy before the age of 20 years old.

Family ID 402

The proband, a 76 years old female (II:1; **Figure 1D**), onset at 56, presented with progressive ankle instability and gait

TABLE 1 | Target genes included in NGS Panel.

Gene	Ref sequence	MIM
AARS	NM_001605.2	601065
ALT1	NM_005309	138200
ARHGEF10	NM_001308152	608236
ATP7A	NM_000052.6	300011
BSCL2	NM_001122955.3	606158
CCT5	NM_012073.5	610150
DMN2	NM_001005360.2	602378
DYNC1H1	NM_001376.4	600112
EGR2	NM_000399.3	129010
FGD4	NM_139241.2	611104
FIUREG4	NM_014845.5	609390
GARS	NM_002047.2	600287
GDAP1	NM_018972.2	606598
GJ B1	NM_000166.5	304040
HSPB1	NM_001540.3	602195
HSPB8	NM_014365.2	608014
IFRD1	NM_003640.5	603502
IKBKAP	NM_001197080.1	603722
KIF1A	NM_001244008.2	601255
KIF1B	NM_015074.3	605995
KIF5A	NM_004984.2	602821
LITAF	NM_004862.3	603795
LMNA	NM_170707.2	150330
LRSAM1	NM_138361.5	610933
MED25	NM_030973.3	610197
MFN2	NM_014874.3	608507
MPZ	NM_000530.6	159440
MTMR2	NM_016156.5	603557
NDRG1	NM_001135242.1	605262
NEFL	NM_006158	162280
PMP22	NM_000304.2	601097
POLG	NM_001126131	174763
PRPS1	NM_002764.3	311850
PRX	NM_181882	605725
RAB7A	NM_004637.5	602298
REEP1	NM_004637	602298
SBF2	NM_030962.3	607697
SEPT9	NM_001113491	604061
SETX	NM_00135152	608465
SH3TC2	NM_024577.3	608206
SLC12A6	NM_00104497.2	604878
SMAD1	NM_005900.3	601595
SOD1	NM_00045	147450
SP110	NM_004509	604457
TDP1	NM_001008744.2	607198
SURF1	NNM_003172.4	185620
TGFB1	NM_000660	190180
TRPV4	NM_021625.4	605427
TTR	NM_0209556	176300

For each gene is reported: (i) name (GeneCards: The Human Gene Database), (ii) Ref sequence (NCBI Reference Sequence Database), and (iii) MIM (Gene/Locus MIM number).

difficulties with and foot drop. She had a diagnosis of axonal neuropathy with an intermediate conduction velocity (38 m/s). Molecular analysis identified the known heterozygous missense mutation: NM_001540.3:c.[570G > C], NP_001531.1:p.(Gln190His) in *HSPB1* (MIM 602195) considered as likely pathogenic (Capponi et al., 2016) and thus responsible for CMT disease, axonal, and type 2F autosomal dominant (MIM 606595). Family members were not available for testing.

Causative Variants Identified as Familiar

Family ID 184

The proband, a 67 years old male (II:1; **Figure 1E**), onset at 45 years old, presented with neuropathy, mild hyposthenia of the distal musculature, and hearing loss. He showed a slow conduction velocity (25 m/s) and signs of pyramidal tract dysfunction. Molecular analysis identified a new heterozygous missense mutation: NM_001244008.2:c.[5332C > T], NP_ 001230937:p.(Arg1778Trp), (rs765668490: C > T) in *KIF1A* (MIM 601255), initially considered as VoUS, but classified as likely pathogenic because of its presence in his affected brother (II:2). This variant is responsible for neuropathy, hereditary sensory, type IIC, and autosomal recessive (MIM 614213).

Family ID 1196

The proband, 69 years old male (II:1; **Figure 1F**), onset at 57 years old, presented with axonal neuropathy. He showed a normal conduction velocity. Molecular analysis identified a new heterozygous stop mutation: NM_018972.2: c.[140delA], NP_061845.2: p.(Lys47ArgfsTer3) in *GADP1* (MIM 606598) considered as pathogenic and thus responsible for CMT disease axonal, autosomal dominant, and type 2K (MIM 607831). His sister (II:9) has a similar condition but was not available for testing.

Family ID 125

The proband, a 45 years old male (II:1; **Figure 1G**), onset at 20 years old, presented with early onset demyelinating neuropathy with slow conduction velocity (18 m/s). Molecular analysis identified the known homozygous splicing mutations: NM_024577.3: c.[805 + 2T > C], (rs139052887T > C) in *SH3TC2* (MIM 608206) considered as pathogenic (Piscosquito et al., 2016) and thus responsible for CMT disease, autosomal recessive, and type 4C (MIM 601596). Both parents (I:1 and I:2) are heterozygote healthy carriers of this variant.

Family ID 1141

The proband a 60 years old male (II:1; **Figure 1H**), onset at 40 years old, presented with neuropathy and spastic paraplegia. He showed an intermediate conduction velocity (40 m/s). Molecular analysis identified a new heterozygous deletion mutation: NM_004984.2:c.[2868_2870delTCT], NP_004975.2:p.(Leu957del), (rs575223790) in *KIF5A* (MIM 602821) considered as likely pathogenic and thus responsible for spastic paraplegia 10 with neuropathy and autosomal dominant (MIM 604187). This variant is familiar since it is present in his affected sister (II:2) and absent in his healthy sister (II:4).

Family ID 580

The proband is a 30 years old female (II:1; **Figure 1I**), onset at 22, presented with axonal neuropathy and an intermediate conduction velocity (<35 m/s). Molecular analysis identified the new homozygous missense mutation: NM_016156.5:c.[463T > C], NP_057240.3:p.(Cys155Arg) in *MTMR2* (MIM 603557) initially considered as VoUS, but classified as likely

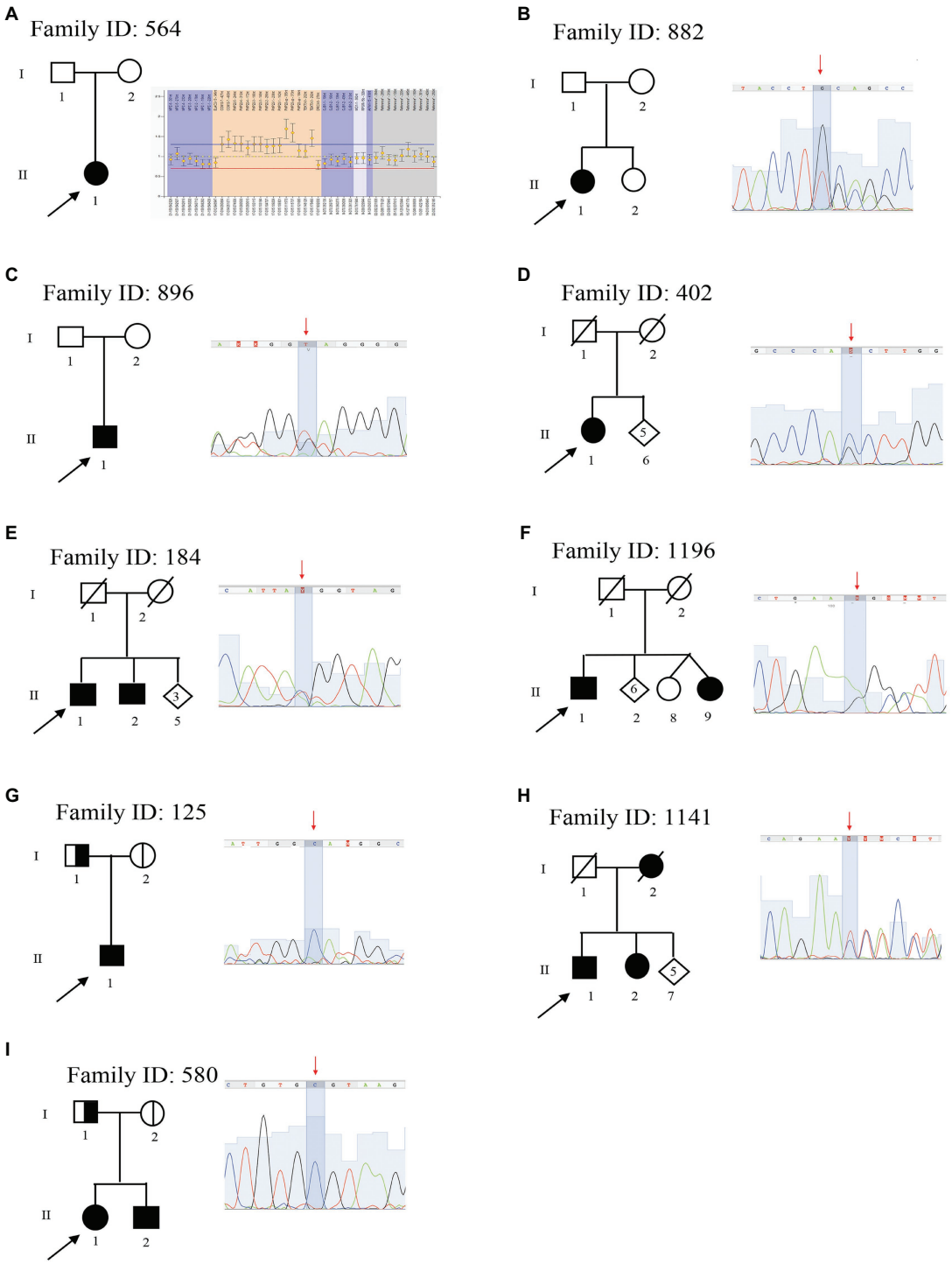


FIGURE 1 | Pedigree and electropherograms of class 4 and 5 mutations identified by next-generation sequencing (NGS). **(A)** Family ID: 564, MPLA analysis and identified a mosaic duplication of 1.5-Mb (17p11.2–12); **(B)** Family ID: 882, heterozygous stop mutation in mitofusin-2 (*MFN2*) gene: NM_001127660.1:c.[2258dupT], NP_001121132.1:p.(Gln754AlafsTer9; rs773371488); **(C)** Family ID: 896, heterozygous stop mutation in *MPZ* gene: NM_000530.8: c.[306delA], NP_000521: p.(Asp104ThrfsTer14), (rs281865125); **(D)** Family ID: 402, heterozygous missense mutation in *HSPB1* gene: NM_001540.3:[c.570G > C], NP_001531.1:p.(Gln190His); **(E)** Family ID: 184, heterozygous missense mutation in *KIF1A* gene: NM_001244008.2:c.[5332C > T], NP_ p.(Arg1778Trp), (rs765668490); **(F)** Family ID: 1196, heterozygous stop mutation in *GADP1* gene: NM_018972.2: c.[140delA], NP_061845.2: p.(Lys47ArgfsTer3); **(G)** Family ID: 1251, homozygous splicing mutations in *SH3TC2* gene: NM_024577.3: c.[805 + 2T > C], (rs139052887); **(H)** Family ID: 1141, heterozygous deletion mutation in *KIF5A* gene: NM_004984.2:c.[2868_2870delTCT], NP_004975.2:p.(Leu957del), (rs575223790); and **(I)** Family ID: 580, homozygous missense mutation in *MTMR2* gene: NM_016156.5:c.[463T > C], NP_057240.3:p.(Cys155Arg).

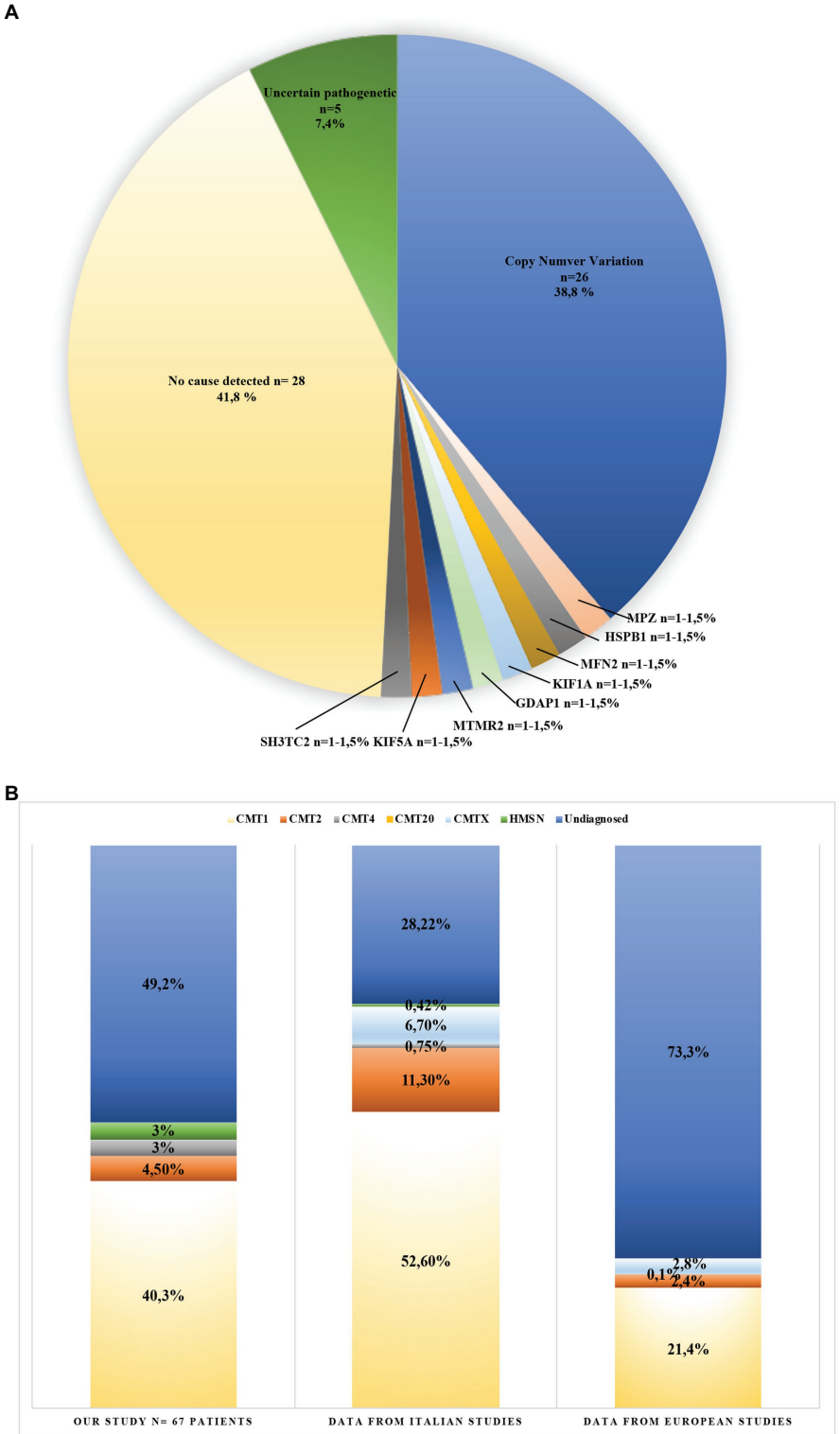


FIGURE 2 | Genotype and phenotype of Charcot-Marie-Tooth (CMT) patients from this study **(A)**, and CMT patients reviewed from the literature **(B)**. **(A)** Variants identified in 67 CMT patients (blue: copy number variation 38.8%; orange: *MPZ* mutation 1.5%; gray: *HSPB1* mutation 1.5%; light yellow: *MFN2* mutation 1.5%; light blue: *KIF1A* mutation 1.5%; green: *GAPD1* mutation 1.5%; dark blue: *MTMR2* mutation 1.5%; brown: *KIF5A* mutation 1.5%; dark gray: *SH3TC2* mutation 1.5%; yellow: no cause detected 41.8%; and dark green: uncertain pathogenetic 7.4%). **(B)** This graphic compares detection yield in CMT phenotypes obtained from the literature. (light yellow: CMT1; orange: CMT2; gray: CMT4; light blue: CMTX; green: HMSN; and blue: undiagnosed). Bar 1: data obtained in this study; Bar 2: data obtained from three Italian studies; and Bar 3: data obtained from 20 European studies (see **Supplementary Material**).

TABLE 2 | Genetic and clinical data in the cohort of patients analyzed. Clinical (A) and genetic (B) features of the 15 CMT patients (16 variants) identified by NGS.

FAMID	GENE	HGVSc	HGVSp	dpSNP ID	GENOTYPE	PHENOTYPE MIM	ACMG	Clin Var	ACMG	AY	AO	Family history	CMT subtype	CV	CV subtype	Reference
882	*MFN2*	NM_001127660.1: c.[2258dupT]	NP_001121132.1: p.(Gln754AlafsTer9)	rs773371488: dupT	Heterozygous	609260	5	SCV001424047	PVS1-PM1-PM2-PP3	45	7	NO	2	>45	Normal	Our study
896	*MPZ*	NM_000530.8: c.[306delA]	NP_000521: p.(Asp104ThrfsTer14)	rs281865125 delA	Heterozygous	607791	5	SCV000928852	PVS1-PM1-PM2	45	20	NO	1 + 2	>15	Very slow	Warner et al., 1996
1196	*GDAP1*	NM_018972.2: c.[140delA]	NP_061845.2: p.(Lys47ArgfsTer3)		Heterozygous	607831	5	SCV001424519	PVS1-PM1-PM2	69	57	YES	2	>45	Normal	Our study
1251	*SH3TC2*	NM_024577.3: c.[805 + 2T > C]		rs139052887 T > C	Homozygous	601596	5	SCV001249582.3, SCV001388540.1	PVS1-PM2-PP3	45	20	YES	1	18	Slow	Piscosquito et al., 2016
564	*PMP22*	NM_153321: g.dup(17)p12			Heterozygous	118220	5	SCV001424528	PS3-PM1-PM2-PP3-PP4	44	41	NO	1	<35	Slow	Rautenstrauss et al., 1998
402	*HSPB1*	NM_001540.3: c.[570G > C]	NP_001531.1: p.(Gln190His)		Heterozygous	606595	4	SCV001424520	PM1-PM2-PP2-PP3	80	56	NO	2	38	Intermediate	Capponi et al., 2016
1141	*KIF5A*	NM_004984.2: c.[2868_2870delTCT]	NP_004975.2: p.(Leu957del)	rs575223790 del TCT	Heterozygous	604187	4	SCV001424521	PM1-PM2-PM4-PP3	60	40	YES	5	40	Intermediate	Our study
580	*MTMR2*	NM_016156.5: c.[463T > C]	NP_057240.3: p.(Cys155Arg)		Heterozygous	601382	4	SCV001424522	PM2-PM3-PP1-PP3-PP4	30	22	YES	2	<35	Intermediate	Our study
184	*KIF1A*	NM_001244008.2: c.[5332C > T]	NP_001230937: p.(Arg1778Trp)	rs765668490: C > T	Heterozygous	614213	4	SCV000952011.2	PM2-PP2-PP3-PP4	67	45	NO	5	25	Slow	Our study
721	*AGRN*	NM_198576.4: c.[5851C > T]	NP_940978.2: p.(Arg1951Cys)	rs746117937: C > T	Heterozygous	615120	3	SCV001377285.1	PM2-PP3-BP1	59	55	NO	2	>15	Very slow	Our study
721	*SCP2*	NM_002979.5: c.[886C > T]	NP_001317516.1: p.(Pro296Ser)		Heterozygous	613724	3	SCV001424523	PM2-PP3	59	55	NO	2			Our study
856	*DNM2*	NM_001005360.2: c.[890G > A]	NP_001005360.1: p.(Arg297His)	rs763894364 G > A	Heterozygous	6482	3	SCV000762731.1	PM2-PP2-PP3	67	NA	NO	1 + 2	<35	Slow	Our study
608	*MED25*	NM_030973.3: c.[949G > T]	NP_112235.2: p.(Gly317Cys)	rs1280659782 G > T	Heterozygous	605589	3	SCV001424524	PM2-PP3-BP1	53	50	NO	2	<35	Intermediate	Our study
962	*DYNC1H1*	NM_001376.4: c.[9919G > T]	NP_001367.2: p.(Val3307Leu)		Heterozygous	614228	3	SCV001424525	PM1-PM2-BP4	36	34	NO	1	20	Slow	Our study
731	*DYNC1H1*	NM_001376.4: c.[6743A > G]	NP_001367.2: p.(Glu2248Gly)		Heterozygous	614228	3	SCV001424526	PM2-PP3	63	NA		1	<35	Slow	Our study
1261	*BSCL2*	NM_001122955.3: c.[124C > T]	NP_001116427.1: p.(Arg42Cys)	rs201493373 C > T	Heterozygous	600794	3	SCV001148309.4	PP2-BS1	48	42	NO	1	>15	Very slow	Our study

FAM ID, Family identification; HGVSc, Nomenclature human genome variation society coding DNA; HGVSp, Nomenclature human genome variation society protein; dbSNP ID, Single nucleotide polymorphism database; ACMG, American college of medical genetics guideline; AY, Years at evaluation; AO, Age at onset; CMT subtype, Demyelinating CMT=1, Axonal CMT =2, dHMN=3, HSN=4, and Possible CMT=5; CV, Conduction velocity m/s; CV subtype, CMT subtype classification for conduction velocity; PVS1, Null variant (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single or multiexon deletion) in a gene where LOF is a known mechanism of disease (pathogenic and very strong); PM1, Located in a mutational hot spot and/or critical and well-established functional domain (e.g., active site of an enzyme) without benign variation (pathogenic and moderate); PM2, Absent from controls (or at extremely low frequency if recessive) in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium (pathogenic and moderate); PM3, For recessive disorders, detected in trans with a pathogenic variant (pathogenic and moderate); PM4, Protein length changes as a result of in-frame deletions/insertions in a non-repeat region or stop-loss variants (pathogenic and moderate); PP1, Cosegregation with disease in multiple affected family members in a gene definitively known to cause the disease (pathogenic and supporting); PP2, Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease Pathogenic and supporting); PP3, Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.; pathogenic and supporting); PP4, Patient's phenotype or family history is highly specific for a disease with a single genetic etiology (pathogenic and supporting); BP1, Missense variant in a gene for which primarily truncating variants are known to cause disease (benign and supporting); BP4, Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.; benign and supporting).

pathogenic because of its presence in homozygous in her affected brother (II:2), and in heterozygous in her parents (I:1 and I:2). Thus, this variant is responsible for CMT disease, type 4B1, and autosomal recessive (MIM 601382).

Variants Identified as VoUs

Seven VoUS were found in 6/67 patients (7%; **Table 2**; **Figure 2A**). Heterozygous variants identified in *AGRN*, *SCP2*, *DNM2*, *MED25*, *DYNC1H1*, and *BSCL2* genes associated with autosomal dominant and recessive CMT.

Family ID 721

The proband a 59 years old male (II:1; **Figure 3A**), onset at 55 years old, presented with axonal neuropathy. He showed a very slow conduction velocity (>15 m/s). He had clear psychic slowness, cognitive impairment, and floating paresthesia in the upper limbs. Molecular analysis identified two new heterozygous missense mutations: (a) in *AGRN* (MIM 103320) NM_198576.4:c.[5851C > T], NP_940978.2: p.(Arg1951Cys), (rs746117937) responsible for myasthenic syndrome, congenital, 8, with pre- and postsynaptic defects, autosomal recessive (MIM 615120), and *SCP2* (MIM 184755) NM_002979.5 c.[886C > T] and NP_001317516.1:p.(Pro296Ser) responsible for leukoencephalopathy with dystonia and motor neuropathy, autosomal recessive (MIM 613724). Both variants are not present in the healthy sister (II:2), but these data are not sufficient to consider this variant as likely pathogenic.

Family ID 856

The proband, a 67 years old female (II:1; **Figure 3B**), presented with axonal neuropathy. She showed a slow conduction velocity (<35 m/s). Molecular analysis identified a new heterozygous missense mutation: NM_001005360.2:c.[890G > A], NP_001005360.1:p.(Arg297His; rs763894364) in *DNM2* (MIM 602378) responsible for CMT disease, axonal type 2M, and autosomal dominant (MIM 606482). Family members were not available for testing.

Family ID 608

The proband, a 53 years old female (II:1; **Figure 3C**), onset at 50, presented with axonal neuropathy. She showed an intermediate conduction velocity (<35 m/s). Molecular analysis identified a new heterozygous missense mutation: NM_030973.3;c.[949G > T], NP_112235.2: p.(Gly317Cys; rs1280659782 G > T) in *MED25* (MIM 610197) responsible for CMT disease, type 2B2, and autosomal recessive (MIM 605589). Family members were not available for testing.

Family ID 962

The proband, 36 years old male (II:1; **Figure 3D**), onset at 34, presented with demyelinating neuropathy. He showed a slow conduction velocity (20 m/s). Molecular analysis identified a new heterozygous missense mutation: NM_001376.4:c.[9919G > T], NP_001367.2:p.(Val3307Leu) in *DYNC1H1* (MIM 600112) responsible for CMT disease, type 20, and autosomal dominant (MIM 614228). Family members were not available for testing.

Family ID 731

The proband, 63 years old female (II:1; **Figure 3E**), presented with axonal neuropathy. She showed a slow conduction velocity (<35 m/s). Molecular analysis identified a new heterozygous missense mutation: NM_001376.4:c.[6743A > G], NP_001367.2:p.(Glu2248Gly) in *DYNC1H1* (MIM 600112) responsible for CMT disease, type 20, and autosomal dominant (MIM 614228). Family members were not available for testing.

Family ID 1261

The proband, 48 years old male (II:1; **Figure 3F**), onset a 42 years old, presented with demyelinating neuropathy. He showed a very slow conduction velocity (>15 m/s) and hollow foot. Molecular analysis identified a new heterozygous missense mutation: NM_001122955.3:c.[124C > T], NP_001116427.1:p.(Arg42Cys); (rs201493373) in *BSCL2* (MIM 606158), responsible for neuropathy, distal hereditary motor, and type VA (MIM 600794) autosomal dominant. Family members were not available for testing.

DISCUSSION

In this study, 67 patients with a clinical diagnosis of CMT were selected for molecular analysis of genes known to be related to the disease. Although the genetic of CMT is featured by a strong genetic heterogeneity, most of the patients carry 17p CNV of *PMP22*. Among this CNV, almost all cases carry the whole gene duplication of *PMP22*. The prevalence of the duplication varies among populations, ranging from 15% in Norwegian studies, up to more than 50% in Italy and Spain (Abe et al., 2011; Saporta et al., 2011; Murphy et al., 2012; Østern et al., 2013; Sivera et al., 2013; Høyer et al., 2014). In this study, only 30% of Italian patients (20/67) carry the whole gene duplication of *PMP22*, thus under-representing other Italian studies though still being in line with prevalence measured in European studies (**Figure 2B**).

When ruling out carriers of 17p CNV in *PMP22*, NGS identified the causative variants in 8/41 patients, which correspond only to 12% of the patients analyzed in this study (8 out of 67), being over 50% of the total of causative variants confirming the clinical diagnosis detected in the present analysis. Thus, combining various genetic tools using a target panel of 49 genes allowed to cover to a greater extent the complex clinical setting.

Several cohort studies tried to identify the diagnostic yield in CMT patients. To this aim, we conducted a systematic review of 23 cohort analysis studies consisting of Italian and European CMT patients in the period between 1997 and 2020. This data, summarized in **Figure 2B** and **Supplementary data**, identified a diagnostic yield of 71.7% in Italy (Mostacciuolo et al., 2001; Manganelli et al., 2014; Gentile et al., 2020) and 26.7% in Europe (Bort et al., 1997; Saporta et al., 2011; Murphy et al., 2012; Gess et al., 2013; Østern et al., 2013; Sivera et al., 2013; Høyer et al., 2014; Antoniadi et al., 2015; Laššuthová et al., 2016; Lupo et al., 2016; Dohrn et al., 2017; Marttila et al., 2017; Bacquet et al., 2018; Hoebeke et al., 2018; Milley et al., 2018; Nicolas et al., 2018; Lerat et al., 2019; Vaeth et al., 2019;

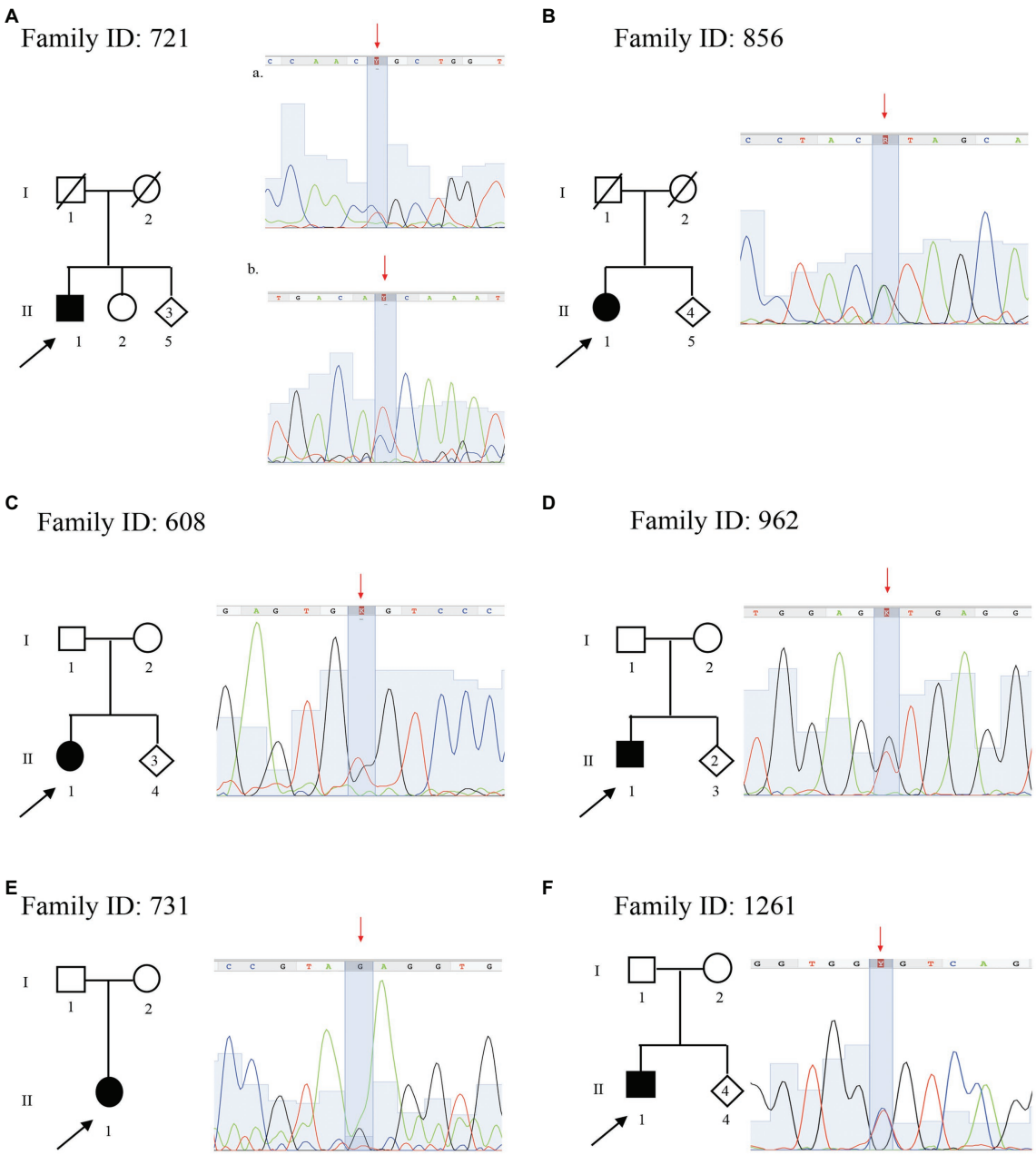


FIGURE 3 | Pedigree and electropherograms of variants of uncertain significance identified by NGS. **(A)** Family ID: 721, two missense variants in (a) *AGRN* gene NM_198576.4:c.[5851C > T], NP_940978.2:p.(Arg1951Cys), (rs746117937) and (b) *SCP2* gene NM_002979.5 c.[886C > T], p.(Pro296Ser); **(B)** Family ID: 856, missense variant in *DNM2* gene NM_001005360.2:c.[890G > A], NP_001005360.1:p.(Arg297His); (rs763894364); **(C)** Family ID: 608, missense variant in *MED25* gene NM_030973.3:c.[949G > T], NP_112235.2: p.(Gly317Cys); (rs1280659782); **(D)** Family ID: 962, N missense variant in *DYNC1H1* gene NM_001376.4:c.[9919G > T], NP_001367.2:p.(Val3307Leu); **(E)** Family ID: 731, missense variant in *DYNC1H1* gene NM_001376.4:c.[6743A > G], NP_001367.2:p.(Glu2248Gly); and **(F)** Family ID: 1261, missense variant in *BSCL2* gene NM_001122955.3:c.[124C > T], NP_001116427.1:p.(Arg42Cys); (rs201493373).

Cortese et al., 2020; **Figure 2B**). Therefore, the prevalence of genetic cases of CMT provided by the present data represents an average compared with the general prevalence measured in Italy and Europe. The differences in the diagnostic rate may be explained by the differences in the specific features of the cohorts being analyzed, the number of demyelinating CMT cases enrolled, and the heterogeneous inclusion criteria considering the common causative, which were genes adopted by the previous MLPA and Sanger sequencing.

In our cohort, NGS approach identified variants in eight different genes: *MFN2*, *MPZ*, *GDAP1*, *SH3TC2*, *HSPB1*, *KIF5A*, *MTMR2*, and *KIF1A* (**Table 2**), five responsible for demyelinating CMT, and three for axonal or intermediate CMT. These NGS-identified variants appeared as sporadic (3/8) or familial (5/8).

A frequency of sporadic or *de novo* mutations is reported in 28–34% (Marques et al., 2005; Ando et al., 2017), although this percentage may vary and some studies report up to a half of the patients as sporadic (Ando et al., 2017). In this study, variants considered as sporadic are reported in three out of eight patients with a NGS genetic diagnosis (37.5%). This percentage, in line with other studies, may represent an overestimation since this trend cannot be verified in a few families where a full familial anamnesis is lacking. This point is crucial since it demonstrates how a careful, multifaceted genetic approach with a complete familial anamnesis and genetic segregation analysis could provide a reduction of sporadic CMT.

Among sporadic variants, the proband Family ID 564 (**Figure 1A**) is a carrier of a whole gene mosaic in the *PMP22* gene. This is consistent with the high instability of the *PMP22* gene which undergoes *de novo PMP22* duplications in up to 90% of sporadic CMT1 (Boerkoel et al., 2002; Marques et al., 2005).

The other probands carry sporadic variants in the gene *MFN2*, *MPZ*, and *HSPB1*. Mutations in *MFN2* cause CMT type 2A by altering mitochondrial fusion and trafficking along with the axonal microtubule system (Guerriero et al., 2020). Pathogenic variants in *MFN2* are typically inherited as autosomal dominant and are usually missense variants (Di Vincenzo et al., 2014).

The sporadic variant identified, Family ID 883 (**Figure 1B**), consists of a single base insertion in the coiled-coil domains producing a stop mutation p.(Gln754AlafsTer9). This could be a *de novo* variant. *De novo MFN2* mutations are regularly found in patients with a classical CMT2 phenotype (Østern et al., 2013), and some *MFN2* mutations have also been reported as *de novo* in several patients (Züchner et al., 2004; Chung et al., 2006; Verhoeven et al., 2006; Cho et al., 2007). The early onset of the proband, at 7 years old, is in line with data reporting that mutations in *MPZ*, *MFN2*, or *NEFL* are the most frequent disease causes of patients with infantile-onset CMT (Hsu et al., 2019).

Thanks to NGS studies, several data are broadening the phenotype spectrum produced by mutations in some CMT genes. For example, mutations in the N-terminal motor domain of *KIF5A* are responsible for hereditary spastic paraplegia (*SPG10* MIM 604187) and CMT type 2 (CMT2), clustered in the switch regions SWI (199–204) and SWII (232–237) necessary for microtubules interaction (Filosto et al., 2018). On the other hand, mutations in the C-terminal cargo-binding tail domain (Tail domain) are related to a specific ALS phenotype (ASL MIM 617921), with an early disease onset but a longer survival compared with typical amyotrophic lateral sclerosis (ALS; Brenner et al., 2018; Nicolas et al., 2018). Beyond these main phenotypes, other complex and overlapping phenotypes are emerging, which demonstrates that *KIF5A* mutations may be responsible for a wide range and heterogeneous disease spectrum (Schüle et al., 2008; Tessa et al., 2008). For example, in a patient, a mixed slowly progressive disease was described resembling ALS as well as HSP and axonal neuropathy. This is caused by a mutation within the terminal region of the stalk domain (Cho et al., 2007). Additionally, the patient Family ID 1141 (**Figure 1H**) reported in this paper represents the first case of a familial mutation, the new heterozygous deletion p.Leu957del, falling in the tail domain of *KIF5A* producing spastic paraplegia 10 with autosomal dominant neuropathy (MIM 604187).

Taken together, these results broaden the phenotypic spectrum of *KIF5A* mutations and confirm the importance of cytoskeletal defects in the pathogenesis of both ALS and CMT.

In line with this, the phenotype related to heat shock protein 27 (HSP27) is broadening. For instance, mutations in the *HSPB1* gene have been reported to cause autosomal dominant CMT with minimal sensory involvement (CMT 2F MIM 606595). Then, two independent studies reported different *HSPB1* mutations in two sporadic cases and one consanguineous family with ALS, suggesting that the disease spectrum of *HSP27* may not be limited to CMT2/dHMN (Scarlato et al., 2015; Capponi et al., 2016; Adriaenssens et al., 2017). In this study, we identified the known heterozygous missense mutation p.Gln190H in Family ID 402 (**Figure 1D**), which falls in the C-terminal domain involved in the control of its chaperone-like activity, responsible for CMT, axonal, and type 2F autosomal dominant (MIM 606595; Lelj-Garolla and Mauk, 2012). This variant falls in the C-terminal domain involved in the control of chaperone-like activity (Lelj-Garolla and Mauk, 2012; Capponi et al., 2016). It has been previously reported in a sporadic ALS patient with onset at 58 yo, and axonal neuropathy with an intermediate conduction velocity (38 m/s; Capponi et al., 2016). Therefore, considering this study in which the same variant has been identified in a CMT patient, the genotype-phenotype correlation is very complex. Functional characterization of different mutations located in the C-terminal domain did not show a clear correlation between the location of the mutation and associated cellular phenotype. This suggested that the genomic variant, more than its location, determines the cell pathology and phenotype, However, the present study contradicts such a conclusion which is not consistent with the present data which indicate how the same substitution has been identified in patients owing to a different disease (ALS in one case CMT in another; Capponi et al., 2016). This suggests the importance of how specific genetic backgrounds may alter the expression of the genetic variants while explaining how rare variants, like VoUS may be responsible for the disease.

In this study, seven VoUS were identified in six (7%) patients. These variants, all in heterozygous status, all fall in one of those genes associated with autosomal dominant and recessive CMT: *AGRN*, *SCP2*, *DNM2*, *MED25*, *DYNC1H1*, and *BSCL2*. This high number of VoUS is in line with the previous studies reporting a higher number of VoUS in neuropathy-associated genes, including single mutations in autosomal recessive CMT genes compared with the control population, with some patients presenting more than one VoUS in different CMT genes (Gonzaga-Jauregui et al., 2015). For instance, here, we report multiple VoUS in the proband of Family ID 721 (**Figure 3A**) characterized by two VoUS in *AGRN* and *SCP2*. A recent hypothesis suggests that a combinatorial effect of rare variants contributes to the disease burden in CMT and partly explains its various phenotypes. This confirms what was demonstrated by using *in vivo*

experimental models in zebrafish (Cortese et al., 2020). Therefore, the interpretation of VoUS remains a key diagnostic challenge in the current NGS era.

CONCLUSION

This cohort analysis demonstrates the importance of combining different molecular approaches to identify the causative variant in CMT patients. The use of NGS target panel consisting of 49 genes identified the causative variants in eight patients, improving the detection rate to 50.8%. Although this seems to represent only a small fraction of patients, the identification of rare mutations allows dissecting of unexpected phenotypes associated with previously known or unknown genotypes, thus broadening the phenotype expression produced by variants. This is the example of a patient carrying a novel familial mutation in the tail domain of *KIF5A* (a protein domain previously associated with familial ALS), and a CMT patient carrying an *HSPB1* mutation, previously reported in ALS.

In this cohort, the higher frequency of VoUS identified is in line with the previous studies, confirming that the interpretation of VoUS remains a key diagnostic challenge in the current NGS era. In line with this, in this cohort, the segregation analysis allowed to correctly interpret two variants, initially reported as VoUS, but re-classified as pathological.

To improve the detection rate in CMT patients, whole-genome sequencing (WGS) and whole-exome sequencing (WES) are strongly required. While the use of WES is now accepted for diagnostic purposes, WGS is able to both identify novel and rare variants in coding as well as noncoding regions. Therefore, these approaches allow the identification of new genes and rare variants, thus improving the genetic detection rate.

DATA AVAILABILITY STATEMENT

The datasets for this article are not publicly available due to concerns regarding participant/patient anonymity. Requests to access the datasets should be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by IRCCS Neuromed Ethical Committees. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

CD, FB, DC, GL, CZ, SZ, AS, FF, VC, EM, EG, MM, and LS performed the recruitment and clinical evaluations of patients. RF, RC, and SS performed the genetic analyses. DC, FF, GN, MS, and SG supervised the work. RF and SG wrote the first draft of the manuscript. All authors read the manuscript, contributed to the manuscript revision, and approved the final version.

FUNDING

This work was supported by the Italian Ministry of Health (current research 2019–2023: Identification of New Variants and/or New Genes Responsible for Ataxia and Spastic Paraplegia), (5XMille – 2018).

ACKNOWLEDGMENTS

The authors are grateful to the patients and their relatives participating in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at https://www.frontiersin.org/articles/10.3389/fgene.2021.682050/full#supplementary-material

REFERENCES

Abe, A., Numakura, C., Kijima, K., Hayashi, M., Hashimoto, T., and Hayasaka, K. (2011). Molecular diagnosis and clinical onset of Charcot-Marie-Tooth disease in Japan. *J. Hum. Genet.* 56, 364–368. doi: 10.1038/jhg.2011.20

Adriaenssens, E., Geuens, T., Baets, J., Echaniz-Laguna, A., and Timmerman, V. (2017). Novel insights in the disease biology of mutant small heat shock proteins in neuromuscular diseases. *Brain* 140, 2541–2549. doi: 10.1093/brain/awx187

Ando, M., Hashiguchi, A., Okamoto, Y., Yoshimura, A., Hiramatsu, Y., Yuan, J., et al. (2017). Clinical and genetic diversities of Charcot-Marie-Tooth disease with MFN2 mutations in a large case study. *J. Peripher. Nerv. Syst.* 22, 191–199. doi: 10.1111/jns.12228

Antoniadi, T., Buxton, C., Dennis, G., Forrester, N., Smith, D., Lunt, P., et al. (2015). Application of targeted multi-gene panel testing for the diagnosis of inherited peripheral neuropathy provides a high diagnostic yield with unexpected phenotype-genotype variability. *BMC Med. Genet.* 16:84. doi: 10.1186/s12881-015-0224-8

Azzedine, H., Senderek, J., Rivolta, C., and Chrast, R. (2012). Molecular genetics of Charcot-Marie-Tooth disease: from genes to genomes. *Mol. Syndromol.* 3, 204–214. doi: 10.1159/000343487

Bacquet, J., Stojkovic, T., Boyer, A., Martini, N., Audic, F., Chabrol, B., et al. (2018). Molecular diagnosis of inherited peripheral neuropathies by targeted next-generation sequencing: molecular spectrum delineation. *BMJ Open* 8:e021632. doi: 10.1136/bmjopen-2018-021632

Berciano, J., Sevilla, T., Casasnovas, C., Sivera, R., Vílchez, J. J., Infante, J., et al. (2012). Illa I; Programa 3 (Enfermedades Neuromusculares) del Centro de Investigación Biomédica en Red de Enfermedades Neurodegenerativas (CIBERNED) del Instituto de Salud Carlos III. Guía diagnóstica en el paciente con enfermedad de Charcot-Marie-Tooth [Guidelines for molecular diagnosis of Charcot-Marie-Tooth disease]. *Neurologia* 27, 169–178. doi: 10.1016/j.nrl.2011.04.015

Boerkoel, C. F., Takashima, H., Garcia, C. A., Olney, R. K., Johnson, J., Berry, K., et al. (2002). Charcot-Marie-Tooth disease and related neuropathies: mutation distribution and genotype-phenotype correlation. *Ann. Neurol.* 51, 190–201. doi: 10.1002/ana.10089

Bort, S., Nelis, E., Timmerman, V., Sevilla, T., Cruz-Martínez, A., Martínez, F., et al. (1997). Mutational analysis of the MPZ, PMP22 and Cx32 genes in patients of Spanish ancestry with Charcot-Marie-Tooth disease and hereditary neuropathy with liability to pressure palsies. *Hum. Genet.* 99, 746–754. doi: 10.1007/s004390050442

Braathen, G. J., Sand, J. C., Lobato, A., Høyer, H., and Russell, M. B. (2011). Genetic epidemiology of Charcot-Marie-Tooth in the general population. *Eur. J. Neurol.* 18, 39–48. doi: 10.1111/j.1468-1331.2010.03037.x

Brenner, D., Yilmaz, R., Müller, K., Grehl, T., Petri, S., Meyer, T., et al. (2018). Hot-spot KIF5A mutations cause familial ALS. *Brain* 141, 688–697. doi: 10.1093/brain/awx370

Capponi, S., Geuens, T., Geroldi, A., Origone, P., Verdiani, S., Cichero, E., et al. (2016). Molecular chaperones in the pathogenesis of amyotrophic lateral sclerosis: the role of HSPB1. *Hum. Mutat.* 37, 1202–1208. doi: 10.1002/humu.23062

Cho, H. J., Sung, D. H., Kim, B. J., and Ki, C. S. (2007). Mitochondrial GTPase mitofusin 2 mutations in Korean patients with Charcot-Marie-Tooth neuropathy type 2. *Clin. Genet.* 71, 267–272. doi: 10.1111/j.1399-0004.2007.00763.x

Chung, K. W., Kim, S. B., Park, K. D., Choi, K. G., and Lee, J. H. (2006). Early onset severe and late-onset mild Charcot-Marie-Tooth disease with mitofusin 2 (MFN2) mutations. *Brain* 129, 2103–2118. doi: 10.1093/brain/awl174

Cortese, A., Wilcox, J. E., Polke, J. M., Poh, R., and Skorupinska, M. (2020). Targeted next-generation sequencing panels in the diagnosis of Charcot-Marie-Tooth disease. *Neurology* 94, e51–e61. doi: 10.1212/WNL.0000000000008672

Di Vincenzo, C., Elzinga, C. D., Medeiros, A. C., Karbassi, I., and Jones, J. R. (2014). The allelic spectrum of Charcot-Marie-Tooth disease in over 17,000 individuals with neuropathy. *Mol. Genet. Genomic. Med.* 2, 522–529. doi: 10.1002/mgg3.106

Dohrn, M. F., Glöckle, N., Mulahasanovic, L., Heller, C., Mohr, J., Bauer, C., et al. (2017). Frequent genes in rare diseases: panel-based next generation sequencing to disclose causal mutations in hereditary neuropathies. *J. Neurochem.* 143, 507–522. doi: 10.1111/jnc.14217

Filosto, M., Piccinelli, S. C., Palmieri, I., Necchini, N., and Valente, M. (2018). A novel mutation in the stalk domain of KIF5A causes a slowly progressive atypical motor syndrome. *J. Clin. Med.* 8:17. doi: 10.3390/jcm8010017

Gentile, L., Russo, M., Fabrizi, G. M., Taioli, F., Ferrarini, M., Testi, S., et al. (2020). Charcot-Marie-Tooth disease: experience from a large Italian tertiary neuromuscular center. *Neurol. Sci.* 41, 1239–1243. doi: 10.1007/s10072-019-04219-1

Gess, B., Schirmacher, A., Boentert, M., and Young, P. (2013). Charcot-Marie-Tooth disease: frequency of genetic subtypes in a German neuromuscular center population. *Neuromuscul. Disord.* 23, 647–651. doi: 10.1016/j.nmd.2013.05.005

Gonzaga-Jauregui, C., Harel, T., Gambin, T., Kousi, M., Griffin, L. B., Francescatto, L., et al. (2015). Exome sequence analysis suggests that genetic burden contributes to phenotypic variability and complex neuropathy. *Cell Rep.* 12, 1169–1183. doi: 10.1016/j.celrep.2015.07.023

Guerriero, S., D'Oria, F., Rossetti, G., Favale, R. A., Zoccolella, S., Alessio, G., et al. (2020). CMT2A harboring mitofusin 2 mutation with optic nerve atrophy and normal visual acuity. *Int. Med. Case Rep. J.* 13, 41–45. doi: 10.2147/IMCRJ.S237620

Hoebeke, C., Bonello-Palot, N., Audic, F., Boulay, C., Tufod, D., Attarian, S., et al. (2018). Retrospective study of 75 children with peripheral inherited neuropathy: genotype-phenotype correlations. *Arch. Pediatr.* 25, 452–458. doi: 10.1016/j.arcped.2018.09.006

Høyer, H., Braathen, G. J., Busk, Ø. L., Holla, Ø. L., Svendsen, M., Hilmarsen, H. T., et al. (2014). Genetic diagnosis of Charcot-Marie-Tooth disease in a population by next-generation sequencing. *Biomed. Res. Int.* 2014:210401. doi: 10.1155/2014/210401

Hsu, Y. H., Lin, K. P., Guo, Y. C., Tsai, Y. S., Liao, Y. C., and Lee, Y. C. (2019). Mutation spectrum of Charcot-Marie-Tooth disease among the Han Chinese in Taiwan. *Ann. Clin. Transl. Neurol.* 6, 1090–1101. doi: 10.1002/acn3.50797

IPNMDB (2014). The mutation database of inherited peripheral neuropathies. Available at: http://www.molgen.ua.ac.be/CMTMutations/Home/Default.cfm

Laššuthová, P., Šafka Brožková, D., Krůtová, M., Neupauerová, J., Haberlová, J., Mazanec, R., et al. (2016). Improving diagnosis of inherited peripheral neuropathies through gene panel analysis. *Orphanet J. Rare Dis.* 11:118. doi: 10.1186/s13023-016-0500-5

Lelj-Garolla, B., and Mauk, A. G. (2012). Roles of the N- and C-terminal sequences in Hsp27 self-association and chaperone activity. *Protein Sci.* 21, 122–133. doi: 10.1002/pro.761

Lerat, J., Magdelaine, C., Roux, A. F., Darnaud, L., Beauvais-Dzugan, H., Naud, S., et al. (2019). Hearing loss in inherited peripheral neuropathies: molecular diagnosis by NGS in a French series. *Mol. Genet. Genomic. Med.* 7:e839. doi: 10.1002/mgg3.839

Li, M. M., Datto, M., and Duncavage, E. J. (2017). Standards and guidelines for the interpretation and reporting of sequence variants in cancer: a joint consensus recommendation of the association for molecular pathology, American Society of Clinical Oncology, and College of American Pathologists. *J. Mol. Diagn.* 19, 4–23. doi: 10.1016/j.jmoldx.2016.10.002

Lupo, V., García-García, F., Sancho, P., Tello, C., García-Romero, M., Villarreal, L., et al. (2016). Assessment of targeted next-generation sequencing as a tool for the diagnosis of Charcot-Marie-Tooth disease and hereditary motor neuropathy. *J. Mol. Diagn.* 18, 225–234. doi: 10.1016/j.jmoldx.2015.10.005

Manganelli, F., Tozza, S., Pisciotta, C., Bellone, E., Iodice, R., Nolano, M., et al. (2014). Charcot-Marie-Tooth disease: frequency of genetic subtypes in a Southern Italy population. *J. Peripher. Nerv. Syst.* 19, 292–298. doi: 10.1111/jns.12092

Marques, W. J., Freitas, M. R., Oliveira, A. S. B., Calia, L., Melo, A., Lucena, R., et al. (2005). 17p duplicated Charcot–Marie–Tooth 1A: characteristics of a new population. *J. Neurol.* 252, 972–979. doi: 10.1007/s00415-005-0797-9

Marttila, M., Kytövuori, L., Helisalmi, S., Kallio, M., Laitinen, M., Hiltunen, M., et al. (2017). Molecular epidemiology of Charcot-Marie-Tooth disease in Northern Ostrobothnia, Finland: a population-based study. *Neuroepidemiology* 49, 34–39. doi: 10.1159/000478860

Milley, G. M., Varga, E. T., Grosz, Z., Nemes, C., Arányi, Z., Boczán, J., et al. (2018). Genotypic and phenotypic spectrum of the most common causative genes of Charcot-Marie-Tooth disease in Hungarian patients. *Neuromuscul. Disord.* 28, 38–43. doi: 10.1016/j.nmd.2017.08.007

Mostacciuolo, M. L., Righetti, E., Zortea, M., Bosello, V., and Schiavon, F. (2001). Charcot-Marie-Tooth disease type I and related demyelinating neuropathies: mutation analysis in a large cohort of Italian families. *Hum. Mutat.* 18, 32–41. doi: 10.1002/humu.1147

Murphy, S. M., Laura, M., Fawcett, K., Pandraud, A., Liu, Y. T., Davidson, G. L., et al. (2012). Charcot-Marie-Tooth disease: frequency of genetic subtypes and guidelines for genetic testing. *J. Neurol. Neurosurg. Psychiatry* 83, 706–710. doi: 10.1136/jnnp-2012-302451

Nicolas, A., Kenna, K. P., Renton, A. E., Ticozzi, N., Faghri, F., Chia, R., et al. (2018). Genome-wide analyses identify KIF5A as a novel ALS gene. *Neuron* 97, 1268–1283. doi: 10.1016/j.neuron.2018.02.027

OMIM (2014). Online Mendelian Inheritance in Man, OMIM (TM). Available at: http://www.ncbi.nlm.nih.gov/omim/

Østern, R., Fagerheim, T., Hjellnes, H., Nygård, B., Mellgren, S. I., and Nilssen, Ø. (2013). Diagnostic laboratory testing for Charcot Marie Tooth disease (CMT): the spectrum of gene defects in Norwegian patients with CMT and its implications for future genetic test strategies. *BMC Med. Genet.* 14:94. doi: 10.1186/1471-2350-14-94

Piscosquito, G., Saveri, P., and Magri, S. (2016). Screening for SH3TC2 gene mutations in a series of demyelinating recessive Charcot-Marie-Tooth disease (CMT4). *J. Peripher. Nerv. Syst.* 21, 142–149. doi: 10.1111/jns.12175

Rautenstrauss, B., Liehr, T., Fuchs, C., Bevot, A., Bornemann, A., Postler, E., et al. (1998). Mosaicism for Charcot-Marie-Tooth disease type 1A: onset in childhood suggests somatic reversion in early developmental stages. *Int. J. Mol. Med.* 1, 333–337. doi: 10.3892/ijmm.1.2.333

Reilly, M. M., and Shy, M. E. (2009). Diagnosis and new treatments in genetic neuropathies. *J. Neurol. Neurosurg. Psychiatry* 80, 1304–1314. doi: 10.1136/jnnp.2008.158295

Rossor, A. M., Polke, J. M., Houlden, H., and Reilly, M. M. (2013). Clinical implications of genetic advances in Charcot-Marie-Tooth disease. *Nat. Rev. Neurol.* 9, 562–571. doi: 10.1038/nrneurol.2013.179

Saifi, G. M., Szigeti, K., Snipes, G. J., Garcia, C. A., and Lupski, J. R. (2003). Molecular mechanisms, diagnosis, and rational approaches to management of and therapy for Charcot-Marie-Tooth disease and related peripheral neuropathies. *J. Investig. Med.* 51, 261–283. doi: 10.1136/jim-51-05-14

Saporta, A. S., Sottile, S. L., Miller, L. J., Feely, S. M., Siskind, C. E., and Shy, M. E. (2011). Charcot-Marie-Tooth disease subtypes and genetic testing strategies. *Ann. Neurol.* 69, 22–33. doi: 10.1002/ana.22166

Scarlato, M., Viganò, F., Carrera, P., Previtali, S. C., and Bolino, A. (2015). A novel heat shock protein 27 homozygous mutation: widening of the continuum between MND/dHMN/CMT2. *J. Peripher. Nerv. Syst.* 20, 419–421. doi: 10.1111/jns.12139

Schüle, R., Kremer, B. P., Kassubek, J., Auer-Grumbach, M., and Kostic, V. (2008). SPG10 is a rare cause of spastic paraplegia in European families. *J. Neurol. Neurosurg. Psychiatry* 79, 584–587. doi: 10.1136/jnnp.2007.137596

Shy, M. E., Lupski, J. R., and Chance, P. F. (2005). "Hereditary motor and sensory neuropathies: an overview of clinical, genetic, electrophysiologic and pathologic features," in *Peripheral Neuropathy. 4th Edn.* eds. P. J. Dyck and P. K. Thomas (Philadelphia: Elsevier Saunders), 1623–1658.

Sivera, R., Sevilla, T., Vílchez, J. J., Martínez-Rubio, D., and Chumillas, M. J. (2013). Charcot-Marie-Tooth disease: genetic and clinical spectrum in a Spanish clinical series. *Neurology* 81, 1617–1625. doi: 10.1212/WNL.0b013e3182a9f56a

Tessa, A., Silvestri, G., de Leva, M. F., Modoni, A., Denora, P. S., and Masciullo, M. (2008). A novel KIF5A/SPG10 mutation in spastic paraplegia associated with axonal neuropathy. *J. Neurol.* 255, 1090–1092. doi: 10.1007/s00415-008-0840-8

Vaeth, S., Christensen, R., Dunø, M., Lildballe, D. L., and Thorsen, K. (2019). Genetic analysis of Charcot-Marie-Tooth disease in Denmark and the implementation of a next generation sequencing platform. *Eur. J. Med. Genet.* 62, 1–8. doi: 10.1016/j.ejmg.2018.04.003

Vallat, J. M., Mathis, S., and Funalot, B. (2013). The various Charcot-Marie-Tooth diseases. *Curr. Opin. Neurol.* 26, 473–480. doi: 10.1097/WCO.0b013e328364c04b

Verhoeven, K., Claeys, K. G., Züchner, S., Schröder, J. M., Weis, J., Ceuterick, C., et al. (2006). MFN2 mutation distribution and genotype/phenotype correlation in Charcot-Marie-Tooth type 2. *Brain* 129, 2093–2102. doi: 10.1093/brain/awl126

Warner, L. E., Hilz, M. J., Appel, S. H., Killian, J. M., Kolodry, E. H., Karpati, G., et al. (1996). Clinical phenotypes of different MPZ (P0) mutations may include Charcot-Marie-Tooth type 1B, Dejerine-Sottas, and congenital hypomyelination. *Neuron* 17, 451–460. doi: 10.1016/S0896-6273(00)80177-4

Washington University (2014). Neuromuscular disease center: hereditary motor sensory neuropathies. Available at: http://neuromuscular.wustl.edu/time/hmsn.html

Wiszniewski, W., Szigeti, K., and Lupski, J. R. (2013). "Chapter 126 – hereditary motor and sensory neuropathies," in *Emery and Rimoin's Principles and Practice of Medical Genetics. 6th Edn.* eds. L. R. David, E. P. Reed and K. Bruce (Oxford, UK: Academic Press), 1–24.

Züchner, S., Mersiyanova, I. V., Muglia, M., Bissar-Tadmouri, N., Rochelle, J., Dadali, E. L., et al. (2004). Mutations in the mitochondrial GTPase mitofusin 2 cause Charcot-Marie-Tooth neuropathy type 2A. *Nat. Genet.* 36, 449–451. doi: 10.1038/ng1341

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.

GLOSSARY

FAM ID	Family identification
HGVSc	Nomenclature human genome variation society coding DNA
HGVSp	Nomenclature human genome variation society protein
dbSNP ID	Single-nucleotide polymorphism database
ACMG	American College of Medical Genetics Guideline
AY	Years at evaluation
AO	Age at onset
CMT subtype	Demyelinating CMT = 1, axonal CMT = 2, dHMN = 3, HSN = 4, and possible CMT = 5
CV	Conduction velocity m/s
CV subtype	CMT subtype classification for conduction velocity
PVS1	Null variant (nonsense, frameshift, canonical ± 1 or 2 splice sites, initiation codon, single, or multiexon deletion) in a gene where LOF is a known mechanism of disease (pathogenic and very strong)
PM1	Located in a mutational hot spot and/or critical and well-established functional domain (e.g., active site of an enzyme) without benign variation (pathogenic and moderate)
PM2	Absent from controls (or at extremely low frequency if recessive) in Exome Sequencing Project, 1,000 Genomes Project, or Exome Aggregation Consortium (pathogenic and moderate)
PM3	For recessive disorders, detected in trans with a pathogenic variant (pathogenic and moderate)
PM4	Protein length changes as a result of in-frame deletions/insertions in a nonrepeat region or stop-loss variants (pathogenic and moderate)
PP1	Cosegregation with disease in multiple affected family members in a gene definitively known to cause the disease (pathogenic and supporting)
PP2	Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease (pathogenic and supporting)
PP3	Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.; pathogenic and supporting)
PP4	Patient's phenotype or family history is highly specific for a disease with a single genetic etiology (pathogenic and supporting)
BP1	Missense variant in a gene for which primarily truncating variants is known to cause disease (benign and supporting)
BP4	Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.; benign and supporting).

frontiers in Genetics

ORIGINAL RESEARCH
published: 19 July 2021
doi: 10.3389/fgene.2021.689868

Identification of Variants Associated With Rare Hematological Disorder Erythrocytosis Using Targeted Next-Generation Sequencing Analysis

*Aleša Kristan[1], Tadej Pajič[2,3,4], Aleš Maver[3], Tadeja Režen[5], Tanja Kunej[6], Rok Količ[7], Andrej Vuga[7], Martina Fink[1,2], Špela Žula[2], Helena Podgornik[2,8], Saša Anžej Doma[2,9], Irena Preložnik Zupan[2,9], Damjana Rozman[5] and Nataša Debeljak[1]**

[1] Medical Centre for Molecular Biology, Institute of Biochemistry and Molecular Genetics, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia, [2] Department of Hematology, University Medical Centre Ljubljana, Ljubljana, Slovenia, [3] Clinical Institute of Genomic Medicine, University Medical Centre Ljubljana, Ljubljana, Slovenia, [4] Clinical Biochemistry, Faculty of Medicine, University of Maribor, Maribor, Slovenia, [5] Centre for Functional Genomics and Bio-Chips, Institute of Biochemistry and Molecular Genetics, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia, [6] Department of Animal Science, Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia, [7] Kemomed Research and Development, Kemomed Ltd., Ljubljana, Slovenia, [8] Clinical Biochemistry, Faculty of Pharmacy, University of Ljubljana, Ljubljana, Slovenia, [9] Department of Internal Medicine, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia

OPEN ACCESS

Edited by:
Yanling Yang,
Peking University First Hospital, China

Reviewed by:
Nancy Monroy-Jaramillo,
National Institute of Neurology and Neurosurgery, Mexico
Prashanth N. Suravajhala,
Birla Institute of Scientific Research, India

***Correspondence:**
Nataša Debeljak
natasa.debeljak@mf.uni-lj.si

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

Received: *01 April 2021*
Accepted: *16 June 2021*
Published: *19 July 2021*

Citation:
Kristan A, Pajič T, Maver A, Režen T, Kunej T, Količ R, Vuga A, Fink M, Žula Š, Podgornik H, Anžej Doma S, Preložnik Zupan I, Rozman D and Debeljak N (2021) Identification of Variants Associated With Rare Hematological Disorder Erythrocytosis Using Targeted Next-Generation Sequencing Analysis. Front. Genet. 12:689868. doi: 10.3389/fgene.2021.689868

An erythrocytosis is present when the red blood cell mass is increased, demonstrated as elevated hemoglobin and hematocrit in the laboratory evaluation. Congenital predispositions for erythrocytosis are rare, with germline variants in several genes involved in oxygen sensing (*VHL*, *EGLN1*, and *EPAS1*), signaling for hematopoietic cell maturation (*EPOR* and *EPO*), and oxygen transfer (*HBB*, *HBA1*, *HBA2*, and *BPGM*) that were already associated with the eight congenital types (ECYT1–8). Screening for variants in known congenital erythrocytosis genes with classical sequencing approach gives a correct diagnosis for only up to one-third of the patients. The genetic background of erythrocytosis is more heterogeneous, and additional genes involved in erythropoiesis and iron metabolism could have a putative effect on the development of erythrocytosis. This study aimed to detect variants in patients with yet unexplained erythrocytosis using the next-generation sequencing (NGS) approach, targeting genes associated with erythrocytosis and increased iron uptake and implementing the diagnostics of congenital erythrocytosis in Slovenia. Selected 25 patients with high hemoglobin, high hematocrit, and no acquired causes were screened for variants in the 39 candidate genes. We identified one pathogenic variant in *EPAS1* gene and three novel variants with yet unknown significance in genes *EPAS1*, *JAK2*, and *SH2B3.* Interestingly, a high proportion of patients were heterozygous carriers for two variants in *HFE* gene, otherwise pathogenic for the condition of iron overload. The association between the *HFE* variants and the development of erythrocytosis is not clearly understood. With a targeted NGS approach, we determined an actual genetic cause for the erythrocytosis in one patient and contributed to better management of the disease for the patient and his

family. The effect of variants of unknown significance on the enhanced production of red blood cells needs to be further explored with functional analysis. This study is of great significance for the improvement of diagnosis of Slovenian patients with unexplained erythrocytosis and future research on the etiology of this rare hematological disorder.

Keywords: NGS – next-generation sequencing, erythrocytosis, targeted panel sequencing, iron metabolism, diagnostics, rare disease (RD)

INTRODUCTION

Erythrocytes or red blood cells are the most abundant cells in the blood with the main function of tissue oxygen delivery (Lee and Percy, 2011). An increase in red blood cell mass (RCM) for more than 125% of predicted for age and sex is defined as absolute erythrocytosis. Increased RCM leads to raised blood viscosity and manifests as elevated hemoglobin (Hb) and hematocrit (Ht). Different criteria of Hb and Ht have been published for the diagnosis of erythrocytosis, with the consensus that Hb > 185 g/L and Ht > 0.52 in men and Hb > 165 g/L and Ht > 0.48 in women in at least two separate blood counts at 2 months apart requires further investigations (Patnaik and Tefferi, 2009; Keohane et al., 2013; McMullin, 2014, 2016; Bento, 2018). Clinical signs are associated with increased blood viscosity and include non-specific signs like dizziness, itching, facial plethora, redness of the hands, pulmonary hypertension, and, in some patients, serious complications like thromboembolic events and death (Lee and Arcasoy, 2015; McMullin et al., 2019a,b).

Several factors could lead to increased RCM. The cause of erythrocytosis could be congenital or acquired and can be further divided into primary and secondary. An erythrocytosis is classified as primary, when there is an intrinsic defect in the erythroid progenitor cells, and secondary when the defect outside the erythroid compartment is driving the bone marrow to produce more red blood cells (McMullin, 2016; Bento, 2018). The most common reasons for erythrocytosis are acquired, due to somatic variants or various extrinsic factors that lead to reduced oxygen supply and thus stimulation of erythropoiesis, such as chronic pulmonary, cardiac, renal, hepatic diseases, erythropoietin (EPO)-secreting tumors, high-altitude living, smoking, sleep apnea, recombinant EPO, and androgen administration (McMullin, 2008; Lee and Percy, 2011; Bento, 2018). Genetic defects in several pathways and mechanisms that regulate erythropoiesis lead to erythrocytosis, including oxygen-sensing pathway and regulation of EPO transcription, EPO signal transduction mediated *via* EPOR-JAK2 signaling cascade, and regulation of hemoglobin-oxygen affinity (Gaspersic et al., 2020). Primary acquired erythrocytosis, i.e., somatic erythrocytosis (OMIM ID: 133100), is the consequence of somatic genetic variants in Janus kinase 2 (*JAK2*) gene and SH2B adaptor protein 3 (*SH2B3*) gene, which lead to constant activation of EPO signaling pathway (McMullin and Cario, 2016; Maslah et al., 2017; Bento, 2018). Somatic variant p.Val617Phe and variants in exon 12 of the *JAK2* gene are the reason for the development of somatic erythrocytosis also termed polycythemia vera (PV) (OMIM ID: 263300) in 98% of erythrocytosis cases (Baxter et al., 2005; Scott et al., 2007; Bento, 2018). Cases with congenital erythrocytosis due to germline variants are much rarer. Based on the genetic origin, there are eight types of congenital (familiar) erythrocytosis, namely, ECYT1–8. Heterozygous variants in the EPO receptor gene (*EPOR*) are responsible for the development of ECYT1 (OMIM ID: 133100). The mechanism underlying ECYT2–5 (OMIM IDs: 263400, 609820, 611783, and 617907) is impairment of genes von Hippel–Lindau tumor suppressor (*VHL*), egl-9 family hypoxia-inducible factor 1 (*EGLN1*), endothelial PAS domain protein 1 (*EPAS1*), and *EPO* involved in the oxygen sensing. ECYT6–8 (OMIM IDs: 617980, 617981, and 222800) are developed due to variants in hemoglobin genes *HBB*, *HBA1*, and *HBA2* and biphosphoglycerate mutase (*BPGM*), which lead to increased oxygen affinity of hemoglobin (**Figure 1**; McMullin, 2016; Bento, 2018). Except for VHL and BPGM-associated erythrocytosis, all other familial types have shown an autosomal-dominant type of inheritance (Amberger et al., 2019). Besides the above-mentioned genes, also other genes are included in the pathways of red blood cell production but were not yet associated with the clinical outcome of erythrocytosis (Gaspersic et al., 2020).

It is increasingly evident that the same genes and pathways could be involved in the pathology of different diseases. Hereditary hemochromatosis (HH) is a disorder of systemic iron overload, due to the deficiency in iron regulatory hormone hepcidin. Variants in genes encoding HH protein (HFE), hemojuvelin (HJV), hepcidin (HAMP), transferrin receptor protein 2 (TFR2), and ferroportin (SLC40A1) could lead to loss of hepcidin transcription or activity and development of five HH types. The most common causes for HH are homozygous or compound heterozygous defects (p.Cys282Tyr, p.His63Asp, and p.Ser65Cys) in gene *HFE*, resulting in autosomal-recessive HH type 1 (Lanktree et al., 2017; Brissot et al., 2018). Recently, several research groups found *HFE* variants at higher frequencies among patients with unexplained erythrocytosis, i.e., idiopathic erythrocytosis (IE), compared with the general population (Biagetti et al., 2018; Burlet et al., 2019; Gurnari et al., 2019). Similarly, patients with the *HFE* variants had significantly higher Hb and Ht values than normal (Barton et al., 2000; Asif et al., 2019). It was suggested that *HFE* mutations could induce erythropoiesis through higher iron bioavailability, since iron is dedicated to the synthesis of Hb (Hentze et al., 2010; Biagetti et al., 2018; Burlet et al., 2019). Several gene panels of hemochromatosis and iron metabolism-associated genes have been developed (Badar et al., 2016; Faria et al., 2016; Ferbo et al., 2016; Lanktree et al., 2017).

In the past, most of the laboratories and diagnostic centers included only genes for ECYT1–8 in their routine laboratory evaluation for congenital erythrocytosis, using Sanger sequencing

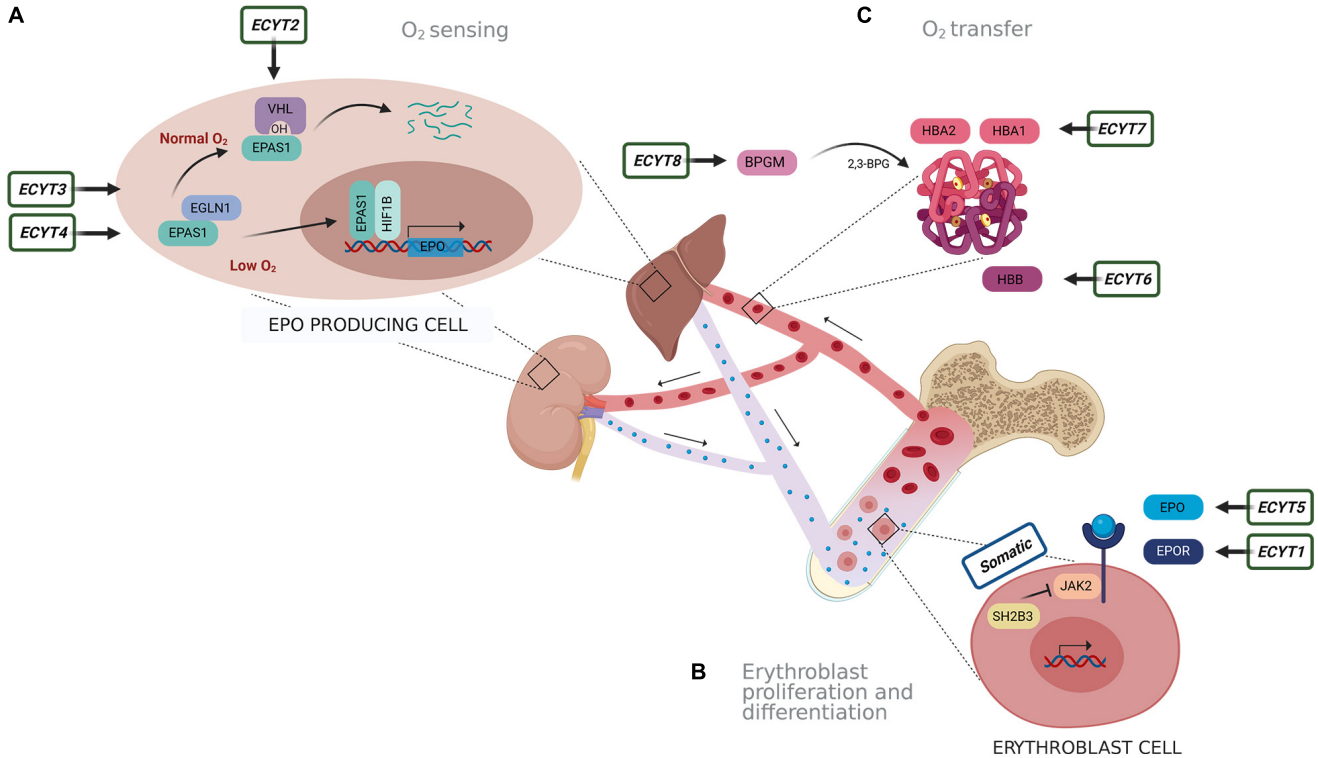


FIGURE 1 | Schematic presentation of genes responsible for eight types of congenital erythrocytosis (ECYT1–8) and somatic erythrocytosis. Genes are involved in pathways of **(A)** oxygen sensing, **(B)** signaling for proliferation and differentiation of erythroblasts and **(C)** oxygen transfer. **(A)** Oxygen is detected in the liver and kidney. In the normal oxygen conditions, hypoxia-inducible transcription factor (HIF) (the main isoform associated with the erythropoiesis is EPAS1) is hydroxylated by prolyl hydroxylase EGLN1, following binding of *VHL*. This results in ubiquitination and degradation of HIF. Under hypoxic conditions (e.g., 1% O_2), oxygen availability is limited; thus, hydroxylation is diminished. This results in the stabilization of HIF that is transported to the nucleus where it forms a dimer complex with its beta subunit HIF1B (officially termed ARNT) and acts as transcription factor on numerous target genes, including *EPO*. **(B)** EPO is transported to the bone marrow, where it binds with its receptor EPOR on the erythroblast cells. This subsequently activates JAK2 signaling cascade for proliferation and differentiation of erythroblasts. SH2B3 is an inhibitor of JAK2 kinase. **(C)** Hemoglobin alpha (HBA1 and HBA2) and beta (HBB) are important binding proteins for oxygen and are involved in oxygen transfer. The enzyme BPGM forms an allosteric effector 2,3-biphosphoglycerate (2,3-BPG) for regulation of hemoglobin oxygen affinity.

as a predominant method (Gaspersic et al., 2020). However, the practice has been so far that only about 20–30% of patients received a proper diagnosis with screening for known variants associated with erythrocytosis. Therefore, the majority of patients remained idiopathic (Bento et al., 2013; Camps et al., 2016; Girodon et al., 2017; Bento, 2018). This implicates that several other genes and mechanisms must be involved in the disease development. In recent years, the next-generation sequencing (NGS) approach was introduced into diagnostics of patients with IE or suspected congenital erythrocytosis. The biggest advantage of NGS sequencing in rare disease diagnostics is that it allows a massive parallel sequencing of multiple genomic regions of interest at once (Fernandez-Marmiesse et al., 2018). Several researchers used whole-genome sequencing (WGS) to identify novel candidate genes and variants involved in the development of congenital erythrocytosis (Taylor et al., 2015; Lenglet et al., 2018). To overcome the drawbacks of WGS, like high cost and processing a huge amount of data, targeted NGS panels were established. Camps et al. (2016) developed the erythrocytosis gene panel of 21 candidate genes, which are involved in key disease-driven pathways or were identified in prior WGS projects (Taylor et al., 2015; Camps et al., 2016). Targeted NGS of erythrocytosis-associated genes was later also used by a French research group studying erythrocytosis (Girodon et al., 2017; Filser et al., 2021), and erythrocytosis gene panels are gradually introduced into the routine clinical practice (Gaspersic et al., 2020).

The diagnostic procedures for the evaluation of patients with abnormally high Hb and Ht had been insufficient in Slovenia so far (Mlakar, 2008). A previous genetic analysis of patients with signs of erythrocytosis included only the examination of the *JAK2* gene for variants causative for PV, with allele-specific PCR, high-resolution melting (HRM) analysis, and Sanger sequencing (Baxter et al., 2005; Ugo et al., 2010; Geay et al., 2020). A huge proportion of patients with excluded PV were undiagnosed and remained idiopathic. It was urgent to modernize the diagnostic algorithm for erythrocytosis and to extend the genetic analysis for other known and novel candidate genes associated with the development of erythrocytosis. We have recently developed a national diagnostic algorithm for erythrocytosis, differentiating patients with absolute erythrocytosis, in whom acquired causes and PV are excluded (Anzej Doma et al., 2021a). The precise

selection of patients to the point where genetic testing is considered is important to improve the diagnostic yield.

In the previous study, we have described and verified the targeted NGS method, which we developed for the genetic characterization of IE (Kristan et al., 2021). In the same gene panel, we included for the first time genes associated with erythrocytosis and HH. This was an important improvement to the erythrocytosis gene panels already used in international clinical practice since erythropoiesis and iron metabolism pathways are clearly intertwined. In the present study, we used targeted NGS to diagnose a group of patients with IE, which were selected over a period of 8 years (Anzej Doma et al., 2021a). The targeted NGS approach, in combination with the national diagnostic algorithm, improved the diagnostic accuracy of the Slovenian patients with IE, as we provided new explanations for the development of erythrocytosis to the patients.

MATERIALS AND METHODS

Patients

Adult patients with erythrocytosis of unknown cause were selected based on the diagnostic algorithm for erythrocytosis (Anzej Doma et al., 2021a) from individuals followed up at University Medical Centre Ljubljana (UMCLj) between 2011 and 2019. All patients were of Slovenian ethnic origin. A complete hemogram was performed for the patients, with additional relevant hematological parameters like ferritin values and saturation of transferrin. The inclusion criteria were (a) confirmed absolute erythrocytosis with hemoglobin >185 g/L for men and 165 g/L for women or hematocrit >0.52 for men and >0.48 for women twice over at least 2 months; (b) absence of pathogenic *JAK2* variants p.Val617Phe and exon 12 variants for PV; and (c) absence of any defined cause of secondary acquired erythrocytosis. Participants gave informed consent; the study was approved by the National Medical Ethics Committee, Ministry of Health of the Republic of Slovenia, approval no. 115/07/15 (0120-198/2015-4, 0120-287/2019-4). Patients were also reviewed for a family history of erythrocytosis. Several patients reported similar symptoms of erythrocytosis in families, and we were able to gather additional samples of other family members. Overall, 25 patients were selected for the genetic analysis with targeted NGS: 21 unrelated patients and two families, each with two participating members. Additionally, a family member without erythrocytosis from one of the families was included. DNA samples were extracted from granulocytes as described earlier (Kristan et al., 2021). We also included commercial reference DNA control NA12878, obtained from the Coriell Institute.

Targeted Next-Generation Sequencing and Data Analysis

We used targeted NGS to analyze samples from selected patients, one family member without erythrocytosis, and a reference DNA control. Library preparation and enrichment of 39 genes involved in erythropoiesis and hemochromatosis development were performed as described previously (Kristan et al., 2021). For enrichment, sample libraries were combined into multiplex pools of 12 samples pooled by mass. Enriched libraries were sequenced on MiniSeq sequencer (Illumina, San Diego, CA, United States) in 2 × 150 cycles. After duplicates were removed, the alignment of reads to UCSC hg19 reference assembly was done using BWA algorithm (v0.6.3), and variant calling was done using GATK framework (v2.8). Only variants exceeding the quality score of 30.0 and depth of 5 were used for downstream analysis. Variant annotation was performed using ANNOVAR and snpEff algorithms, with pathogenicity predictions in dbNSFPv2 database. Reference gene models and transcript sequences are based on the RefSeq database. Variants with population frequency exceeding 1% in 1000 genomes and ESP6500, synonymous variants, intronic variants, and variants outside the clinical target were filtered out during analyses (Bergant et al., 2018). Separately, we also looked for clinically significant variants in *HFE* gene, which has a higher population frequency than 1%. The interpretation of sequence variants was based on ACMG/AMP standards and guidelines (Richards et al., 2015), modified in accordance with the ACGS recommendations (Ellard et al., 2019). Variants are classified as follows: class 1, benign variants; class 2, possibly benign variants; class 3, variants of uncertain significance, not enough evidence; class 4, possibly pathogenic variants; and class 5, pathogenic variants. Classification of variants is supported by the evidence categories (Richards et al., 2015). A list of *in silico* tools with deleterious predictions of variants was obtained from the Varsome database[1] (Kopanos et al., 2019). A variant was predicted to have a deleterious effect on the gene or gene product (evidence category PP3), when the majority of computational predictors or the majority of meta-predictors (REVEL, METALR, METASVM, and CADD) supported a deleterious effect. The cutoff value on deleteriousness for CADD (Combined Annotation Dependent Depletion) score was set to >20 (Rentzsch et al., 2019). The frequencies of identified variants in the general population were gathered from the GnomAD browser, version v2.1.1[2]. We reviewed databases ClinVar (Landrum et al., 2018), LOVD (Fokkema et al., 2011), HGMD (Stenson et al., 2003), COSMIC (Tate et al., 2019), and CIViC (Griffith et al., 2017) to search for clinically relevant data of variants. Gene abbreviations are according to the official symbols in HUGO Gene Nomenclature Committee (Tweedie et al., 2021). The purpose of the reference DNA sample was to serve as a control for sequencing performance. For benchmarking the method sensitivity, reference DNA NA12878 was sequenced, and variant calls were compared against the high-confidence variant calls provided by the NIST Genome in a bottle consortium (v.3.3.2).

Validation Using Sanger Sequencing

All identified pathogenic variants were further confirmed by Sanger sequencing (GATC Biotech, Konstanz, Germany). Sanger sequencing and prior PCR amplification were performed with custom-designed primers (IDT) and are available upon request.

[1]https://varsome.com/

[2]https://gnomad.broadinstitute.org/

RESULTS

Sequencing Performance

All 27 samples, including patients, one individual without erythrocytosis, and a reference DNA control, were sequenced by the targeted NGS. On average, 87% of mapped reads were on target regions, which indicates a successful enrichment. Targeted sequenced bases had high quality, as over 92% of aligned bases were sequenced with the minimum quality of Q30. With sequencing, we reached the median coverage of 230× (**Figure 2A**). The capture design encompassed 377 regions, of which we attained an average coverage of at least 10× for 362 regions (96%) across the sequenced samples, indicating that the majority of captured regions were sequenced at sufficient coverage for sensitive variant detection (**Figure 2B**). Only 15 target regions (4.0%) had an average coverage lower than 10× across samples, probably due to the high guanine–cytosine (GC) content of these regions (**Supplementary Table 1**). We cross-referenced the regions with poor coverage against the published pathogenic or likely pathogenic variants in the ClinVar database and found that these regions do not contain any known erythrocytosis-associated clinically relevant variants.

Using our analysis approach, we detected 47 out of 51 variants within the targeted region of the NA12878 reference sample. Of

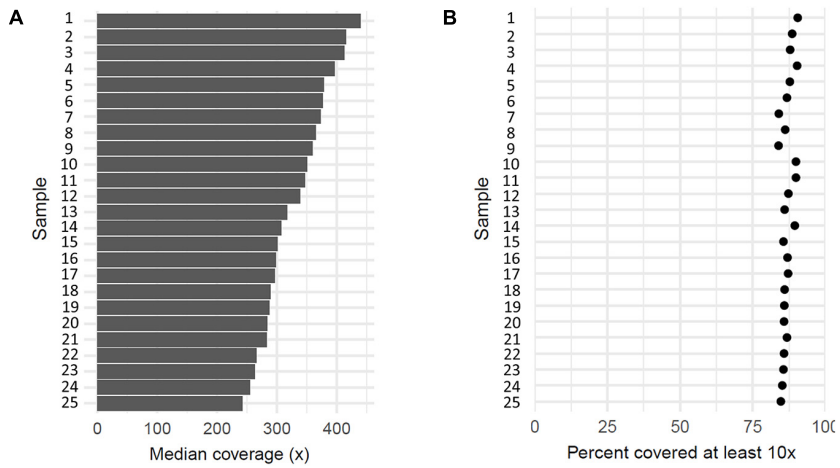


FIGURE 2 | Coverage of 25 sequenced patient's samples. **(A)** Each bar represents median coverage depth for an individual sample. **(B)** Each dot represents the percentage of target regions with coverage at least 10× for an individual sample.

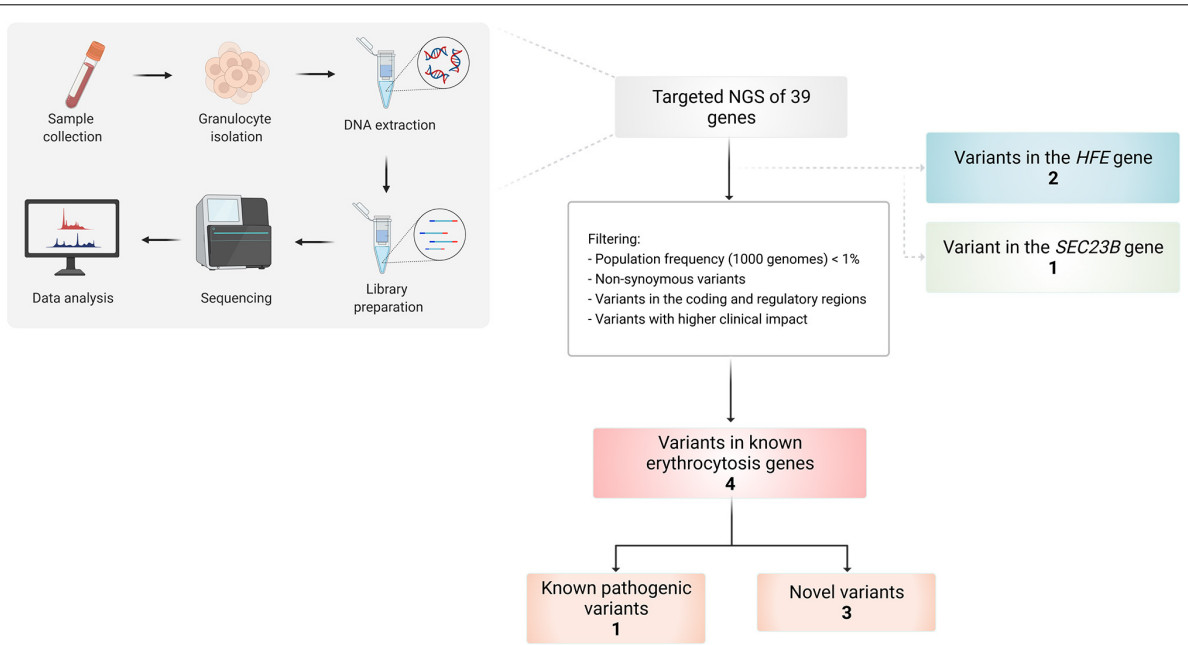


FIGURE 3 | Workflow of the study and overview of the identified variants in selected patients using targeted NGS of 39 genes.

these, three undetected variants were located within the known regions of poor coverage and were thus not considered in the sensitivity calculation. Apart from that, only a single variant was not detected in the sensitivity experiment due to the low coverage of the region in the control sample, corresponding to an estimated sensitivity of 97.9% for our sequencing approach.

Identified Variants

Several variants were identified using the erythrocytosis and hemochromatosis gene panel (**Figure 3**). Four variants were identified in the known erythrocytosis-causing genes *EPAS1*, *JAK2*, and *SH2B3* (**Table 1**). One variant in *EPAS1* gene had been already reported in the literature as causing erythrocytosis, while the remaining variants were not yet clinically associated with erythrocytosis. In addition, we also identified variants from the hemochromatosis set of genes, two *HFE* variants (**Table 2**) and one in the SEC23 homolog B (*SEC23B*) gene, which was found to be associated with a different condition. All the identified variants were located in the coding regions of the analyzed genes.

Known Pathogenic Variants

Heterozygous variant c.1609G>A in *EPAS1* gene is a pathogenic variant that represents an established cause of familial erythrocytosis type 4 (ECYT4, OMIM ID 611783). This variant has been previously reported as pathogenic in numerous unrelated individuals with erythrocytosis (Gale et al., 2008; Percy et al., 2008a; Perrotta et al., 2013; Liu et al., 2017; Oliveira et al., 2018). In addition, it has been reported to co-segregate with erythrocytosis in multiple affected family members (Gale et al., 2008; Percy et al., 2008a; Liu et al., 2017). The variant is also classified as pathogenic in the ClinVar, LOVD, and HGMD databases (ClinVar ID: 6469, LOVD ID: EPAS1_000002, HGMD ID: CM081583); and the deleterious effect on the protein was confirmed in functional studies (Gale et al., 2008; Furlow et al., 2009; Perrotta et al., 2013; Tarade et al., 2018). The clinical presentation of the patient was highly consistent with the disease.

Novel Variants in Known Candidate Genes

Three heterozygous variants c.2120A>C, c.1767C>A, and c.901G>A with unknown clinical impact were identified in known erythrocytosis-associated genes *EPAS1*, *JAK2*, and *SH2B3* genes, respectively.

Heterozygous variant c.901G>A in *SH2B3* gene has been already recorded in the HGMD database (ID: CM1614293). It was previously reported in one patient with IE without a family history of erythrocytosis (Camps et al., 2016).

On the contrary, heterozygous variant c.2120A>C in *EPAS1* gene has not yet been reported in association with human diseases in the literature and is not recorded among the clinically relevant variants in ClinVar, LOVD, and HGMD databases at the time of data analysis.

Similarly, heterozygous variant c.1767C>A in *JAK2* gene has not been yet recorded among the clinically relevant variants in ClinVar, HGMD, COSMIC, and CIViC databases. The variant in *JAK2* gene was not found in other affected or unaffected relatives, and the allele fraction [variant allele frequency (VAF)] was calculated to be 22%, which indicates its possible somatic origin.

We classified all three variants as variants with unknown significance (VUS) due to low population frequency in the GnomAD project or high pathogenicity. Variants in *EPAS1* and *JAK2* genes were observed at extremely low frequencies in the gnomAD project or were absent from the GnomAD controls. Variant *SH2B3* c.901G>A was predicted to have a deleterious effect on the protein as the majority of meta-predictors (REVEL, METASVM, and CADD) supported its deleteriousness.

Variants in the Hemochromatosis Genes

With the targeted NGS analysis, we identified two heterozygous variants p.(Cys282Tyr) and p.(His63Asp) in *HFE* gene. Out of 25 patients with IE, 10 patients were carriers for variant p.(Cys282Tyr) or variant p.(His63Asp). Comparing the observed VAFs with the allele frequencies in the Slovenian population (Cukjati et al., 2007), IE patients had a higher incidence of variant p.(Cys282Tyr). Only a few patients with identified heterozygous *HFE* variants had elevated ferritin and saturation of transferrin > 45%, which is indicative of iron overload (**Table 2**).

Additionally, we identified a carriership of a pathogenic missense c.40C>T (NM_006363.4) variant in *SEC23B* gene, associated with the congenital dyserythropoietic anemia, type 2 (CDAII; OMIM:224100) in two unrelated patients. The significance of this finding in the development of erythrocytosis is not clear and is less likely causative concerning a high frequency of heterozygous variant carriers in the population. This variant was therefore excluded from the further report and interpretation of variants.

DISCUSSION

With previously established targeted NGS of 39 erythrocytosis and hemochromatosis-associated genes, we identified high-confidence variants among patients with IE and discovered few potential disease-causing variants in patients who previously lacked a specific diagnosis. We were able to provide an actual genetic cause (variant p.Gly537Arg in *EPAS1* gene) for erythrocytosis in one patient, which was also validated by Sanger sequencing. This is the first patient in Slovenia with a conclusive diagnosis of congenital erythrocytosis type 4 (ECYT4). Relatives of this patient will be further included in the study and genetically tested, to provide proper prognosis and genetic counseling to the family. Three novel candidate variants in the genes known to be associated with erythrocytosis, i.e., *EPAS1*, *JAK2*, and *SH2B3*, were identified. Interestingly, another research group studying erythrocytosis identified the majority of variants in genes *EPAS1*, *SH2B3*, *JAK2*, and *EGLN1* (Girodon et al., 2021). Variant c.901G>A (NM_005475.2) in *SH2B3* gene was the only one already identified in the erythrocytosis patient without myeloproliferative neoplasm (MPN) (Camps et al., 2016). Nevertheless, further segregation analyses and functional studies are necessary to support and ascertain its pathogenic nature.

Among all identified variants, the variants located within or in the vicinity of the important regions for protein function will have a higher likelihood of causality. The identified pathogenic

TABLE 1 | Known and novel variants detected in the known erythrocytosis genes.

Genomic location on hg19	Gene	Coding DNA change (RefSeq transcript)	Protein change	RS number	Allele frequency (GnomAD)	*In silico* tools with deleterious predictions[1] (CADD value)	Patient information[2]	Genotype	Classification[3] (evidence categories)
Chr.2:46607420	*EPAS1*	c.1609G>A (NM_001430.5)	p.(Gly537Arg)	rs137853036	NA	18/22: BayesDel addAF, BayesDel noAF, DANN, DEOGEN2, EIGEN, EIGEN PC, FATHMM-MKL, FATHMM-XF, LIST-S2, LRT, MVP, MutPred, Mutation assessor, MutationTaster, PROVEAN, PrimateAI, SIFT, SIFT4G; meta predictors: CADD (29.7)	Female; age, 21 years; Hb, 197 g/L; Ht, 0,59; RBC 6,35 × 10^{12}/L; platelets, 184 × 10^{9}/L; EPO, 73.4 IU/L; smoking; several thromboembolic events; pulmonary hypertension; erythrocytosis since childhood; phlebotomy; anticoagulation therapy	Het	Pathogenic (PS1, PS3, PS4, PM1, PM2, PM5, PP1_STR, PP4)
Chr.9:5072617	*JAK2*	c.1767C>A (NM_004972.4)	p.(Asn589Lys)	rs1362123436	0.000004	12/22: DANN, DEOGEN2, FATHMM, FATHMM-MKL, LIST-S2, LRT, Mutation assessor, MutationTaster, PROVEAN, PrimateAI, SIFT, SIFT4G; meta predictors: CADD (23.6)	Male; age, 62 years; Hb, 221 g/L; Ht, 0.65; RBC, 6.36 × 10^{12}; EPO 11.4 IU/L; platelets, 211 × 10^{9}/L; smoking; elevated blood pressure and heart rate; phlebotomy; positive family history, son also affected but without variant	Het	VUS (PM2)
Chr. 2:46608809	*EPAS1*	c.2120A>C (NM_001430.4)	p.(Lys707Thr)	rs950180639	NA	12/21: BayesDel addAF, BayesDel noAF, DANN, DEOGEN2, EIGEN PC, FATHMM-MKL, FATHMM-XF, LIST-S2, Mutation assessor, MutationTaster, SIFT, SIFT4G; meta predictors: CADD (26)	Male; age, 54 years; Hb, 182 g/L; Ht, 0.53; RBC, 6.24 × 10^{12}; EPO 6.1 IU/L; platelets, 271 × 10^{9}/L; smoking; mild sleep apnea; asthma; Hashimoto's thyroiditis, phlebotomy; possible positive family history, brother also similar symptoms	Het	VUS (PM2)
Chr. 12:111884812	*SH2B3*	c.901G>A (NM_005475.3)	p.(Glu301Lys)	rs374278232	0.00004	15/21: BayesDel noAF, DANN, DEOGEN2, EIGEN, EIGEN PC, FATHMM-MKL, FATHMM-XF, LRT, MVP, MetaLR, MetaSVM, MutationTaster, PROVEAN, PrimateAI, SIFT, SIFT4G; meta predictors: MetaSVM, CADD (28.4)	Male; age, 57 years; Hb, 181 g/L; Ht, 0.53; RBC, 6.05 × 10^{12}; EPO 10.1 IU/L; platelets, 226 × 10^{9}/L; possible sleep apnea; aspirin; no positive family history	Het	VUS (PP3)

VUS, variant of unknown significance; Het, heterozygous; Hb, hemoglobin; Ht, hematocrit; RBC, red blood cells; EPO, erythropoietin; NA, not available; hg19, genome assembly GRCh37/hg19.

[1]Number of pathogenicity predictions tools out of all indicated in silico tools with deleterious effect were collected from the Varsome database.

[2]The indicated hematological parameters (Hb, Ht, RBC, and platelets) and age were registered at the first examination; however, additional measurements have been made later, and Hb and Ht were elevated at least twice over > 2 months but are not reported here.

[3]Classification of variants based on ACMG/AMP standards.

Hb normal range: 130–170 g/L for male and 120–150 g/L for female; Ht normal range: 0.40–0.50 for male and 0.36–0.46 for female; RBC normal range: 4.5–5.5 × 10^{12}/L for male and 3.8–4.8 × 10^{12}/L for female; platelet normal range: 150–410 × 10^{9}/L; EPO normal range: 3.3–16.6 IU/L; ferritin normal range: 20–300 μg/L for male and 10–120 μg/L for female; transferrin saturation normal range: 0.15–0.45.

EPAS1 variant c.1609G>A (p.Gly537Arg) is located six amino acids from the hydroxylation site p.Pro531, crucial for the oxygen-dependent regulation of EPAS1 stability (Kristan et al., 2019). Another amino acid change (c.1609G>T; p.Gly537Trp) (Percy et al., 2008b) (evidence category PM5) and the same amino acid change caused by a different base substitution (c.1609G>C; p.Gly537Arg) (Oliveira et al., 2018) (evidence category PS1) have been reported as pathogenic in patients with erythrocytosis. Furthermore, several amino acid substitutions c.1601C>G (p.Pro534Arg), c.1601C>T (p.Pro534Leu), c.1603A>G (p.Met535Val), c.1603A>T (p.Met535Leu), c.1604T>C (p.Met535Thr), c.1615G>A (p.Asp539Asn), c.1620C>A (p.Phe540Leu), and c.1631C>G (p.Pro544Arg) in the immediate vicinity of the identified variant have been reported in patients with erythrocytosis as pathogenic (Bento, 2018; Kristan et al., 2019), suggesting the critical importance of the identified region for the function of EPAS1 protein and severe effect on patient's phenotype (Anzej Doma et al., 2021b). Some previous studies have shown that identified variant c.1609G>A (p.Gly537Arg), similar to other variants located C-terminal to p.Pro531, inhibits the binding of the EGLN1 hydroxylase to the EPAS1 protein. This results in impaired degradation and thus enhanced stabilization of EPAS1, which affects increased expression of target genes (Gale et al., 2008; Furlow et al., 2009; Perrotta et al., 2013; Tarade et al., 2018). Another identified variant in *EPAS1* gene, c.2120A>C (p.Lys707Thr), is located far downstream of the primary hydroxylation site, unlikely to have an effect on hydroxylation and degradation. However, the variant is located within the nuclear localization signal (NLS) at the C-terminus of the EPAS1 protein. Two basic NLS motifs were shown to be present in region 705–742 amino acids and are required for localization of transcription factor EPAS1 to the nucleus under hypoxia and normoxia. Furthermore, the introduced variant p.Lys709Thr partly impaired the nuclear accumulation of EPAS1 induced by hypoxia (Luo and Shibuya, 2001). This, together with the low frequency of variant among GnomAD controls, implies the potential effect of variant p.Lys707Thr on protein function and activity. Further functional studies are needed to provide further evidence of variant causality. No variants in the NLS region were previously associated with ECYT4, indicating that analysis of this region may be of importance in further diagnoses.

Gene *JAK2* codes for a non-receptor tyrosine kinase, in which its roles are to phosphorylate tyrosine residues in the EPO receptor and signal transducers and activators of transcription (STATs) and to activate the transcription of target genes for proliferation and differentiation of hematopoietic stem cells. The protein JAK2 has four major domains, with two Jak homology (JH) domains important for downstream signaling. The C-terminal JH1 domain (849–1124 aa) is a kinase domain responsible for phosphorylation, while structurally similar pseudo-kinase domain (JH2) (545–809 aa) lacks catalytic activity, and it is involved in auto-inhibition of JAK2 in the absence of ligand–receptor interaction (Gnanasambandan and Sayeski, 2011; Seif et al., 2017; UniProt Consortium., 2021). Variants in exons 12–15 were found in patients with various hematological disorders, and those exons (except for the exon 12) are located in the JH2 domain (Ma et al., 2009; Gnanasambandan and Sayeski, 2011). The most common somatic pathogenic variant in the exon 14, p.Val617Phe, which is responsible for the development of PV and other MPNs, allows the kinase to evade negative regulation and confer constitutive activation of JAK2 (Gnanasambandan and Sayeski, 2011). Identified variant p.(Asn589Lys) in the exon 13 of *JAK2* gene is also located in the auto-inhibitory domain JH2; therefore, the potential pathophysiological mechanism for enhanced erythrocytes production could be the same.

SH2B3 protein (also termed LNK) is a negative regulator of erythropoiesis, as it binds to the cytokines and JAK2 and inhibits downstream signaling pathways (Maslah et al., 2017). SH2B3 has three domains: a dimerization domain, central pleckstrin homology domain (PH), and a Src homology 2 (SH2) domain. The PH domain (194–307 aa) has a role in binding the adapter protein SH2B3 to the plasma membrane, which is important for early inhibitory action during cytokine signaling. A mutational hotspot in the SH2B3 protein is present in the PH domain, encompassing exons 2, 3, and 4. Identified *SH2B3* variant in our study was located in exon 4, at the C-terminus of the PH domain. The majority of the *SH2B3* variants found in patients with PV and also in patients with p.Val617Phe-negative erythrocytosis were located in the PH domain, with the amino acid residue 208 (exon 2) as a preferential mutational target site (Spolverini et al., 2013; McMullin and Cario, 2016; Maslah et al., 2017).

TABLE 2 | Distribution of *HFE* variants among IE patients and number of variant carriers with elevated ferritin and transferrin saturation.

Genomic location on hg19	Coding DNA change (RefSeq transcript)	Protein change	RS number	Genotype	Number of patients (*N* = 25)	Proportion of patients (%)	Allele frequencies	Allele frequencies in Slovenian population[1]	Variant carriers with elevated ferritin	Variant carriers with elevated transferrin saturation
Chr. 6:26093141	c.845G>A (NM_000410.3)	p.(Cys282Tyr)	rs1800562	Het	4	0.16	0.08	0.036	1	2
Chr. 6:26091179	c.187C>G (NM_000410.3)	p.(His63Asp)	rs1799945	Het	6	0.24	0.12	0.128	0	1

Het, heterozygous; hg19, genome assembly GRCh37/hg19.
[1]Allele frequencies were obtained from the 1,282 Slovenian blood donors tested for the HFE variants (Cukjati et al., 2007).
Ferritin normal range: 20–300 μg/L for male and 10–120 μg/L for female. Transferrin saturation normal range: 0.15–0.45.

Somatic variants in *JAK2* and *SH2B3* genes are the cause for somatic erythrocytosis (OMIM ID: 133100), somatic myelofibrosis (OMIM ID: 254450), and thrombocythemia (OMIM ID: 614521, 187950) (McMullin and Cario, 2016; Amberger et al., 2019). In addition, germline variants in the *JAK2* gene are causative for thrombocythemia 3 (OMIM ID: 614521) (Mead et al., 2012; Amberger et al., 2019); and in several cases, germline *JAK2* variants, also in combination with other variants, were responsible for the activation of JAK2/STAT signaling in hereditary erythrocytosis or PV patients (Kapralova et al., 2014, 2016; Milosevic Feenstra et al., 2016; Wu et al., 2018). Heterozygous variant identified in the *JAK2* gene was present in only one affected family member and was observed in 22% of the sequence reads, as VAF was 0.22. The expected VAF for heterozygous germline variants is around 50%, with the distribution from 40 to 60%. For the somatic variants, VAF deviates from the germline and is usually lower (<40%), because the acquired variant is not present in all cells (Montgomery et al., 2018; Baer et al., 2019). This together indicates that JAK2 p.(Asn589Lys) could be of somatic origin; however, further testing of different type of sample (e.g., buccal swab) is required to confirm that. We already reported about two additional variants in the same individual with the JAK2 p.(Asn589Lys) variant, missense in the *EGLN1* gene and intronic in the *JAK2*, that showed segregation in the family (Kristan et al., 2021). In the present study, we used different filtration parameters for the selection of variants as in the previous study by Kristan et al. (2021); therefore, those two variants were filtered out. None of the two variants were directly involved in the development of erythrocytosis; however, the correlation between the variants, together with the novel, potential somatic variant in the *JAK2* gene, should be tested.

Besides known erythrocytosis genes, additional genes involved in the HIF and other erythropoiesis and iron metabolism pathways were included in the gene panel, in an attempt to discover novel candidate genes for the development of erythrocytosis. Unfortunately, we did not identify any novel disease-driver genes to be implicated in the erythropoiesis; however, we recognized a high incidence of heterozygous variants in *HFE* gene among the studied group of IE patients. This confirmed the observations of other researchers (Biagetti et al., 2018; Burlet et al., 2019; Gurnari et al., 2019). The difference of allele frequencies among our population of IE patients and the general population was larger for variant p.(Cys282Tyr). However, due to the small number of patients, this finding is of limited importance. Although homozygous variants p.(Cys282Tyr) have a high prevalence, the penetrance is low, as only 25–60% of patients develop clinical signs and only 55–82% of patients have increased serum ferritin level (Brissot et al., 2018). Low penetrance of causal *HFE* variants could indeed explain elevated ferritin values and saturation of transferrin in a small number of patients from our cohort. It would be useful to compare ferritin and saturation of transferrin values between larger cohorts of IE patients with *HFE* variants and IE *HFE*-wt patients, to see if there is any statistical significant difference. Biagetti et al. (2018) showed a higher frequency of elevated ferritin in mutated patients; however, the difference was not statistically significant due to a small group of patients (Biagetti et al., 2018). The pathophysiological mechanism between *HFE* variants and enhanced erythropoiesis is not clearly understood. In patients with high ferritin and transferrin values, this correlation could be explained by iron overload. However, this is not the case in patients with normal values; therefore, additional mechanisms must be involved. One possible explanation could be the linkage between the HIF pathway and iron metabolism. Schwartz et al. (2019) showed that low levels of hepatic iron regulator hepcidin lead to iron uptake and stabilization of EPAS1, through decreased prolyl-hydroxylase EGLN1 activity, as iron is an important co-factor for hydroxylation. Stabilized duodenal EPAS1 than activates target genes which regulate iron efflux (Schwartz et al., 2019). However, Schwartz et al. (2019) focused on the regulation of EPAS1 in duodenal enterocytes, and it would be interesting to study this mechanism also in other cell types and potentially find a connection with increased erythropoiesis.

With our sequencing approach, we achieved a high accuracy of sequenced bases, and high percentage of targeted regions had sufficient depth of coverage for germline variant calling. The sensitivity of our approach was high, as we accurately detected approximately 98% of variants calls in the reference NA12878 sample. Nevertheless, some limitations of the method were recognized and should be taken into consideration for future applications. The biggest limitation was insufficient coverage of some target regions. The majority of the target regions with poor coverage were seen in the first coding exons, as those regions are GC rich. Inadequate coverage in the first exon regions with high GC content is a common problem in the NGS approach, and usually other methods are applied to sufficiently sequence those regions (Chen et al., 2013). Another shortcoming of the applied method was the detection of variants in *HBA1* and *HBA2* genes, which was also observed by other researchers (Girodon et al., 2017; Filser et al., 2021). NGS, particularly read alignment for these two genes, is challenging, due to high sequence similarity. For future analysis of genes with high homology, a classical sequencing approach is suggested. One limitation of the study was also a low number of studied patients, but this was expected, as congenital erythrocytosis is a rare disorder. We identified only a few pathogenic variants and VUS with a likely causative mechanism in 16% (four out of 25) of patients with high Hb and Ht included in the study. This percentage is similar to the proportion of identified variants observed by other researchers, despite differences in the gene panel (Girodon et al., 2021). This indicates that the number of genes and gene regions included in the targeted panel must be extended. Also, we should consider the possibility that we missed larger structural variants, for instance, larger deletions, insertions, translocations, or copy number variations, that are harder to detect with limited targeted NGS or Sanger sequencing. To bypass this limitation, whole-exome sequencing (WES) or WGS needs to be applied in the future, especially the long-read sequencing approach for more sensitive and specific identification of larger variants.

CONCLUSION

Through the study, we examined the patients with suspected congenital erythrocytosis for the first time with the established targeted NGS in the Slovenian clinical setting. As far as we know, this is the first established erythrocytosis gene panel for NGS, which includes genes associated with enhanced erythropoiesis and iron uptake. We showed that targeted NGS was indeed useful to explore variants in cases with suspected congenital erythrocytosis; however, in only approximately 15% of the patients, clinically interesting variants were identified. We identified ECYT4 with a known pathogenic variant in *EPAS1* gene and provided a diagnosis for one patient enabling proper decision regarding prognosis, genetic counseling, and treatment. Three VUS were identified in the known erythrocytosis genes, i.e., *EPAS1*, *JAK2*, and *SH2B3*. Further functional studies are necessary in order to elucidate the mechanism of action and explore the effect of variants on the development of erythrocytosis. Additionally, the germline or somatic origin should be clarified for the variants in *JAK2* and *SH2B3*, as those two genes are commonly associated with the somatic type of erythrocytosis.

In several patients, we detected two heterozygous variants in *HFE* gene, which confirmed the observations of other researchers. However, for the comprehensive study of the involvement of the *HFE* variants in erythropoiesis, we need to perform additional research on a larger group of patients.

For the remaining patients with no identified variants, broader approaches like WGS and WES should be applied, to detect the possible genetic cause of congenital erythrocytosis.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because due to concerns regarding participant/patient anonymity. Requests to access the datasets should be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Slovenian National Medical Ethics Committee, Ministry of Health of the Republic of Slovenia. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AK, ND, and IPZ contributed to conception and design of the study. IPZ, SAD, MF, TP, and HP revised clinical data. AK, TR, TK, RK, AV, and ND contributed with design of targeted NGS approach. AK, TP, AM, and ŠŽ performed the NGS analysis. AK wrote the first draft of the manuscript. TP, AM, and ND wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This study was supported by the Slovenian Research Agency, Grant Nos. L3-9279 and P1-0390 and Young Researcher funding to AK and by University Medical Centre, Ljubljana, Grant Nos. 20170073 and 20200231.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.689868/full#supplementary-material

REFERENCES

Amberger, J. S., Bocchini, C. A., Scott, A. F., and Hamosh, A. (2019). OMIM.org: leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Res.* 47, D1038–D1043. doi: 10.1093/nar/gky1151

Anzej Doma, S., Drnovsek, E., Kristan, A., Fink, M., Sever, M., Podgornik, H., et al. (2021a). Diagnosis and management of non-clonal erythrocytosis remains challenging: a single centre clinical experience. *Ann. Hematol.* doi: 10.1007/s00277-021-04546-4544 [Epub ahead of print].

Anzej Doma, S., Kristan, A., Debeljak, N., and Preložnik Zupan, I. (2021b). Congenital erythrocytosis – a condition behind recurrent thromboses: a case report and literature review. *Clin. Hemorheol. Microcirculat*. doi: 10.3233/CH-211120 [Epub ahead of print].

Asif, S., Begemann, M., and Raza, S. (2019). Polycythemia in patients with hereditary hemochromatosis: real or myth? *J. Clin. Med. Res.* 11, 422–427. doi: 10.14740/jocmr3816

Badar, S., Busti, F., Ferrarini, A., Xumerle, L., Bozzini, P., Capelli, P., et al. (2016). Identification of novel mutations in hemochromatosis genes by targeted next generation sequencing in Italian patients with unexplained iron overload. *Am. J. Hematol.* 91, 420–425. doi: 10.1002/ajh.24304

Baer, C., Walter, W., Hutter, S., Twardziok, S., Meggendorfer, M., Kern, W., et al. (2019). "Somatic" and "pathogenic" - is the classification strategy applicable in times of large-scale sequencing? *Haematologica* 104, 1515–1520. doi: 10.3324/haematol.2019.218917

Barton, J. C., Bertoli, L. F., and Rothenberg, B. E. (2000). Peripheral blood erythrocyte parameters in hemochromatosis: evidence for increased erythrocyte hemoglobin content. *J. Lab. Clin. Med.* 135, 96–104. doi: 10.1016/s0022-2143(00)70026-70026

Baxter, E. J., Scott, L. M., Campbell, P. J., East, C., Fourouclas, N., Swanton, S., et al. (2005). Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet* 365, 1054–1061. doi: 10.1016/S0140-6736(05)71142-71149

Bento, C. (2018). Genetic basis of congenital erythrocytosis. *Int. J. Lab. Hematol.* 40, 62–67. doi: 10.1111/ijlh.12828

Bento, C., Almeida, H., Maia, T. M., Relvas, L., Oliveira, A. C., Rossi, C., et al. (2013). Molecular study of congenital erythrocytosis in 70 unrelated patients revealed a potential causal mutation in less than half of the cases (Where is/are the missing gene(s)?). *Eur. J. Haematol.* 91, 361–368. doi: 10.1111/ejh.12170

Bergant, G., Maver, A., Lovrecic, L., Cuturilo, G., Hodzic, A., and Peterlin, B. (2018). Comprehensive use of extended exome analysis improves diagnostic

yield in rare disease: a retrospective survey in 1,059 cases. *Genet. Med.* 20, 303–312. doi: 10.1038/gim.2017.142

Biagetti, G., Catherwood, M., Robson, N., Bertozzi, I., Cosi, E., McMullin, M. F., et al. (2018). HFE mutations in idiopathic erythrocytosis. *Br. J. Haematol.* 181, 270–272. doi: 10.1111/bjh.14555

Brissot, P., Pietrangelo, A., Adams, P. C., de Graaff, B., McLaren, C. E., and Loreal, O. (2018). Haemochromatosis. *Nat. Rev. Dis. Primers* 4:18016. doi: 10.1038/nrdp.2018.16

Burlet, B., Bourgeois, V., Buriller, C., Aral, B., Airaud, F., Garrec, C., et al. (2019). High HFE mutation incidence in idiopathic erythrocytosis. *Br. J. Haematol.* 185, 794–795. doi: 10.1111/bjh.15631

Camps, C., Petousi, N., Bento, C., Cario, H., Copley, R. R., McMullin, M. F., et al. (2016). Gene panel sequencing improves the diagnostic work-up of patients with idiopathic erythrocytosis and identifies new mutations. *Haematologica* 101, 1306–1318. doi: 10.3324/haematol.2016.144063

Chen, Y. C., Liu, T., Yu, C. H., Chiang, T. Y., and Hwang, C. C. (2013). Effects of GC bias in next-generation-sequencing data on de novo genome assembly. *PLoS One* 8:e62856. doi: 10.1371/journal.pone.0062856

Cukjati, M., Vaupotic, T., Rupreht, R., and Curin-Serbec, V. (2007). Prevalence of H63D, S65C and C282Y hereditary hemochromatosis gene mutations in Slovenian population by an improved high-throughput genotyping assay. *BMC Med. Genet.* 8:69. doi: 10.1186/1471-2350-8-69

Ellard, S., Baple, E. L., Berry, I., Forrester, N., Turnbull, C., Owens, M., et al. (2019). *ACGS Best Practice Guidelines for Variant Classification 2019.* Available online at: https://www.acgs.uk.com/media/11285/uk-practiceguidelines-for-variant-classification-2019-v1-0-3.pdf

Faria, R., Silva, B., Silva, C., Loureiro, P., Queiroz, A., Fraga, S., et al. (2016). Next-generation sequencing of hereditary hemochromatosis-related genes: novel likely pathogenic variants found in the Portuguese population. *Blood Cells Mol. Dis.* 61, 10–15. doi: 10.1016/j.bcmd.2016.07.004

Ferbo, L., Manzini, P. M., Badar, S., Campostrini, N., Ferrarini, A., Delledonne, M., et al. (2016). Detection of a rare mutation in the ferroportin gene through targeted next generation sequencing. *Blood Transfus* 14, 531–534. doi: 10.2450/2016.0286-215

Fernandez-Marmiesse, A., Gouveia, S., and Couce, M. L. (2018). NGS technologies as a turning point in rare disease research, diagnosis and treatment. *Curr. Med. Chem.* 25, 404–432. doi: 10.2174/0929867324666170718101946

Filser, M., Aral, B., Airaud, F., Chauveau, A., Bruce, A., Polfrit, Y., et al. (2021). Low incidence of EPOR mutations in idiopathic erythrocytosis. *Haematologica* 106, 299–301. doi: 10.3324/haematol.2019.244160

Fokkema, I. F., Taschner, P. E., Schaafsma, G. C., Celli, J., Laros, J. F., and den Dunnen, J. T. (2011). LOVD v.2.0: the next generation in gene variant databases. *Hum. Mutat.* 32, 557–563. doi: 10.1002/humu.21438

Furlow, P. W., Percy, M. J., Sutherland, S., Bierl, C., McMullin, M. F., Master, S. R., et al. (2009). Erythrocytosis-associated HIF-2α mutations demonstrate a critical role for residues C-terminal to the hydroxylacceptor proline. *J. Biol. Chem.* 284, 9050–9058. doi: 10.1074/jbc.M808737200

Gale, D. P., Harten, S. K., Reid, C. D. L., Tuddenham, E. G. D., and Maxwell, P. H. (2008). Autosomal dominant erythrocytosis and pulmonary arterial hypertension associated with an activating HIF2 (Schwartz et al.) mutation. *Blood* 112, 919–921. doi: 10.1182/blood-2008-04-153718

Gaspersic, J., Kristan, A., Kunej, T., Zupan, I. P., and Debeljak, N. (2020). Erythrocytosis: genes and pathways involved in disease development. *Blood Transfus* doi: 10.2450/2020.0197-120 Online ahead of print.

Geay, A., Aral, B., Bourgeois, V., Martin, P., Airaud, F., Garrec, C., et al. (2020). Diagnosis of exon 12-positive polycythemia vera rescued by NGS. *Clin. Case Rep.* 8, 790–792. doi: 10.1002/ccr3.2720

Girodon, F., Airaud, F., Garrec, C., Bezieau, S., and Gardie, B. (2017). Gene panel sequencing in idiopathic erythrocytosis. *Haematologica* 102:e30. doi: 10.3324/haematol.2016.158337

Girodon, F., Aral, B., Airaud, F., Garrec, C., Bezieau, S., and Gardie, B. (2021). "NGS analysis in hereditary erythrocytosis: a french experience," in *Proceedings of the MPN&MPNr-EuroNet Fifteenth Meeting: Diagnosis and therapy of Myeloproliferative Neoplasms and Related Hereditary Diseases*. Available online at: http://mpneuronet.com/events/ljubljana/

Gnanasambandan, K., and Sayeski, P. P. (2011). A structure-function perspective of Jak2 mutations and implications for alternate drug design strategies: the road not taken. *Curr. Med. Chem.* 18, 4659–4673. doi: 10.2174/092986711797379267

Griffith, M., Spies, N. C., Krysiak, K., McMichael, J. F., Coffman, A. C., Danos, A. M., et al. (2017). CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nat. Genet.* 49, 170–174. doi: 10.1038/ng.3774

Gurnari, C., Lombardi, A. M., Cosi, E., Biagetti, G., Buccisano, F., Franceschini, L., et al. (2019). Genetic analysis of erythrocytosis reveals possible causative and modifier gene mutations. *Br. J. Haematol.* 186, e100–e103. doi: 10.1111/bjh.15931

Hentze, M. W., Muckenthaler, M. U., Galy, B., and Camaschella, C. (2010). Two to tango: regulation of Mammalian iron metabolism. *Cell* 142, 24–38. doi: 10.1016/j.cell.2010.06.028

Kapralova, K., Horvathova, M., Kucerova, J., Pospisilova, D., Leroy, E., Constantinescu, S. N., et al. (2014). JAK2 E846D germline mutation associated with erythrocytosis causes increased and prolonged epo-induced activation of STAT5. *Blood* 124, 4008–4008. doi: 10.1182/blood.V124.21.4008.4008

Kapralova, K., Horvathova, M., Pecquet, C., Fialova Kucerova, J., Pospisilova, D., Leroy, E., et al. (2016). Cooperation of germ line JAK2 mutations E846D and R1063H in hereditary erythrocytosis with megakaryocytic atypia. *Blood* 128, 1418–1423. doi: 10.1182/blood-2016-02-698951

Keohane, C., McMullin, M. F., and Harrison, C. (2013). The diagnosis and management of erythrocytosis. *BMJ* 347:f6667. doi: 10.1136/bmj.f6667

Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Albarca Aguilera, M., Meyer, R., et al. (2019). VarSome: the human genomic variant search engine. *Bioinformatics* 35, 1978–1980. doi: 10.1093/bioinformatics/bty897

Kristan, A., Debeljak, N., and Kunej, T. (2019). Genetic variability of hypoxia-inducible factor alpha (HIFA) genes in familial erythrocytosis: analysis of the literature and genome databases. *Eur. J. Haematol.* 103, 287–299. doi: 10.1111/ejh.13304

Kristan, A., Gaspersic, J., Rezen, T., Kunej, T., Kolic, R., Vuga, A., et al. (2021). Genetic analysis of 39 erythrocytosis and hereditary hemochromatosis-associated genes in the Slovenian family with idiopathic erythrocytosis. *J. Clin. Lab. Anal.* 35:e23715. doi: 10.1002/jcla.23715

Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., et al. (2018). ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* 46, D1062–D1067. doi: 10.1093/nar/gkx1153

Lanktree, M. B., Sadikovic, B., Waye, J. S., Levstik, A., Lanktree, B. B., Yudin, J., et al. (2017). Clinical evaluation of a hemochromatosis next-generation sequencing gene panel. *Eur. J. Haematol.* 98, 228–234. doi: 10.1111/ejh.12820

Lee, F. S., and Percy, M. J. (2011). The HIF pathway and erythrocytosis. *Annu. Rev. Pathol.* 6, 165–192. doi: 10.1146/annurev-pathol-011110-130321

Lee, G., and Arcasoy, M. O. (2015). The clinical and laboratory evaluation of the patient with erythrocytosis. *Eur. J. Intern. Med.* 26, 297–302. doi: 10.1016/j.ejim.2015.03.007

Lenglet, M., Robriquet, F., Schwarz, K., Camps, C., Couturier, A., Hoogewijs, D., et al. (2018). Identification of a new VHL exon and complex splicing alterations in familial erythrocytosis or von Hippel-Lindau disease. *Blood* 132, 469–483. doi: 10.1182/blood-2018-03-838235

Liu, Q., Tong, D., Liu, G., Yi, Y., Zhang, D., Zhang, J., et al. (2017). HIF2A germline-mutation-induced polycythemia in a patient with VHL-associated renal-cell carcinoma. *Cancer Biol. Ther.* 18, 944–947. doi: 10.1080/15384047.2017.1394553

Luo, J. C., and Shibuya, M. (2001). A variant of nuclear localization signal of bipartite-type is required for the nuclear translocation of hypoxia inducible factors (1alpha, 2alpha and 3alpha). *Oncogene* 20, 1435–1444. doi: 10.1038/sj.onc.1204228

Ma, W., Kantarjian, H., Zhang, X., Yeh, C. H., Zhang, Z. J., Verstovsek, S., et al. (2009). Mutation profile of JAK2 transcripts in patients with chronic myeloproliferative neoplasias. *J. Mol. Diagn.* 11, 49–53. doi: 10.2353/jmoldx.2009.080114

Maslah, N., Cassinat, B., Verger, E., Kiladjian, J. J., and Velazquez, L. (2017). The role of LNK/SH2B3 genetic alterations in myeloproliferative neoplasms and other hematological disorders. *Leukemia* 31, 1661–1670. doi: 10.1038/leu.2017.139

McMullin, M. F. (2008). The classification and diagnosis of erythrocytosis. *Int. J. Lab. Hematol.* 30, 447–459. doi: 10.1111/j.1751-553X.2008.01102.x

McMullin, M. F. (2014). Secondary erythrocytosis. *Hematology* 19, 183–184. doi: 10.1179/1024533214z.000000000263

McMullin, M. F. (2016). Investigation and management of erythrocytosis. *Curr. Hematol. Malig. Rep.* 11, 342–347. doi: 10.1007/s11899-016-0334-331

McMullin, M. F., and Cario, H. (2016). LNK mutations and myeloproliferative disorders. *Am. J. Hematol.* 91, 248–251. doi: 10.1002/ajh.24259

McMullin, M. F., Harrison, C. N., Ali, S., Cargo, C., Chen, F., Ewing, J., et al. (2019a). A guideline for the diagnosis and management of polycythaemia vera. a British society for haematology guideline. *Br. J. Haematol.* 184, 176–191. doi: 10.1111/bjh.15648

McMullin, M. F. F., Mead, A. J., Ali, S., Cargo, C., Chen, F., Ewing, J., et al. (2019b). A guideline for the management of specific situations in polycythaemia vera and secondary erythrocytosis: a British society for haematology guideline. *Br. J. Haematol.* 184, 161–175. doi: 10.1111/bjh.15647

Mead, A. J., Rugless, M. J., Jacobsen, S. E., and Schuh, A. (2012). Germline JAK2 mutation in a family with hereditary thrombocytosis. *N. Engl. J. Med.* 366, 967–969. doi: 10.1056/NEJMc1200349

Milosevic Feenstra, J. D., Nivarthi, H., Gisslinger, H., Leroy, E., Rumi, E., Chachoua, I., et al. (2016). Whole-exome sequencing identifies novel MPL and JAK2 mutations in triple-negative myeloproliferative neoplasms. *Blood* 127, 325–332. doi: 10.1182/blood-2015-07-661835

Mlakar, U. (2008). Smernice za odkrivanje in Zdravljenje prave policitemije. *Zdrav Vestn.* 77, 11–14.

Montgomery, N. D., Selitsky, S. R., Patel, N. M., Hayes, D. N., Parker, J. S., and Weck, K. E. (2018). Identification of germline variants in tumor genomic sequencing analysis. *J. Mol. Diagn.* 20, 123–125. doi: 10.1016/j.jmoldx.2017.09.008

Oliveira, J. L., Coon, L. M., Frederick, L. A., Hein, M., Swanson, K. C., Savedra, M. E., et al. (2018). Genotype–phenotype correlation of hereditary erythrocytosis mutations, a single center experience. *Am. J. Hematol.* doi: 10.1002/ajh.25150 Online ahead of print.

Patnaik, M. M., and Tefferi, A. (2009). The complete evaluation of erythrocytosis: congenital and acquired. *Leukemia* 23, 834–844. doi: 10.1038/leu.2009.54

Percy, M. J., Beer, P. A., Campbell, G., Dekker, A. W., Green, A. R., Oscier, D., et al. (2008a). Novel exon 12 mutations in the HIF2A gene associated with erythrocytosis. *Blood* 111, 5400–5402. doi: 10.1182/blood-2008-02-137703

Percy, M. J., Furlow, P. W., Lucas, G. S., Li, X., Lappin, T. R. J., McMullin, M. F., et al. (2008b). A Gain-of-Function mutation in the HIF2A gene in familial erythrocytosis. *New Engl. J. Med.* 358, 162–168. doi: 10.1056/nejmoa073123

Perrotta, S., Stiehl, D. P., Punzo, F., Scianguetta, S., Borriello, A., Bencivenga, D., et al. (2013). Congenital erythrocytosis associated with gain-of-function HIF2A gene mutations and erythropoietin levels in the normal range. *Haematologica* 98, 1624–1632. doi: 10.3324/haematol.2013.088369

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., and Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* 47, D886–D894. doi: 10.1093/nar/gky1016

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. *Genet. Med.* 17, 405–424. doi: 10.1038/gim.2015.30

Schwartz, A. J., Das, N. K., Ramakrishnan, S. K., Jain, C., Jurkovic, M. T., Wu, J., et al. (2019). Hepatic hepcidin/intestinal HIF-2alpha axis maintains iron absorption during iron deficiency and overload. *J. Clin. Invest.* 129, 336–348. doi: 10.1172/JCI122359

Scott, L. M., Tong, W., Levine, R. L., Scott, M. A., Beer, P. A., Stratton, M. R., et al. (2007). JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. *N. Engl. J. Med.* 356, 459–468. doi: 10.1056/NEJMoa065202

Seif, F., Khoshmirsafa, M., Aazami, H., Mohsenzadegan, M., Sedighi, G., and Bahar, M. (2017). The role of JAK-STAT signaling pathway and its regulators in the fate of T helper cells. *Cell Commun. Signal.* 15:23. doi: 10.1186/s12964-017-0177-y

Spolverini, A., Pieri, L., Guglielmelli, P., Pancrazzi, A., Fanelli, T., Paoli, C., et al. (2013). Infrequent occurrence of mutations in the PH domain of LNK in patients with JAK2 mutation-negative 'idiopathic' erythrocytosis. *Haematologica* 98, e101–e102. doi: 10.3324/haematol.2013.090175

Stenson, P. D., Ball, E. V., Mort, M., Phillips, A. D., Shiel, J. A., Thomas, N. S., et al. (2003). Human Gene Mutation Database (HGMD): 2003 update. *Hum. Mutat.* 21, 577–581. doi: 10.1002/humu.10212

Tarade, D., Robinson, C. M., Lee, J. E., and Ohh, M. (2018). HIF-2alpha-pVHL complex reveals broad genotype-phenotype correlations in HIF-2alpha-driven disease. *Nat. Commun.* 9:3359. doi: 10.1038/s41467-018-05554-5551

Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., et al. (2019). COSMIC: the catalogue of somatic mutations in Cancer. *Nucleic Acids Res.* 47, D941–D947. doi: 10.1093/nar/gky1015

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

Tweedie, S., Braschi, B., Gray, K., Jones, T. E. M., Seal, R. L., Yates, B., et al. (2021). Genenames.org: the HGNC and VGNC resources in 2021. *Nucleic Acids Res.* 49, D939–D946. doi: 10.1093/nar/gkaa980

Ugo, V., Tondeur, S., Menot, M. L., Bonnin, N., Le Gac, G., Tonetti, C., et al. (2010). Interlaboratory development and validation of a HRM method applied to the detection of JAK2 exon 12 mutations in polycythemia vera patients. *PLoS One* 5:e8893. doi: 10.1371/journal.pone.0008893

UniProt Consortium. (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* 49, D480–D489. doi: 10.1093/nar/gkaa1100

Wu, Q. Y., Ma, M. M., Fu, L., Zhu, Y. Y., Liu, Y., Cao, J., et al. (2018). Roles of germline JAK2 activation mutation JAK2 V625F in the pathology of myeloproliferative neoplasms. *Int. J. Biol. Macromol.* 116, 1064–1073. doi: 10.1016/j.ijbiomac.2018.05.120

Conflict of Interest: RK and AV are employees of the Kemomed Ltd., Kemomed Research and Development.

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

frontiers in Genetics

CASE REPORT
published: 05 August 2021
doi: 10.3389/fgene.2021.684555

Case Report: Whole Exome Sequencing Revealed Two Novel Mutations of *PIEZO1* Implicated in Nonimmune Hydrops Fetalis

Yuan Chen[1†], *Ying Jiang*[1†], *Bangwu Chen*[2], *Yeqing Qian*[3], *Jiao Liu*[4], *Mengmeng Yang*[1], *Baihui Zhao*[1] *and Qiong Luo*[1*]

[1] Department of Obstetrics, Women's Hospital, Zhejiang University School of Medicine, Hangzhou, China, [2] Department of Obstetrics, Ninghai Maternal and Child health Care Hospital, Ningbo, China, [3] Department of Reproductive Genetics, Women's Hospital Zhejiang University School of Medicine, Hangzhou, China, [4] Department of Obstetrics, Lishui Maternal and Child Health Care Hospital, Lishui, China

Nonimmune hydrops fetalis (NIHF) is a serious and complex fetal condition. Prenatal diagnosis of hydrops fetalis is not difficult by ultrasound. However, determining the underlying etiology of NIHF remains a challenge which is essential to address for prenatal counseling. We extracted DNA from a proband prenatally diagnosed unexplained NIHF. Trio-whole exome sequencing (WES) was performed to filter candidate causative variants. Two gene mutations were identified as a compound heterozygous state in the proband. Both variants located on the *PIEZO1* gene: c.3895C > T, a missense mutation in exon 27 paternally inherited; c.4030_4032del, a maternally inherited in-frame deletion in exon 28. Both variants were first reported to be related to NIHF. *PIEZO1* gene mutations, leading to an autosomal recessive congenital lymphatic dysplasia, which can present as NIHF and partial or complete resolution postnatally. In conclusion, WES can aid in the elucidation of the genetic cause of NIHF and has a positive effect on the assessment of prognosis.

Keywords: whole exome sequencing, nonimmune hydrops fetalis, PIEZO1 gene, sanger sequencing, case report

OPEN ACCESS

Edited by:
Hu Hao,
The Sixth Affiliated Hospital of Sun Yat-sen University, China

Reviewed by:
Said El Shamieh,
Beirut Arab University, Lebanon
Shruti Marwaha,
Stanford University, United States

***Correspondence:**
Qiong Luo
luoq@zju.edu.cn

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

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

Received: *23 March 2021*
Accepted: *27 May 2021*
Published: *05 August 2021*

Citation:
Chen Y, Jiang Y, Chen B, Qian Y, Liu J, Yang M, Zhao B and Luo Q (2021) Case Report: Whole Exome Sequencing Revealed Two Novel Mutations of PIEZO1 Implicated in Nonimmune Hydrops Fetalis. Front. Genet. 12:684555. doi: 10.3389/fgene.2021.684555

INTRODUCTION

Nonimmune hydrops fetalis (NIHF) is a rare and serious congenital disease, accounting for about 90% of hydrops fetalis cases, with a prevalence estimated to be ∼0.03%∼0.06% (Norton et al., 2015). The underlying pathogenesis of NIHF is complicated, ranging from cardiovascular, chromosomal, hematologic, infectious, thoracic, twin-twin transfusion, urinary tract abnormalities, gastrointestinal, lymphatic dysplasia, tumors, skeletal dysplasias, syndromic, inborn errors of metabolism, miscellaneous, and other unknown causes. The prognosis of NIHF appears poor generally (mortality reported is over 50%) (Norton et al., 2015). Despite the rapid development of intrauterine therapy, the decisions and survival rates of prenatal management depend on the definable and treatable causes. Therefore, identifying the underlying cause of NIHF is particularly important for prenatal counseling. The Society of Maternal-Fetal Medicine (SMFM) published clinical guidelines for NIHF in 2015. Targeted ultrasonic evaluation, middle cerebral artery (MCA) Doppler, fetal karyotype or chromosomal microarray analysis, and polymerase chain reaction (PCR) for cytomegalovirus, parvovirus, toxoplasmosis are recommended as the first-line workup of NIHF (Norton et al., 2015). Even with standard evaluations, the cause remains unknown in 20% to 68% of cases as reported (Sparks et al., 2019).

In recent years, genetic testing has been widely used to identify the causes of fetal anomalies. Traditional chromosome karyotyping by G-banding technique accounted for 9%∼19% of fetal structural abnormalities. Chromosomal microarray analysis (CMA) further reveals additional findings in 6% of cases (Staebler et al., 2005; Wapner et al., 2012; Norton et al., 2015; Levy and Wapner, 2018). In a large and multicenter cohort study (65 NIHF cases were included), the underlying etiology of prenatally diagnosed NIHF was determined in 44% of cases, and a genetic cause was identified in 25% of those receiving standard genetic testing (Sparks et al., 2019). More recently, WES has been applied to explain those cases with normal results of karyotype and CMA. In a prospective cohort study enrolling 234 parents-fetus trios with structural anomalies but were negative in karyotype and chromosomal microarray findings. The genetic diagnostic yield of WES was approximately 10% (Petrovski et al., 2019). In another large case series of unexplained NIHF (127 cases were included), diagnostic genetic variants were identified in 29% of cases by WES (Sparks et al., 2020). A total of 131 genes were summarized with strong evidence for a relationship with NIHF and further categorized as follows: cardiovascular, inborn metabolic issues, hematological, lymphatic, skeletal, neuromuscular, syndromic disorders. Lymphatic disorders play an important part in the etiology of NIHF and comprise 15% of NIHF (Bellini et al., 2015). Based on strong and emerging evidence genes, the genetic variants associated with lymphatic malformations account for approximately 5.1% of NIHF, including *FLT4*, *CCBE1*, *ADAMITS3*, *PIEZO1*, *EPHB4*, *SOX18,* and *FOXC2* (Quinn et al., 2021).

PIEZO1 is known to encode a mechanically activated ion channel in the plasma membrane of several types of cells (including hepatic erythroblasts, fetal splenic plasma cells, and lymphatic vessels of fetal peritoneum) and highly conserved in mammalia (Coste et al., 2010; Andolfo et al., 2013). *PIEZO1* gain of function mutations is identified to be the cause of autosomal dominant dehydrated hereditary stomatocytosis (DHS) (OMIM 194380), which is a rare genetic cause of hemolytic anemia. Biallelic *PIEZO1* loss-of-function mutations were first reported in 2015 to relate with other different phenotypes. These include autosomal recessive congenital lymphatic dysplasia (GLD) (OMIM 616843), which is one rare cause of NIHF *in utero* and can present as generalized lymphoedema involving hydrothorax, hydropericardium, chylothorax, facial and extremities lymphoedema, intestinal or pulmonary lymphangiectasia (Fotiou et al., 2015; Lukacs et al., 2015).

More and more genetic defects are revealed to have a relationship with NIHF (Norton et al., 2015; Sparks et al., 2020). Some cases may be accompanied by characteristic manifestations which can be screened by prenatal sonographic scannings. Yet this is not always the case, such as in lymphatic disorders. In those circumstances, WES is considered a potential technique to aid in the systematic evaluation of unexplained NIHF cases (Sparks et al., 2020; Quinn et al., 2021).

In our study, we performed whole exome sequencing in a case of prenatally diagnosed NIHF. Two novel homozygous variants in *PIEZO1* were identified which were suspected to be the pathogenesis of lymphatic dysplasia with NIHF.

MATERIALS AND METHODS

Patient Information

A 33-year-old primipara presented at 30 weeks gestation with fetal hydrothorax examined by routine prenatal ultrasound. These parents were referred to our prenatal diagnosis center for further confirmation. No remarkable abnormalities of prior prenatal examinations during pregnancy were found, including normal nuchal fold thickness and low risk prenatal serum screening results (chromosomes 18, 21, and NTD). Both of the parents denied a family history of hydrops fetalis, congenital metabolic diseases, and lymphoedema. At 31 weeks gestation, a thorough fetal systematic ultrasound was performed by prenatal diagnostic ultrasound specialists and showed massive bilateral pleural effusion, generalized skin edema (skin thickness of head and neck region was 18 mm), minor ascites, and polyhydramnios (amniotic fluid index was 28.5 cm) (**Figures 1A–C**). No other structural anomalies were found. Then systemic evaluations about hydrops fetalis showed Rh-positive blood type, negative serologic testing for TORCH, normal blood urine routine examinations, hepatorenal function, and fetal echocardiogram. After comprehensive counseling, the couples declined *utero* fetal therapy including fetal thoracentesis or placement of thoracoamniotic shunt for drainage even though the underlying etiology of NIHF was unclear. For fear of poor prognosis, the parents terminated the pregnancy at 31 weeks. A female fetus was delivered with a birth weight of 2,610 g and presented as anasarca in appearance. Further autopsy and pathology information for the fetus was not available.

Ethics Approval

After giving full authorization and signing a written informed consent by both of the parents, whole exome sequencing (WES) (trio analysis of the proband, mother, and father) was carried out. Our study was approved by the Ethics Committee of the Women's Hospital, Zhejiang University, School of Medicine (approval number: IRB-20210124-R) and conformed with ethical standards of experiments on human subjects.

DNA Extraction

For DNA sampling, 5 ml peripheral whole blood was collected from the biological parents, and a medial thigh muscle tissue ($1 \times 1 \times 1$ cubic centimeter) was obtained from the aborted fetus. Then EDTA was added for anticoagulation. A Qiagen Blood DNA mini kit (Qiagen®, Germany) was applied to extract genomic DNA from 2 ml peripheral whole blood of the parents. In addition, a Genomic DNA Purification Kit (*Invitrogen*, cat. K0512, USA) was used for DNA extraction from the muscle sample. Then all of the DNA samples were preserved at −20°C. The remaining samples were stored at −80°C refrigeration.

Whole Exome Sequencing and Data Analysis

The detailed exome sequencing process for the proband sample and data filtering protocol have been described previously (Jiang et al., 2019). In brief, for targeted enrichment of target region sequences, the precapture library was enriched by multiple probe hybridization using Agilent SureSelect Human

Exon Sequence Capture Kit (Agilent, USA). The target region included the coding sequence of the target gene and its neighboring 10 bp intron region. The quality control testing of the sequencing data was achieved by Qubit (Agilent High Sensitivity DNA Kit, Agilent). The postcapture sequencing libraries were analyzed by Illumina DNA Standa ds and Primer Premix Kit (Kapa Biosystems, USA). All filtered variants by new-generation sequencing (NGS) using Illumina HiSeq 2,500 platform (Illumina, USA) were further validated by Sanger sequencing in the proband and parental samples. Fastp (v.0.12.0) was used for quality control of fastq files by using the -N 15-L 30 parameters and low quality reads were filtered out from all sequencing dates using a quality score of 20 (Q20).

The sequencing data were compared to the GRCh37/hG19 reference sequence of the human genome using BWA (v.0.7.17-r1188) with MEM align method. Then variants were called by GATK (v.4.0) and annotated to public databases by Annovar. Pre-treatment for variant calling included: 1) sorting the matched SAM file according to the chromosomal position, and converting it into the compressed binary file BAM by Picard-SortsAM. 2) Adding a sample ID for matched reads by Picard-Add Or Replace Read Groups. 3) Marking the same sequence created by PCR duplication, optical duplication by Picard-Mark Duplicates. 4) GATK-Baserecalibrator and GATK-ApplyBQSR were used for site quality correction according to the known SNP and InDel sites of 1,000 Genomes Project. 5) GATK-Haplotypecaller was used for variant identification and forming a VCF file. Susceptibility genes, mitochondrial genes, genes related to polygenetic diseases, and complex diseases were not included in this analysis.

Bioinformatics Analysis of Variants

The comparison for all filtered single nucleotide variants (SNVs) and small insertions/deletions with the population were conducted on several public databases, including Human Gene Mutation Database (HGMD) (http://www.hgmd.org), Clinvar (http://www.ncbi.nlm.nih.gov/clinvar), Exome Aggregation Consortium (ExAC) (http://exac.broadinstitute.org), Genome Aggregation Database (gnomAD), 1,000 Genomes (http://exac.broadinstitute.org) database, and Genome Asia 1,000k (http://www.genomeasia100k.com/) database, to exclude the mutation sites as known polymorphic sites. The genetic effects and evolutionary conservation of novel variants were further predicted and estimated using SIFT (http://sift.jcvi.org), PROVEAN, Polyphen2 (http://genetics.bwh.harvard.edu/pph2) and Mutation Taster (http://www.mutationtaster.org) programs. The domain information of the protein was retrieved from UniProt (https://www.uniprot.org/). The pathogenicity of identified variants was interpreted according to the American College of Medical Genetics and Genomics (ACMG) standards and guidelines (Richards et al., 2015). Multiple sequence alignment was conducted using NCBI HomoloGene and DNAMAN program to compare *PIEZO1* protein sequences with different species including *H. sapiens*, *P. troglodytes*, *M. mulatta*, *C. lupus*, *B. taurus*, *M. musculus*, and *R. norvegicus*. The variants interpretation and ACMG classification are based on the current understanding of the disease and variants at the time of documenting.

RESULTS

Coverage information for WES is provided in **Supplementary Table 1**. The trio-based WES detected two novel heterozygous mutations in the *PIEZO1* gene (NM_001142864.3) of the proband: a missense variant (exon 27, Chr16:88792765: c.3895C>T) and an in-frame deletion (exon 28, Chr16:88792029-88792031: c.4030_4032del) (**Table 1**). The family pedigree is presented in **Figure 1D**.

The c.3895C > T variant generated a single amino acid substitution from Arg to Cys at position 1,299 (p.R1299C) in exon 27 of the *PIEZO1* gene located at chromosome 16, which was paternally inherited (**Figure 2A**). To date, the variant c.3895C > T is not present in HGMD, Clinvar, ExAC, gnomAD, or the 1000 Genomes database (PM2). Bioinformatics analysis predicted the novel variant (NM_001142864.3: c.3895C > T, p.R1299C) to be deleterious by SIFT (score 0.000) and PROVEAN (score −6.84). Mutation Taster suggested that this mutation was located in a highly evolutionarily conserved site, leading to a change of amino acid sequence. Protein features might also be affected (PP3).

The c.4030_4032del variant detected from the proband was located in exon 28 of the *PIEZO1* gene, resulting in an in-frame deletion of glutamic acid at position 1,344 (p. E1344del) and a change in protein length (PM4). According to the results by Sanger sequencing, we found that the in-frame deletion variant was maternally inherited and the unaffected mother was heterozygous for this mutation (**Figure 2B**). c.4030_4032del have not been reported in the HGMD and Clinvar databases, although the gnomAD database reported an extremely minor population frequency of the mutation as 0.00002568 (PM2). However, no homozygote has been reported, and no carriers of the mutation have been found in East Asian populations. According to the information obtained from the Uniprot website, the amino acid at position 1,344 is located on a coiled coil domain (1,339–1,368) of the PIEZO1 ion channel pore. The coiled coil is a highly versatile folding motif that may determine many biological functions involving signal transduction, molecular recognition, protein refolding, and ion channel formation (Burkhard et al., 2001). One amino acid deletion was supposed to affect the stability or rigidity of the coiled coils motif. p.Ile1357Val, p.Ile1357Met, p.R1358P, and p.Q1361R, which are adjacent to position 1,344, have been previously described as pathogenic mutations leading to autosomal dominant dehydrated stomatocytosis (Albuisson et al., 2013; Kager et al., 2018; Picard et al., 2019). The two mutations associated with amino acids positions were both located in conserved functional domains, which were deeply conserved across species (**Figures 2C,D**).

DISCUSSION

By whole exome sequencing, we identified two heterozygous *de novo* variants of *PIEZO1* in an unexplained NIHF case by routine ultrasound examinations prenatally. One was a novel missense mutation in exon 27, inherited from the father of proband. The other was a maternally inherited in-frame deletion in exon 28. Both variants were firstly reported in the PIEZO1

TABLE 1 | Diagnostic variants of the proband and parents.

Sample	Gene	Exome	Chromosome	HGVS DNA	HGVS Protein	Heterozygosity	Inheritance	Classification
Proband	*PIEZO1*	Exon27	Chr16:88792765	c.3895C > T	p.R1299C	Heterozygous	paternal	VUS
	PIEZO1	exon28	Chr16:88792029-88792031	c.4030_4032del	p.E1344del	Heterozygous	Maternal	VUS
Father	*PIEZO1*	exon27	Chr16:88792765	c.3895C > T	p.R1299C	Heterozygous	–	VUS
	PIEZO1	exon28	Chr16:88792029-88792031	c.4030_4032del	p.E1344del	Not detected	–	–
Mather	*PIEZO1*	exon27	Chr16:88792765	c.3895C > T	p.R1299C	Not detected	–	–
	PIEZO1	exon28	Chr16:88792029-88792031	c.4030_4032del	p.E1344del	Heterozygous	–	VUS

HGVS, human genome variation society; VUS, variants of uncertain significance.

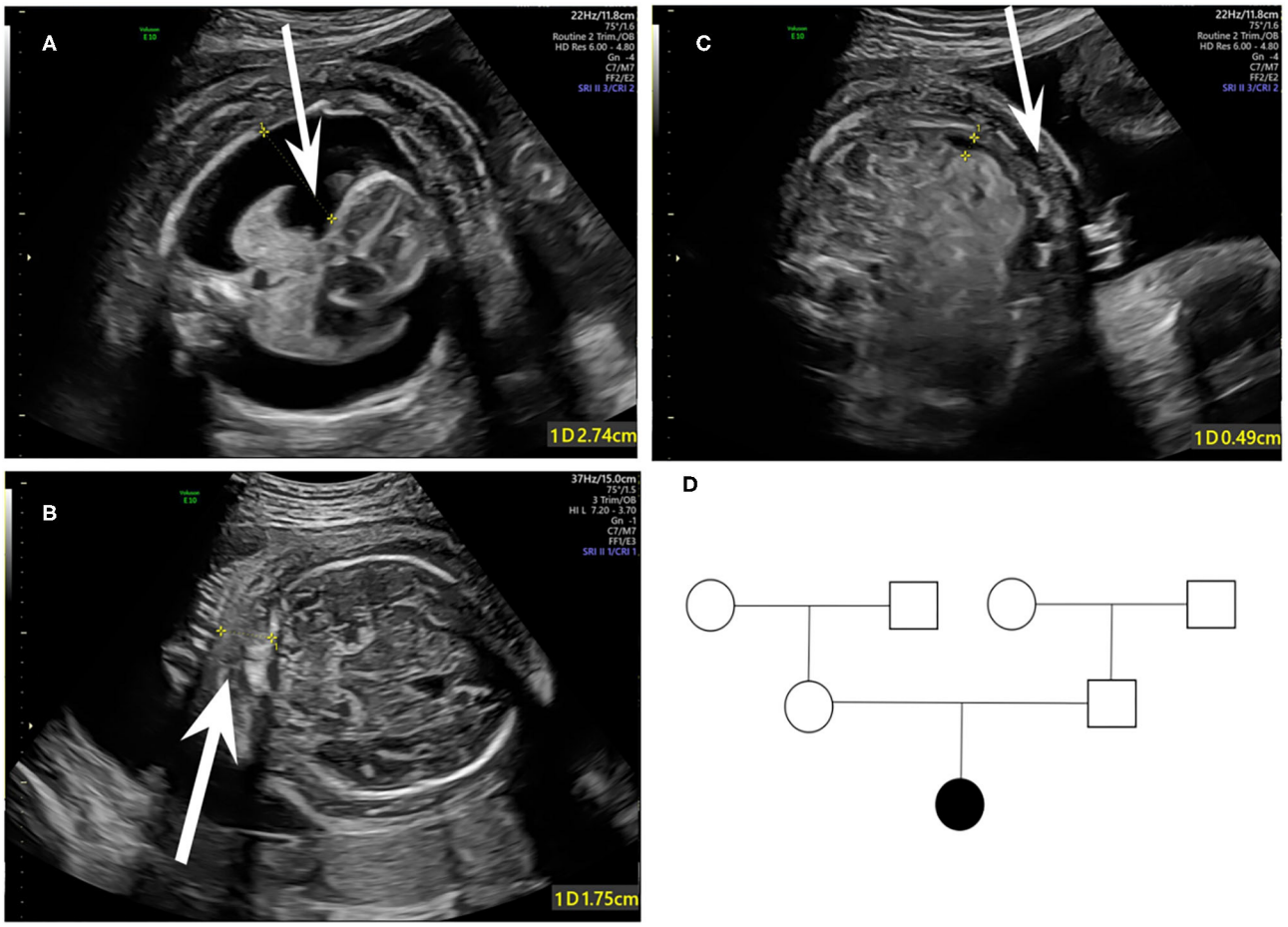


FIGURE 1 | Fetal sonographic findings at 31 gestation in the NIHF case with *PIEZO1* mutations: **(A)** Bilateral pleural effusion on axial view. **(B)** Skin edema on axial view. **(C)** Peritoneal effusion on axial view. **(D)** Pedigree of the family. The affected fetus is marked in black.

gene, and PIEZO1-associated autosomal recessive congenital lymphatic dysplasia is one rare cause of NIHF in *utero*.

Hydrops fetalis is defined as the presence of two or more abnormal fluid collections in the fetus by prenatal ultrasonography. The sonographic features include ascites, pleural effusions, pericardial effusions, and generalized skin edema (defined as skin thickness over 5 mm) (Norton et al., 2015). Hydrops fetalis is usually considered a lethal fetal condition and the affected mother may suffer from mirror syndrome (Désilets et al., 2018). According to the pathophysiology of hydrops fetalis, the etiology is classified as immune and nonimmune (Norton et al., 2015). With the routine examination of maternal Rh(D) type and application of preventive Rh(D) immunoglobulin, the incidence rate of red cell alloimmunization associated with hydrops fetalis is decreased. Combining thorough ultrasound scanning with fetal MRI, the majority of structural abnormalities associated with NIHF can be prenatally diagnosed (Norton et al., 2015). Even though, it is

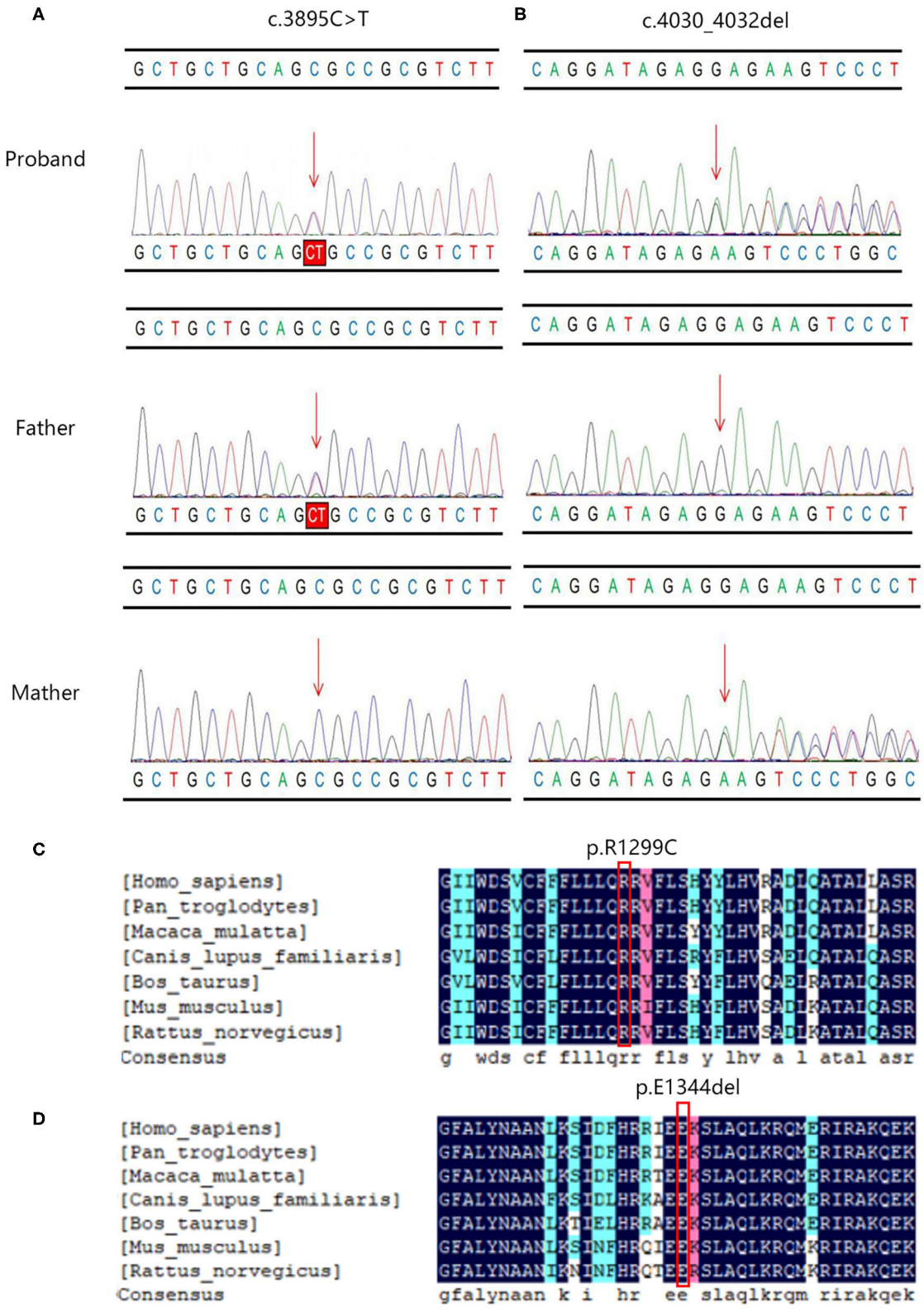


FIGURE 2 | The Sanger sequencing and conservation analysis of *PIEZO1* variants. The chromatograms demonstrated the compound heterozygous status of **(A)** c.3895C > T and **(B)** c.4030_4032del in the *PIEZO1* gene in the proband, proband's father, and proband's mother. The position of the variant or the corresponding wild type nucleotide are labeled with red arrows. Conservation analysis shows that the affected amino acids at positions 1,229 **(C)** and 1,344 **(D)** are located in a highly conserved region across species.

still a challenging task to elucidate the exact etiology of NIHF, which is essential for prenatal counseling and determination of the appropriate therapy if available (Norton et al., 2015). In recent years, whole exome sequencing (WES) has been gradually applied to identify *de-novo* mutations for evaluation of fetal abnormalities, which is insufficient by conventional molecular approaches such as karyotyping and chromosomal microarray analysis (CMA), including NIHF (Sparks et al., 2020; Quinn et al., 2021).

In our study, the affected proband is compound heterozygotes for a missense variant and a single amino acid in-frame deletion in the *PIEZO1* gene, both of which were not previously reported. Up to now, approximately 13 mutations in the *PIEZO1* gene have been reported to present as GLD associated with NIHF, which consist of six missense variants, four nonsense mutations, and three splicing mutations (Yates et al., 2017; More et al., 2020). The effects of *PIEZO1* variants present significantly more severe prenatally, while the complete or partial resolution of the edema postnatally has been reported in some patients. Fotiou et al. identified 10 *PIEZO1* mutations in six families which cosegregate with the GLD phenotype. The affected individuals suffer from *in utero* demise, complete resolution after birth, the reappearance of lymphoedema in peripheries at early childhood, or recurrent facial cellulitis (Quinn et al., 2021).

Along with development, the junctions of lymphatic endothelial cells transform from zippers into buttons and buttons initially appear at birth, which means a significant change occurs in lymphatics during prenatal and postnatal development. In addition, dexamethasone has promoted button formation during the early development period after birth (Yates et al., 2017). Another study analyzed *PIEZO1* protein expression in human fetal tissues. An obvious positive result was observed in fetal lymphatic vasculature while the immunoreactivity of *PIEZO1* protein was completely negative in healthy adult subjects (Andolfo et al., 2013). Based on these findings *in vitro*, one reasonable inference is that *PIEZO1* plays an important role in maintaining the lymphatic system *in utero* state, but has a decreased significance at neonate state. The clinical outcomes of *PIEZO1* associated NIHF range from fetal demise to a complete resolution after birth (Martin-Almedina et al., 2018). Whether various clinical outcomes are related to different mutations of the *PIEZO1* gene remains unclear. Questions remain regarding the kinds of variants in *PIEZO1* that result in poor outcomes or which can even be lethal, and which kind of variants have a good prognosis and are worthy of aggressive treatment.

In our report, the clinical manifestation of proband was severe hydrops fetalis without accompanied fetal structural abnormalities. This was determined by searching population frequency databases and undertaking bioinformatics analysis using several software packages, c.3895C > T and c.4030_4032del mutations are not suspected to be benign variations. Due to the lack of an identified human PIEZO1 channel structure by experimental methods, high-resolution cryo-electron microscopy (cryo-EM) structure and the mechanogating mechanism of the mouse PIEZO1 channel have been determined (Zhao et al., 2018). The residues H1300 to S1362 are assigned to the beam structure, which is considered to act as a lever-like mechanotransduction apparatus for converting the tension-induced conformational change of the peripheral blade to the intracellular gate. The two mutations are located beside and in the key structure and are supposed to affect the conductance and mechanosensitivity of the PIEZO1 channel. However, the exact biological effect needs to be further verified according to more functional experiments. In addition, when combined with the carrier state of the parents, as confirmed by Sanger sequencing, the inherited pattern is consistent with the *PIEZO1* associated autosomal recessive congenital lymphatic dysplasia. Thus, the phenotype of the patient could be considered as supporting evidence (PP4). Since well-established functional studies *in vivo* or *in vitro* of a damaging effect on the mutations were not available and reported information was rare, both variants were classified as a variant of unknown clinical significance (VUS) based on ACMG guidelines (c.3895C > T: PM2 + PP3 + PP4; c.4030_4032del: PM2+PM4+PP4).

In conclusion, our report suggested that two novel variants (c.3895C > T and c.4030_4032del) identified in *PIEZO1* could be associated with NIHF, which might be a severe phenotype of GLD *in utero*. These results provide additional information that reveals the specific mechanisms underlying PIEZO1-related lymphatic development and function. Our data also advocates the application of WES to further detect the underlying genetic cause of unexplained NIHF by standard genetic testing, which is important for making an accurate diagnosis for the precise evaluation of the recurrence risk and prognosis, detailed prenatal counseling, and high-quality perinatal and postnatal management.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the Women's Hospital, Zhejiang University, School of Medicine. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

YC, YJ, and QL designed the study and wrote the manuscript. MY and BC undertook patients care, collecting samples and clinical information. YC, YQ, and BZ undertook

bioinformatics analysis. JL and QL revised the manuscript. All authors reviewed the manuscript and approved the final version.

FUNDING

This work was supported by the Natural Science Foundation of Zhejiang Province (LY20H040009, LQ20H040008), and the Scientific Research Foundation of the National Health Commission (WKJ-ZJ-2126).

ACKNOWLEDGMENTS

The authors would like to thank the staff at the Women's Hospital, Zhejiang University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.684555/full#supplementary-material

REFERENCES

Albuisson, J., Murthy, S. E., Bandell, M., Coste, B., Louis-Dit-Picard, H., Mathur, J., et al. (2013). Dehydrated hereditary stomatocytosis linked to gain-of-function mutations in mechanically activated PIEZO1 ion channels. *Nat. Commun.* 4:1884. doi: 10.1038/ncomms3440

Andolfo, I., Alper, S. L., De Franceschi, L., Auriemma, C., Russo, R., De Falco, L., et al. (2013). Multiple clinical forms of dehydrated hereditary stomatocytosis arise from mutations in PIEZO1. *Blood* 121, 3925–3935, s3921–3912. doi: 10.1182/blood-2013-02-482489

Bellini, C., Donarini, G., Paladini, D., Calevo, M. G., Bellini, T., Ramenghi, L. A., et al. (2015). Etiology of nonimmune hydrops fetalis: An update. *Am. J. Med. Genet. A.* 167a, 1082–1088. doi: 10.1002/ajmg.a.36988

Burkhard, P., Stetefeld, J., and Strelkov, S. V. (2001). Coiled coils: a highly versatile protein folding motif. *Trends Cell. Biol.* 11, 82–88. doi: 10.1016/S0962-8924(00)01898-5

Coste, B., Mathur, J., Schmidt, M., Earley, T. J., Ranade, S., Petrus, M. J., et al. (2010). Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels. *Science* 330, 55–60. doi: 10.1126/science.1193270

Désilets, V., De Bie, I., and Audibert, F. (2018). No. 363-investigation and management of nonimmune fetal hydrops. *J. Obstet. Gynaecol. Can.* 40, 1077–1090. doi: 10.1016/j.jogc.2017.12.011

Fotiou, E., Martin-Almedina, S., Simpson, M. A., Lin, S., Gordon, K., Brice, G., et al. (2015). Novel mutations in PIEZO1 cause an autosomal recessive generalized lymphatic dysplasia with nonimmune hydrops fetalis. *Nat. Commun.* 6:8085. doi: 10.1038/ncomms9085

Jiang, Y., Qian, Y. Q., Yang, M. M., Zhan, Q. T., Chen, Y., Xi, F. F., et al. (2019). Whole-exome sequencing revealed mutations of MED12 and EFNB1 in fetal agenesis of the corpus callosum. *Front. Genet.* 10:1201. doi: 10.3389/fgene.2019.01201

Kager, L., Jimenez Heredia, R., Hirschmugl, T., Dmytrus, J., Krolo, A., Müller, H., et al. (2018). Targeted mutation screening of 292 candidate genes in 38 children with inborn haematological cytopenias efficiently identifies novel disease-causing mutations. *Br. J. Haematol.* 182, 251–258. doi: 10.1111/bjh.15389

Levy, B., and Wapner, R. (2018). Prenatal diagnosis by chromosomal microarray analysis. *Fertil. Steril.* 109, 201–212. doi: 10.1016/j.fertnstert.2018.01.005

Lukacs, V., Mathur, J., Mao, R., Bayrak-Toydemir, P., Procter, M., Cahalan, S. M., et al. (2015). Impaired PIEZO1 function in patients with a novel autosomal recessive congenital lymphatic dysplasia. *Nat. Commun.* 6:8329. doi: 10.1038/ncomms9329

Martin-Almedina, S., Mansour, S., and Ostergaard, P. (2018). Human phenotypes caused by PIEZO1 mutations; one gene, two overlapping phenotypes? *J. Physiol.* 596, 985–992. doi: 10.1113/JP275718

More, T. A., Dongerdiye, R., Devendra, R., Warang, P. P., and Kedar, P. S. (2020). Mechanosensitive Piezo1 ion channel protein (PIEZO1 gene): update and extended mutation analysis of hereditary xerocytosis in India. *Ann. Hematol.* 99, 715–727. doi: 10.1007/s00277-020-03955-1

Norton, M. E., Chauhan, S. P., and Dashe, J. S. (2015). Society for maternal-fetal medicine (SMFM) clinical guideline #7: nonimmune hydrops fetalis. *Am. J. Obstet. Gynecol.* 212, 127–139. doi: 10.1016/j.ajog.2014.12.018

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

Picard, V., Guitton, C., Thuret, I., Rose, C., Bendelac, L., Ghazal, K., et al. (2019). Clinical and biological features in PIEZO1-hereditary xerocytosis and Gardos channelopathy: a retrospective series of 126 patients. *Haematologica* 104, 1554–1564. doi: 10.3324/haematol.2018.205328

Quinn, A. M., Valcarcel, B. N., Makhamreh, M. M., Al-Kouatly, H. B., and Berger, S. I. (2021). A systematic review of monogenic etiologies of nonimmune hydrops fetalis. *Genet. Med.* 23, 3–12. doi: 10.1038/s41436-020-00967-0

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. *Genet. Med.* 17, 405–424. doi: 10.1038/gim.2015.30

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

Sparks, T. N., Thao, K., Lianoglou, B. R., Boe, N. M., Bruce, K. G., Datkhaeva, I., et al. (2019). Nonimmune hydrops fetalis: identifying the underlying genetic etiology. *Genet. Med.* 21, 1339–1344. doi: 10.1038/s41436-018-0352-6

Staebler, M., Donner, C., Van Regemorter, N., Duprez, L., De Maertelaer, V., Devreker, F., et al. (2005). Should determination of the karyotype be systematic for all malformations detected by obstetrical ultrasound? *Prenat. Diagn.* 25, 567–573. doi: 10.1002/pd.1187

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. doi: 10.1056/NEJMoa1203382

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

Zhao, Q., Zhou, H., Chi, S., Wang, Y., Wang, J., Geng, J., et al. (2018). Structure and mechanogating mechanism of the Piezo1 channel. *Nature* 554, 487–492. doi: 10.1038/nature25743

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