Zhou *et al. BMC Medical Genomics* (2024) 17:230
https://doi.org/10.1186/s12920-024-02005-3

BMC Medical Genomics

RESEARCH Open Access

Whole exome sequencing analysis of 167 men with primary infertility

Haiyan Zhou[1†], Zhaochu Yin[1†], Bin Ni[1], Jiwu Lin[1], Shuwei Luo[2,3*†] and Wanqin Xie[1,3*†]

Abstract

Background Spermatogenic failure is one of the leading causes of male infertility and its genetic etiology has not yet been fully understood.

Methods The study screened a cohort of patients ($n = 167$) with primary male infertility in contrast to 210 normally fertile men using whole exome sequencing (WES). The expression analysis of the candidate genes based on public single cell sequencing data was performed using the R language Seurat package.

Results No pathogenic copy number variations (CNVs) related to male infertility were identified using the the GATK-gCNV tool. Accordingly, variants of 17 known causative (five X-linked and twelve autosomal) genes, including *ACTRT1*, *ADAD2*, *AR*, *BCORL1*, *CFAP47*, *CFAP54*, *DNAH17*, *DNAH6*, *DNAH7*, *DNAH8*, *DNAH9*, *FSIP2*, *MSH4*, *SLC9C1*, *TDRD9*, *TTC21A*, and *WNK3*, were identified in 23 patients. Variants of 12 candidate (seven X-linked and five autosomal) genes were identified, among which *CHTF18*, *DDB1*, *DNAH12*, *FANCB*, *GALNT3*, *OPHN1*, *SCML2*, *UPF3A*, and *ZMYM3* had altered fertility and semen characteristics in previously described knockout mouse models, whereas *MAGEC1*,*RBMXL3*, and *ZNF185* were recurrently detected in patients with male factor infertility. The human testis single cell-sequencing database reveals that *CHTF18*, *DDB1* and *MAGEC1* are preferentially expressed in spermatogonial stem cells. *DNAH12* and *GALNT3* are found primarily in spermatocytes and early spermatids. *UPF3A* is present at a high level throughout spermatogenesis except in elongating spermatids. The testicular expression profiles of these candidate genes underlie their potential roles in spermatogenesis and the pathogenesis of male infertility.

Conclusion WES is an effective tool in the genetic diagnosis of primary male infertility. Our findings provide useful information on precise treatment, genetic counseling, and birth defect prevention for male factor infertility.

Keywords Spermatogenic failure, Whole exome sequencing, Male infertility, Azoospermia, Spermatogenesis

[†]Haiyan Zhou and Zhaochu Yin share first-authorship.

[†]Shuwei Luo and Wanqin Xie share senior authorship.

*Correspondence:
Shuwei Luo
luobosing@gmail.com
Wanqin Xie
wanqinxie@126.com

[1]National Health Commission Key Laboratory of Birth Defects for Research and Prevention, Hunan Provincial Maternal and Child Health Care Hospital, Changsha, Hunan 410008, China
[2]Center for Reproductive Medicine, Hunan Provincial Maternal and Child Health Care Hospital, Changsha, Hunan 410008, China
[3]Hunan Provincial Maternal and Child Health Care Hospital, No. 53 Xiangchun Road, Changsha, Hunan 410008, China

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Introduction

Worldwide, infertility affects 8–12% of fertile-aged couples, and male factor infertility accounts for approximately 50% of the condition [1, 2]. Primary spermatogenic failure is a major cause of male infertility [3]. Taking advantage of next-generation sequencing technologies (e.g., whole exome sequencing, whole genome sequencing), a growing number of monogenic variants have been associated with human spermatogenic failure [4, 5].

Programmed differentiation of spermatogonial stem cells (SSCs) involves over 2,000 genes enriched in human testis [6]. However, only a small proportion of these genes have been linked to spermatogenic failure/defects [7]. The genetic causes of male factor infertility remain largely unknown.

In the present study, we retrospectively analyzed the blood samples of 167 men with primary infertility using whole exome sequencing (WES) combined with gene expression analysis based on single cell sequencing data [8]. The results are expected to increase our knowledge on monogenic variants associated with spermatogenic failure/ defects, as well as to provide new information on gene diagnosis, genetic counseling, and individualized treatment for primary male infertility.

Materials and methods

Clinical samples

Patients with primary male infertility and altered semen parameters (e.g., azoospermia, asthenozoospermia, oligozoospermia, oligoasthenoteratozoospermia) according to the WHO guidelines (2010) were enrolled for the study at the cytogenetics laboratory of a tertiary medical center. The eligible patients had a normal karyotype and a negative result of AZF deletion test. Patients with a medical history of ejaculatory duct obstructions, cryptorchidism, epididymis and vas deferens, orchitis, hypogonadism, chemotherapy or radiation therapy were excluded for the study.The blood samples collected from the veins were stored at -80 C until use. .

WES

According to the manufacturer's instructions, genomic DNA was isolated from peripheral blood leukocytes of all participants using a QIAamp mini DNA kit (Qiagen, Hilden, Germany) for library construction. WES was performed at the BGI Clinical Testing Center (Shenzhen, Guangdong, China). Exons were captured using the BGI-Exome kit V4 and sequenced by BGI-seq 500 with 100 bp paired end reads at an average sequencing depth of 100× and average coverage of 99.5%. Low-quality reads were removed by SOAPnuke, and then the reads were mapped to the human genome reference (GRCh38/hg38) by the Burrows-Wheeler aligner (BWA-MEM, version 0.7.10). Variants Calling was performed using the Genome Analysis Tool Kit (GATK, version 3.3). The ANNOVAR tool was used to annotate and classify all variants.

Variants filtering and prioritization

All variants were filtered based on their frequency in public databases including gnomAD (http://gnomad.broadinstitute.org/), the 1000 Genome Project (http://browser.1000genomes.org), dbSNP (http://www.ncbi.nlm.nih.gov/snp) and BGI internal database; and variants with minor allele frequency (MAF) < 0.01 were retained. Autosomal homozygous or double heterozygous variants, and X-Chromosomal variants were considered in this study. Prioritization of candidate variants was based on a recent review on 657 gene-disease relationships (GDR) for monogenic causes of human male infertility [4], and the genes associated with male infertility in the Mouse Genome Informatics database (MGI: http://www.informatics.jax.org/). Online tools including SIFT (http://provean.jcvi.org/), PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/), MutationTaster (http://www.mutationtaster.org/), CADD (https://cadd.gs.washington.edu/snv), and GERP (http://mendel.stanford.edu/SidowLab/downloads/gerp/) were used to predict the damage of candidate variants.The exome data of 210 normally fertile men served as control to further screen the candidate variants.

CNV analysis using WES data

The GATK-gCNV pipeline [9] was used to analyze the germline CNVs from WES data of 167 patients in the study. CNVs with a quality score (QS) > 1000 were considered for annotation using DECIPHER database (https://decipher.sanger.ac.uk/). In order to identify potential male infertility-related genes within the pathogenic CNVs, coding genes covered by pathogenic CNVs were filtered based on the human candidate genes for male infertility [4].

Sanger sequencing

Forward and reverse primers for PCR amplification were designed using the Primer3 program based on the human genome reference sequence GRCh38/hg38. The purified PCR product was loaded for capillary electrophoresis on an ABI 3730XL DNA analyzer (ABI PRISM, Foster City, CA, USA).

Differential expression analysis of candidate genes

Single cell sequencing data (https://ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE109037) were downloaded and subjected to quality control and filtering following the scripts provided by the submitter [8]. Gene expression analysis was performed using the R language (Version 4.2.3) Seurat package (V4.0).

Results

Clinical features

The blood samples of a total of 167 patients, with a mean age ± standard deviation of 30.4 ± 4.8 yrs and a range of 22 to 49 years, were retrospectively analyzed. Based on medical records, semen characteristics of these patients include non-obstructive azoospermia (NOA) (n=58, 34.7%), asthenozoospermia (n=27, 16.2%), oligoasthenzoospermia (n=63, 37.7%), oligoasthenoteratozoospermia (n=6, 3.6%), oligozoospermia (n=10, 6.0%), teratozoospermia (n=3, 1.8%). All patients had a normal karyotype and no Y-chromosome infertility.

Genetic findings

In this study, variants of the genes that have been substantially demonstrated as causative genes of human male infertility were selected in priority for diagnostic purpose. Based on a recent review on 657 gene-disease relationships (GDR) for monogenic causes of human male infertility [4] and a complementary literature search, we identified deleterious variants of 17 known causative (five X-linked and twelve autosomal) genes, including *ACTRT1*, *ADAD2*, *AR*, *BCORL1*, *CFAP47*, *CFAP54*, *DNAH17*, *DNAH6*, *DNAH7*, *DNAH8 DNAH9*, *FSIP2*, *MSH4*, *SLC9C1*, *TDRD9*, *TTC21A*, *and WNK3* in 23 patients (Table 1). The mutation spectrum of these known causative genes in patients included missense (64.1%), splice site (17.9%), frameshift (5.1%), nonsense (5.1%), start codon (2.6%), and in-frame deletion (2.6%) mutations, with a predominance of missense mutations resulting in amino acid substitution. We used the exome data from 210 normally fertile men as control. X-linked variants that were present in both male factor infertility patients and normal fertile controls were discarded. With respect to autosomal recessive variants *DNAH17* c.12488T > G p.(Phe4163Cys), c.5278G > T p.(Asp1760Tyr), and c.3601 C > T p.(Arg1201Trp), *DNAH8* c.6565 A > G p.(Met2189Val), and *FSIP* c.3365 C > G p.(Ser1122Cys), though one or two heterozygous carriers of these variants in the control group, they were still selected as causal variants for patients with male factor infertility in this study (Table 1).

Based on the genes that have been linked to male infertility in mouse models or recurrently affected in patients, we selected 12 candidate genes for further investigation, involving seven X-linked and five autosomal genes, namely *CHTF18*, *DDB1*, *DNAH12*, *FANCB*, *GALNT3*, *MAGEC1*, *OPHN1*, *RBMXL3*, *SCML2*, *UPF3A*, *ZMYM3*, *ZNF185* (Table 2). The mutation spectrum of candidate genes in patients included missense (50%), nonsense (15.4%), frameshift (11.5%), splice site (19.2%) and intron/exon boundary deletion (3.8%) mutations. These variants were not found in the 210 fertile men, except two variants *DNAH12* c.6286 C > T p.(Gln2096X) and: c.6021−3 A > G, which had one heterozygous carrier, respectively. Considering the autosomal recessive inheritance pattern of the *DNAH12* gene, the two variants were still considered as candidate causal variants for male factor infertility.

The results of the Sanger sequencing validation for the identified variants are shown in the Supplementary Data Sheet 1.

A total of 29 CNVs were identified using the GATK-gCNV tool, and none of them were directly related to male infertility in DECIPHER (accessed on August 27th, 2024; Supplementary Data Sheet 2). In one patient (M1241), a heterozygous deletion of chromosome 10q26.3 resulted in loss of a part of the spermatogenic failure gene *SYCE1*. This CNV was not considered to be the cause of infertility in the patient based on the recessive inheritance pattern of *SYCE1*.

Differential expression of candidate genes

We explored the differential expression of candidate genes at the developmental stages of spermatogenesis by referencing the human testis single cell-sequencing database [8]. Of the 12 candidate genes, *CHTF18*, *DDB1* and *MAGEC1* were preferentially expressed in spermatogonial stem cells (Fig. 1B, C and G). *DNAH12* and *GALNT3* were found primarily in spermatocytes and early spermatids (Fig. 1D and F). Obviously, *UPF3A* was present at a high level throughout spermatogenesis except in elongating spermatids (Fig. 1K). The testicular expression profiles of the genes mentioned above underlie their potential roles in spermatogenesis and the pathophysiology of male infertility. Other candidate genes were found to be expressed only in a small part of the cells during spermatogenesis and remain to be further investigated (Fig. 1).

Discussion

Several previous studies have explored the genetic etiology of spermatogenic failure using WES, with a diagnostic yield ranging from 3.2% (10 out of 314 patients) to 47.6% (10 out of 21 patients) [31, 37–41]. In a more recent study that screened likely pathogenic and pathogenic (LP/P) variants of a male infertility gene panel in 521 patients with idiopathic primary spermatogenic failure, a molecular diagnosis rate of 12% (n=64) was achieved [42]. In the present study, we identified 23 out of 167 patients with primary male infertility to carry candidate variants of known causative genes for spermatogenic failure, giving a diagnostic rate of 13.8%. We reasoned that performance of WES for detection of causal mutations of spermatogenic failure might be affected by the genetic background and inclusion criteria (e.g. pathological type, severity) of patients. Definitely, accumulating

Table 1 Variants of known causative genes for male infertility detected in the study

Gene	Patient	Semen characteristics	Variation	Zygosity	SIFT/PolyPhen/MutationTaster/CADD/GERP	Reference
ACTRT1	M1555	Oligoasthenoteratozoosp-ermia (OAT)	NM_138289.4:c.169G > A;p.(Val57Met)	Hemi	D/P/N/14.14/-0.162/	Sha et al., 2021 [10]
ADAD2	M1292	Non-obstructive azoospermia (NOA)	NM_001145400.2:c.1459 C > T;p.(Leu487Phe)	Het	D/D/D/27/4.36/	Dai et al., 2023 [11]
			NM_001145400.2:c.1557 C > G;p.(Cys519Trp)	Het	D/D/D/24.3/-0.113/	
AR	M1472	NOA	NM_001011645.3:c.2T > G;p.(M1?)	Hemi	D/B/D/8.163/-0.63/	Giwercman et al., 2001 [12]
BCORL1	M1201	Oliogasthenozoospermia	NM_001184772.3:c.56G > A;p.(Arg19Gln)	Hemi	D/D/D/25.6/4.72/	Lu et al., 2021 [13]
CFAP47	M1294	NOA	NM_001304548.2:c.1744G > T;p.(Asp582Tyr)	Hemi	D/D/N/22/-0.786/	Liu et al., 2021 [14]
	M1568	OAT	NM_001304548.2:c.1763 C > T;p.(Pro588Leu)	Hemi	D/D/D/26.1/5.35/	
	M850	Oliogasthenozoospermia	NM_001304548.2:c.2609G > A;p.(Arg870Gln)	Hemi	D/D/N/13.88/3.25/	
DNAH17	M1233	NOA	[a] NM_173628.4:c.12488T > G;p(.Phe4163Cys)	Het	NA/P/D/28.3/5.07/	Whitfield et al., 2019 [15]
			NM_173628.4:c.3644G > C;p.(Arg1215Thr)	Het	D/NA/D/24/4.37/	
	M1252	NOA	NM_173628.4c.11,920 C > A;p.(Pro3974Thr)	Het	D/P/D/20.6/4.23/	
			NM_173628.4:c.10,331 C > Al; p.(Ala3444Asp)	Het	NA/P/D/25.6/5.11/	
	M1824	NOA	[a]NM_173628.4:c.5278G > T;p.(Asp1760Tyr)	Het	NA/NA/D/34/4.65/	
			[a]NM_173628.4:c.3601 C > T;p.(Arg1201Trp)	Het	D/NA/D/27.4/-2.85/	
DNAH6	M1705	Oliogasthenozoospermia	NM_001370.2:c.8888G > A;p.(Arg2963His)	Het	D/D/D/32/5.58/	Tu et al., 2019 [16]
			NM_001370.2:c.10025G > A;p.(Arg3342His)	Het	D/D/D/28.9/5.86/	
DNAH7	M807	Oliogasthenozoospermia	NM_018897.3:c.11,018 A > G;p.(Tyr3673Cys)	Het	D/P/D/24.4/4.98/	Wei et al., 2021 [17]
			NM_018897.3:c.7264 C > A;p.(Pro2422Thr)	Het	D/D/D/26.8/5.34/	
DNAH8	M1237	Oliogasthenozoospermia	[a]NM_001206927.2:c.6565 A > G;p.(Met2189Val)	Het	D/D/D/26.1/5.87/	Liu et al., 2020 [18]
			NM_001206927.2:c.12,496 C > T;p.(Arg4166X)	Het	NA/NA/A/48/3.19/	
	M1761	Asthenozoospermia	NM_001206927.2:c.391−2 A > C	Het	NA/NA/D/24.6/5.66/	
			NM_001206927.2:c.9964-1G > A	Het	NA/NA/D/25/5.37/	
DNAH9	M1470	NOA	NM_004662.2:c.112 C > T;p.(Arg38Trp)	Het	D/D/D/27.5/3.17/	Tang et al., 2021 [19]
			NM_004662.2:c.1016 A > T;p.(His339Leu)	Het	D/D/D/24.3/5.19/	
FSIP2	M1596	Oliogasthenozoospermia	NM_173651.4:c.17G > A;p.(Gly6Asp)	Het	D/NA/N/19.9/3.66/	Martinez et al.,2018 [20]
			[a]NM_173651.4:c.3365 C > G;p.(Ser1122Cys)	Het	D/NA/N/20.6/3.44/	
	M1793	Oliogasthenozoospermia	NM_173651.4:c.1347 + 1G > A	Het	NA/NA/N/14.1/2.52/	
			NM_173651.4:c.5235del; p.(Glu1747del)	Het	NA/NA/NA/NA/NA/	
MSH4	M1295	NOA	NM_002440.4:c.2107 + 5G > A	Hom	NA/NA/NA/NA/NA/	Wyrwoll et al.,2021 [21]
SLC9C1	M1497	Asthenozoospermia	NM_183061.3:c.1711 C > T;p.(Gln571X)	Het	NA/NA/A/34/1.12/	Cavarocchi et al., 2021 [22]
			NM_183061.3:c.1012_1015del; p.(Ile338Phe fs*2)	Het	NA/NA/NA/NA/NA/	
TDRD9	M852	NOA	NM_153046.3:c.2737T > C;p.(Phe913Leu)	Hom	D/D/D/32/5.09/	Arafat et al., 2017 [23]
TTC21A	M1717	Asthenozoospermia	NM_001105513.3:c.1427T > A;p.(Leu476Gln)	Het	D/D/D/27.2/4.74/	Liu et al., 2019 [24]
			NM_001105513.3:c.2822−2 A > C	Het	NA/NA/D/23.4/4.79/	
WNK3	M858	Oliogasthenozoospermia	NM_001002838.4:c.2746 C > T;p.(Arg916Trp)	Hemi	D/P/N/24.8/5.02/	Fakhro et al., 2018 [25]
CFAP54	M827	Oliogasthenozoospermia	NM_001306084.2:c.1540−4 C > T	Het	NA/NA/NA/NA/NA	Tian et al., 2023 [26]
			NM_001306084.2:c.8642_8643del; p.(Leu2882Phefs*7)	Het	NA/NA/NA/NA/NA	

[a] the variant had one or two heterozygous carriers in the 210 normally fertile men serving as control. "NA", not applicable

Table 2 Variants of candidate genes for male infertility

Gene	Patients	Semen characteristics	Variation	Zygosity	SIFT/PolyPhen/MutationTaster/CADD/GERP	Phenotypes of KO mice (or patients) from literature
CHTF18	M1286	NOA	NM_022092.3:c.1439G > A;p.(Gly480Glu)	Het	T/P/N/1.258/1.41/	Oligospermia, subfertility, abnormal morphology of spermatogenic cells [27]
			NM_022092.3: c.1901G > C;p.(Arg634Pro)	Het	D/D/N/23.7/0.644/	
DDB1	M1809	NOA	NM_001923.3:c.1550 A > G;p.(Tyr517Cys)	Het	T/D/D/23.4/5.67	Decreased Sertoli cell proliferation [28]
			NM_001923.3:c.762 + 6 C > T	Het	NA/NA/NA/NA/NA	
DNAH12	M842	Asthenozoospermia	NM_001366028.2c.6286 C > T;p.(Gln2096X)	Hom	NA/NA/A/38/1.27/	Patients with multiple morphological abnormalities of the sperm flagella (MMAF) [29, 30] and one patient with sperm motility disorder [31]
	M867	NOA	NM_001366028.2:c.9675-1G > A	Het	NA/NA/D/26.9/5.78/	
			NM_001366028.2:c.6021–3 A > G	Het	NA/NA/NA/NA/NA/	
	M889	OAT	NM_001366028.2:c.10763G > A;p.(Arg3588His)	Het	D/P/D/35/5.34/	
			NM_001366028.2:c.898–5 C > T	Het	NA/NA/NA/NA/NA/	
FANCB	M1782	Teratozoospermia	NM_152633.3:c.2165G > C;p.(Arg722Thr)	Hemi	D/D/D/26.3/5.34	Infertility, reduction of undifferentiated spermatogonia [32]
GALNT3	M865	Oliogasthenozoosper-mia	NM_004482.4:c.1696 C > T;p.(Gln566Ter)	Hom	NA/NA/A/35/1.61	Infertility, oligoasthenoteratozoospermia [33]
MAGEC1	M1528	Oligoszoopermia	NM_005462.5:c.392 C > T;p.(Pro131Leu)	Hemi	D/P/D/23/NA/	
	M898, M1600	Asthenozoospermia	NM_005462.5:c.1754_1755del; p.(Gln585Argfs*61)	Hemi	NA/NA/NA/NA/NA	
OPHN1	M875	Oliogasthenozoosper-mia	NM_002547.3:c.487-5_487-3del	Hemi	NA/NA/NA/NA/NA	No reproduction and lack of motility of most sperm (https://www.informatics.jax.org/downloads/Reference_texts/J123336.pdf)
RBMXL3	M1222	Asthenozoospermia	NM_001145346.2:c.101T > A;p.(Ile34Asn)	Hemi	D/D/N/23.3/0.69	Candidate gene for NOA [7]
	M839	Asthenozoospermia	NM_001145346.2:c.548 C > T;p.(Pro183Leu)	Hemi	T/D/N/10.97/1.29	
SCML2	M1204, M1862	Asthenozoospermia	NM_006089.3:c.1244 C > G;p.(Pro415Arg)	Hemi	D/P/N/23.2/4.47	ICSI using testicular sperm rescued infertility [34]
UPF3A	M847	Asthenozoospermia	NM_001353644.2:c.187dupA; p.(Glu66Argfs*4)	Hom	NA/NA/NA/NA/NA	Reduction in sperm count, a defect in spermatocyte progression [35]
ZMYM3	M1562.	Oliogasthenozoosper-mia	NM_001171162.1:c.3766 + 4 C > T	Hemi	NA/NA/NA/NA/NA	Infertility, meiotic division arrest [36]
	M1765	NOA	NM_001171162.1:c.3770 C > T;p.(Thr1257Met)	Hemi	T/D/D/8.703/3.04	
ZNF185	M1569	NOA	NM_007150.3:c.1981T > C;p.(Cys661Arg)	Hemi	D/D/D/26.2/5.25/	
	M1799	NOA	NM_007150.3:c.1924 A > G;p.(Ile642Val)	Hemi	D/D/D/25.9/4.42/	
	M1813	NOA	NM_007150.3:c.661G > A;p.(Val221Met)	Hemi	T/D/N/26.7/3.83/	

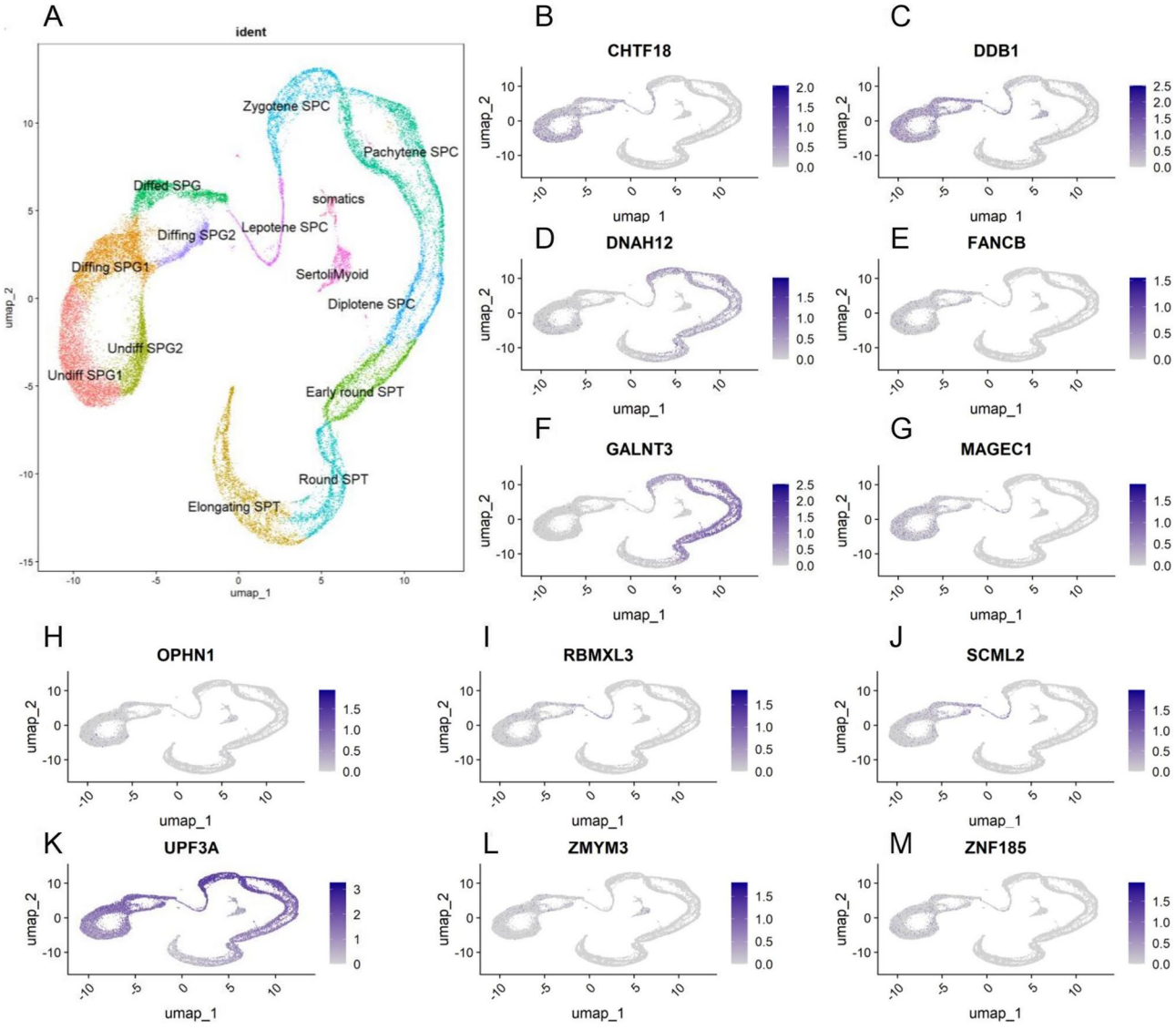

Fig. 1 (A) Uniform Manifold Approximation and Projection (UMAP) plot of testicular cells. The cells are colored based on cell types. SPG, SPC and SPT stand for spermatogonia, spermatocytes and spermatids, respectively. (B-M) Expression patterns of the candidate genes on the UMAP space

evidence on causative genes for spermatogenic failure would improve the diagnostic performance of WES.

Among the 17 known causative genes for male infertility identified in this study, *AR*, *MSH4*, and *TDRD9* have been reported to be associated with azoospermia [12, 23, 43, 44]. The human androgen receptor (AR) gene, located on X chromosome q11-12, encodes a ligand-activated transcription factor that plays a key role in spermatogenesis [45]. Recently, a novel missense mutation [NM_000044.6: c.2051G>C p.(Gly684Ala)] within the ligand-binding domain of AR in azoospermic individuals from a Chinese family has been identified using WES [44]. In this study, we identified a start-loss variant of *AR* in a patient with NOA. *MSH4* (MutShhomologue 4) is a member of the mammalian mismatch repair (MMR) gene family, and plays a vital role in meiotic chromosome synapsis [46]. Homozygous missense [NM_002440.4:c.1913 C>T p.(Pro638Leu); c.2261 C>T p. (Ser754Leu)], homozygous deletion frameshift [c.805_812del p.(Val269Glnfs*15); c.2220_2223del p.(Lys741Argfs*2)], and compound heterozygous [c.1950G>A p.(Trp650X) and c.2179delG p.(Asp727Metfs*11); c.244G>A p.(Gly82Ser) and c.670delT p.(Leu224Cysfs*3)] variants of *MSH4* have been reported in patients with NOA [37, 43]. Here, we have identified a homozygous splice-site variant in a patient with NOA. The mouse *Tdrd9* gene encodes an ATPase/DExH-type helicase, which has been linked to male sterility and meiotic failure [47]. A previous study has identified a homozygous 4 bp deletion frameshift mutation in *TDRD9* [c.720_723 del TAGT p.(Ser241Profs*4)] as the causative mutation in five azoospermic infertile men of a large consanguineous

Bedouin family [23]. In our study, we identified a homozygous missense variant of *TDRD9* in a NOA patient. Together, our findings expand the mutation spectrum of these azoopsermia genes. Interestingly, it has been shown that biallelic deleterious mutations of *CFAP54* can induce severe MMAF and NOA in humans [26].With respect to the recently identified MMAF causative gene *CFAP47* [14], one patient (M1294) hemizygous for a missense variant of the gene had NOA from our study. This finding may expand the phenotype of male infertility caused by *CFAP47*. However, further evidence is required to confirm the relationship between *CFAP47* and NOA.

Five of the 17 known causative genes, namely *DNAH17*, *DNAH6*, *DNAH7*, *DNAH8*, and *DNAH9*, belong to the DNAH family, members of which encode the axonemal dynein heavy chains [48]. Dyneins are components of the inner and outer dynein arms attached to the doublets of the axonemal microtubules in the motile cilia and sperm flagella. Mutations in the DNAH genes are known to be associated with cilia-related disorders, such as immotile cilia syndrome [49].Using WES, a recent study analyzed 90 cases of Chinese patients with MMAF and identified two patients with biallelic heterozygous missense variants of *DNAH8* [c.11771 C>T p.(Thr3924Met) plus c.6689 A>G p.(Lys2230Arg); c.9427 C>T p.(Arg3143Cys) plus c.12721G>A p.(Ala4241Thr)]. Meanwhile, reanalysis of the exome sequencing data of 167 MMAF patients from France, Iran, and North Africa identified a patient of homozygous frameshift mutation [c.6962_6968del p.(His2321Profs*4)] in *DNAH*8 [18]. In this study, we identified two patients (M1237 and M1761) with asthenozoopsermia carrying double heterozygous variants of *DNAH8*. In the study by Whitfield et al., five patients with asthenozoospermia were found to carry variants of *DNAH17* [15]. In our study, three patients with azoospermia were found to carry double heterozygous missense mutations in *DNAH17.* In line with our finding, a most recent study also detected multiple-heterozygous damaging variants in *DNAH17* in Japanese men with isolated non-obstructive azoospermia [50]. Regarding *DNAH9*, we noticed the difference between the NOA phenotype of the patient M1470 in this study and the asthenozoospermia phenotype of the patients described previously [19]. Further investigation is needed to clarify this issue.

In the present study, we have identified 12 candidate genes for human spermatogenic failure/defects. Based on testicular single cell sequencing data, six of the 12 candidate genes (*CHTF18*, *DDB1*, *DNAH12*, *GALNT3*, *MAGEC1*, and *UPF3A*) are deferentially expressed at spermatogenic stages, suggestive of distinct roles of these genes in spermatogenesis. Variants of *DNAH12* were previously identified in patients with MMAF or sperm motility disorder [29, 31]. In accordance with these findings, a homozygous nonsense variant of *DNAH12* was identified in a patient (M842) with asthenozoospermia in our study, adding evidence to *DNAH12* being a causative gene for male infertility. However, the patient M867 who carried double heterozygous splice-site variants in *DNAH12* had azoospermia by routine semen analysis, seeming to be inconsistent with the asthenozoospermia and MMAF phenotype of patients with homozygous or compound heterozygous variants in *DNAH12* [30]. We reason that the possibility of *DNAH12* leading to azoospermia could not be excluded, considering that single cell sequencing data revealed its expression in spermatocytes and early spermatids, and that Dnah12$^{-/-}$ male mice had a dramatically decreased sperm number compared with the wild type [30]. In addition, the MMAF causative genes *DNAH6* and *DNAH17* have also been associated with azoospermia [50, 51]. To confirm whether *DNAH12* is a causative for NOA, more azoopsermic patients with pathogenic variants in *DNAH12* are anticipated to be identified in future studies.

Interestingly, loss of Chtf18 in mice leads to oligospermia due to defective meiotic recombination [26]; but the patient (M1286) double heterozygous for missense variants in *CHTF18* from our study was diagnosed as NOA. The Upf3a knockout mice were described to have significant reduction in sperm count and a defect in spermatocyte progression [35], whereas the patient M847 carrying a homozygous frameshift variant of *UPF3A* from our study had a semen characteristic of asthenozoospermia. The phenotypic differences may be attributed to genetic differences between humans and mouse strains. Though our findings provide important information on the roles of *CHTF18* and *UPF3A* in human spermatogenesis, whether they are casual genes for spermatogenic failure needs to be further verified.

There are some limitations to the study. First, the sample size of the current study was relatively small. Secondly, this study was a retrospective study and parental samples were not available. Therefore, the maternal and paternal origins of the candidate variants identified in patients are unclear. Furthermore, the patients did not undergo testicular puncture or biopsy, so it is impossible to assess the pathological types in the testes and detect the expression of related genes.

Taken together, our study shows that WES is an effective tool in the genetic diagnosis of primary male infertility. Our study increases a comprehensive and in-depth understanding of the genetic factors involved in spermatogenesis and has meaning for the diagnosis, clinical treatment and genetic counseling of male infertility.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12920-024-02005-3.

Supplementary Material 1. Supplementary Data Sheet 1 Results of Sanger sequencing validation.

Supplementary Material 2. Supplementary Data Sheet 2. Results of CNV analysis using WES data.

Acknowledgements
We thank the patients for their cooperation and continued support for this study.

Author contributions
HZ collected and prepared the samples for exome sequencing, and wrote the first version of the manuscript. ZY interpreted the WES data. BN and JL collected the clinical information of the patients. SL performed the gene expression analysis and wrote the manuscript. WX designed the study and wrote the manuscript and take responsibility for the integrity of the work. Each author believes that the manuscript represents honest work. All authors read and approved the final version of manuscript.

Funding
This research was supported by Natural Science Foundation of Hunan Province (2022JJ30322), and Hunan Provincial Hospital of Maternal and Child Health Care's High-Level Talent Development Scheme.

Data availability
The data that support the findings of this study have been deposited into CNGB Sequence Archive (CNSA) of China National GeneBank DataBase (CNGBdb) with accession number CNP0005865.

Declarations

Ethics approval and consent to participate
The present study was approved by the Medical Ethics Committee of Hunan Provincial Maternal and Child Health Care Hospital (No. K202031), and written informed consent was obtained from all study participants.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 28 June 2024 / Accepted: 4 September 2024
Published online: 12 September 2024

References

1. Agarwal A, Baskaran S, Parekh N, Cho CL, Henkel R, Vij S, et al. Male infertility. Lancet. 2021;397(10271):319–33.
2. Minhas S, Bettocchi C, Boeri L, Capogrosso P, Carvalho J, Cilesiz NC, et al. European Association of Urology Guidelines on male sexual and Reproductive Health: 2021 update on male infertility. Eur Urol. 2021;80(5):603–20.
3. Cham TC, Chen X, Honaramooz A. Current progress, challenges, and future prospects of testis organoidsdagger. Biol Reprod. 2021;104(5):942–61.
4. Houston BJ, Riera-Escamilla A, Wyrwoll MJ, Salas-Huetos A, Xavier MJ, Nagirnaja L, et al. A systematic review of the validated monogenic causes of human male infertility: 2020 update and a discussion of emerging gene-disease relationships. Hum Reprod Update. 2021;28(1):15–29.
5. Nagirnaja L, Lopes AM, Charng WL, Miller B, Stakaitis R, Golubickaite I, et al. Diverse monogenic subforms of human spermatogenic failure. Nat Commun. 2022;13(1):7953.
6. Krausz C, Riera-Escamilla A. Genetics of male infertility. Nat Rev Urol. 2018;15(6):369–84.
7. Malcher A, Stokowy T, Berman A, Olszewska M, Jedrzejczak P, Sielski D, et al. Whole-genome sequencing identifies new candidate genes for nonobstructive azoospermia. Andrology. 2022;10(8):1605–24.
8. Hermann BP, Cheng K, Singh A, Roa-De La Cruz L, Mutoji KN, Chen IC, et al. The mammalian spermatogenesis single-cell transcriptome, from Spermatogonial Stem cells to Spermatids. Cell Rep. 2018;25(6):1650–e16671658.
9. Babadi M, Fu JM, Lee SK, Smirnov AN, Gauthier LD, Walker M, et al. GATK-gCNV enables the discovery of rare copy number variants from exome sequencing data. Nat Genet. 2023;55(9):1589–97.
10. Sha Y, Liu W, Li L, Serafimovski M, Isachenko V, Li Y, et al. Pathogenic variants in ACTRT1 cause Acephalic Spermatozoa Syndrome. Front Cell Dev Biol. 2021;9:676246.
11. Dai S, Liu M, Liu M, Jiang C, Yang Y, Han H, et al. Population-based genetic analysis in infertile men reveals novel mutations of ADAD family members in patients with impaired spermatogenesis. Hum Mol Genet. 2023;32(11):1814–25.
12. Giwercman YL, Nikoshkov A, Bystrom B, Pousette A, Arver S, Wedell A. A novel mutation (N233K) in the transactivating domain and the N756S mutation in the ligand binding domain of the androgen receptor gene are associated with male infertility. Clin Endocrinol (Oxf). 2001;54(6):827–34.
13. Lu C, Zhang Y, Qin Y, Xu Q, Zhou R, Cui Y, et al. Human X chromosome exome sequencing identifies BCORL1 as contributor to spermatogenesis. J Med Genet. 2021;58(1):56–65.
14. Liu C, Tu C, Wang L, Wu H, Houston BJ, Mastrorosa FK, et al. Deleterious variants in X-linked CFAP47 induce asthenoteratozoospermia and primary male infertility. Am J Hum Genet. 2021;108(2):309–23.
15. Whitfield M, Thomas L, Bequignon E, Schmitt A, Stouvenel L, Montantin G, et al. Mutations in DNAH17, encoding a sperm-specific axonemal outer Dynein Arm Heavy Chain, cause isolated male infertility due to Asthenozoospermia. Am J Hum Genet. 2019;105(1):198–212.
16. Tu C, Nie H, Meng L, Yuan S, He W, Luo A, et al. Identification of DNAH6 mutations in infertile men with multiple morphological abnormalities of the sperm flagella. Sci Rep. 2019;9(1):15864.
17. Wei X, Sha Y, Wei Z, Zhu X, He F, Zhang X, et al. Bi-allelic mutations in DNAH7 cause asthenozoospermia by impairing the integrality of axoneme structure. Acta Biochim Biophys Sin (Shanghai). 2021;53(10):1300–9.
18. Liu C, Miyata H, Gao Y, Sha Y, Tang S, Xu Z, et al. Bi-allelic DNAH8 variants lead to multiple morphological abnormalities of the sperm flagella and primary male infertility. Am J Hum Genet. 2020;107(2):330–41.
19. Tang D, Sha Y, Gao Y, Zhang J, Cheng H, Zhang J, et al. Novel variants in DNAH9 lead to nonsyndromic severe asthenozoospermia. Reprod Biol Endocrinol. 2021;19(1):27.
20. Martinez G, Kherraf ZE, Zouari R, Fourati Ben Mustapha S, Saut A, Pernet-Gallay K, et al. Whole-exome sequencing identifies mutations in FSIP2 as a recurrent cause of multiple morphological abnormalities of the sperm flagella. Hum Reprod. 2018;33(10):1973–84.
21. Wyrwoll MJ, van Walree ES, Hamer G, Rotte N, Motazacker MM, Meijers-Heijboer H, et al. Bi-allelic variants in DNA mismatch repair proteins MutS Homolog MSH4 and MSH5 cause infertility in both sexes. Hum Reprod. 2021;37(1):178–89.
22. Cavarocchi E, Whitfield M, Chargui A, Stouvenel L, Lores P, Coutton C, et al. The sodium/proton exchanger SLC9C1 (sNHE) is essential for human sperm motility and fertility. Clin Genet. 2021;99(5):684–93.
23. Arafat M, Har-Vardi I, Harlev A, Levitas E, Zeadna A, Abofoul-Azab M, et al. Mutation in TDRD9 causes non-obstructive azoospermia in infertile men. J Med Genet. 2017;54(9):633–9.
24. Liu W, He X, Yang S, Zouari R, Wang J, Wu H, et al. Bi-allelic mutations in TTC21A induce asthenoteratospermia in humans and mice. Am J Hum Genet. 2019;104(4):738–48.
25. Fakhro KA, Elbardisi H, Arafa M, Robay A, Rodriguez-Flores JL, Al-Shakaki A, et al. Point-of-care whole-exome sequencing of idiopathic male infertility. Genet Med. 2018;20(11):1365–73.
26. Tian S, Tu C, He X, Meng L, Wang J, Tang S, et al. Biallelic mutations in CFAP54 cause male infertility with severe MMAF and NOA. J Med Genet. 2023;60(8):827–34.
27. Berkowitz KM, Sowash AR, Koenig LR, Urcuyo D, Khan F, Yang F, et al. Disruption of CHTF18 causes defective meiotic recombination in male mice. PLoS Genet. 2012;8(11):e1002996.
28. Zheng W, Nazish J, Wahab F, Khan R, Jiang X, Shi Q. DDB1 regulates sertoli cell proliferation and testis cord remodeling by TGFbeta Pathway. Genes (Basel) 2019, 10(12).
29. Li Y, Wang Y, Wen Y, Zhang T, Wang X, Jiang C, et al. Whole-exome sequencing of a cohort of infertile men reveals novel causative genes in teratozoospermia that are chiefly related to sperm head defects. Hum Reprod. 2021;37(1):152–77.
30. Geng H, Wang K, Liang D, Ni X, Yu H, Tang D, et al. Further evidence from DNAH12 supports favorable fertility outcomes of infertile males with dynein axonemal heavy chain gene family variants. iScience. 2024;27(7):110366.

31. Oud MS, Houston BJ, Volozonoka L, Mastrorosa FK, Holt GS, Alobaidi BKS, et al. Exome sequencing reveals variants in known and novel candidate genes for severe sperm motility disorders. Hum Reprod. 2021;36(9):2597–611.
32. Kato Y, Alavattam KG, Sin HS, Meetei AR, Pang Q, Andreassen PR, et al. FANCB is essential in the male germline and regulates H3K9 methylation on the sex chromosomes during meiosis. Hum Mol Genet. 2015;24(18):5234–49.
33. Miyazaki T, Mori M, Yoshida CA, Ito C, Yamatoya K, Moriishi T, et al. Galnt3 deficiency disrupts acrosome formation and leads to oligoasthenoteratozoospermia. Histochem Cell Biol. 2013;139(2):339–54.
34. Sakashita A, Ooga M, Otsuka K, Maezawa S, Takeuchi C, Wakayama S, et al. Polycomb protein SCML2 mediates paternal epigenetic inheritance through sperm chromatin. Nucleic Acids Res. 2023;51(13):6668–83.
35. Shum EY, Jones SH, Shao A, Chousal JN, Krause MD, Chan WK, et al. The antagonistic gene paralogs Upf3a and Upf3b govern nonsense-mediated RNA decay. Cell. 2016;165(2):382–95.
36. Hu X, Shen B, Liao S, Ning Y, Ma L, Chen J, et al. Gene knockout of Zmym3 in mice arrests spermatogenesis at meiotic metaphase with defects in spindle assembly checkpoint. Cell Death Dis. 2017;8(6):e2910.
37. Krausz C, Riera-Escamilla A, Moreno-Mendoza D, Holleman K, Cioppi F, Algaba F, et al. Genetic dissection of spermatogenic arrest through exome analysis: clinical implications for the management of azoospermic men. Genet Med. 2020;22(12):1956–66.
38. Chen S, Wang G, Zheng X, Ge S, Dai Y, Ping P, et al. Whole-exome sequencing of a large Chinese azoospermia and severe oligospermia cohort identifies novel infertility causative variants and genes. Hum Mol Genet. 2020;29(14):2451–9.
39. Tang D, Li K, Geng H, Xu C, Lv M, Gao Y, et al. Identification of deleterious variants in patients with male infertility due to idiopathic non-obstructive azoospermia. Reprod Biol Endocrinol. 2022;20(1):63.
40. Zhang H, Li W, Jiang Y, Li J, Chen M, Wang R, et al. Whole exome sequencing identifies genes Associated with Non-obstructive Azoospermia. Front Genet. 2022;13:872179.
41. Kherraf ZE, Cazin C, Bouker A, Fourati Ben Mustapha S, Hennebicq S, Septier A, et al. Whole-exome sequencing improves the diagnosis and care of men with non-obstructive azoospermia. Am J Hum Genet. 2022;109(3):508–17.
42. Lillepea K, Juchnewitsch AG, Kasak L, Valkna A, Dutta A, Pomm K, et al. Toward clinical exomes in diagnostics and management of male infertility. Am J Hum Genet. 2024;111(5):877–95.
43. Li P, Ji Z, Zhi E, Zhang Y, Han S, Zhao L, et al. Novel bi-allelic MSH4 variants causes meiotic arrest and non-obstructive azoospermia. Reprod Biol Endocrinol. 2022;20(1):21.
44. Yuan Y, Xu WQ, Chen Y, Luo T, Chen HY. A Gly684Ala substitution in the androgen receptor is the cause for azoospermia in a Chinese family with mild androgen insensitivity syndrome and normal hormone levels. Front Genet. 2022;13:988202.
45. Smith LB, Walker WH. The regulation of spermatogenesis by androgens. Semin Cell Dev Biol. 2014;30:2–13.
46. Neyton S, Lespinasse F, Moens PB, Paul R, Gaudray P, Paquis-Flucklinger V, et al. Association between MSH4 (MutS homologue 4) and the DNA strand-exchange RAD51 and DMC1 proteins during mammalian meiosis. Mol Hum Reprod. 2004;10(12):917–24.
47. Shoji M, Tanaka T, Hosokawa M, Reuter M, Stark A, Kato Y, et al. The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. Dev Cell. 2009;17(6):775–87.
48. Silvanovich A, Li MG, Serr M, Mische S, Hays TS. The third P-loop domain in cytoplasmic dynein heavy chain is essential for dynein motor function and ATP-sensitive microtubule binding. Mol Biol Cell. 2003;14(4):1355–65.
49. Inaba K, Mizuno K. Sperm dysfunction and ciliopathy. Reprod Med Biol. 2016;15(2):77–94.
50. Muranishi Y, Kobori Y, Katoh-Fukui Y, Tamaoka S, Hattori A, Osaka A, et al. Systematic molecular analyses for 115 karyotypically normal men with isolated non-obstructive azoospermia. Hum Reprod. 2024;39(5):1131–40.
51. Gershoni M, Hauser R, Yogev L, Lehavi O, Azem F, Yavetz H, et al. A familial study of azoospermic men identifies three novel causative mutations in three new human azoospermia genes. Genet Med. 2017;19(9):998–1006.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.