Preprints.org

Article

A Comprehensive Genetic Study of Microtubule-Associated Genes Clusters for Male Infertility in a Taiwanese Cohort

Chying-Chyuan Chan , Te-Hsin Yen , Hao-Chen Tseng , Brang Mai , Pin-Kuan Ho , Jian-Liang Chou , Gwo-Jang Wu [*] , Yu-Chuan Huang[*]

Article

A Comprehensive Genetic Study of Microtubule-Associated Genes Clusters for Male Infertility in a Taiwanese Cohort

Short Title: Microtubule-associated genes clusters for male infertility

Chying-Chyuan Chan [1,2]**, Te-Hsin Yen** [2]**, Hao-Chen Tseng** [3]**, Brang Mai** [3]**, Pin Kuan Ho** [4]**, Jian-Liang Chou** [1,5]**, Gwo-Jang Wu** [1,6,]*** and Yu Chuan Huang** [1,3,5,7,]*****

[1] Graduate Institute of Medical Sciences, National Defense Medical Center, Taipei, Taiwan.
[2] Department of Obstetrics and Gynecology, Taipei City Hospital-Renai Branch, Taipei, Taiwan.
[3] School of Pharmacy, National Defense Medical Center, Taipei, Taiwan.
[4] School of Dentistry, National Defense Medical Center, Taipei, Taiwan.
[5] Department of Research and Development, National Defense Medical Center, Taipei, Taiwan.
[6] Reproductive Medical Center, Dept. of Ob/Gyn, Tri-Service General Hospital, No. 325, Sec. 2, Chenggong Rd, Neihu Dist., Taipei, 11490, Taiwan, Republic of China.
[7] Graduate Institute of Medical Sciences & School of Pharmacy, National Defense Medical Center, No. 161, Sec 6, Minquan E. Road, Taipei 11490, Taiwan
*   Correspondence: gwojang@yahoo.com (G.-J.W.); yuh@mail.ndmctsgh.edu.tw (Y.C.H.);
    Tel.: 886-2-87927205 (G.-J.W.); 886-2-87923100 (Y.C.H.); Fax: 886-2-87927207 (G.-J.W.);
    886-2-87923111 (Y.C.H.)
†   Yu Chuan Huang and Gwo-Jang Wu contributed equally to this study.

Abstract: Background: Identifying genetic mutations that lead to spermatogenic impairment is essential for offering patients informed genetic counseling and preventing the transmission of genetic defects to their offspring, often through the use of advanced reproductive technologies. **Objectives:** The purpose of this study was to identify potential single nucleotide polymorphisms (SNPs) associated with male infertility. **Methods and Materials:** Genetic variants that may cause infertility are identified by combining targeted next generation sequencing (NGS) panel and whole exome sequencing (WES). The validation step of Sanger sequencing adds confidence to the identified variants. **Results:** Our analysis revealed five distinct affected genes covering seven SNPs: based on targeted NGS panel and WES data, namely *SPATA16* (rs16846616, 1515442, 1515441), *CFTR* (rs213950), *KIF6* (rs2273063), *STPG2* (r2903150) and *DRC7* (rs3809611). Infertile men have a higher mutation rate than fertile men, especially those with azoospermia. **Discussion and Conclusions:** These findings strongly support the hypothesis that dysfunction of microtubule-related and spermatogenesis-specific genes contributes to idiopathic male infertility. The *SPATA16, CFTR, KIF6, STPG2* and *DRC7* mutations are associated with male infertility, specifically azoospermia, further examination of this genetic function is required.

Keywords: male infertility; azoospermia; oligozoospermia; next-generation sequencing; single nucleotide polymorphism; microtubule-associated genes

Introduction

Infertility is a reproductive disorder defined as the failure to achieve clinical pregnancy 12 months after unprotected intercourse [1,2]. The World Health Organization estimates that 9% of couples worldwide struggle with fertility problems, 50% of which are caused by male factors.

Microtubules (MTs) are the structural components of cells designed to participate in many essential processes such as cell division, intracellular transport and cell shape maintenance, and play a vital role in ensuring the normal formation and function of sperm. In spermatogenic cells, MTs are

required for several processes, including the assembly of flagella in spermatids, and the production and maintenance of mature sperm motility [3].

Motor proteins in MTs are specialized molecular motors that use adenosine triphosphate (ATP) as an energy source. These motors are responsible for transporting various cellular cargoes along MTs tracks, ensuring delivery of these cargoes to specific destinations within the cell, examples of such motor proteins include kinesins and dyneins [4]. Movement in the outward or anterograde direction (from sperm head to tail) is directly facilitated by the microtubule motor protein kinesin-2. In contrast, cytoplasmic dynein 2 (also known as dynein 1b) promotes movement in an inward or retrograde direction [5].

Furthermore, MTs ensure the successful division of cells during spermatogenesis [6]. As the sperm tail develops, the matrix built within the centrosome of the spermatid anchors the MTs. Within this region, MTs play a crucial role in facilitating the movement of vesicles from the Golgi to the acrosome [7]. MTs are also constituent elements of the sperm tail, the elongated flagella whose central axoneme protrudes from the basal body located posterior to the nucleus. To enable sperm motility, the movement of the inner and outer duplex MTs of the sperm flagella is dependent on the energy generated by the hydrolysis of ATP [8].

Genetic mutations that disrupt MTs function during sperm development may cause male infertility [9]. Mutations in *TSARG4* (Testis and spermatogenesis-associated gene 4) capable of impairing sperm head formation [10]. Mutations in *SPATA6* can aberrantly remodel the sperm cell head while interfering with head-tail assembly [11]. Mutations in the protein of katanin reveal a defective formation of the central microtubules in the flagellar axoneme, which is critical for male fertility [9,12]. Mutations of the N-terminal and C-terminal of *DNAH1* can have different effects on the axoneme structure of human spermatozoa. [13]. More studies also have shown that the regulation of MT dynamics is critical for male infertility (azoospermia and oligospermia) [14,15].

This study utilized targeted NGS panel and WES to explore male infertility (MI)-associated SNPs to understand its causes and search for human azoospermia and oligospermia gene clusters that lead to abnormal sperm function. The aim of this study was to identify microtuble-related genes (such as *KIF6* and *DRC7*) and their properties in sperm cell function, motility, intracellular trafficking, differentiation and cell division. Assessment of microtubules dynamics is an important aspect to fully understand the correlation between genes and male infertility. Further research in this area may improve diagnosis, treatment, and interventions for male infertility.

Materials and Methods

Patients and controls

From July 2018 to Jan 2020, 31 infertile men between 30 and 45 years old were recruited during routine infertility treatment at the Reproductive Medical Center, Tri-Service General Hospital and Taipei City Hospital-Renai Branch (Taipei, Taiwan). The infertile men were included and their semen analyzed (Table 1), we divided them into two groups based on sperm concentration in semen as azoospermia (<0.1million/ml, n=15), and oligozoospermia (0.1-15 million/ml, n=16). The fertile controls (n=9) were men who had children in the previous 3 years. The study protocol was approved by the Institutional Review Board of Taipei City Hospital Research Ethics Committee (protocol Ver2.0-1090414) (Taipei, Taiwan) and all the patients provided written consent prior to enrollment in the study.

Table 1. Thirty-one (31) infertile men with no or low semen sperm count (<15 million/mL) and nine fertile men as controls were included in this study.

Group by phenotype	Patient ID	volume(ml)	Sperm counts (M)	Age	Motility (%)	Morphology (%)
CRT Fertile (n=9)	CRT-1	3.2	normal	40	45	60
	N1	2.1	normal	32	35	36
	N2	nd	normal	33	nd	nd
	N3	nd	normal	35	nd	nd
	N4	nd	normal	42	nd	nd
	N5	nd	normal	38	nd	nd
	N8	nd	normal	29	nd	nd
	N27(NN7)	nd	normal	40	nd	nd
	N31	nd	normal	50	61	61
Oligo <=15 million (n=16)	S01	1.9	15	42	13	58
	S03	1.8	11.6	40	65	72
	S04	1.9	11.9	36	47	52
	S05	0.5	5.6	37	45	60
	S06	3.4	2.1	32	90	30
	S07	4.5	4.5	41	20	58
	S08	4.1	1.4	45	21	40
	S11	1.5	0.3	36	67	61
	S15	0.7	3.6	38	38	55
	S18	1.5	7.7	44	71	53
	S20	2.6	4.2	34	52	45
	CA21	3.2	4.4	32	52	33
	CA22	3.5	7.1	36	81	53
	CA24	1.6	7	39	54	45
	CA25	1.2	6.5	41	57	51
	S31	2.1	14	32	35	45
Azoo <0.1 million (n=15)	P1 (S-2017)	3.5	<0.1	40	nd	nd
	P2	4	<0.1	38	nd	nd
	S02	0.5	<0.1	43	50	50
	S09	1.3	<0.1	32	50	50
	S10	0.5	<0.1	30	0	30
	S13	0.9	<0.1	42	0	nd
	S14	<0.1	<0.1	39	0	0
	S16	6.7	<0.1	39	42	30
	S17	0.3	<0.1	35	nd	nd
	S26	1.2	<0.1	36	57	52
	S32	1.5	<0.1	43	nd	nd
	CA23	0.4	<0.1	39	nd	nd
	CA26	0.9	<0.1	36	nd	nd
	TCA26	0.7	<0.1	30	nd	nd
	TS1	0.6	<0.1	34	nd	nd

Semen analysis

Patients refrained from sexual activity for 3 to 5 days, then semen samples were collected in sterile containers by masturbation. The semen is kept at room temperature for 15 to 30 minutes to allow for it to liquefy. After the process of liquefaction, the total volume of the semen and the sperm concentration, motility, and shape were documented in accordance with the World Health Organization Laboratory Manual for the Examination and Processing of Human Semen, Fifth Edition. Semen volume less than 1.5 mL, sperm concentration less than <15 million/mL, motility less than 40% or progressive motility (PR) plus non-progressive motility (NP) less than 32%, and normal morphology in less than 30% were considered abnormal

Y chromosome microdeletion (YCMD) examination

The Y chromosome's long arms (Yq) harbor numerous coding genes responsible for regulating spermatogenesis and testes development. Microdeletions within the AZF (azoospermia factor) region can lead to a diverse range of infertility phenotypes. To investigate these deletions, Yq microdeletion analysis was conducted by amplifying the AZFa, AZFb, and AZFc loci along with their associated

sequence-tagged sites (STSs) markers. YCMD are identified in approximately 13% of men with nonobstructive azoospermia and 5% of men with severe oligozoospermia. Employing the YCMD assay (Promega), a PCR-based blood test, we assessed the presence or absence of sequence-tagged sites (STSs) as well as clinically relevant microdeletions (Figure S1).

YCMD testing is a prevalent method for evaluating male infertility. The Y chromosome's long arm (Yp) comprises of three azoospermia factor regions prone to deletions: AZFa, AZFb, and AZFc. Many Y chromosome genes are exclusive to spermatogenesis, and deletions within these regions, particularly at Yp's AZFc, can disrupt sperm formation. While Y-chromosome microdeletions and the genetic underpinnings of male infertility have garnered significant research attention, the clinical implications within this realm remain scattered and lack a unified consensus.

Targeted next-generation sequencing (NGS) panel

The amplicon libraries of 8 infertile males and 2 fertile controls were constructed using the Ion AmpliSeq™ Library Kit v2.0 (ThermoFisher Scientific) according to the manufacturer's instructions, and using the Ion Xpress™ Barcode Adapter Kit (ThermoFisher Scientific) for barcode adapter ligation. The patient's genomic DNA was extracted from peripheral blood leukocytes using the MagPurix DNA extraction kit (Taipei, Taiwan) according to the manufacturer's instructions. The custom Ion Ampliseq NGS panel was designed with 257 amplicons (IAD142966_182, 131+126 primer pairs containing 27,707bps) to ensure comprehensive coverage of 15 gene targets previously associated with male infertility (Table 2). Massively parallel sequencing was performed in the GeneStudio S5 Sequencing System (ThermoFisher Scientific) to cover exonic regions of genes, increasing the average sequencing depth by more than 1000-fold. All reads were further analyzed by Torrent Suite™ software (ThermoFisher Scientific) with the human reference genome (GRCh37.p5/hg19). All synonymous and non-altered protein splice site variants were then removed, leaving only the coding gene for comparison.

Table 2. The customized Ion Ampliseq NGS panel of infertility included 15 spermatogenesis-related genes with 257 amplicons (IAD142966_182, 131+126 primers). We performed a massive parallel sequencing on an Ion S5 semiconductor sequencer (S5, ThermoFisher Scientific).

Gene ID	Gene name	Chromosome	# of Amplicons	Total Bp	Covered Bp	Missed Bp	Coverage
G1	FSHB	chr11	3	410	410	0	1
G2	AURKC	chr19	9	1011	1011	0	1
G3	LHB	chr19	3	456	191	265	0.419
G4	PICK1	chr22	17	1368	1368	0	1
G5	SPATA16	chr3	16	1810	1810	0	1
G6	DAZL	chr3	15	1071	1071	0	1
G7	ESR1	chr6	15	1874	1874	0	1
G8	CFTR	chr7	46	4713	4709	4	0.999
G9	KAL2	chr8	28	2825	2825	0	1
G10	GNRH1	chr8	4	321	321	0	1
G11	NR5A1	chr9	14	1446	1424	22	0.985
G12	AR	chrx	19	2873	2873	0	1
G13	TEX11	chrx	34	3160	3068	92	0.971
G14	USP26	chrx	15	2752	2752	0	1
G15	KAL1 (ANOS1)	chrx	19	2183	2000	183	0.916

Whole-exome sequencing (WES) and Variant analysis

WES is a widely used next-generation sequencing (NGS) method that involves sequencing the protein-coding regions of the genome. We selected 15 samples of participants (8 cases and 7 controls) extracted from peripheral blood leukocytes by using the MagPurix DNA extraction kit, processed according to Illumnia TruSeq DNA Exome Kit guidelines and fragmented 100ng of genomic DNA into 150bp inserts by Covaris M22. Briefly, fragment gDNA is ligated with adapters and enriched by PCR. After PCR amplification, the probe captures the fragments and is then amplified to create the WES library. The Nextseq 550 (Illumnia) was used for WES. Then, those sequenced reads were aligned to the human reference genome (hg38) by using NGS Core Tools/FASTQ mapping with CLC

Genome Workbench 23.0.4. We filtered out and compared those detected variants with SNPs database(p<0.05). SNPs profiles were assessed by using resequencing analysis/variant comparison module of CLCs. The cases (n=8) were compared with controls (n=7), and we put emphasis on those genes significantly expressed in testis tissue on the basis of NCBI HPA RNA-seq normal tissue.

Validation by Sanger sequencing

After the genes of interest were identified using targeted NGS sequencing and WES, Sanger sequencing was applied to examine the potential SNPs of the *CFTR, SPATA16, KIF6, STPG2, DRC7* genes. The product of targeted genes was gained by using PCR (Proflex, ThermoFisher Scientific) with a pair of primers, the sequences are listed in Table S1. Sanger sequencing of Applied Biosystems Genetic analyzer (3130X, MA, USA) was performed on the infertile cases and fertile controls to validate the findings of the candidate SNPs by WES and targeted NGS. Sanger sequencing was also performed on an additional group of infertile men and fertile controls.

In silico evaluation workflow

The *in silico* evaluation of the pathogenicity of nucleotide changes in exons was performed using CADD (Combined Annotation Dependent Depletion), SIFT (Sorting Intolerant from Tolerant), and PolyPhen2 (Polymorphism Phenotyping) (applied in Table 5, Table S3). CADD is a tool for scoring the deleteriousness of single-nucleotide variants in the human genome, with a cutoff score of 15 and higher values indicating more deleterious conditions. The PolyPhen-2 score predicts the likely impact of amino acid substitutions on human protein structure and function, with scores ranging from 0.0 (tolerated) to 1.0 (deleterious). SIFT predicts whether amino acid substitutions are likely to affect protein function based on scores and qualitative predictions (either 'tolerated' or 'deleterious'). Minor allele frequencies (MAFs) were examined in the genome aggregation database gnomAD and Taiwan Biobank (TWB). When discussing MAF, we should recognize that variant allele frequency is basically a fraction, the variant positivity rate divided by the total number of alleles screened.

Table 5. The occurrence of mutations on the *STPG2* and *KIF6* genes was validated through Sanger sequencing. Thirty-one (31) infertile men with no or low semen sperm count (<15 million/mL), twenty ordinary men and nine fertile men.

Phenotype	IRB	STPG2	KIF6
Ordinary Men n=20	N1	A/G	X
	N2	X	G/A
	N3	A/G	G/A
	N4	X	X
	N5	X	X
	N6	A/G	X
	N7	X	X
	N8	X	X
	N9	A/G	X
	N10	A/G	G/A
	N11	X	X
	N12	X	G/A
	N13	A/G	X
	N14	X	X
	N15	X	X
	N16	X	X
	N17	X	X
	N18	A/G	X
	N19	X	X
	N20	A/G	X
CRT Fertile n=8	PN2	X	G/A
	PN3	X	X
	PN4	X	X
	PN5	A/G	X
	PN6	X	X
	PN7	A/G	X
	NN7	X	X
Oligozoospermia n=16	S01	A/G	G/A
	S03	A/G	G/A
	S04	A/G	G/A
	S05	X	X
	S06	X	G/A
	S7	A/G	G/A
	S8	A/G	A/A
	S11	X	G/A
	S12	A/G	X
	S15	A/G	G/A
	S18	X	G/A
	S20	X	X
	CA22	X	X
	CA24	X	X
	CA25	A/G	A/A
	S31	X	X
Azoospermia n=15	P1	A/G	G/A
	P2	A/G	G/A
	S02	A/G	G/A
	S09	A/G	G/A
	S10	A/G	G/A
	S13	A/G	G/A
	S14	A/G	G/A
	S16	A/G	G/A
	S17	X	X
	S26	X	X
	S32	A/G	G/A
	CA23	A/G	X
	CA26	A/G	G/A
	TCA26	X	G/A
	TS1	A/G	G/A

Results

Semen analysis

All samples were examined in accordance to the World Health Organization guidelines. After macroscopic and microscopic observations on sperm parameters, only sperm concentration is an
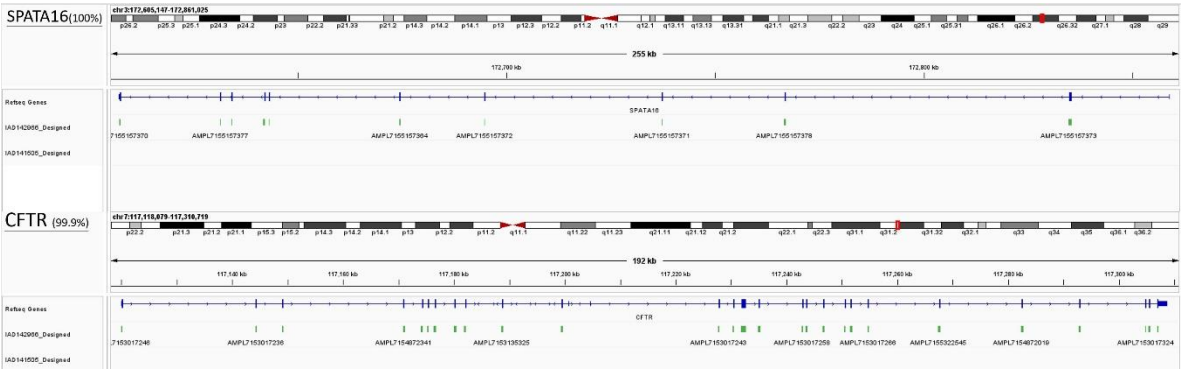
evident factor. Patients were divided into three groups based on sperm concentration, fertile controls (>15 million/mL), oligozoospermia (<15 million/mL), and azoospermia (<0.1 million/mL). The sperm of infertile men and fertile controls were examined and summarized in Table 1. There were 16 patients categorized into the oligozoospermia group, 15 patients categorized into the azoospermia group, and 9 fertile controls. The mean age of fertile controls, oligozoospermic, and azoospermic patients were 37, 38, and 37 years old, respectively. Semen volumes in oligozoospermia and azoospermia were abnormal, especially in azoospermia (1.64+/-1.8, n=15).

Y chromosome microdeletion

By employing the YCMD assay (Promega), a PCR-based blood test, the presence or absence of sequence-tagged sites (STSs) became assessable alongside clinically relevant microdeletions (Figure S1). The Y chromosome's long arms (Yq) contain numerous coding genes responsible for regulating spermatogenesis and testes development. Notably, Y chromosome microdeletions (YCMD) are detected in around 13% of men with nonobstructive azoospermia and approximately 5% of men experiencing severe oligozoospermia. Within the framework of this study, YCMD analysis was carried out on a group of 8 infertile men, comprising 4 with oligozoospermia and 4 with azoospermia (Figure S2). Interestingly, notable AZFc deletions were observed solely in two azoospermic patients (S02 and S-2017). Microdeletion of AZF-STS was not identified in 2 of the infertile patients, while others exhibited certain degrees of microdeletion (Table S2). The extent was insufficient to imply a direct cause-and-effect relationship.

SNPs of spermatogenesis-related genes in targeted NGS

In this study, spermatogenesis-related genes were used to design a male infertility **targeted NGS panel** which included *SPATA16, AURKC, CFTR, ESR1, TEX11, LHB, USP26, AR, GNRH1, NR5A1, KAL2, DAZL, PICK1, FSHB,* and *ANOS1.* Their coverage and the number of amplicons per gene in the targeted NGS panel are shown in Table 2. Ten of the 15 genes have 100% coverage, the designed amplicons of *SPATA16* and *CFTR* are shown in Figure 1 and other genes in Figures S2A, S2B, and S2C.

Figure 1. The designed amplicons (green) covered the *SPATA16* (chr3 q26.31) and *CFTR* (chr7 q31.2) regions of a 15-gene targeted NGS panel (Ampliseq), with other genes shown in the supplementary Figure S3A, S3B & S3C.

Variant's analysis and Validation in targeted NGS

Targeted sequencing was performed on 8 infertile men and 2 fertile controls. SNPs were identified from the 15 spermatogenesis-related genes (Table 2). Primers of *SPATA16, CFTR* and *ESR1* genes were used for Sanger validation of NGS variants, the sequences are shown in Table S1. *SPATA16* mutations at three SNPs (*rs*16846616), (*rs*1515442), and (rs1515441) were identified in all azoospermic patients and one oligozoospermic patient, while no mutation displayed in fertile (Table 3). *CFTR* mutation at one SNP (*rs*213950) was identified in 7 of 8 infertile men and 1 of 2 fertile

controls, though when verified with Sanger sequencing, no fertile controls displayed this *CFTR* mutation (Table 4). The remaining genes seemed to correlate more poorly. The mutations at SNPs of *TEX11* (*rs*4844247), *LHB* (*rs*4146251380), *USP26* (*rs*61741870) and (*rs*41299088), and *ANOS1* (*rs*2229013) were displayed in only 1 of 8 infertile men. *NR5A1* (*rs*1110061) and *ANOS1* (*rs*808119) mutations were present in both fertile controls. *TEX11* (*rs*6525433) displayed mutations in only 2 of 8 infertile men. GNRH1 displayed mutations in 5 of 8 infertile men and 1 of 2 fertile controls, however this was not validated with Sanger sequencing (Table 4). *ESR1* displayed mutations in 5 of 8 infertile men and none in fertile controls and could be further explored as a gene target for male infertility. The results of the Sanger sequencing examining the *SPATA16* and *CFTR* genes are presented on Table 4 and Figure 2. No fertile controls displayed the *CFTR* mutation while 6 of 16 oligozoospermic and 10 of 15 azoospermic patients did, suggesting that *CFTR* may be associated with male infertility (oligozoospermia and azoospermia). The mutations of *SPATA16* at two SNPs (rs16846616) and (rs1515411) occurred simultaneously, affecting 12 of 15 azoospermic, 6 of 16 oligozoospermic, and 2 of 8 fertile control patients. The mutation of *SPATA16* at one SNP(*rs*1515442) affected 9 of 15 azoospermic, 1 of 16 oligozoospermic, and none of the fertile control patients. The *SPATA16* mutations appear to be associated with azoospermia. Three nearby SNPs (rs1515442), (rs1515441), and (rs16846616) in *SPATA16* were shown to have 4~9.6-fold higher mutation incidence in azoospermia compared to oligozoospermia while *CFTR* was only 1.8-fold higher (Figure 2).

Using CADD, SIFT, and PolyPhen2, we performed an *in silico* evaluation of the predicted pathogenicity of *SPATA16* and *CFTR* mutations. In terms of CADD, the *SPATA16* SNP was identified as a deleterious variant. Their values were 15.67, 19.12, and 22.2, which surpassed the deleterious variant cutoff of 15 (Table S3). They are also deleterious in SIFT. Only one nucleotide variant of *SPATA16* (rs1515442) appeared benign in PolyPhen2. Overall, *SPATA16* mutations at two SNPs (rs16846616 and rs1515441) are expected to be deleterious. MAF was examined in gnomAD and TWB, and the results are shown in Table S3. Two SNPs (rs16846616) and (rs1515411) of *SPATA16* had lower MAF frequencies, 13.1%, and 11.9%, respectively in genomAD compared with TWB (32.8 and 32.5%). Interestingly, in our study in Taiwan, *SPATA16* SNPs (rs16846616) and (rs1515441) co-occur at an 80% higher frequency in azoospermic patients. (Table S3).
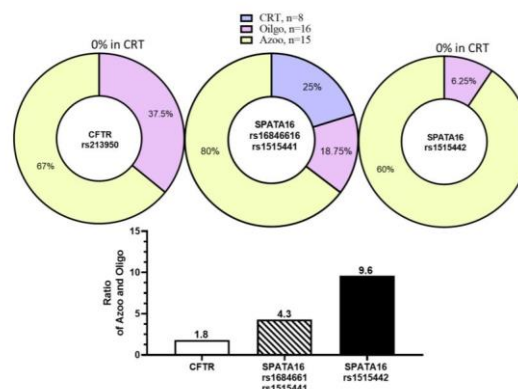

Figure 2. Three nearby SNPs (rs1515442, rs1515441 and rs16846616) in *SPATA16* were shown 4~9.6-fold mutation incidence in azoospermia compared to oligozoospermia while *CFTR* only 1.8-fold higher.

Table 3. The point mutation of genes in infertile men showed in red square (▦) ; the point mutation of genes in fertile men showed in black square (▦). (Azoo: azoospermia; Oligo: oligozoospermia)

Gene	Chr.	SNP ID	region (hg19)	Oligo S06	Oligo S07	Oligo S08	Oligo S11	Azoo S02	Azoo S09	Azoo S10	Azoo S-2017	Fertile N1	Fertile CRT-1
SPATA16	3	rs 16846616	174835082			■		■	■	■	■		
		rs 1515442	174835125			■		■	■	■	■		
		rs 1515441	174835290			■		■	■	■	■		
CFTR	7	rs 213950	117199533	■	■	■	■		■	■	■		■
ESR1	6	rs 17847065	152129484	■	■		■		■	■			
TEX11	X	rs 6525433	70073159	■			■						
		rs 4844247	68990301				■						
LHB	19	rs 4146251380	49519489								■		
USP26	X	rs 61741870	132161755					■					
		rs 41299088	132160826					■					
AR	X	rs 777131133	66765516					■				■	
GNRH1	8	rs 6185	25280800		■	■	■			■	■	■	
NR5A1	9	rs 1110061	127262802	■	■		■			■		■	■
ANOS1	X	rs 808119	8504833	■	■	■						■	■
		rs 2229013	8503786					■					

Table 4. The incidence of rs1515442, rs1515441 and rs16846616 on the *SPATA16* gene by using Sanger sequencing.

Phenotype	Patient ID	CFTR	SPATA16		
		RS213950 G→A	RS16846616 T→C	RS1515442 C→T	RS1515441 C→T
CRT Fertile (n=9)	CRT-1				
	N1				
	N2				
	N3		T→C		G→A
	N4		T→C		G→A
	N5				
	N8				
	N27(NN7)				
	N31(sample 1)				
Oligo <=15 million (n=16)	S01				
	S03	C→T	T→C		G→A
	S04				
	S05				
	S06	C→T			
	S07	C→T			
	S08	C→T	T→C	C→T	G→A
	S11	C→T			
	S15				
	S18				
	S20				
	CA21				
	CA22	C→T	T→C		G→A
	CA24				
	CA25				
	S31				
Azoo <0.1 million (n=15)	P1 (S-2017)	C→T	T→C	C→T	G→A
	P2		T→C	C→T	G→A
	S02	C→T	T→C	C→T	G→A
	S09	C→T	T→C	C→T	G→A
	S10	C→T	T→C	C→T	G→A
	S13		T→C		G→A
	S14				
	S16		T→C		G→A
	S17	C→T			
	S26	C→T	T→C	C→T	G→A
	S32	C→T	T→C	C→T	G→A
	CA23		T→C		G→A
	CA26	C→T			
	TCA26	C→T	T→C	C→T	G→A
	TS1	C→T	T→C	C→T	G→A

SNPs of microtubule-associated genes in WES

Based on the variant comparison data (p<0.05) from CLC cases and controls, we have provided a detailed list of 15 SNPS in 13 candidate genes with significant expression in testicular tissues, including *KIF6, STPG2, DRC7, NEK2, TRIM49, CATSPER2, CMTM2, SART3, DYNC2H1, CCDC168, BORCS5, TPTE, and RADIL*. Among these, certain genes exhibit a high global Minor Allele Frequency (MAF) (>0.3) and a higher mutation rate in controls (>0.3). The genes with high MAF and mutation rate in controls, include *TRIM49, CCDC168, TPTE, and RADIL*, which might lead to an apparent

elevated mutation rate in controls as well. The mutation rates and MAF for these genes in both cases and controls are presented in Table S4.

Variant analysis and Validation in WES

We selected two candidate genes, *KIF6* and *STPG2*, from the pool of 60 candidates for Sanger sequencing validation. Sanger sequencing was employed to confirm the presence of SNPs on *KIF6*, and *STPG2*, facilitating a gene comparison between infertile and fertile men (Table 5). In the case of *KIF6*, 1 out of 8 controls exhibited mutation, whereas 22 out of 31 cases harbored mutations. Similarly, for *STPG2*, there were 2 mutations among the 8 controls and 20 mutations among the 31 cases. Furthermore, in the *KIF6* gene, both A/A and G/A types of SNPs were observed (Figure S4).   Upon comparing infertile men to fertile men (fertile within 3 years), the occurrence of SNPs on the *KIF6* gene appeared to be notably higher by a multiple of 5 (as shown in Figure 3).

Using CADD, SIFT, and PolyPhen2, we conducted an in silico evaluation of the predicted pathogenicity of 13 mutations. In terms of CADD scores, 5 SNPs surpassed the deleterious variant cutoff of 15 (Table S4), namely *STPG2* (rs17558183), *NEK2, DYNC2H1, BORCS5*, and *TPTE*. Among them, *BORCS5* (CADD score: 29.9) and *STPG2* (CADD score: 25.6) exhibited the highest scores. When considering SIFT predictions, *BORCS5, STPG2, DRC7*, and *CATSPER2* were classified as deleterious variants. Within the framework of PolyPhen2 analysis, *STPG2* (rs17558183), *TRIM49, BORCS5*, and *TPTE* were designated as possibly damaging, suggesting that alterations in these amino acids might disrupt protein structure. Minor allele frequency (MAF) was assessed both globally and specifically in East Asian populations (MAF_eas), with the results also summarized in Table S4. Notably, *KIF6* (MAF: 0.04), *BORCS5* (MAF: 0.06), *DRC7* (MAF: 0.24), and *STPG2* (MAF: 0.39) were listed in global MAF. It is worth highlighting that *STPG2* exhibited a nearly 40% mutation rate, despite demonstrating a favorable case-to-control ratio (88/14) in whole exome sequencing (WES) data. However, this high mutation frequency in *STPG2* could impact its potential as a candidate diagnostic marker. Additionally, it's noteworthy that *KIF6* displayed significant differences in both global MAF (4%) and MAF_eas (25%), indicating a higher mutation rate in East Asian male populations. Interestingly, in our Taiwan study, the *KIF6* SNP (rs2273063) and *STPG2* (rs3809611) co-occur in 80% mutation rate in azoospermia and oligozoospermia patients with 20% MAF in ordinary men of Taiwan by Sanger sequencing.
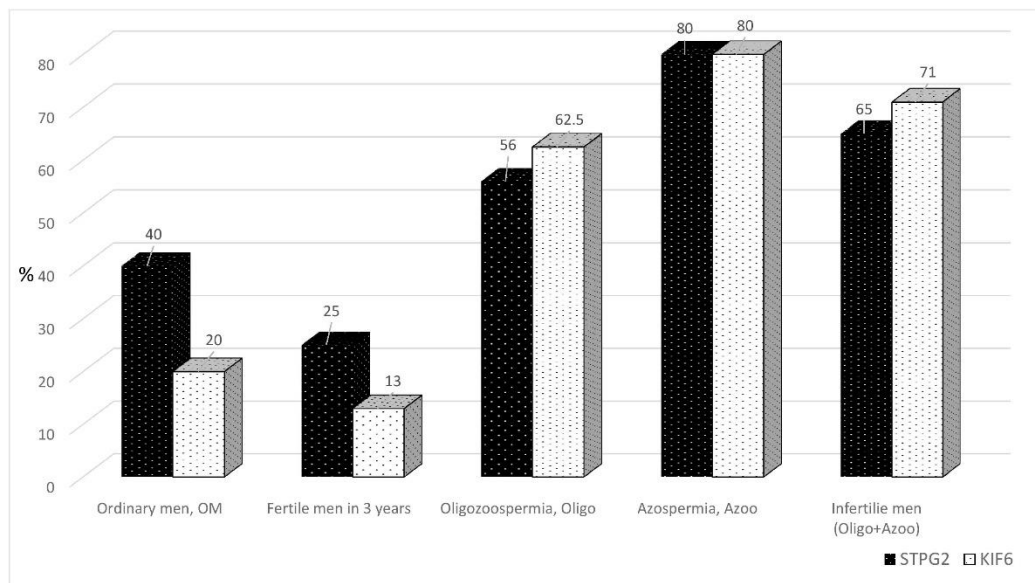
Figure 3. The incidence of SNPs on *KIF6* gene in infertile men (azoospermia) showed more than 6 times higher compared to fertile men by Sanger sequencing.

Discussion

Genetic screening to identify genetic mechanisms of spermatogenesis failure in infertile men has become of clinical importance. These results not only allow to determine the etiology but also prevent the iatrogenic transmission of genetic defects to offspring through assisted reproductive techniques. These goals pose enormous challenges to reproductive medicine. In this study, we found some MI-related genes associated with male infertility. Through targeted NGS, we extensively examined 15 candidate genes involved in spermatogenesis (257 amplicons, 27707bp) within a cohort of unrelated Taiwanese infertile men (n = 8) (see Table 2). Building upon previous studies, we confirmed the critical roles of *SPATA16* in globozoospermia and *CFTR* in obstructive oligozoospermia or azoospermia. Notably, our findings provide further support for mutations in *SPATA16* and *CFTR* genes being associated with non-obstructive oligozoospermia and azoospermia with normal morphology. Intriguingly, we established a strong correlation between *SPATA16* and azoospermia in our patient group. Additionally, our study raises questions about the implications of ESR1 and GnRH1 gene mutations in male infertility, which warrant more in-depth investigation.

Microtubule-associated genes affect spermatogenesis by variants analysis

KIF6 (kinesin family member 6) gene encodes motor proteins, including kinesins and dyneins, that play crucial roles in intracellular transport and cellular movement. Their activity is particularly important in cellular divisions during spermatogenesis and sperm motility. [16,17]. The two C-terminal tail domains interact with transported cargo through adapters and are connected to the head by a filamentous coiled stem that oligomerizes and regulates dynamics of microtubule [18–21]. A single nucleotide polymorphism (SNP, rs2273063) in the *KIF6* gene was identified through Illumina TruSeq whole exome sequencing (WES). This SNP had a higher prevalence in cases (87.5%) compared to controls (0%). Consequently, we validated SNPs by Sanger sequencing, which demonstrated that

mutations in infertile cases were five times higher (0.65; 13/20) than in fertile controls (0.13; 1/8). Additionally, we performed a random sampling in ordinary men to estimate the mutation frequency of typical men in Taiwan (0.2; 4/20), contrasting it with the Minor Allele Frequency (MAF) (0.04; Table5). This reveals that the incidence of *KIF6* SNPs in collected infertile men in Taiwan is 16.25 times higher than the global MAF, and 3.25 times higher than the ordinary men in Taiwan. Based on Sanger sequencing data, the frequencies observed in our cases and controls suggest that *KIF6* mutations account for 60% of azoospermia cases and 50% of oligospermia cases, respectively. This underscores a greater occurrence of infertility among men (with azoospermia and oligospermia) compared to fertile men (Table4). Notably, its variant SNPs (A/G and A/A) may exhibit potential for genetic diagnosis and serve as markers for the progression of spermatogenesis (Fig. 2).

DRC7 (dynein regulatory complex subunit 7) is predicted to play a role in flagellated sperm motility, which aligns with its involvement in regulating ciliary motility [22,23]. Moreover, *DRC7* in flagellated sperm motility, as well as its potential involvement in spermatogenesis and the development of sperm cells [24]. A single nucleotide polymorphism (SNP, rs3809611) in the *DRC7* gene was identified through Illumina TruSeq whole exome sequencing (WES). The mutation rate of the *DRC7* gene is reported to be 75%, which is lower than the mutation rate of the *KIF6* gene you mentioned earlier. The *DRC7* and *KIF6* genes were the only two genes within the control group that exhibited no mutations. The *DRC7* gene has the lowest global Minor Allele Frequency (MAF) among all candidate genes, with MAF values of 0.24 (global) and 0.16 (east asia). This indicates that the SNP is relatively uncommon in these populations. Both *DRC7* and *KIF6* stand out as promising candidates for meaningful genetic testing in the context of male infertility.

STPG2 (sperm tail PG-rich repeat containing 2) contains a PG-rich motif characterized by a five-residue pattern: P-G-P-x-Y, and forms a similar structure bound to the outer junction of microtubule [25]. The expression profile of STPG2 protein suggests that it might be important in both testicular development and spermatogenesis and its deletion could impair spermatogenesis [26]. A single nucleotide polymorphism (SNP, rs3809611) in the *STPG2* gene was identified through Illumina TruSeq whole exome sequencing (WES). The frequency of mutations among infertile patients was 87.5% (7/8), while the control group with recent fertility exhibited a mutation frequency of 14.3% (1/7). To validate SNP by Sanger sequencing, the ratio of *STPG2*-mutated cases (20/31) to controls (2/8) is approximately only 2.6. Furthermore, it's noteworthy that our random sampling in ordinary men displayed a similar mutation frequency of 40%, which matches with the MAF. Of particular interest, the Azoospermia subgroup showed a mutation rate of 80% (12/15), as did the oligospermia group with 50% mutations (8/16). However, the *STPG2* gene itself presents an almost 40% mutation rate globally, whereas in East Asian populations, this rate drops to 30%. Consequently, when evaluating its potential relevance to male infertility, it's less representative compared to *KIF6* due to its already elevated mutation rate in ordinary men.

NEK2 (NIMA-related kinase 2) NEK2 (NIMA-associated kinase 2) is involved in centrosome duplication, a key process that ensures the correct organization of the microtubule organizing center of the cell [27]. Accurate cell division is essential for proper sperm development. Any disruption caused by abnormal *NEK2* activity could result in damaged sperm cells. These disturbances may lead to diseases such as male infertility [28]. *NEK2* has a global MAF of 0.16 and an East Asian MAF of 0.22. Its relatively low mutation rate can be used as a strong reference index for male infertility. A single nucleotide polymorphism (SNP, rs2230489) in the *NEK2* gene was identified by Illumina TruSeq whole-exome sequencing (WES), and the mutation rate of the cases was 75% (6/8), while the mutation rate of the control group was 14.3 %(1/7).

Other genes affect spermatogenesis by variants analysis

SPATA16 (spermatogenesis-related 16, also known as NYD-SP12) is highly expressed in human testis and localized to the Golgi apparatus and pro-acrosomal vesicles [29], which fuse to form the acrosome during spermatogenesis [30,31]. It was identified as the first autosomal gene and demonstrated that a homozygous mutation (c.848G→A) led to globozoospermia, the production of round headed and acrosomeless spermatozoa. In this study, we report that three closely homozygous

mutations in *SPATA16* (rs16846616), (rs1515442) and (rs1515441) are not only associated with male infertility, but especially azoospermia in our Taiwan cohort. In particular, *SPATA16* (rs16846616) and (rs1515441) co-occur with a higher frequency (80%) in our azoospermic patients of Taiwan.

CFTR (cystic fibrosis transmembrane conductance regulator) is one of the most common genetic mutations leading to azoospermia, which can lead to abnormalities in the male reproductive tract and ultimately result in infertility [32,33]. In our Sanger sequencing study, you found that 51.6% (16 out of 31) of infertile men had *CFTR* mutations. These mutations were absent in ordinary men. Among the infertile men with *CFTR* mutations, 67% of those with azoospermia (lack of sperm in semen) and 37.5% of those with oligospermia (low sperm count) had these mutations. Based on our findings, we establish a strong association between mutations in the *CFTR* gene, specifically the rs213950 mutation, and the occurrence of oligospermia and azoospermia in this cohort of Taiwanese patients.

CMTM2 (CKLF-Like Marvel Transmembrane Domain Containing 2) plays a critical role in spermiogenesis in mice has been studied extensively34].*CMTM2* has been extensively studied and is found to play a crucial role in spermiogenesis in mice. This gene's significance is particularly pronounced during the essential stages of sperm development. *CMTM2-/-* mice were unable to produce sperm, while *CMTM2+/-* mice exhibited a significant reduction in sperm count and motility. A single nucleotide polymorphism (SNP, rs2290182) in the *CMTM2* gene was identified by Illumina TruSeq whole-exome sequencing (WES), and the mutation rate of the cases was 75% (6/8), while the mutation rate of the control group was 14.3 %(1/7). The Global Minor Allele Frequency (GMAF) and East Asian Minor Allele Frequency (MAF) were determined to be 0.17 and 0.28, respectively, indicating a higher prevalence within the East Asian population.

In this study, Ampiseq targeted NGS panel and Illumina TruSeq WES were both used to pinpoint 6 specific genes (*KIF6, DRC7, STPG2, SPATA16, CFTR, CMTM2*) that play a role in microtubule association and spermiogenesis, which are crucial factors in understanding the causes of male infertility.

Conclusions

Male infertility-associated genes can be roughly divided into three affected clusters: tubulin-associated genes (dynein and kinesin), outer junction proteins- associated genes, and spermatogenesis-associated genes. In our experimental results provide pinpointing evidence for a potential link between genetic variations in the *KIF6, DRC7,* and *STPG2* genes and male infertility (Figure 4). The fact that the prevalence of these SNPs was found to be significantly higher in infertile men compared to fertile men suggests that these genetic variations could indeed be associated with the development of male infertility. The alignment of our findings with the significance of microtubule-related genes, particularly in the context of sperm cell structure and function, adds an additional layer of biological plausibility to the observed associations. Microtubule is indeed essential components of the cytoskeleton and play crucial roles in cellular processes, including the formation of sperm heads and tails. The potential for genetic diagnosis based on these findings is promising, as it suggests that certain genetic variations could serve as markers for identifying individuals at a higher risk of male infertility. Our further approach also includes CRISPR gene knockout experiments to gain deeper insights into the precise roles of these candidates and their potential associations with male infertility. By involving larger azoospermic and oligozoospermic cohorts, we anticipate accelerating the identification of novel genes contributing to these phenotypes, ultimately contributing to a more comprehensive understanding of male infertility.
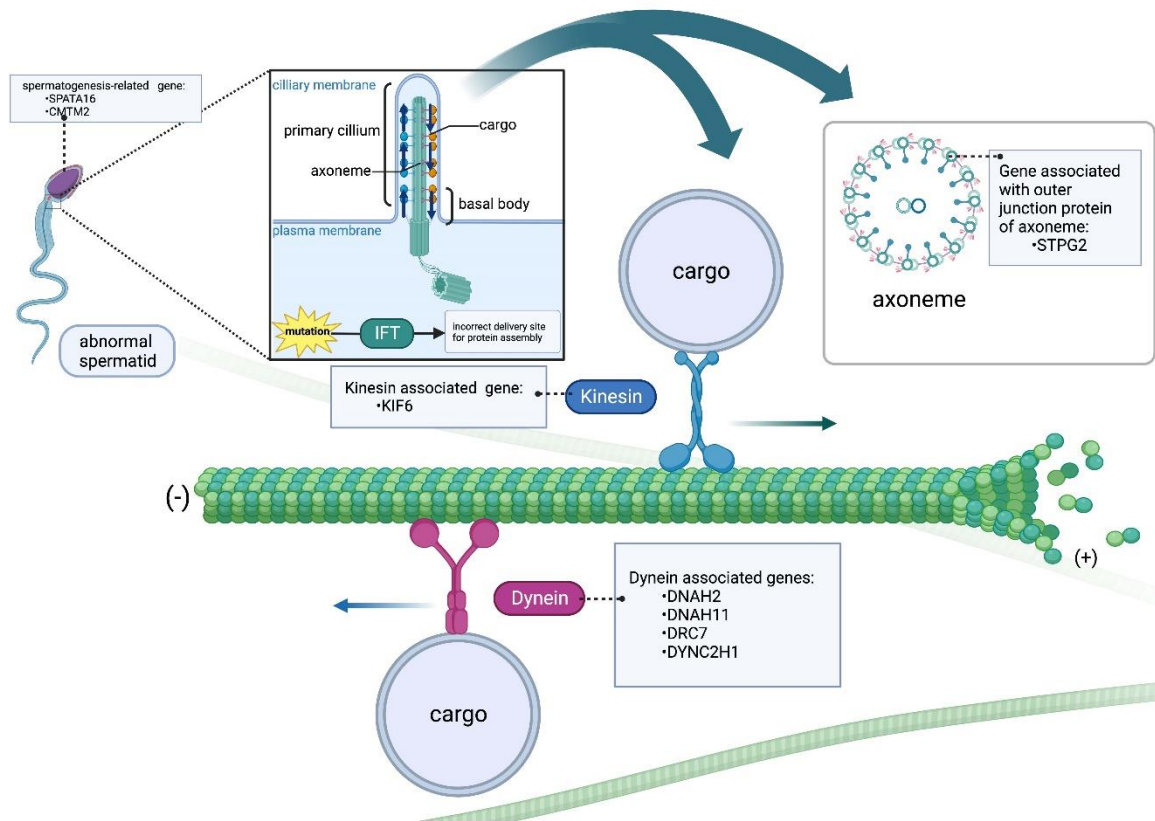
Figure 4. Male infertility-associated genes can be roughly divided into three affected clusters: microtubule-associated genes (dynein and kinesin), outer junction proteins- associated genes, and spermatogenesis-associated genes.

Disclosures: THY, HCT, PKH, JLC, BM&YCH performed the research; CCC, HCT, PKH, GJW&YCH analyzed the data and wrote the paper; THY, CCC&GJW executed the IRB for the study and YCH, BM, CCC&GJW designed the research study and wrote the paper.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations

azoospermia (Azoo); oligozoospermia (Oligo); Y chromosome microdeletion (YCMD); targeted next-generation sequencing (targeted NGS); Illumnia Truseq DNA whole-exome sequencing (WES); minor allele frequency (MAF); CLC Genome Workbench 23.0.4 (CLC) spermatogenesis-related 16 (*SPATA16*); cystic fibrosis transmembrane conductance regulator (*CFTR*); kinesin family member 6 (*KIF6*) ; sperm tail PG-rich repeat containing 2 (*STPG2*); dynein regulatory complex subunit 7 (*DRC7*); NIMA-related kinase 2 (*NEK2*); CKLF-Like Marvel Transmembrane Domain Containing 2 (*CMTM2*); intra-flagellar transport (IFT)

References

1. Smits RM, van de Ven PM, et al. Defining infertility - a systematic review of prevalence studies. Hum Reprod Update. 2019;25(2):209-223.

2. Tahmasbpour E, Balayla J. Natural Conception Rates in Subfertile Couples Following Ovulation Induction and Intrauterine Insemination. J Reprod Infertil. 2014;15(1):17-22.

3. Gunes S, Seli E. Inhibition of microtubules and dynein rescues human sperm function. Fertil Steril. 2020;113(2):338-345.

4. Vale RD, Reese TS, et al. The mechanism of dynein motility: insight from crystal structures of the motor domain. BioEssays. 2003;25(10):961-968.

5. Lehti MS, Sironen A. Formation and function of the manchette and flagellum during spermatogenesis. Reproduction. 2016;151(1):R43-R54.

6. Helmke KJ, Heald R, et al. A novel RII-binding protein (RACK1) that interacts with the protein kinase CβII subfamily. J Biol Chem. 2013;278(8):5143-5153.

7. Kierszenbaum AL, Rivkin E, et al. TNP1 haploinsufficiency in humans leads to a diminished sperm phenotype. Biol Reprod. 2011;85(5):121-126.

8. Mohri H, Ohtake H, et al. Submicromolar concentration of ATP as an energy source for mammalian sperm motility. J Biol Chem. 2012;287(1):7456-7467.

9. O'Donnell L, Stanton PG, et al. Identification of atypical aspermatogenesis in testis biopsies from infertile men. Hum Reprod. 2012;27(12):3331-3337.

10. Zhu F, Liu C, et al. Mutations in PMFBP1 cause acephalic spermatozoa syndrome. Am J Hum Genet. 2016;99(1):1-10.

11. Liska F, Gosele C, et al. Disruption of axonemal microtubule organization causes sperm tail defects in an infertile ccdc39 mouse model. J Cell Sci. 2009;122(21):3316-3324.

12. Pleuger C, Fietz D, et al. The role of fibrous sheath interacting protein 1 in spermiogenesis. Cell Tissue Res. 2016;366(1):135-146.

13. Khan R, Zaman Q, Chen J, et al. Novel Loss-of-Function Mutations in DNAH1 Displayed Different Phenotypic Spectrum in Humans and Mice. Front Endocrinol (Lausanne). 2021;12:765639. Published 2021 Nov 17. doi:10.3389/fendo.2021.765639

14. O'Donnell L, O'Bryan MK. Microtubules and spermatogenesis. Semin Cell Dev Biol. 2014;30:45-54.

15. Dunleavy JEM, O'Bryan MK, et al. Regulation of microtubules in the seminiferous epithelium. Microtubule Dynamics: Methods and Protocols. 2019;1926:119-134.

16. Kato T, Kuramochi-Miyagawa S, et al. Gene knockout of Tptr2, encoding a testosterone-responsive protein, results in impaired sperm differentiation and male fertility in mice. Proc Natl Acad Sci U S A. 2018;115(46):11775-11780.

17. Sweeney HL, Holzbaur ELF. Motor proteins. Cold Spring Harb Perspect Biol. 2018;10(5):a021931.

18. Miki H, Okada Y, et al. Mu2-adaptin dissociation from non-clathrin-coated vesicles is regulated by large Arf1 GTPase-activating proteins. J Cell Biol. 2001;155(2):305-318.

19. Marx A, Nugoor C, et al. Structure of the actin-depolymerizing factor homology domain in complex with actin. J Cell Biol. 2009;186(3):333-340.

20. Hirokawa N, Tanaka Y. Kinesin superfamily proteins (KIFs): Various functions and their relevance for important phenomena in life and diseases. Exp Cell Res. 2015;334(1):16-25.

21. Hirokawa N, Noda Y, et al. KIF1 motor proteins in axonal transport of synaptic vesicles. Cell Mol Life Sci. 2020;77(1):17-31.

22. Smith EF, Lefebvre PA. The role of central apparatus components in flagellar motility and microtubule assembly. Cell Motil Cytoskeleton. 2020;77(2):38-49.

23. Morohoshi A, Aiso S, et al. Mechanism of spermatogenesis and fertility in Drc7, a gene responsible for immotile cilia syndrome. Mol Biol Cell. 2020;31(19):2111-2122.

24. Jones C, van Blerkom J, et al. The dynein regulatory complex is required for ciliary motility and otolith biogenesis in the inner ear. Nature. 2018;553(7689):354-357.

25. Kubo N, Imai F, et al. Mechanism of STPG2-mediated microtubule stabilization in cilia. Nat Commun. 2023;14(1):343.

26. Yakut T, Over V, et al. A novel locus on the X chromosome is linked to dizygotic male infertility. J Med Genet. 2013;50(5):345-353.

27. Singharajkomron N, Yodsurang V, Seephan S, et al. Evaluating the Expression and Prognostic Value of Genes Encoding Microtubule-Associated Proteins in Lung Cancer. Int J Mol Sci. 2022;23(23):14724. Published 2022 Nov 25. doi:10.3390/ijms232314724

28. Ge ZJ, Liang QX, et al. Association study between polymorphisms in NEK2 gene and male infertility. Gene. 2020;731:144347.

29. Ghédir H, Ibala-Romdhane S, et al. Homozygous mutation of SPATA16 is associated with male infertility in human globozoospermia. Am J Hum Genet. 2019;105(2):413-420.

30. Xu H, Zeng X, et al. Testis-specific expression of the mouse gene for the axonemal protein sperm-associated antigen 6 (Spag6). Gene Expr Patterns. 2003;3(2):231-236.

31. ElInati E, Lefièvre L, et al. Globozoospermia is mainly due to DPY19L2 deletion via non-allelic homologous recombination involving two recombination hotspots. Hum Mol Genet. 2012;21(16):3695-3702.

32. Yu Y, Fuscoe JC, et al. A rat gene catalog based on microarray data. Physiol Genomics. 2004;16(3):249-260.

33. Bieniek JM, Drabovich AP, et al. Seminal plasma proteins as potential markers of cystic fibrosis transmembrane conductance regulator (CFTR) function. Mol Cell Proteomics. 2021;20:100107.

34. Zhang T, Li Y, et al. CKLF-like MARVEL transmembrane domain-containing member 2 deficiency impairs spermatogenesis and male fertility in mice† Biology of Reproduction. 2016;94(6):123.