

Mutation analysis in patients with total sperm immotility

Rute Pereira^{1,2} · Jorge Oliveira^{3,4} · Luis Ferraz⁵ · Alberto Barros^{6,7} ·
Rosário Santos^{3,4} · Mário Sousa^{1,3}

Received: 29 January 2015 / Accepted: 1 April 2015 / Published online: 16 April 2015
© Springer Science+Business Media New York 2015

Abstract

Purpose Perform the genetic characterization of five patients with total sperm immotility using Sanger sequencing and Whole Exome Sequencing (WES), in order to increase the knowledge on the genetics of sperm immotility and, ultimately, allow the identification of potential genetic markers for infertility.

Capsule We studied four patients with fibrous sheath dysplasia and one with situs inversus totalis, all with total sperm immotility. Nine new DNA sequence variants were identified. Whole Exome Sequencing (WES) revealed to be the most efficient procedure for the genetic analysis.

Electronic supplementary material The online version of this article (doi:10.1007/s10815-015-0474-6) contains supplementary material, which is available to authorized users.

✉ Mário Sousa
msousa@icbas.up.pt

- ¹ Laboratory of Cell Biology, Department of Microscopy, Institute of Biomedical Sciences Abel Salazar (ICBAS), University of Porto (UP), Rua Jorge Viterbo Ferreira, 228, 4050-313 Porto, Portugal
- ² Department of Biology, Faculty of Sciences, University of Porto, Rua do Campo Alegre 1021/1055, 4169-007 Porto, Portugal
- ³ Multidisciplinary Unit for Biomedical Research-UMIB, ICBAS-UP, Porto, Portugal
- ⁴ Molecular Genetics Unit, Centre of Medical Genetics Dr. Jacinto Magalhães (CGMJM), Hospital Centre of Porto (CHP), Praça Pedro Nunes, 88, 4099-028 Porto, Portugal
- ⁵ Department of Urology, Hospital Eduardo Santos Silva, Hospital Centre of Vila Nova de Gaia (CHVNG), Espinho, E.P.E., Rua Conceição Fernandes, 4434-502 Vila Nova de Gaia, Portugal
- ⁶ Centre of Reproductive Genetics Alberto Barros (CGR), Av. do Bessa, 240, 1º Dto. Frente, 4100-012 Porto, Portugal
- ⁷ Department of Genetics, Faculty of Medicine, University of Porto, Alameda Prof. Hernâni Monteiro, 4200-319 Porto, Portugal

Methods Prospective study at a University Medical school. We analysed five men with total sperm immotility, four with dysplasia of the fibrous sheath (DFS), associated with disruption of several axonemal structures, and one patient with situs inversus totalis, which showed absence of dynein arms (DA) and nexin bridges. We screened 7 genes by Sanger sequencing, involved in sperm motility and associated to ultrastructural defects found in these patients (*CCDC39*, *CCDC40*, *DNAH5*, *DNAI1*, *RSPH1*, *AKAP3* and *AKAP4*). Additionally, we performed WES analysis in the patient with situs inversus.

Results We identified nine new DNA sequence variants by WES. Two of these variants were considered particularly relevant: a homozygous missense change in *CCDC103* gene (c.104G > C, p.R35P) probably related with absence of dynein arms; the other in the *INSL6* gene (c.262_263delCC) is thought to be also involved in sperm immotility.

Conclusions Our work suggests that WES is an effective strategy, especially as compared with conventional sequencing, to study highly heterogenic genetic diseases, such as sperm immotility. For future work we expect to expand the analysis of WES to the other four patients and complement findings with expression analysis or functional studies to determine the impact of the novel variants.

Keywords Dysplasia of the fibrous sheath (DFS) · Genetic diagnosis · situs inversus totalis · Sperm immotility · Whole Exome Sequencing (WES)

Introduction

The axoneme (Ax) is the flagellar motor of the sperm cell and extends throughout the flagellum. The human sperm Ax contains nine peripheral microtubule doublets (A, B) and a central

pair of microtubules. Doublets are linked by nexin bridges, present dynein arms (DA) (ODA: outer and IDA: inner) and are linked to the central microtubule pair by radial spokes (RS). The central pair, with its surrounding fibrillar sheath and central bridge is named the central pair complex (CPC). The Ax is surrounded by the outer dense fibres (ODF) and by the fibrous sheath (FS) [1]. The FS is an important element since influences/modulates the bending and the flagellar beat and in addition provides energy for motility [2]. Sperm motility depends on DA motor complexes [3], and their ATPase activity is regulated by a nexin-dynein regulatory complex (DRC) that links doublets; and on RS, which provides a radial scaffold and facilitates signal transduction from the CPC to DA, thus governing flagellum beat and waveform [4, 5].

Absolute asthenozoospermia, that is total sperm immotility, is reported at frequency of 1 of 5000 men, with the two major causes being the ultrastructural defects in the sperm flagellum (due to both genetically inherited and congenital defects) and necrozoospermia [6]. Primary ciliary dyskinesia and dysplasia of the FS are two main disorders associated with sperm immotility.

Primary ciliary dyskinesia (PCD, OMIM: #244400) is a heterogeneous, autosomal recessive disease characterized by immotile cilia, that result in chronic infections of the upper respiratory tract. Kartagener syndrome (KS) is characterized by a combination of *situs inversus*, chronic sinusitis and bronchiectasis. It occurs in about 50 % of PCD patients and the majority of these patients present severe sperm immotility [7, 8]. The characteristic structural feature of PCD is absence of DA, consequently research on the genetic basis of PCD started focusing on DA [9]. Nonetheless, several genes associated with more flagellar structures are currently associated with PCD [10]. Mutations in *DNAI1* and *DNAH5* genes lead to ODA defects and are the major cause of PCD [9, 11]. Mutations in *CCDC39* and *CCDC40* were associated with defects in IDA, nexin links, RS, CPC and doublets defects [12, 13], and a recent study found mutations in these genes among 69 % of individuals (37/54 of families) with PCD [14]. Finally, mutations in the *RSPH1* gene were also found associated with PCD in patients with defects in RS and CPC [15].

Dysplasia of the FS (DFS) is diagnosed by marked hypertrophy and hyperplasia of the FS in association with deficiency or absence of the annulus, mitochondria, the CPC (50 % of the cases) and/or of the DA [16]. The incidence reported in familial studies and the association with dynein deficiency in some cases, suggests that the DFS may have an autosomal recessive inheritance [17]. The major components of the FS are the A-kinase anchor proteins AKAP4 and AKAP3, to which several important roles have been attributed [18]. The AKAP4 gene encodes a protein involved in FS assembly [19], whereas AKAP3 is involved in FS structure organization [20].

Although mutations in these genes were suggested to cause DFS [21, 22] there is no evidence for their involvement in the pathogenesis of DFS.

In the present study, we analysed five patients with total sperm immotility, four cases with DFS and one patient presenting *situs inversus totalis*. Given the specific axoneme changes observed in these patients, the genes *CCDC39*, *CCDC40*, *DNAH5*, *DNAI1*, *RSPH1*, *AKAP3* and *AKAP4* were screened by Sanger sequencing to identify possible genetic alterations that could explain patient phenotypes. Additionally, in the patient with *situs-inversus* Whole Exome Sequencing (WES) was performed. The objective of the present work was to increase the knowledge on the genetics of sperm immotility to allow future identification of potential genetic biomarkers and treatments.

Material and methods

Ethical considerations

Ethical guidelines were followed in the conduct of research, with written informed consent having been obtained before the beginning of the present work. This work did not involve human or animal experiments. An approval by an Ethics Committee and the provisions of the Declaration of Helsinki as revised in Tokyo 2004 on human experimentation does not apply to this kind of work. According to the National Law on Medically Assisted Procreation (Law 32/2006) and the National Council on Medically Assisted Procreation guidelines (CNPMA, 2008), databases, semen and blood samples from the five individuals with total absence of sperm motility were processed and analysed without the need of further ethical obligations. Patient characteristics and the ultrastructural analysis of sperm will be presented elsewhere.

DNA extraction

Patient genomic DNA (gDNA) was extracted from peripheral blood [23].

Genes and primer design

Seven different genes that code for essential proteins of sperm flagellum were selected: *AKAP3* (NM_001278309.1; 12p13.3), *AKAP4* (NM_003886.2; Xp11.2), *CCDC39* (NM_181426.1; 3q26.33), *CCDC40* (NM_017950.3; 17q25.3), *DNAH5* (NM_001369.2; 17q25.3), *DNAI1* (NM_001281428.1; 9p13.3) and *RSPH1* (NM_080860.2, 21q22.3). All exonic regions were included in the analysis, with the exception of the *DNAH5* gene. Considering the size of this gene, screening was carried out on only 40 of the 79

exons, known to harbour the majority of the reported mutations [11].

The reference sequences of these genes were retrieved using the National Center for Biotechnology Information (NCBI: <http://www.ncbi.nlm.nih.gov/>) and the UCSC (University of California Santa Cruz, California, USA) Genome Browsers (<http://genome.ucsc.edu/index.html>). Primer Express software (Life Technologies; Foster City, California, USA) was used for design primers. Parameters used for primer design were as follows: primer length: 18–25 bp; GC content: 40–60 %; amplicon length: 400–650 bp; and melting temperature: 59–62°C. Each designed primer pair was tested for the presence of dimer formation using the FastPCR software (version 3.7.7; Institute of Biotechnology, University of Helsinki, Finland) and for their specificity towards the regions of interest using the PrimerBlast tool (NCBI, Bethesda, USA) [24].

Polymerase chain reaction

The Polymerase chain reaction (PCR) mixture used for standard conditions (30 µl) contained: 15 µl of PCR Master Mix (Promega, Madison, USA); 12 µl of sterile bidistilled water, 1 µl of each primer at 10 pmol/µl each (Thermo Fisher Scientific; Einsteinstrasse, Germany) and 1 µl of DNA at 100 ng/ul. To amplify genomic regions with higher complexity, Master Mix and slight different PCR conditions were used. PCR conditions were optimized for each primer pair, using a DNA sample from a healthy fertile male. PCR was performed in a 9700, 9800 or VERITI (Life Technologies) thermocycler. PCR products were analyzed by 2 % agarose gel electrophoresis: mix of TAE 1x SeaKem LE Agarose (Lonza, Rockland, USA), and 5 µl/100 ml of GelRed Nucleic Acid Gel Stain, 10,000X in water (Biotium, California, USA).

Sanger DNA sequencing

Following successful PCR amplification, reactions were enzymatically purified using Illustra ExoStar kit (GE Healthcare, Buckinghamshire, UK). New asymmetric PCR reactions were prepared based on the cycle sequencing method, using the BigDye Terminator v1.1 Cycle Sequencing Kit (Life Technologies). Obtained PCR products were purified using Performa DTR (Dye Terminator Removal) (EdgeBio, Gaithersburg, USA) and analysed by capillary electrophoresis in a 3130xl genetic analyzer (Life Technologies). Sequencing data analysis was performed using SeqScape V2.5 software (Life Technologies). Variants were interpreted using in-house software: Variobox [25]; and external on-line resources: Exome variant server (EVS) (<http://evs.gs.washington.edu/EVS/>) and NCBI database of single nucleotide polymorphism (dbSNP) (<http://www.ncbi.nlm.nih.gov/projects/SNP/>). To perform further bioinformatics analyses

of the rare or new sequence variants, the following bioinformatic tools were used: Polyphen-2 (Polymorphism Phenotyping v2) [26]; SIFT [27]; MutationTaster (MT) [28]; Human Splicing Finder [29]. Some of these tools are also integrated in the commercial software Alamut Visual V2.4 software (Interactive Biosoftware, Rouen, France).

Whole-exome sequencing

The exome of the patient with KS was sequenced according to Oliveira and co-workers [30], using the AmpliSeq strategy on an Ion Proton next-generation sequencing (NGS) platform (Life Technologies). Variant calling was performed by running Torrent Variant Caller plugin version 4.0, using the optimized parameters for WES recommended for AmpliSeq sequencing (Life Technologies). The obtained Variant Call Format (VCF) file was analysed with GEMINI software [31] for annotation and prioritization. Variant filtering was performed using a list of 67 candidate genes (Suppl. data 1) and variants matching the autosomal recessive disease model (either homozygous or two heterozygous changes in the same gene). The Ion Reporter™ Software (Life technologies) was also used to filter rare variants by Gene Ontology [32], using the keywords: sperm, flagellar, motility, axoneme and dynein. Candidate variants were manually checked on the Binary Alignment Map (BAM) file through GenomeBrowse version 2.0.2 (Golden Helix, Bozeman, USA). Seven variants were selected and submitted to PCR amplification and Sanger sequencing for confirmation, as previously described.

Results

Genetic analysis by Sanger sequencing

Seven candidate genes (*CCDC39*, *CCDC40*, *DNAH5*, *DNAI1*, *RSPH1*, *AKAP3* and *AKAP4*) were analysed by Sanger sequencing, in order to identify sequence variants that could have correlation with patient phenotypes. Novel variants (reported here for the first time) and rare variants (whose frequency on the population is unknown or lower than 1 %, the cut-off value to distinguish between a rare variant and a polymorphism) are listed in Table 1 together with their bioinformatic analysis (Electropherogram of variants are shown in Suppl. Fig. 1).

The sperm of patient 1 presented DFS with loss of the CPC and RS A total of 48 DNA sequence variants were detected, with 43 being considered polymorphic, due to their high frequency on a control population (based on EVS and dbSNP). Of the remainder, 4 were rare variants and 1 variant was novel. Regarding the rare variants, 3 were found in the *DNAH5* gene and one in the *DNAI1* gene. Concerning the rare variants in

Table 1 Rare and novel sequence variants found by Sanger Sequencing and their respective bioinformatic analysis

Exonic DNA variants		Bioinformatic analysis						
Gene	Patient	DNA change (Location)	Freq.	Protein change	PolyPhen 2	SIFT	HSF	Mutation taster
<i>CCDC39</i>	3 (He)	c.233G > A (exon 3)	0.56 %	p.R78H	Benign (<i>s</i> =0.004)	Tolerated (<i>s</i> =0.28)	No effect on Splice	P (<i>p</i> =0.90)
	1 (He)	c.2540A > G (exon 18)	New variant	p.E847G	Benign (<i>s</i> =0.165)	Tolerated (<i>s</i> =0.06)	ESE (wt: 72.9 -mt: 82.9)	Disease causing (<i>p</i>=0.69)
<i>CCDC40</i>	4 (He)	c.2682G > A (exon 16)	0.61 %	P (=)	–	–	No effect on Splice	P (<i>p</i> =0.90)
Intronic DNA variants		Bioinformatic analysis						
Gene	Patient	DNA change (Location)	Freq.	Protein change	Splice Algorithms	Splice Algorithms	Mutation taster	
<i>CCDC40</i>	3 (He)	c.2620-92C > T (Intron 15)	New variant		No effect on Splice	No effect on Splice	P (<i>p</i> =0.9)	
<i>DNAH5</i>	3 (Ho)	c.1537-102 T > A (Intron11)	ND		No effect on Splice	No effect on Splice	P (<i>p</i> =0.9)	
	4 (Ho)	c.1537-100_1537-99delTT (Intron 11)	ND					
	1 (He)							
	1–4 (He)	c.3835-3delT (intron24)	0.1 %		↓ ASS	nt:92.6 – mt:89.1 (SpliceSiteFinder-like)	Disease causing (<i>p</i>=1)	
						nt:12.8 – mt:10.4 (MaxEntScan)		
						nt:11.9 – mt:9.8 (GeneSplicer)		
						nt:85.8– mt:85.3 (HSF)		
						No effect on Splice	P (<i>p</i> =0.9)	
	3 (He)	c.5882 + 133A > G (Intron 35)	New variant					
	2 (Ho) 4,5 (He)	c.7408-84_7408-83delAT (intron 44)	ND					
	1 (He)	c.10282-81delT (Intron60)	ND					
	3 (He)	c.10872 + 84 T > C (Intron63)	ND					
	3 (He)	c.11570 + 124G > C (Intron67)	New variant					
<i>DNAH1</i>	1 (He)	c.81 + 61A > G (Intron2)	0.3 %		↑ ASS	wt:54.03/ mt:82.97; nt: 83.19(HSF)	Disease causing (<i>P</i>=0.98)	

Human Splice Finder (HSF): <http://www.umd.be/HSF/>; PolyPhen-2: <http://genetics.bwh.harvard.edu/pph2/>; SIFT: <http://sift.jevl.org/>; and the MutationTaster: <http://www.mutationtaster.org/>. Scores higher than 0.5 in Polyphen are considered damaging and in SIFT scores less than 0.05 are considered deleterious. In MutationTaster a probability (p) closer to 1 indicates a high confidence in the prediction. At bold are the variants with predict pathogenic impact

Freq. Frequency within control population (based on Exome variant server and SNP database from NCBI), *He* Heterozygous, *Ho* Homozygous, *ND* Variant listed in databases but with undetermined frequency on population, *R* Arginine, *H* Histidine, *E* Glutamic acid, *G* Glycine, *s* score, *p* probability, *P* Polymorphism, *ASS* Acceptor Splice Site, *wt* reference score (i.e., the score of non-mutated sequence), *mt* mutant score (i.e., the score of mutated sequence), *nt* native splice site (i.e., normally occurring splice site), *ESE* Exonic Splice Enhancer

DNAH5 gene the variants c.1537-100_1537-99delTT and c.10282-81delT were predicted as non-pathogenic; and the heterozygous variant c.3835-3delT was detected in all DFS patients (1–4). Consequently, was analysed by four splice algorithms, incorporated in Alamut Visual V2.4 software (Interactive Biosoftware), and was predicted to lead to a slight reduction of the score of native (nt, i.e., normally occurring) splice site acceptor (SSA): from 92.6 to 89.1 in SpliceSiteFinder-like; 12.8 to 10.4 in MaxEntScan; 11.9 to 9.8 in GeneSplicer and 85.85 to 85.29 in HSF. In addition, the MutationTaster attributed a probability of being pathogenic ($p=1$). Regarding the other rare variant **c.81 + 61A > G** located in intron 2 of the *DNAH1* gene, the bioinformatic analysis predicted an increase of 53.56 % of the score of SSA, making the mutant score (mt, i.e., the score of mutated sequence) almost equivalent to the score of nt (mt: 82.97; nt: 83.19). This variant is also predicted to be pathogenic by the MutationTaster ($p=0.98$). The new variant is a missense change, located in exon 18 of the *CCDC39* gene (**c.2540A > G**), predictably leading to a change in the protein (p.E847G). However both Polyphen-2 and SIFT predict this amino-acid change as non-pathogenic, respectively with scores of 0.17 and 0.066. In turn, the MutationTaster tool predicts that this variant may have a low probability of being disease causing ($p=0.69$). According to HSF analysis, there is a marginal increase of the score of the alternative splicing factor SF2/ASF, an exonic splicing enhancer (ESE) motif (wt score: 72.9 and mt score: 82.9).

Sperm from patient 2 presented DFS with a variable number of doublets and RS. He had 50 DNA variants: 48 were polymorphisms and 2 were rare variants in the *DNAH5* gene: the variant c.3835-3delT and the homozygous variant c.7408-84_7408-83delAT. The latter was classified as a polymorphism.

Sperm of patient 3 presented DFS with absence of DA, nexin bridges and CPC The patient presented 59 DNA variants, with 52 polymorphisms, 4 rare variants and 3 new variants. The heterozygous variant c.233G > A in *CCDC39* has a lower frequency (0.6 %) in the control population and leads to an amino acid change (p.R78H), but no pathogenic impact is associated. The variant c.3835-3delT in *DNAH5* gene is also present in this patient. Three new variants (1: **c.2620-92C > T** in *CCDC40* and 2 in *DNAH5*: **c.5822 + 133 A > G** and **c.11570 + 124 G > C**) and 2 rare variants (both in *DNAH5*: c.1537-102 T > A and 10872 + 84 T > C) are foreseen as probably non-pathogenic.

Sperm of patient 4 presented DFS with absence of the CPC and RS He had 51 variants were identified: 47 frequent variants and 4 rare variants. Three rare heterozygous variants were detected in *DNAH5* (c.1537-100_1537-99delTT, c.3835-3delT and c.7408-84_7408-83delAT), also identified

in other patients of this study; and an additional rare variant, was found in *CCDC40*: c.2682G > A. With the exception of c.3835-3delT, all rare variants detected in this patient were polymorphisms without predicted pathogenic impact.

Patient 5 has *situs inversus totalis*, and sperm showed absence of DA and nexin links He presented 43 DNA variants none of which with predicted pathogenicity.

Whole exome sequencing

The results in the previous section were considered insufficient to identify potential genetic markers which may explain the phenotype displayed by our patients. Considering the morphological defects in complex structures with an important role in sperm motility, it is conceivable that several genes might be involved. It was therefore decided to perform WES analysis of patient 5 (with KS) with the aim of identifying the genetic cause underlying this phenotype

The WES run generated 37,737,917 sequence reads, 99 % of which were efficiently aligned against the hg19 human reference genome. The exome's target regions were on average covered 108.9 times, with 93 % having coverage of over 20 %. In addition, 84 % of all sequenced bases had a Phred quality score above Q20 [33, 34]. These values are above the generally defined standard threshold and thus acceptable for this study. Analysis identified 49,230 single nucleotide variants (SNV) and 3479 insertions/deletions that were listed in a VCF file.

In order to find a possible genetic explanation for total sperm immotility in more than 50,000 DNA sequence variants obtained, we applied two filter approaches. As an initial approach, the GEMINI database framework [31] was used to filter the variants matching the autosomal recessive disease model, using the list of 67 candidate genes previously selected reported in literature as being associated with sperm immotility due to flagellar abnormalities. This filtering narrowed down the analysis to 258 sequence variants. The possible involvement of other *loci* in the patient's phenotype was also verified. To that end, the Ion Reporter Software (Life technologies) was used to filter the rare variants by Gene Ontology. This generated a further 46 DNA variants. With the application of these two approaches, the total number of variants was reduced from 52,709 to 304. The filtering steps gave priority to rare variants located in exonic regions and those placed within intronic regions 8 bp adjacent to the exons. Therefore, 167 intronic variants and 53 variants with a frequency above 1 % were excluded. Of the remaining 84 variants, 51 were synonymous, 30 were missense and 3 were frame-shifts. The 51 synonymous variants were evaluated using bioinformatic tools and subsequently excluded from further analysis since no effects on splicing were predicted. 17 of the 30 missense variants belonged to the gene *HYDIN*

located in chromosome 16q22.2, which has an identical paralogous segment inserted in chromosome 1q21.1 [35]. The high number of novel variants found in this gene may be attributed to this duplication, as previously suggested [36], and therefore these changes were excluded from further analysis.

The remaining 16 variants were manually checked on the BAM file using the software GenomeBrowse version 2.0.2. (Golden Helix). 9 variants were considered sequencing artifacts and the remaining 7 were selected for confirmation by Sanger sequencing (Suppl. Fig. 2). 6 of these were confirmed as true variants (5 missense changes and 1 frame-shift deletion), with 5 being novel (Table 2; Fig. 1).

The **novel homozygous variant c.104G > C** located in exon 2 of the **CCDC103** gene (NM_001258398.1; 17q21.31) leads to a substitution of a conserved amino-acid arginine (R) by a proline (P) at position 35 (Fig. 1a). Three different bioinformatic tools suggested that this variant is potentially pathogenic, with the scores of 0 in SIFT, and 0.99 in Polyphen-2 and MutationTaster. The heterozygous deletion of two C nucleotides at position 262 of exon 1 in **INSL6** (NM_007179.2; 9p24), alters the gene's open reading frame and may result in a completely different translation (Fig. 1b). MutationTaster evaluates this variant as disease-causing ($p=1$).

Two different heterozygous missense variants were identified in genes that encode proteins of the dynein heavy chain (HC): **DNAH10** (NM_207437.3; 12q24.31) and **DNAH6** (NM_001370.1; 2p11.2) (Fig. 1 c, d). According to Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) [37], these two genes may interact with a score of 0.540. The heterozygous missense variant c.7895C > T in **DNAH10** has a very low frequency (0.13 %). This, aligned to the fact that this gene encodes part of the DA HC, prompted further analysis of this variant, predicted to cause a change of the polypeptide sequence (p.T2632M), which is considered to be damaging by SIFT ($s=0.02$) and Polyphen-2 ($s=0.98$) and as disease-causing ($p=0.99$) by MutationTaster. The second variant, c.3167A > T in the **DNAH6** gene, is firstly identified and is also a heterozygous missense change in another conserved amino-acid (p.D1056V) that is predicted to be damaging in all bioinformatic tools used.

The two final missense heterozygous variants p.N276K and p.R1482L were found in the genes **GAS8** (NM_001481.2; 16q24.3) and **SPAG17** (NM_206996.2; 1p12), respectively (Fig. 1e, f). The change of the amino-acid Asparagine (N) by the amino-acid lysine (K) at position 276 in GAS8, is predicted as being possibly pathogenic by MutationTaster and by Polyphen-2 ($p=0.99$ and $s=0.72$, respectively); but, according to SIFT ($s=0.38$) no pathogenic impact is expected. For the **SPAG17** variant (c.4445G > T), no pathogenic impact is predicted, despite the change of an R to a leucine (L) at position 1482.

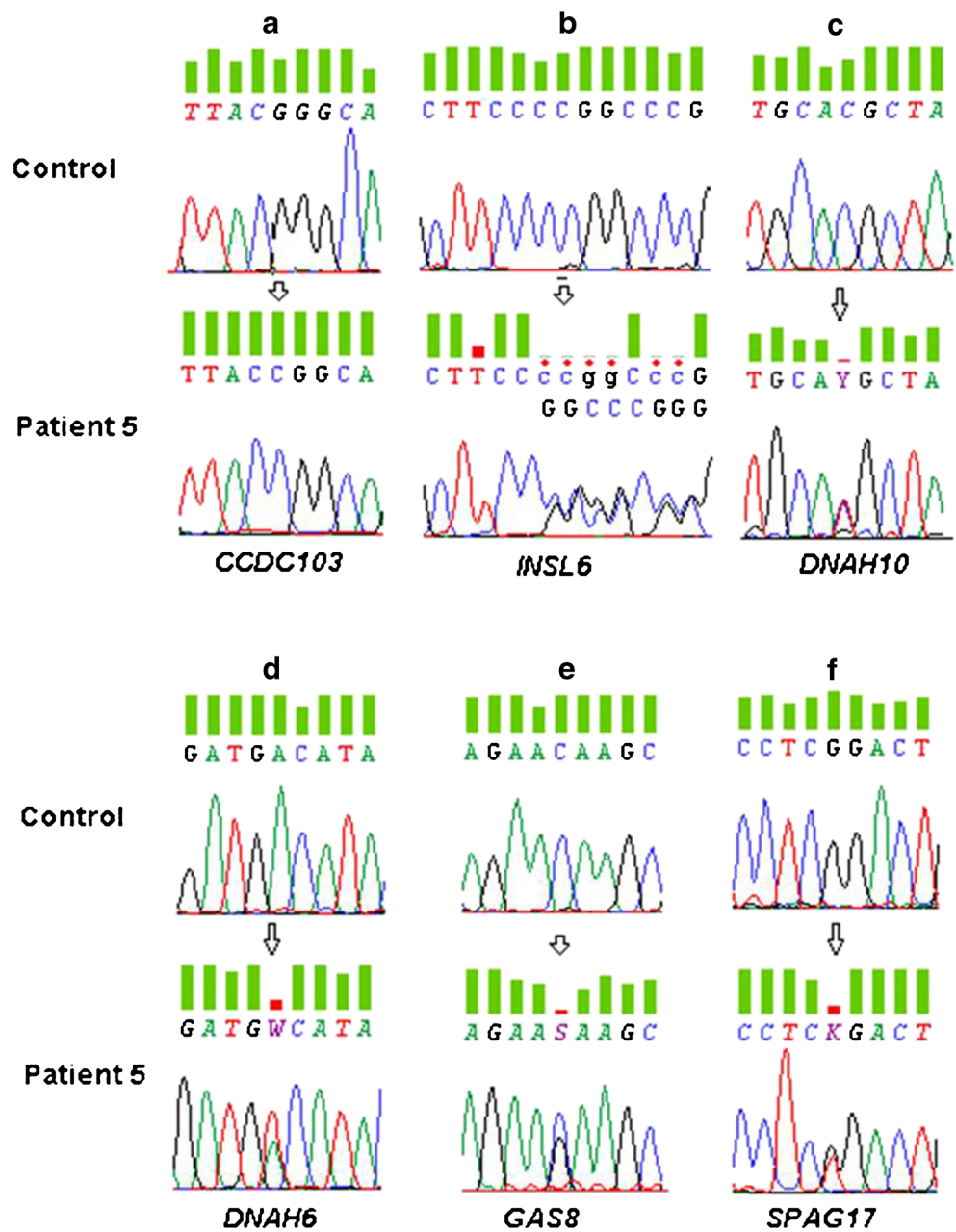
Table 2 List of candidate variants detected by exome sequencing and their respective bioinformatic analysis

Genomic localization	Gene	Transcript	Exon	DNA change/ Zygosity	Protein change	Freq.	SIFT	PolyPhen2	MutationTaster	Phenotype ^a
chr17:42978470	CCDC103	NM_001258398.1	2	c.104G > C Homozygous	p.R35P	New variant	Damaging ($s=0$)	Damaging ($s=0.99$)	Disease Causing ($p=0.99$)	PCD [50]
chr9:5185339	INSL6	NM_007179.2	1	c.262_263delCC Heterozygous	p.P88Gfs*27	New variant	na	na	Disease Causing ($p=1$)	Spermatogenic Failure [52, 54]
chr12:124362332	DNAH10	NM_207437.3	47	c.7895C > T Heterozygous	p.T2632M	0.13 %	Damaging ($s=0.02$)	Damaging ($s=0.98$)	Disease Causing ($p=0.99$)	no phenotype associated
chr2:84832709	DNAH6	NM_001370.1	20	c.3167A > T Heterozygous	p.D1056V	New variant	Damaging ($s=0$)	Damaging ($s=0.98$)	Disease Causing ($p=0.99$)	no phenotype associated
chr16:90103711	GAS8	NM_001481.2	7	c.828C > G Heterozygous	p.N276K	New variant	Tolerated ($s=0.38$)	Possibly Damaging ($s=0.72$)	Disease Causing ($p=0.99$)	Sperm immotility/ PCD [45]
chr1:118550809	SPAG17	NM_206996.2	31	c.4445G > T Heterozygous	p.R1482L	New variant	Tolerated ($s=0.23$)	BENIGN ($s=0.005$)	Polymorphism ($p=0.9$)	PCD [58]

SIFT: <http://sift.jevl.org>; PolyPhen-2: <http://genetics.bwh.harvard.edu/pph2/> and the MutationTaster: <http://doro.charite.de/MutationTaster/>. Scores higher than 0.5 in Polyphen are considered as damaging and in SIFT scores less than 0.05 are considered deleterious. In MutationTaster a probability (p) closer to 1 indicates a high confidence in the prediction. At bold are the variants with pathogenic impact. *Freq.* Frequency within control population (based on Exome variant server (<http://evs.gs.washington.edu>) and SNP database from NCBI (<http://www.ncbi.nlm.nih.gov/snp>), *D* Aspartic acid, *L* Leucine, *G* Glycine, *K* Lysine, *M* Methionine, *N* Asparagine, *R* Arginine, *P* Proline, *T* Threonine, *V* Valine, *s* score, *p* probability, *PCD* primary ciliary dyskinesia, *na* not applicable

^a Previously described phenotypes associated to gene defects

Fig. 1 a-f. Sanger sequencing electropherograms of the variants selected from exome sequencing analysis from patient 5. **a.** variant c. 104G > C in *CCDC103* gene, **b.** variant c.262_263delCC in *INSL6* gene, **c.** variant c. 7895 C > T in *DNAH10*, reported in public variant databases (frequency on control population is 0.0013), **d.** variant c.3167A > T in *DNAH6* gene, **e.** variant c.828C > G in *GAS8* gene and **f.** variant c.4445G > T in *SPAG17* gene



Discussion

We report the identification of nine novel DNA sequence variants (<http://www.ncbi.nlm.nih.gov/clinvar/submitters/505199/>) and ten rare variants, of which 5 novel (c.2540A > G, c.104G > C, c.3167A > T, c.828C > G, c.262_263delCC) and 3 rare (c.3835-3delT, c.7895C > T and c.81 + 61A > G) have been predicted to be deleterious.

The novel variant c.2540A > G in the *CCDC39* gene and the rare variant c.81 + 61A > G in the *DNAH11* gene, are both predicted as disease causing by MutationTaster and to possibly influence splicing, which could lead to a competition for the splice site and, ultimately, have effects on protein's

function. The heterozygous deletion in the *DNAH5* gene (c.3835-3delT) was present in all patients with the diagnosis of DFS. Nevertheless, this variant has a very low frequency (<0.1 %) and may also affect splicing. The gene *DNAH5* codes for an axonemal HC dynein protein [38] and assuming the hypothesis of its pathogenicity, this heterozygous mutation may not be sufficient to cause the absence of DA, but can have an effect on the flagellar movement. Two previous works detected mutations in another gene - *DNAH11*- that encodes a HC dynein protein in patients with abnormal ciliary beating pattern but with a typical Ax ultrastructure [39, 40]. Notwithstanding the robustness of bioinformatic tools, expression analysis are required to evaluate their effect on

splicing. Unfortunately, these studies were not possible due to inaccessibility to RNA from the patient.

In patients 1 to 4, mutations in *AKAP3* and *AKAP4* were expected, given their roles in sperm motility [18]. Previously, it was demonstrated that male mice lacking *AKAP4* were infertile due to reduced sperm motility, and that the FS was incompletely developed [41]. Other studies detected deletions in *AKAP3/4* genes and a moderate diffuse signal on immunostaining for human AKAP4 protein in DFS patients [21, 22]. However, our work did not add further evidence towards the hypothesis that *AKAP3/4* are the genetic causes of the DFS phenotype, which are in agreement with other two reports where gene mutations were also not detected in these genes [42, 43]. This demonstrates that the role of *AKAP3/4* is still unclear and leads us to hypothesize that DFS is a multigenic disease that may be caused by not yet identified genes related with FS function and/or assembly; or could be caused by spermiogenic genes related to formation, transport or attachment of FS components. Further studies about molecular elements and mechanisms of intraflagellar transport of cell factors involved in human FS development and structure need to be done to fully understand the genetic cause(s) of DFS.

Regarding patient 5, considering his phenotype (lack of DA and nexin bridges), we are not expect to detect only polymorphic variants. However, given the genetic heterogeneity of sperm motility and the complexity of Ax components, is it likely that several genes may be involved. As a result, a WES analysis was performed in this patient. From the 6 candidate variants confirmed by Sanger sequencing, only the heterozygous variant in *SPAG17* was predicted to be a polymorphism by all bioinformatic tools. The *GAS8* was shown to be required for normal motility of cilia [44] and was associated with tail maturation, and to the attachment of the DRC to microtubules [45, 46]. The *GAS8* variant (p.N276K) occurs in a highly conserved central region (pfam number: 13851) and two bioinformatic tools predict this as possibly being pathogenic, which may suggest that may have some association with patient's phenotype.

The missense variants c.7895C > T in *DNAH10* and c.3167A > T in *DNAH6* are both located in genes encoding dynein HC protein. Each HC comprises an N-terminal domain; a C-terminal segment, that includes a highly conserved central section with six AAA modules (AAA for: ATPases associated diverse cellular activities); and a microtubule-binding domain [47]. According to NCBI Conserved Domain Database, the variant in *DNAH10* codes for a residue located in the third AAA module, a highly conserved region, which contains a P-loop (P3) with an ATP binding site. It was showed that mutations affecting P3 appeared to block ATP binding and hydrolysis, which could ultimately may affect sperm motility [48]. The variant in *DNAH6* corresponds to an amino-acid change in a residue located in the N-terminal region, which is important to specify the intracellular location

of the dynein isoform by the binding of accessory proteins to this domain and is required for assembly of the dynein particle [47]. To our knowledge, neither of these two dynein HC genes have been associated with sperm motility in humans. However, RNAi knockdown of *DNAH10* in *Trypanosoma brucei* lead to flagellum immotility and to structural defects in the Ax [49]. The KS has an autosomal recessive inheritance pattern, consequently, these 2 heterozygous variants, by themselves, may not be sufficient to cause the lack of DA. But, if the predicted interaction of these proteins, as suggested by STRING, has a biological meaning, these changes may be sufficient to prevent DA assembly or formation.

In the *CCDC103* gene, the homozygous missense variant, c.104G > C is predicted to be damaging. Multiple alignments of the protein sequence showed that the affected arginine residue is phylogenetically highly conserved. The *CCDC103* gene is a DA attachment factor [50], which reinforces the hypothesis that this variant could be responsible for the lack of DA in our patient. *CCDC103* mutations have been detected in PCD patients with absence of IDA and/or ODA, *situs inversus* and paralysis of respiratory cell cilia beat [50]. However, the referred study did not focus on the patients' fertility, and there appear to be no other studies that relate sperm immotility to mutations in this gene. Therefore, this is the first report of a DNA sequence variant in *CCDC103* with a predicted pathogenic impact associated with sperm motility in a patient with total sperm immotility, absence of DA and *situs inversus totalis*.

The *INSL6* gene, located at chromosome 9p24, belong to the relaxin family of peptide hormones whose members are involved in several reproductive functions [51], and is predominantly expressed in male germ cells [52]. *INSL6* deficiency was shown, in mice, to cause an arrest of spermatogenesis at late stages of meiotic prophase, which lead to a reduction of sperm production and immotility, and thus proposed that *INSL6* gene are required for normal spermatogenesis [53]. Similarly, a heterozygous missense mutation in exon 2 of *INSL6* was found in a patient with spermatogenic failure, and the authors proposed that this mutation might be causal, due to the probable disruption of *INSL6* pro-hormone processing [54]. We identified a novel heterozygous deletion of two C nucleotides in the coding region (c.262_263del). Despite being heterozygous, this variant disrupts the reading frame, thereby predictably affecting the protein function. Thus, ultimately, according to previous reports about its function, this deletion might compromise spermatogenesis and hence cause the sperm motility of our patient.

Overall, both the new homozygous missense variant in *CCDC103* and the frame-shift variant in *INSL6* are the most likely to lead to, respectively, the absence of DA and the total sperm immotility in the patient with *situs inversus*. Nonetheless, to understand the real impact of these variants,

further studies, namely expression analysis and knock-in animal models, are essential.

Considering the genetic heterogeneity of PCD/FSD and the high complexity of sperm motility, WES seems currently the most effective technical approach available. The automated Sanger sequencing method is considered the “gold-standard” for genetic testing, since it is highly accurate; however, it has lower throughput and is thus more laborious and costly than NGS. WES is described as a powerful discovery tool and it has already contributed to the identification of new genetic causes of PCD [55, 56]. However, the use of WES still faces some limitations, such as the considerable rate of false-positive variants generated by misaligned reads or sequencing errors [57] and consequently Sanger sequencing is still needed to confirm the results.

In conclusion, our study highlights the difficulties in the identification the genetic causes of complex diseases using conventional Sanger sequencing, and demonstrates that WES is an efficient approach to increase our knowledge about the genetic causes of sperm immotility and infertility. Further, this study identified nine novel variants, with two being likely to be considered as genetic markers of sperm immotility.

Acknowledgements We would like to acknowledge: Helena Oliveira, MSc, ESHRE Senior clinical embryologist; Ilda Pires, MSc, ESHRE Senior clinical embryologist and Madalena Cabral, BSc, Embryologist, for semen processing at CHVNG; Ana Gonçalves, BSc, and Cláudia Osório, BSc, for semen processing at CGR; Elsa Oliveira, 1st Class Technical Specialist of Pathology, Cytology and Thanatology in the Area of Diagnosis and Therapy and Ângela Alves, Technical assistant teaching and research, for semen processing for electron microscopy at ICBAS-UP.

We also would like to acknowledge Conceição Egas, PhD and Hugo Froufe, MSc (GenoInseq) for performing exome sequencing and assisting the initial variant filtering/analysis.

Funding This work was financed by the Institutions of the authors and in part by UMIB, which is funded by National Funds through FCT-Foundation for Science and Technology, under the Pest-OE/SAU/UI0215/2014.

Disclosure statement The authors have nothing to declare.

References

- Curry MR, Watson PF. Sperm structure and function. In: Grudzinskas JG, Yovich JL, editors. *Gametes - the spermatozoon*. NY: Cambridge University Press; 1995. p. 45–69.
- Eddy EM, Toshimori K, O'Brien DA. Fibrous sheath of mammalian spermatozoa. *Microsc Res Tech*. 2003;61:103–15.
- Burgess SA, Knight PJ. Is the dynein motor a winch? *Curr Opin Struct Biol*. 2004;14:138–46.
- Smith EF, Yang P. The radial spokes and central apparatus: mechano-chemical transducers that regulate flagellar motility. *Cell Motil Cytoskeleton*. 2004;57:8–17.
- Wirschell M, Olbrich H, Werner C, Tritschler D, Bower R, Sale W, et al. The nexin-dynein regulatory complex subunit DRC1 is essential for motile cilia function in algae and humans. *Nat Genet*. 2013;45:262–8.
- Ortega C, Verheyen G, Raick D, Camus M, Devroey P, Tournaye H. Absolute asthenozoospermia and ICSI: what are the options? *Hum Reprod Update*. 2011;17(5):684–92.
- Afzelius BA. The immotile-cilia syndrome: a microtubule-associated defect. *Crit Rev Biochem Mol Biol*. 1985;19:63–87.
- Boon M, Jorissen M, Proesmans M, De Boeck K. Primary ciliary dyskinesia, an orphan disease. *Eur J Pediatr*. 2013;172:151–62.
- Leigh MW, Pittman JE, Carson JL, Ferkol TW, Dell S, Davis SD, et al. Clinical and genetic aspects of primary ciliary dyskinesia/Kartagener syndrome. *Genet Med*. 2009;11:473–87.
- Pereira R, Oliveira J, Sousa M. A molecular approach to sperm immotility in humans: a review. *Med Reprod y Embriol Clín*. 2014;01(01):15–25.
- Djakow J, Svobodová T, Hrach K, Uhlík J, Cinek O, Pohunek P. Effectiveness of sequencing selected exons of DNAH5 and DNA11 in diagnosis of primary ciliary dyskinesia. *Pediatr Pulmonol*. 2012;47:864–75.
- Merveille A-C, Davis EE, Becker-Heck A, Legendre M, Amirav I, Bataille G, et al. CCDC39 is required for assembly of inner dynein arms and the dynein regulatory complex and for normal ciliary motility in humans and dogs. *Nat Genet*. 2011;43:72–8.
- Becker-Heck A, Zohn IE, Okabe N, Pollock A, Lenhart KB, Sullivan-Brown J, et al. The coiled-coil domain containing protein CCDC40 is essential for motile cilia function and left-right axis formation. *Nat Genet*. 2011;43:79–84.
- Antony D, Becker-Heck A, Zariwala MA, Schmidts M, Onoufriadis A, Forouhan M, et al. Mutations in CCDC39 and CCDC40 are the major cause of primary ciliary dyskinesia with axonemal disorganization and absent inner dynein arms. *Hum Mutat*. 2013;34:462–72.
- Kott E, Legendre M, Copin B, Papon J-F, Dastot-Le Moal F, Montantin G, et al. Loss-of-function mutations in *rspH1* cause primary ciliary dyskinesia with central-complex and radial-spoke defects. *Am J Hum Genet*. 2013;93:561–70.
- Chemes HE, Olmedo SB, Carrere C, Osés R, Carizza C, Leisner M, et al. Ultrastructural pathology of the sperm flagellum: association between flagellar pathology and fertility prognosis in severely asthenozoospermic men. *Hum Reprod*. 1998;13:2521–6.
- Chemes HE, Rawe VY. The making of abnormal spermatozoa: cellular and molecular mechanisms underlying pathological spermiogenesis. *Cell Tissue Res*. 2010;341:349–57.
- Luconi M, Cantini G, Baldi E, Forti G. Role of a-kinase anchoring proteins (AKAPs) in reproduction. *Front Biosci*. 2011;16:1315–30.
- Turner RMO, Johnson LR, Haig-Ladewig L, Gerton GL, Moss SB. An X-linked gene encodes a major human sperm fibrous sheath protein, hAKAP82: genomic organization, protein kinase a-rii binding, and distribution of the precursor in the sperm tail. *J Biol Chem*. 1998;273:32135–41.
- Mandal A, Naaby-Hansen S, Wolkowicz MJ, Klotz K, Shetty J, Retief JD, et al. FSP95, a testis-specific 95-kilodalton fibrous sheath antigen that undergoes tyrosine phosphorylation in capacitated human spermatozoa. *Biol Reprod*. 1999;61:1184–97.
- Baccetti B, Collodel G, Gambera L, Moretti E, Serafini F, Piomboni P. Fluorescence in situ hybridization and molecular studies in infertile men with dysplasia of the fibrous sheath. *Fertil Steril*. 2005;84:123–9.
- Baccetti B, Collodel G, Estenoz M, Manca D, Moretti E, Piomboni P. Gene deletions in an infertile man with sperm fibrous sheath dysplasia. *Hum Reprod*. 2005;20:2790–4.
- Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res*. 1988;16:1215.

24. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 2012;13:134.
25. Gaspar P, Lopes P, Oliveira J, Santos R, Dalgleish R, Oliveira JL. Variobox: automatic detection and annotation of human genetic variants. *Hum Mutat*. 2014;35(2):202–7.
26. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7:248–9.
27. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003;31:3812–4.
28. Schwarz JM, Rödelberger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods*. 2010;7:575–6.
29. Desmet F-O, Hamroun D, Lalande M, Collod-Bérout G, Claustres M, Bérout C. Human splicing finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res*. 2009;37:e67.
30. Oliveira J, Negrao L, Fineza I, Taipa R, Melo-Pires M, Fortuna AM, et al. New splicing mutation in the choline kinase beta (CHKB) gene causing a muscular dystrophy detected by whole-exome sequencing. *J Hum Genet*. 2015. doi:10.1038/jhg.2015.20.
31. Paila U, Chapman BA, Kirchner R, Quinlan AR. GEMINI: integrative exploration of genetic variation and genome annotations. *PLoS Comput Biol*. 2013;9, e1003153.
32. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. *Nat Genet*. 2000;25:25–9.
33. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res*. 1998;8:175–85.
34. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res*. 1998;8:186–94.
35. Doggett NA, Xie G, Meincke LJ, Sutherland RD, Mundt MO, Berbari NS, et al. A 360-kb interchromosomal duplication of the human HYDIN locus. *Genomics*. 2006;88:762–71.
36. Berg JS, Evans JP, Leigh MW, Omran H, Bizon C, Mane K, et al. Next generation massively parallel sequencing of targeted exomes to identify genetic mutations in primary ciliary dyskinesia: implications for application to clinical testing. *Genet Med*. 2011;13:218–29.
37. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, et al. STRING v9. 1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res*. 2013;41:D808–15.
38. Hornef N, Olbrich H, Horvath J, Zariwala MA, Fliegau M, Loges NT, et al. DNAH5 mutations are a common cause of primary ciliary dyskinesia with outer dynein arm defects. *Am J Respir Crit Care Med*. 2006;174:120–6.
39. Schwabe GC, Hoffmann K, Loges NT, Birker D, Rossier C, De Santi MM, et al. Primary ciliary dyskinesia associated with normal axoneme ultrastructure is caused by DNAH11 mutations. *Hum Mutat*. 2008;29:289–98.
40. Knowles MR, Leigh MW, Carson JL, Davis SD, Dell SD, Ferkol TW, et al. Mutations of DNAH11 in patients with primary ciliary dyskinesia with normal ciliary ultrastructure. *Thorax*. 2012;67:433–41.
41. Miki K, Willis WD, Brown PR, Goulding EH, Fulcher KD, Eddy EM. Targeted disruption of the Akap4 gene causes defects in sperm flagellum and motility. *Dev Biol*. 2002;248:331–42.
42. Turner RMO, Musse MP, Mandal A, Klotz KEN, Friederike C, Jayes L, et al. Molecular genetic analysis of two human sperm fibrous. *J Androl*. 2001;22:302–15.
43. Moretti E, Scapigliati G, Pascarelli NA, Baccetti B, Collodel G. Localization of AKAP4 and tubulin proteins in sperm with reduced motility. *Asian J Androl*. 2007;9:641–9.
44. Colantonio JR, Vermont J, Wu D, Langenbacher AD, Fraser S, Chen J-N, et al. The dynein regulatory complex is required for ciliary motility and otolith biogenesis in the inner ear. *Nature*. 2008;457:205–9.
45. Yeh S-D, Chen Y-J, Chang ACY, Ray R, She B-R, Lee W-S, et al. Isolation and properties of Gas8, a growth arrest-specific gene regulated during male gametogenesis to produce a protein associated with the sperm motility apparatus. *J Biol Chem*. 2002;277:6311–7.
46. Bekker JM, Colantonio JR, Stephens AD, Clarke WT, King SJ, Hill KL, et al. Direct interaction of Gas11 with microtubules: implications for the dynein regulatory complex. *Cell Motil Cytoskeleton*. 2007;64:461–73.
47. Asai DJ, Koonce MP. The dynein heavy chain: structure, mechanics and evolution. *Trends Cell Biol*. 2001;11:196–202.
48. Silvanovich A, Li M, 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;14:1355–65.
49. Zukas R, Chang AJ, Rice M, Springer AL. Structural analysis of flagellar axonemes from inner arm dynein knockdown strains of *Trypanosoma brucei*. *Biochem J*. 2012;36:133–42.
50. Panizzi JR, Becker-heck A, Castleman VH, Al-mutairi D, Liu Y, Loges NT, et al. CCDC103 mutations cause primary ciliary dyskinesia by disrupting assembly of ciliary dynein arms. *Nat Genet*. 2012;44:714–9.
51. Anand-Ivell R, Dai Y, Ivell R. Neohormones as biomarkers of reproductive health. *Fertil Steril*. 2013;99:1153–60.
52. Lok S, Johnston DS, Conklin D, Lofton-Day CE, Adams RL, Jelmsberg AC, et al. Identification of *INSL6*, a new member of the insulin family that is expressed in the testis of the human and rat. *Biol Reprod*. 2000;62:1593–9.
53. Burnicka-Turek O, Shirmeshan K, Paprotta I, Grzmil P, Meinhardt A, Engel W, et al. Inactivation of insulin-like factor 6 disrupts the progression of spermatogenesis at late meiotic prophase. *Endocrinology*. 2009;150:4348–57.
54. Chen G-W, Luo X, Liu Y-L, Jiang Q, Qian X-M, Guo Z-Y. R171H missense mutation of *INSL6* in a patient with spermatogenic failure. *Eur J Med Genet*. 2011;54:e455–7.
55. Knowles MR, Leigh MW, Ostrowski LE, Huang L, Carson JL, Hazucha MJ, et al. Exome sequencing identifies mutations in *CCDC114as* a cause of primary ciliary dyskinesia. *Am J Hum Genet*. 2013;92:99–106.
56. Onoufriadis A, Shoemark A, Munye MM, James CT, Schmidts M, Patel M, et al. Combined exome and whole-genome sequencing identifies mutations in *ARMC4* as a cause of primary ciliary dyskinesia with defects in the outer dynein arm. *J Med Genet BMJ Publishing Group Ltd*; 2014;51:61–7.
57. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet*. 2011;12:745–55.
58. Teves ME, Zhang Z, Costanzo RM, Henderson SC, Corwin FD, Zweit J, et al. Sperm-associated antigen-17 gene is essential for motile cilia function and neonatal survival. *Am J Respir Cell Mol Biol*. 2013;48:765–72.