Hum. Reprod. Advance Access published October 26, 2016

Human Reproduction, pp. 1–9, 2016

doi:10.1093/humrep/dew262

human reproduction

ORIGINAL ARTICLE ***Reproductive genetics***

Whole-exome sequencing of familial cases of multiple morphological abnormalities of the sperm flagella (MMAF) reveals new *DNAH1* mutations

Amir Amiri-Yekta[1,2,†], Charles Coutton[2,3,†], Zine-Eddine Kherraf[2,4], Thomas Karaouzène[2,4], Pauline Le Tanno[2,3], Mohammad Hossein Sanati[1,5], Marjan Sabbaghian[6], Navid Almadani[1], Mohammad Ali Sadighi Gilani[6], Seyedeh Hanieh Hosseini[7], Salahadin Bahrami[1], Abbas Daneshipour[1], Maurizio Bini[7], Christophe Arnoult[2], Roberto Colombo[8,9], Hamid Gourabi[1,†#], and Pierre F. Ray[2,4,*#]

[1]Department of Genetics, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR, PO Box 16635-148, Tehran, Iran [2]Genetic Epigenetic and Therapies of Infertility, Institute for Advanced Biosciences, INSERM 1209, CNRS UMR 5309, Université Grenoble Alpes, Grenoble F38000, France [3]CHU de Grenoble, UF de Génétique Chromosomique, Grenoble F-38000, France [4]CHU de Grenoble, UF de Biochimie Génétique et Moléculaire, Grenoble F-38000, France [5]Department of Medical Genetics, Institute of Medical Biotechnology, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran [6]Department of Andrology, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR, PO Box 16635-148, Tehran, Iran [7]Center for the Study and Treatment of Fertility Disorders, Niguarda Ca' Granda Metropolitan Hospital, Milan, Italy [8]Faculty of Medicine, Institute of Clinical Biochemistry, Catholic University, Rome, Italy [9]Center for the Study of Rare Hereditary Diseases, Niguarda Ca' Granda Metropolitan Hospital, Milan, Italy

*Correspondence address. UF de Biochimie et Génétique Moléculaire, CHU de Grenoble, 38043 Grenoble cedex 9, France. Tel: +33-4-76-76-55-73; E-mail: pray@chu-grenoble.fr

Submitted on July 15, 2016; resubmitted on September 9, 2016; accepted on September 15, 2016

STUDY QUESTION: Can whole-exome sequencing (WES) of patients with multiple morphological abnormalities of the sperm flagella (MMAF) identify causal mutations in new genes or mutations in the previously identified dynein axonemal heavy chain 1 (*DNAH1*) gene?

SUMMARY ANSWER: WES for six families with men affected by MMAF syndrome allowed the identification of *DNAH1* mutations in four affected men distributed in two out of the six families but no new candidate genes were identified.

WHAT IS KNOWN ALREADY: Mutations in *DNAH1*, an axonemal inner dynein arm heavy chain gene, have been shown to be responsible for male infertility due to a characteristic form of asthenozoospermia called MMAF, defined by the presence in the ejaculate of spermatozoa with a mosaic of flagellar abnormalities including absent, coiled, bent, angulated, irregular and short flagella.

STUDY DESIGN, SIZE, DURATION: This was a retrospective genetics study of patients presenting a MMAF phenotype. Patients were recruited in Iran and Italy between 2008 and 2015.

PARTICIPANTS/MATERIALS, SETTING, METHODS: WES was performed for a total of 10 subjects. All identified variants were confirmed by Sanger sequencing. Two additional affected family members were analyzed by direct Sanger sequencing. To establish the

[†]These authors are equally contributing first authors.
[#]These authors have co-leadership.

prevalence of the *DNAH1* mutation identified in an Iranian family, we carried out targeted sequencing on 38 additional MMAF patients of the same geographical origin. RT-PCR and immunochemistry were performed on sperm samples to assess the effect of the identified mutation on RNA and protein.

MAIN RESULTS AND THE ROLE OF CHANCE: WES in six families identified a causal mutations in two families. Two additional affected family members were confirmed to hold the same homozygous mutation as their sibling. In total, *DNAH1* mutations were identified in 5 out of 12 analyzed subjects (41.7%). If we only include index cases, we detected two mutated subjects out of six (33%) tested MMAF individuals. Furthermore we sequenced one *DNAH1* exon found to be mutated (c.8626-1G > A) in an Iranian family in an additional 38 MMAF patients from Iran. One of these patients carried the variant confirming that this variant is relatively frequent in the Iranian population. The effect of the c.8626-1G > A variant was confirmed by RT-PCR and immunochemistry as no RNA or protein could be observed in sperm from the affected men.

LARGE SCALE DATA: N/A.

LIMITATIONS, REASONS FOR CAUTION: WES allows the amplification of 80–90% of all coding exons. It is possible that some *DNAH1* exons may not have been sequenced and that we may have missed some additional mutations. Also, WES cannot identify deep intronic mutations and it is not efficient for detection of large genomic events (deletions, insertions, inversions). We did not identify any causal mutations in *DNAH1* or in other candidate genes in four out of the six tested families. This indicates that the technique and/or the analysis of our data can be improved to increase the diagnosis efficiency.

WIDER IMPLICATIONS OF THE FINDINGS: Our findings confirm that *DNAH1* is one of the main genes involved in MMAF syndrome. It is a large gene with 78 exons making it challenging and expensive to sequence using the traditional Sanger sequencing methods. We show that WES sequencing is good alternative to Sanger sequencing to reach a genetic diagnosis in patients with severe male infertility phenotypes.

STUDY FUNDING/COMPETING INTEREST(S): This work was supported by following grants: the 'MAS-Flagella' project financed by the French ANR and the DGOS for the program PRTS 2014 and the 'Whole genome sequencing of patients with Flagellar Growth Defects (FGD)' project financed by the Fondation Maladies Rares for the program Séquençage à haut débit 2012. The authors have no conflict of interest.

Key words: male infertility / genetic diagnosis / gene mutations / exome sequencing / teratozoospermia / flagellum / MMAF / DNAH1

Introduction

Men with primary ciliary dyskinesia (PCD), a pathology grouping different phenotypic entities due to molecular defects altering cilia function and affecting mainly the pulmonary function, are often infertile with asthenozoospermia and abnormal flagellar morphology (Storm van's and Omran, 2005). Some patients however present with similar sperm abnormalities with no other associated syndromes, suggesting analogous molecular mechanisms affecting only the sperm flagellum. As this phenotype is restricted to fertility, it could be considered as a new clinical entity and such sperm defects have indeed been described on many occasions since 1984 (Escalier and David, 1984). Patients consistently have had astheno-teratozoospermia characterized by a mosaic of flagellar abnormalities including absent, coiled, bent, angulated, irregular or short flagella mainly due to numerous ultrastructural defects of the axoneme. This phenotype has been reported as dysplasia of the fibrous sheath, short/stump tails or non-specific flagellar anomalies (Chemes *et al.*, 1987; Stalf *et al.*, 1995; Chemes and Rawe, 2003). Until recently, genetic causes of flagellar abnormalities have remained largely unexplained. In 2005, deletions in *AKAP3* and *AKAP4*, two genes encoding proteins of the fibrous sheath, were reported in one patient (Baccetti *et al.*, 2005); these findings however remain to be confirmed. In 2014, our team carried out a large genetics study and we proposed to use the term multiple morphological abnormalities of the sperm flagella (MMAF) which seemed to clearly define the phenotype of the included patients (Ben Khelifa *et al.*, 2014). In this study, homozygosity mapping using a single nucleotide polymorphism (SNP) array allowed the identification of pathogenic mutations in the dynein axonemal heavy chain 1 (*DNAH1*) gene in 28% of the analyzed patients. The identification of *DNAH1* mutations indicated that *DNAH1* is a major gene involved in the MMAF phenotype and is expected to account for up to one-third of cases (Ben Khelifa *et al.*, 2014). This indicates that MMAF is genetically heterogeneous and that other genes are likely involved in this syndrome (Coutton *et al.*, 2015).

It is currently estimated that 1500–2000 genes are involved in the control of spermatogenesis, including 300–600 specifically expressed in male germ cells (Matzuk and Lamb, 2008). The abundance of potential candidate genes makes the identification of pathogenic mutations difficult and complex. Gene identification is however the key to improving our understanding of the pathophysiology of infertility and could open new perspectives for the diagnosis and treatment of infertile patients. In recent years, different promising genomic approaches have catalyzed the identification of new genes involved in male infertility (El *et al.*, 2012). Whole-exome sequencing (WES), the sequencing of the coding sequence located in the exons of the translated genes, now appears as the best strategy to detect disease-causing variations in individuals affected with Mendelian disorders (Bamshad *et al.*, 2011). Analysis of WES data however remains challenging as 20 000–30 000 variants differing from the genomic reference sequence are usually found in any given individual (Gilissen *et al.*, 2012). It is then extremely arduous to identify the causal variant(s) from this large number of variants of usually unknown significance. The number of potentially pathogenic variants can be reduced by analyzing cohorts of affected individuals and looking for variants or defects in the same gene present in several affected

individuals. Alternatively, consanguineous kindreds can be analyzed and this strategy can be particularly successful for genetically heterogeneous pathology such as infertility (Boycott *et al.*, 2013). Indeed, WES of small families has been successfully used for the detection of causal genes in phenotypes such as non-obstructive azoospermia (Ayhan *et al.*, 2014) or sperm fertilization defects (Escoffier *et al.*, 2015). In this study, we wanted to evaluate the efficiency of family-based WES to identify new *DNAH1* mutations or new genetic causes of MMAF syndrome.

Materials and Methods

Patient and control individuals

Five Iranian families (1–5) and one Italian family (6) were included in this study (Fig. 1). Among these 6 families, we included 12 subjects (P1–P12) presenting with asthenozoospermia due to a combination of 5 morphological defects of the sperm flagella including: absent, short, bent, coiled flagella and of irregular width (Table I) without any of the additional symptoms associated with PCD. About 10 individuals originated from Middle East (Iranians) and were treated in Tehran at the Royan Institute, (Reproductive Biomedicine Research Center) for primary infertility from 2008 to 2015. Two brothers of European origin (Southern Italia) consulted for primary infertility at the Center for the Study of Rare Inherited Diseases in Milan, Italia. About 10 of the 12 subjects were born from related parents, usually first cousins. WES was performed for a total of 10 patients and *DNAH1* targeted Sanger sequencing was performed for two additional brothers (P7 and P12) who had not been included in time for WES (Fig. 1). Except for Families 5 and 6, at least two infertile brothers from each family were included and were analyzed by WES (Fig. 1). All subjects had normal somatic karyotypes.

Sperm analysis was carried out in the source laboratories during the course of the routine biological examination of the patient, according to World Health Organization (WHO) guidelines (World Health Organization, 2010). The morphology of patients' sperm was assessed with Papanicolaou staining (Fig. 2). Small variations in protocol might occur between the different laboratories. Subjects were recruited on the basis of the identification of >5% of at least four of the aforementioned flagellar morphological abnormalities (absent, short, coiled, bent and irregular flagella) (Table I). Unfortunately, sperm parameters from patient P2 were not available. All subjects presented with severe asthenozoospermia: seven patients had no motility (≤5%), two had sperm motility < 15% and two (P11 and P12) had approximately 30% motility. These latter two patients were considered to have a milder form of MMAF syndrome.

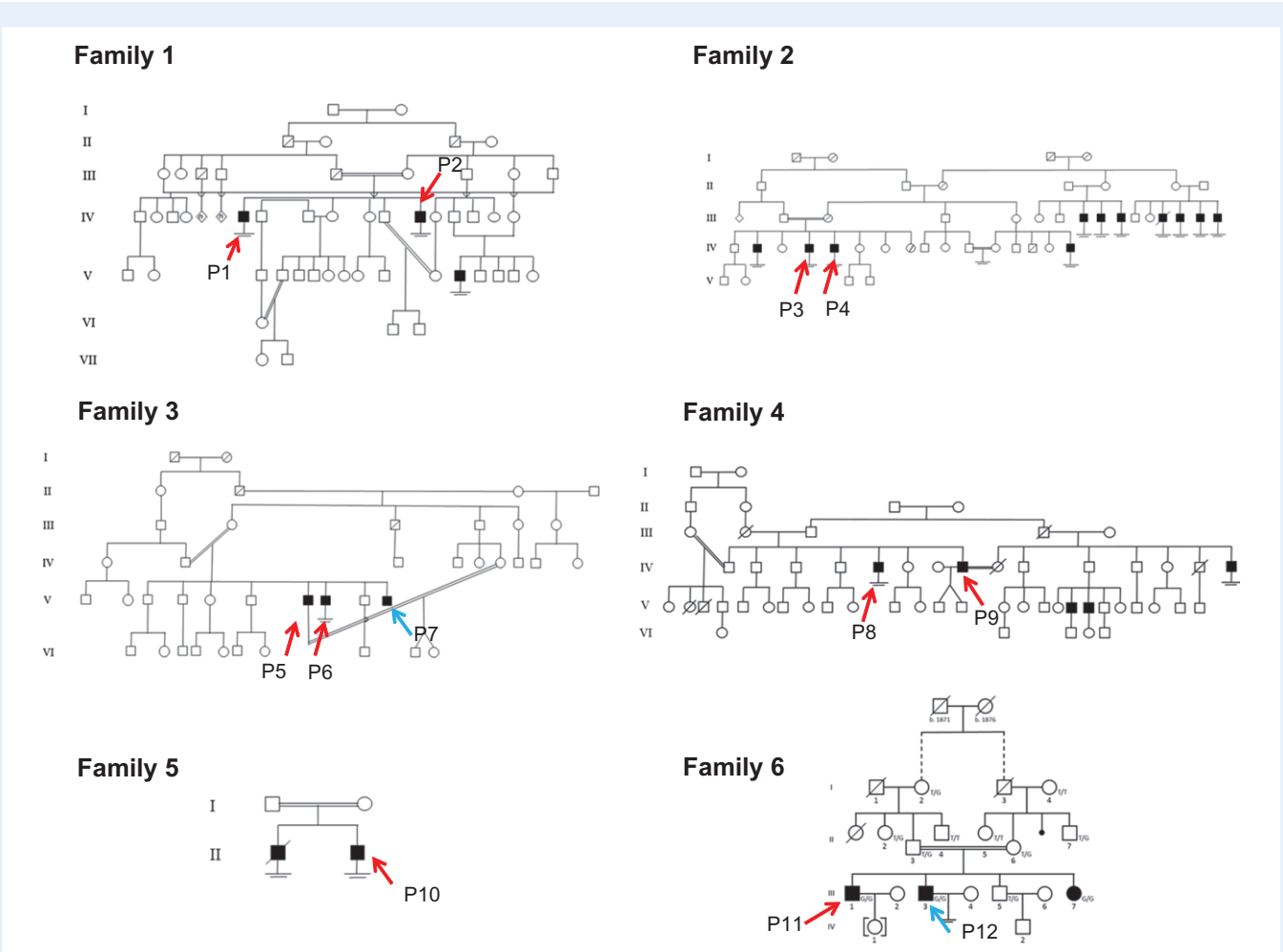

Figure 1 Pedigrees of five Iranian (1–5) and one Italian family (6). Black squares indicate infertile subjects in the family. Red arrows indicate patients for whom WES was performed. Blue arrow indicates patients (P7 and P12) who did not have WES but subsequent exon targeted *DNAH1* Sanger sequencing.

Table I Semen parameters of the 12 patients among whom five subjects (P5, P6, P7, P11 and P12) carried *DNAH1* homozygous variants. Semen parameters from the additional sporadic patient SP4 with a *DNAH1* mutation are also reported. *DNAH1* mutations were identified by WES in patients P5, P6 and P11 and by Sanger sequencing in patients P7, P12 and SP4. Values are the average of two separate analyses.

Semen parameters	Family 1		Family 2		Family 3			Family 4			Family 5	Family 6		SP4
	P1	P2	P3	P4	P5	P6	P7	P8	P9		P10	P11	P12	SP4
DNAH1 mutation	None	None	None	None	c.8626-1G > A	c.8626-1G > A	c.8626-1G > A	None	None		None	c.3860.T > G	c.3860.T > G	c.8626-1G > A
Sperm volume (ml)	4	NA	3.2	3.5	4	4	4	1	0.9		1	3.5	3	4.5
Sperm concentration (10^6/ml)	20	NA	6	16	29	18	37	52	50		21	24	31	13
Motility (A + B) 1 h (%)	0	NA	15	14.3	5	0	0	0	0		0	27	33	0
Vitality (%)	88	NA	NA	NA	97	50	82	92	0		89	74	85	88
Normal spermatozoa (%)	0	NA	0	1	0	0	0	0	0		3	8	11	0
Anomalies of the head (%)	1	NA	6	34	12	33	22	2	5		14	24	18	13
Absent flagella (%)	0	NA	5	1	0	0	0	0	2		15	5	8	0
Short Flagella (%)	90	NA	53	6	85	65	80	70	70		35	41	48	86
Coiled Flagella (%)	+	NA	+	6	8	1	0	5	1		5	6	12	1

NA = not available. + : anomalies reported (>5%) but not accurately quantified.

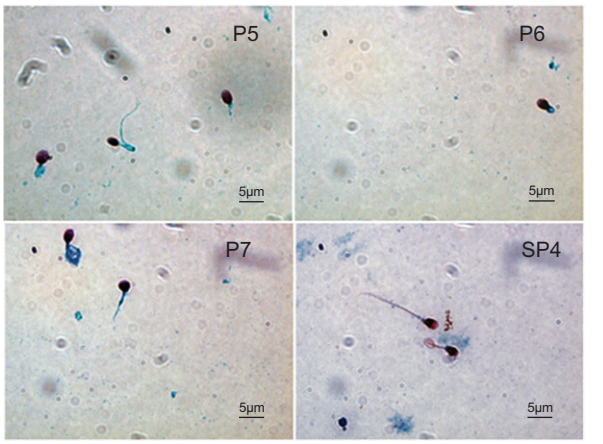

Figure 2 Light microscopy analysis of spermatozoa from three c.8626-1G > A homozygous Iranian brothers (P5, P6, P7) and one sporadic Iranian patient (SP4). Sperm samples were spread over a slide and dried at room temperature for Papanicolaou staining. These images are given as examples of typical spermatozoa observed in mutated patients.

In order to evaluate the incidence of the detected mutation in the general Iranian population, we recruited 38 additional patients (SP1 to SP38) presenting with MMAF phenotype with at less 80% of sperm with flagellar abnormalities (Supplementary Table S1).

Saliva and/or peripheral blood was obtained for all participants. Sperm samples were obtained, following informed consent, from patients P5, P6, P7 and SP4. All subjects answered a health questionnaire focused on PCD manifestations during their medical consultation for infertility. Informed consent was obtained from all the subjects participating in the study according to the local Institutional Review Board protocols and the principles of the Declaration of Helsinki. In addition, the study was approved by local ethic committees.

Molecular analysis

DNA extraction

DNA was extracted from blood or saliva. Blood DNA extraction was carried out from 5 to 10 ml of frozen EDTA blood using the quick guanidium chloride extraction procedure. Saliva was collected with Oragene DNA Self-Collection Kits (DNA genotek Inc., Canada) and DNA extraction was performed following the manufacturer's recommendations.

Exome sequencing and bioinformatics analysis

WES was carried out on DNA extracted from the 10 studied subjects. Coding regions and intron/exon boundaries were enriched using the 'all Exon V5 kit' (Agilent Technologies, Wokingham, UK). DNA sequencing was undertaken at the Plateforme Biopuces et Séquençage IGBMC, Illkirch, France, on the HiSeq 2000 from Illumina®. All steps from sequence mapping to variant selection were performed using the ExSQLibur pipeline (https://github.com/tkaraouzene/ExSQLibur). Sequence reads were aligned to the reference genome (hg19) using MAGIC (SEQC/MAQC-III Consortium, 2014). Duplicate reads and reads that mapped to multiple locations in the exome were excluded from further analysis. Positions whose sequence coverage was below 10 on either the forward or reverse strand were excluded. Single nucleotide variations (SNV) and small insertions/deletions (indels) were identified and quality-filtered using in-house

scripts. The most promising candidate variants were identified using an in-house bioinformatics pipeline. Variants with a minor allele frequency greater than 5% in the NHLBI ESP6500 or in 1000 Genomes Project Phase 1 data sets, or greater than 1% in ExAC were discarded. We also compared these variants to an in-house database of 56 control exomes. All variants present in a homozygous state in this database were excluded. We used Variant Effect Predictor to predict the impact of the selected variants. We only retained variants impacting splice donor/acceptor sites or causing frameshift, inframe insertions/deletions, stop gain, stop loss or missense mutations, except those scored as 'tolerated' by SIFT (sift.jcvi.org) and as 'benign' by Polyphen-2 (genetics.bwh.harvard.edu/pph2).

Sanger sequencing

The two variations identified in *DNAH1* using WES were verified by Sanger sequencing. The coding exon 23 and *intron-exon boundary* adjacent to intron54/*exon55 were amplified as indicated in the* Supplementary Table SII. Sequencing reactions were carried out with BigDye Terminator v3.1 (Applied Biosystems). Sequence analysis were carried out on ABI 3130XL (Applied Biosystems).

RNA extraction

Nucleated cells were isolated from whole blood using Ficoll® 400 (Sigma-Aldrich Corp., St. Louis, MO, USA) following the manufacturer's protocol. RNA extraction was carried out on the isolated white blood cells using Macherey Nagel NucleoSpin® RNA II columns (Macherey Nagel, Hoerdt, France) using the manufacturer's protocol.

RT-PCR

Reverse transcription was carried out in three patients P5, P6, P7 and two healthy controls (C1 and C2) with 5 µl of extracted RNA (approximately 500 ng). Hybridization of the oligo-dT was performed by incubating for 5 min at 65°C and quenching on ice with the following mix: 5 µl of RNA, 3 µl of poly T oligo primers (dT)12–18 (10 mM, Pharmacia), 3 µl of the four dNTPs (0.5 mM, Roche diagnostics) and 2.2 µl of H_2O. Reverse Transcription was then carried out for 30 min at 55°C after the addition of 4 µl of 5× buffer, 0.5 µl RNase inhibitor and 0.5 µl of Transcriptor Reverse transcriptase (Roche Diagnostics). Then 2 µl of the obtained cDNA mix was used for the subsequent PCR. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as a housekeeping gene (internal control). Primers sequences and RT-PCR conditions are indicated in the Supplementary Table SIII.

Immunostaining

Sperm were fixed in PBS/4% paraformaldehyde for 1 min at room temperature. After washing in 1 ml PBS, the sperm suspension was spotted onto 0.1% poly L-lysine pre-coated slides (Thermo Scientific). After attachment, sperm were permeabilized with 0.1% (v/v) Triton X-100—DPBS (Triton X-100; Sigma-Aldrich) for 5 min at room temperature. Slides were then blocked in 5% corresponding normal serum—DPBS (normal goat or donkey serum; GIBCO, Invitrogen) and incubated overnight at 4°C with primary antibodies. Polyclonal DNAH1 antibodies were purchased from Prestige Antibodies® (Sigma-Aldrich, France) (1:50). Monoclonal mouse anti-acetylated-α-tubulin antibodies were purchased from Sigma-Aldrich (1:2500). Washes were performed with PBS containing 0.1% of Tween 20, followed by 1 h incubation at room temperature with Alexa Fluor 555-labeled goat anti-rabbit or Dylight 488-labeled goat anti-rabbit (1:400) secondary antibodies. Appropriate controls were performed, omitting the primary antibodies. Samples were counterstained with 5 mg/ml Hoechst 33342 (Sigma-Aldrich) and mounted with DAKO mounting media (Life Technology). Whole images were reconstructed and projected from Z-stack images using ZEN software.

Results

WES identified two new *DNAH1* mutations

Given the notion of consanguinity in the families we studied, we postulated that the infertility has been likely transmitted by autosomal recessive inheritance and was thus caused by a homozygous mutation. We proceeded to WES to identify a possible genetic defect that could explain the observed MMAF phenotype. After exclusion of frequent variant and applying stringent filters, a limited list of homozygous variants was identified in each proband. Apart from these three variants in *DNAH1*, no variants were present in genes described to be strongly expressed in the testis nor in any gene described to be connected with cilia, the axoneme or the flagellum. Only the *DNAH1* variants were retained as likely causal (Supplementary Table SIV). In cases where the brothers were simultaneously analyzed by WES, only common variants shared by both brothers were retained. Two different pathogenic mutations in the *DNAH1* gene were identified in Families 3 and 6.

In Family 3, the c.8626-1G > A variant was identified in the two MMAF brothers analyzed by WES. This mutation affects the final G nucleotide of *DNAH1* intron 54, one of the consensus splice acceptor bases (Fig. 3). The alternate splicing is predicted to cause a frameshift in the new transcript and to induce a premature stop codon. The c.8626-1G > A variant was absent from over 60 000 individuals described in the ExAC database which confirms that it is not a polymorphism and that a splicing mutation occurring at this localization would be negatively selected throughout evolution. No other variants were identified in the *DNAH1* coding sequence and UTR regions. Sanger sequencing confirmed the splicing mutation for both infertile brothers (Fig. 3) and showed that the third infertile brother has the same mutation while their father and a non-affected brother were heterozygous (data not shown).

In Family 6, a second mutation was found in *DNAH1*. The identified mutation is c.3860 T > G (p.Val1287Gly). This was confirmed in the same subject by bidirectional Sanger sequencing of *DNAH1* exon 23. No other unreported *DNAH1* variant was identified by exome sequencing. No rare variants were present in other genes reported to be associated with male infertility. The presence of the c.3860 T > G mutation was tested in the other family members available for genotyping by restriction analysis and confirmed by Sanger sequencing. In this large Italian family, the other infertile brother and one sister were homozygous for the *DNAH1* mutation (Fig. 3). Six other related fertile subjects of both sexes were heterozygous for the mutation. The sister, also homozygous for p.Val1287Gly mutation, was healthy by general evaluation and had not yet attempted to have children.

Targeted PCR-Sanger sequencing in 38 MMAF Iranian patients

We next assessed 38 Iranian patients (SP1 to SP38) presenting with morphological abnormality of flagella (Supplementary Table S1) for the c.8626-1G > A splicing mutation identified in the Iranian Family 3. Sequencing results showed that one of these affected men (SP4) is homozygous for this new splice site mutation indicating that this mutation segregates in the Iranian population.

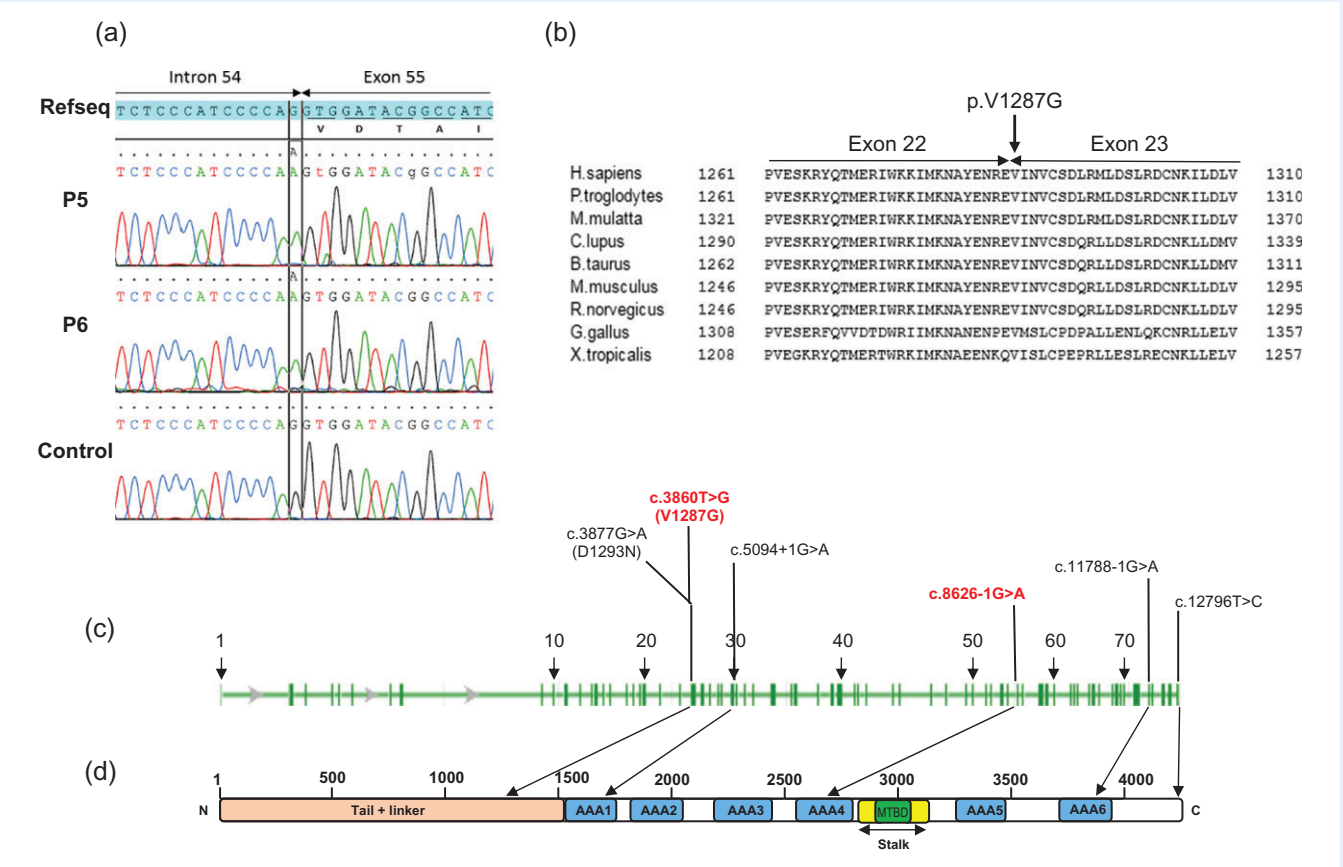

Figure 3 *DNAH1* novel point mutations. (**a**) Electropherogramms of *DNAH1* showing the homozygous splice acceptor site mutation c.8626-1G > A for P5, P6 and a heathy control individual. (**b**) Interspecies alignment. The two Italian brothers carry a homozygous missense mutation c.3860.T > G in *DNAH1* exon 23, leading to the change of the valine in position 1287 into a Glycine. The valine in position 1287 is well conserved between species. Amino acid sequence alignment was realized with homologene (http://www.ncbi.nlm.nih.gov.gate1.inist.fr/homologene/?term=dnah1). (**c**) Localization of the *DNAH1* mutations in the intron-exon structure showing the position of the two novel identified mutations (in red) and the mutations identified in our previous study (in black) (Ben Khelifa *et al.*, 2014). Green boxes indicate coding exons, white boxes indicate noncoding exons. (**d**) Protein domain map of the axonemal dynein heavy chain 1 (DNAH1). The blue boxes indicate the six known AAA-ATPase domains (AAA 1 to 6) as detected by homology (Uniprot server). The microtubule-binding domain (MTBD) lies between AAA4 and AAA5. The N-terminal part of the protein binds to the intermediate, light-intermediate dynein chains. The position of the stalk and the MTBD are indicated.

Detrimental effects of the two identified mutations

To assess the functional impact of the *DNAH1* splice acceptor site mutation c.8626-1G > A, we studied mRNA products isolated from two control and from P5, P6 and P7. RT-PCR of patients' samples yielded no product despite repeated attempts, contrary to what was observed from the controls yielding bands of the expected size (Fig. 4). RT-PCR targeting *GAPDH* confirmed the integrity of patients' RNA (Fig. 4). This suggests a specific degradation of the mutant *DNAH1* transcripts by nonsense mediated mRNA decay (NMD). To further validate the pathogenicity of this variant, we analyzed DNAH1 localization by immunofluorescence on patients' sperm. In control individuals, DNAH1 antisera decorated the full length of the sperm flagellum (Fig. 5). In contrast, in sperm from the three brothers carrying the c.8626-1G > A mutation as well as in sperm from the sporadic case SP4, DNAH1 immunostaining was absent, indicating that the splicing defect induces the degradation of the transcripts by NMD thus precluding protein production (Fig. 5).

Unfortunately, no mRNA analysis or immunostaining could be performed on sperm cells from the Italian patients, P11 and P12. We however found Val1287 to be very well conserved throughout evolution (Fig. 3) and this missense change is also predicted to be likely damaging by SIFT and PolyPhen-2, two prediction software for nonsynonymous SNPs. This variant was also absent from all the control sequence databases (dbSNP v137, 1000 Genomes Project, NHLBI Exome Variant Server).

Discussion

To date, SNP array-based homozygosity mapping has permitted the identification of mutations in three main genes leading to teratozoospermia (Dieterich *et al.*, 2007; Harbuz *et al.*, 2011; Ben Khelifa *et al.*, 2014). This strategy is however time-consuming and requires that several patients

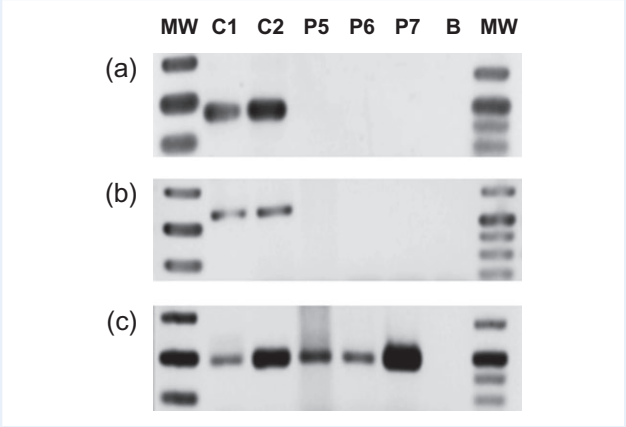

Figure 4 RT-PCR analyses of subjects P5, P6 and P7 (c.8626-1G > A homozygotes) and control individuals from the general population. Electrophoresis showing the RT-PCR amplification of (**a**) *DNAH1* exons 53–56 (**b**) DNAH1 exons 53–57 and (**c**) *GAPDH* amplification used as housekeeping gene. Controls DNA C1 and C2 show good amplification for all tested loci whereas the three patients do not amplify *DNAH1* but only the *GAPDH* control. There is no amplification from the RT-negative blank control (column B).

share a common genetic cause responsible for the phenotype. Due to the high genetic heterogeneity of infertility, SNPs array has now reached its limits and has been supplanted by next generation sequencing and in particular WES. Unfortunately in this study, this strategy has not led to the identification of new reliable candidate genes in MMAF patients. Some deleterious variants have been identified but they do not concern genes that have been described to have a strong connection with spermatogenesis. What appears as a negative result can be explained by: (i) a lack of information regarding a mutated gene explaining why it was not selected; (ii) a false starting assumption based on the hypothesis that the infertility has been likely transmitted by autosomal recessive inheritance; (iii) exclusion criteria being too stringent, eliminating the pathogenic variant (e.g. silent variant modifying splicing sites), (iv) a variant undetectable by the technique used or our current bio-informatic pipeline, e.g., genomic rearrangements (large deletions and duplications), unsequenced deep intronic variants or some exonic variants (as only about 90% of coding nucleotides were covered). These results therefore highlight the fact that WES cannot be expected to provide 100% positive diagnoses. A diagnosis was however reached for 2 out of 6 of the analyzed families (33%). Two new *DNAH1* mutations were identified. These results thus reinforce the fact that *DNAH1* remains the main gene associated with MMAF and corroborate our previous study indicating that approximately 30% of subjects with MMAF are expected to harbor *DNAH1* mutations. These findings raise

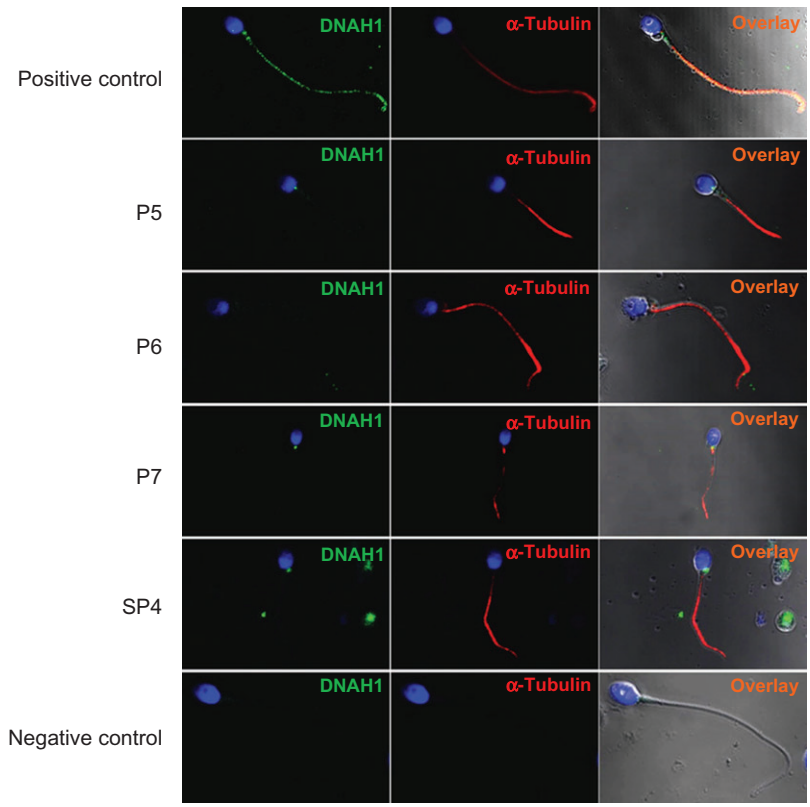

Figure 5 Immunofluorescence images of human spermatozoa from the four patients with the c.8626-1G > A mutation with DNAH1 antibodies and α-Tubulin. DNAH1 staining (green) was observed throughout the flagellum in positive control sperm, whereas it was absent in all sperm samples from patients with the c.8626-1G > A variant. The α-tubulin signal (red) was observed in controls and patients. Sperm were counterstained with Hoechst 33342 (blue) as nuclei marker. These images are given as examples of typical stainings observed in patients.

the question of whether *DNAH1* routine diagnosis should be proposed to all MMAF patients. *DNAH1* is a very large gene (84 Kb and 78 exons) making conventional Sanger sequencing difficult, laborious and costly. In view of our results, we recommend to perform WES in MMAF patients as a first approach. WES in MMAF patients provides a fast way to sequence all *DNAH1* exons while giving the opportunity to identify new gene defects. This strategy, followed by the assessment of ICSI results obtained with patients with different gene defects, could allow improvements in the prediction of ICSI success rates and thus to provide better counseling to MMAF patients. We recently reported that ICSI with spermatozoa from MMAF patients with *DNAH1* mutations had a high pregnancy rate following ICSI (Wambergue *et al.*, 2016). *DNAH1* mutation positive patients identified in this study can thus be encouraged to initiate an IVF/ICSI procedure.

One of the main difficulties associated with WES is the confirmation of the deleterious effect of the identified variant. The effect of the first *DNAH1* variant (c.8626-1G > A) cannot be questioned as: (i) it was identified in four individuals (P5, P6, P7 and SP4); (ii) it affects a consensus splice site known to be essential to the mRNA splicing machinery; and (iii) the mRNA and the protein were shown to be absent from mutated patients. The effect of the second variant, Val1287Gly, is not as easy to predict as it concerns only one amino acid. We however have several arguments in favor of its pathogenicity: (i1) the variant was described as deleterious by two prediction tools; (ii) it was not found in any database now including in excess of 60 000 individuals; and (iii) it affects a conserved residue located in the N-terminal part of the protein known to be important for the structure of the dynein arms (Habura *et al.*, 1999) and it is positioned close to the first missense *DNAH1* mutation (Asp1293Asn) identified previously. The four patients with the severe variant present 0% of morphologically normal spermatozoa with a motility <5% in contrast with the two brothers with the Val1287Gly variant who present a milder phenotype with approximately 30% motility and 10% of morphologically normal spermatozoa (Table I). This suggests that the Val1287Gly mutated protein is likely preserve a residual activity as was previously observed in the patient with the nearby missense mutation. This supports the hypothesis of a phenotype continuum depending on the severity of *DNAH1* mutations (Ben Khelifa *et al.*, 2014). We can therefore expect that individuals harboring homozygous or compound heterozygous *DNAH1* mutations of moderate severity could present with intermediate asthenozoospermia and low levels of morphological anomalies.

When all patients with MMAF analyzed in our previous and present work are pooled, we have a total of 32 patients including 24 index cases. Seven unrelated individuals out of 24 carried a homozygous variant in *DNAH1* (29.2%). Although a large majority of patients are of North African origin, we have now identified *DNAH1* mutations in Middle East and European patients indicating that *DNHA1* diagnosis should not be ruled out for non-Maghrebian individuals. The c.8626-1G > A variant was found in 1/38 sporadic MMAF Iranian cases resulting in a prevalence estimated at 2.6% in this Iranian population. It would be interesting to compare the haplotype of these affected Iranian patients in order to highlight a possible founder effect.

As reported in our previous work, all individuals carrying mutations in *DNAH1* only presented with male infertility and did not report any other symptoms associated with PCD such as an impairment of the respiratory functions (Ben Khelifa *et al.*, 2014). The data reported here is consistent with the hypothesis suggesting that DNAH1 function in cilia may be compensated by other dyneins. The absence of clinical signs does not formally exclude that *DNAH1* mutations could lead to a slight functional impairment of motile cilia function in respiratory epithelium but the impairment would not necessarily lead to lung dysfunction. Unfortunately, family members in which a *DNAH1* mutation was identified did not consent to the investigation of their respiratory function.

The abundance of potential candidate genes makes identification of pathogenic mutations difficult and complex. However, gene identification is the key to improving knowledge of the pathophysiology of MMAF and opens new perspectives for diagnosis and treatment of infertile patients. Further genetic studies are therefore warranted to identify other genes involved in MMAF to better characterize the genetic etiology of the MMAF phenotype and to improve the management of patients diagnosed with flagellar defects.

Supplementary data

Supplementary data are available at http://humrep.oxfordjournals.org/.

Acknowledgments

The authors thank the patients for their interest and cooperation.

Authors' roles

P.F.R., C.C., A.A.Y. and H.G. analyzed the data, wrote the manuscript and had full access to all of the data and take responsibility for the integrity of the data and its accuracy. A.A.Y., Z.E.K., P.L.T. performed the experimental analyses. T.K. performed data mining for the genetic data. M.H.S., M.S., N.A., M.A.S.G., S.H.H., S.B., A.D., C.A., R.C., M.B. and H.G. provided clinical samples and data and supplied biological materials. All authors contributed to the report.

Funding

This work was supported by following grants: the 'MAS-Flagella' project financed by the French ANR and the DGOS for the program PRTS 2014 and the 'Whole genome sequencing of patients with Flagellar Growth Defects (FGD)' project financed by the Fondation Maladies Rares for the program Séquençage à haut débit 2012.

Conflict of interest

None declared.

References

Ayhan O, Balkan M, Guven A, Hazan R, Atar M, Tok A, Tolun A. Truncating mutations in TAF4B and ZMYND15 causing recessive azoospermia. *J Med Genet* 2014;**51**:239–244.

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**:2790–2794.

Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* 2011;**12**:745–755.

Ben Khelifa M, Coutton C, Zouari R, Karaouzene T, Rendu J, Bidart M, Yassine S, Pierre V, Delaroche J, Hennebicq S *et al.* Mutations in DNAH1, which encodes an inner arm heavy chain dynein, lead to male

infertility from multiple morphological abnormalities of the sperm flagella. *Am J Hum Genet* 2014;**94**:95–104.

Boycott KM, Vanstone MR, Bulman DE, MacKenzie AE. Rare-disease genetics in the era of next-generation sequencing: discovery to translation. *Nat Rev Genet* 2013;**14**:681–691.

Chemes EH, Rawe YV. Sperm pathology: a step beyond descriptive morphology. Origin, characterization and fertility potential of abnormal sperm phenotypes in infertile men. *Hum Reprod Update* 2003;**9**:405–428.

Chemes HE, Brugo S, Zanchetti F, Carrere C, Lavieri JC. Dysplasia of the fibrous sheath: an ultrastructural defect of human spermatozoa associated with sperm immotility and primary sterility. *Fertil Steril* 1987;**48**:664–669.

Coutton C, Escoffier J, Martinez G, Arnoult C, Ray PF. Teratozoospermia: spotlight on the main genetic actors in the human. *Hum Reprod Update* 2015;**21**:455–485.

Dieterich K, Soto RR, Faure AK, Hennebicq S, Ben Amar B, Zahi M, Perrin J, Martinez D, Sele B, Jouk PS *et al*. Homozygous mutation of AURKC yields large-headed polyploid spermatozoa and causes male infertility. *Nat Genet* 2007;**39**:661–665.

El IE, Muller J, Viville S. Autosomal mutations and human spermatogenic failure. *Biochim Biophys Acta* 2012;**1822**:1873–1879.

Escalier D, David G. Pathology of the cytoskeleton of the human sperm flagellum: axonemal and peri-axonemal anomalies. *Biol Cell* 1984;**50**:37–52.

Escoffier J, Lee HC, Yassine S, Zouari R, Martinez G, Karaouzene T, Coutton C, Kherraf ZE, Halouani L, Triki C *et al*. Homozygous mutation of PLCZ1 leads to defective human oocyte activation and infertility that is not rescued by the WW-binding protein PAWP. *Hum Mol Genet* 2016;**25**:878–891.

Gilissen C, Hoischen A, Brunner HG, Veltman JA. Disease gene identification strategies for exome sequencing. *Eur J Hum Genet* 2012;**20**:490–497.

Habura A, Tikhonenko I, Chisholm RL, Koonce MP. Interaction mapping of a dynein heavy chain. Identification of dimerization and intermediate-chain binding domains. *J Biol Chem* 1999;**274**:15447–15453.

Harbuz R, Zouari R, Pierre V, Ben Khelifa M, Kharouf M, Coutton C, Merdassi G, Abada F, Escoffier J, Nikas Y *et al*. A recurrent deletion of DPY19L2 causes infertility in man by blocking sperm head elongation and acrosome formation. *Am J Hum Genet* 2011;**88**:351–361.

Matzuk MM, Lamb DJ. The biology of infertility: research advances and clinical challenges. *Nat Med* 2008;**14**:1197–1213.

Stalf T, Sanchez R, Kohn FM, Schalles U, Kleinstein J, Hinz V, Tielsch J, Khanaga O, Turley H, Gips H *et al*. Pregnancy and birth after intracytoplasmic sperm injection with spermatozoa from a patient with tail stump syndrome. *Hum Reprod* 1995;**10**:2112–2114.

Storm van's GK, Omran H. Primary ciliary dyskinesia: clinical presentation, diagnosis and genetics. *Ann Med* 2005;**37**:439–449.

Wambergue C, Zouari R, Fourati Ben MS, Martinez G, Devillard F, Hennebicq S, Satre V, Brouillet S, Halouani L, Marrakchi O *et al*. Patients with multiple morphological abnormalities of the sperm flagella due to DNAH1 mutations have a good prognosis following intracytoplasmic sperm injection. *Hum Reprod* 2016;**31**:1164–1172.

World Health Organization. *WHO Laboratory Manual for the Examination and Processing of Human Semen*, Geneva, 5th edn. World Health Organization, 2010.

Human Reproduction, 2016
doi:10.1093/humrep/dew262

human reproduction

SUPPLEMENTARY DATA

Supplementary Table S1 **Semen parameters of 38 Iranian sporadic patients.**

Patients	Semen volume (ml)	Sperm conc. (10^6/ml)	Total motility 1 h (%)	Vitality (%)	Normal spermatozoa (%)	Anomalies of the flagella (%)			Anomalies of the head (%)			Amorph (%)
						Short	Coiled	Other	Round	Pin	Giant	
SP1	4.50	0.15	0		0	98						2
SP2	5	7	0	60	0	86	2	2		1		9
SP3	0.3	6	0	60	1	92				2		5
SP4*	4.5	13	Few spz	88	0	86	1					13
SP5	2.5	30	0	70	0	98						2
SP6	4	11	Few spz	88	0	89			3	1	2	5
SP7	3	12	0	88	0	92		2				8
SP8	1	5	0	55	1	81				2	6	10
SP9	2.4	6	0	35	0	95				3	2	
SP10	2.1	3	0	82	0	84	1		4			11
SP11	5	4	Few spz	13	0	79	1		1			19
SP12	3.2	0.2	Few spz		0	86	5	2				5
SP13	3	20	0	92	0	96				2		2
SP14	2	20	0	80	0	100						
SP15	2.5	14	Few spz	88	0	86	4	2	2	2		4
SP16	3	1	0	85	0	88	2		4		2	4
SP17	5.5	7	0	88	0	92			3			5
SP18	4	3	0	88	0	90				5		5
SP19	3	16	0	85	0	90	4		2	1		3
SP20	6.5	6	0	80	0	90		3				7
SP21	3.2	5	0	95	0	100						
SP22	4.6	0.5	0	80	0	89	1		7			3
SP23	3.4	25	0	80	0	84	1	8				7
SP24	6.5	4	Few spz	70	0	92						8
SP25	4.2	2	0	35	2	83		2		2		11
SP26	4	5	Few spz	84	1	78		4		2		15
SP27	4.6	25	0	80	0	93	3		2			2
SP28	3	34	0	56	0	71	8	1	1	3		16
SP29	3.4	2.5	Few spz	48	0	79	3	2		1	1	14
SP30	3.5	30	5	75	1	60			5	2		32
SP31	2	5	0	3	0	88	1		1			10
SP32	3	0.25	0		0	86			8	4		2
SP33	5	0.5	Few spz		1	72		2		2		23
SP34	1	8	0	95	0	95						5
SP35	3	1	0		0	80	2		7	1		10
SP36	2.5	7	0	97	0	84	9		2	2		5
SP37	4.5	12	0	84	0	98						2
SP38	2.5	32	0	92	0	83			4	3		10

*Mutated patient.

Supplementary Table SII **Primer sequences used for Sanger sequencing of *DNAH1* exons.**

Primer name	Primer sequence (5′-3′)	Tm
DNAH1-int54F	CACCCCAACTCTCCTTCCAT	58°C
DNAH1-int54R	TCTGGGCATCGTCAGCAATA	
DNAH1-ex23F	TGGGATGAGCCTATCTTGCT	60°C
DNAH1-ex23R	AGCCTTGTGGGCAGACAGT	

Supplementary Table SIII **Primer sequences used in RT-PCR and respective melting temperatures (Tm).**

Primer name	Primer sequence (5′-3′)	Size of amplicons	Tm
RTex55-DNAH1F	GCTTCATATTTTCTCCATCC	55 F/55R1: 584 bp 55 F/55R2: 441 bp	54°C
RTex55-DNAH1R1	CAATGTTGTCCTTGTCAAAC		
RTex55-DNAH1R2	ATGCACACAGCTTCTATGAC		

Supplementary Table SIV **All DNAH1 variations identified by WES for all probands analyzed from six MMAF families.**

Gene	Variant coordinates	Transcript	cDNA variation	Amino acid variation	Prediction	Nationality
DNAH1	Chr3:52420175:G:A	ENST00000420323	c.8626-1G > A	Splice acceptor	Damaging	Iranian
DNAH1	Chr3:52391630:T:G	ENST00000420323	c.3860.T > G	p.Val1287Gly	Possibly_damaging	Italian