Biallelic mutations in *ARMC2* lead to severe astheno-teratozoospermia due to sperm flagellum malformations in human and mouse.

Charles Coutton[1,2,†,*], Guillaume Martinez[1,2,,†], Zine-Eddine Kherraf[1,3], Amir Amiri-Yekta[1,3,4], Magalie Boguenet[1], Antoine Saut[1], Xiaojin He[5], Feng Zhang[6], Marie Cristou-Kent[1], Jessica Escoffier[1], Marie Bidart[7], Véronique Satre[1,2], Béatrice Conne[8], Selima Fourati Ben Mustapha[9], Lazhar Halouani[9], Ouafi Marrakchi[9], Mounir Makni[9], Habib Latrous[9], Mahmoud Kharouf[9], Karin Pernet-Gallay[10], Mélanie Bonhivers[11], Sylviane Hennebicq[1,12], Nathalie Rives[13,14], Emmanuel Dulioust[15,16], Aminata Touré[16,17,18], Hamid Gourabi[4], Yunxia Cao[5], Raoudha Zouari[9], Seyedeh Hanieh Hosseini[19], Serge Nef[8], Nicolas Thierry-Mieg[20], Christophe Arnoult[1,§] and Pierre F. Ray[1,3§,**]

[1] Univ. Grenoble Alpes, INSERM U1209, CNRS UMR 5309, Institute for Advanced Biosciences, Team Genetics Epigenetics and Therapies of Infertility, 38000 Grenoble, France.
[2] CHU Grenoble Alpes, UM de Génétique Chromosomique, 38000 Grenoble, France.
[3] CHU Grenoble Alpes, UM GI-DPI, Grenoble, 38000, France
[4] Department of Genetics, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR, Tehran, Iran, PO Box 16635-148
[5] Reproductive Medicine Center, Department of Obstetrics and Gynecology, the First Affiliated Hospital of Anhui Medical University, Hefei 230022, China
[6] Obstetrics and Gynecology Hospital, Institute of Metabolism and Integrative Biology, School of Life Sciences, Fudan University, Shanghai 200011, China
[7] Clinatec, Pôle Recherche, INSERM UMR 1205, CHU Grenoble, Alpes, Univ. Grenoble Alpes, 38000 Grenoble, France
[8] Department of Genetic Medicine and Development, University of Geneva Medical School, 1211, Geneva, Switzerland.
[9]Polyclinique les Jasmins, Centre d'Aide Médicale à la Procréation, Centre Urbain Nord, 1003 Tunis, Tunisia
[10] Grenoble Neuroscience Institute, INSERM 1216, Grenoble, F38000, France
[11] Université de Bordeaux, Microbiologie Fondamentale et Pathogénicité, CNRS UMR 5234, 33000 Bordeaux, France.
[12] CHU de Grenoble, UF de Biologie de la procréation, F-38000 Grenoble, France.
[13] Normandie Univ, UNIROUEN, EA 4308 'Gametogenesis and Gamete Quality', Rouen University Hospital, Department of Reproductive Biology-CECOS, F 76000 Rouen, France.
[14] Inserm U1085-IRSET, Université de Rennes 1, Rennes, France.
[15] Laboratoire d'Histologie Embryologie - Biologie de la Reproduction, GH Cochin Broca Hôtel Dieu, Assistance Publique-Hôpitaux de Paris, 75014, Paris, France.
[16] Université Paris Descartes, Sorbonne Paris Cité, Faculté de Médecine, Paris 75014, France.
[17] INSERM U1016, Institut Cochin, Paris 75014, France.
[18] Centre National de la Recherche Scientifique UMR8104, Paris 75014, France.
[19] Department of Andrology, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR, Tehran, Iran.
[20] Univ. Grenoble Alpes, CNRS, TIMC-IMAG / BCM, 38000 Grenoble, France

[†, §] These authors contributed equally to this work
*Correspondance : CCoutton@chu-grenoble.fr ; **Correspondance : PRay@chu-grenoble.fr

Abstract

Male infertility is a major health concern. Among its different causes, multiple morphological abnormalities of the flagella (MMAF) induces asthenozoospermia and is one of the most severe forms of qualitative sperm defects. Sperm of affected men display short, coiled, absent and/or irregular flagella. To date, six genes have been found to be recurrently associated with MMAF (*DNAH1*, *CFAP43*, *CFAP44*, *CFAP69*, *FSIP2* and *WDR66)* but more than half of the cases analyzed remain unresolved, suggesting that many yet uncharacterized gene defects account for this phenotype. Here, whole-exome sequencing (WES) was performed on 168 infertile men with a typical MMAF phenotype. Five unrelated affected individuals carried a homozygous deleterious mutation in *ARMC2*, a gene not previously linked to the MMAF phenotype. Using the CRISPR-Cas9 technique, we generated homozygous *Armc2* mutant mice, which also presented a MMAF phenotype, thus confirming the involvement of *ARMC2* in human MMAF. Immunostaining experiments in *AMRC2*-mutated individuals and mutant mice evidenced the absence of the axonemal central pair complex (CPC) proteins, SPAG6 and SPEF2, while the other tested axonemal and peri-axonemal components were present, suggesting that ARMC2 is involved in CPC assembly and/or stability. Overall, we showed that biallelic mutations in *ARMC2* cause male infertility in man and mouse by inducing a typical MMAF phenotype, indicating that this gene is necessary for sperm flagellum structure and assembly.

Text

The characterization of the genetic basis of male infertility represents an important challenge as over 4000 genes are thought to be needed for sperm production[1] and therefore defects in any of these genes can hamper spermatogenesis and induce one of many established sperm phenotypes[2]. Of late, this task has been greatly facilitated by high throughput sequencing technologies which have allowed to identify an increasing number of genes required for sperm production[3]. This is especially true for teratozoospermia, qualitative defects of spermatogenesis leading to morphological sperm abnormalities[4,5]. The MMAF phenotype is one of the most severe forms of qualitative sperm defects responsible for astheno-teratozoospermia[4] characterized by the presence of immotile spermatozoa presenting with a mosaic of sperm flagellum malformations including short, coiled, absent, and/or flagella of irregular caliber[6]. Whole-exome sequencing (WES) analysis revealed that mutations in *DNAH1* (MIM:603332), *CFAP43* (MIM:617558), *CFAP44* (MIM:617559), *CFAP69* (MIM:617949), *FSIP2* (MIM:615796) and *WDR66* (MIM:612573) account for the main genetic causes of MMAF[6–13]. Rarer recessive mutations in *AK7* (MIM:615364), *CEP135* (MIM 611423) and *CFAP65* (MIM:614270) were also recently identified in different familial cases of MMAF[10,14,15]. Despite these recent findings, more than half of the studied MMAF cases remain without a diagnosis demonstrating the high genetic heterogeneity of this phenotype and the need for further genetic explorations[12].

Here, we analysed by whole-exome sequencing (WES) a total of 168 MMAF individuals including 78 MMAF individuals previously reported[9] and an additional 90 unrelated and unpublished MMAF individuals. All but one individual were analysed together with the same analytical pipeline constituting a 167 strong cohort. In this main cohort, 83 individuals were of North African origin and consulted for primary infertility at the Clinique des Jasmins

in Tunis, 52 individuals originated from the Middle East (Iran) and were treated in Tehran at the Royan Institute, (Reproductive Biomedicine Research Center) for primary infertility and 32 subjects were recruited in France: 25 at the Cochin Institute in Paris, 3 in Rouen, 2 in Grenoble, 1 in Lille and Caen. All individuals presented with a typical MMAF phenotype characterized by severe asthenozoospermia (total sperm motility below 10%) with at least three of the following flagellar abnormalities present in >5% of the spermatozoa: short, absent, coiled, bent or irregular flagella (Table 1, Figure 1A-D). All individuals had a normal somatic karyotype (46,XY) with normal bilateral testicular size, normal hormone levels and secondary sexual characteristics. Informed consent was obtained from all the subjects participating in the study according to local protocols and the principles of the Declaration of Helsinki. The study was approved by local ethics committees, and samples were then stored in the CRB Germethèque (certification under ISO-9001 and NF-S 96-900) following a standardized procedure or were part of the Fertithèque collection declared to the French Ministry of health (DC-2015-2580) and the French Data Protection Authority (DR-2016-392).

WES and bioinformatics analysis were performed according to an improved version of our previously described protocol[9] as described in supplemental Material and Methods. Data analysis of the whole cohort of 167 MMAF individuals permitted to identify 54 individuals (32.3 %) with harmful mutations in known MMAF-related genes (Table S1). Previously unreported variants were identified in 15 subjects in *CFAP43*, (n=2), *CFAP44* (n=1), *DNAH1* (n=6), *WDR66* (n=4) and *FSIP2* (n=2), thus confirming the importance of these genes in the etiology of the MMAF syndrome (Table S1). In addition, we identified four individuals (ARMC2$_{1-4}$) with a homozygous variant in *ARMC2*, a gene not previously associated with any pathology, accounting for 2.4% of our cohort. *ARMC2* (NM_032131.5) is located on chromosome 6 and contains 18 exons encoding a predicted 867-amino acid protein

(NP8115507.4 ; Q8NEN0). Three subjects had a loss of functions variant (ARMC2$_{1,3,4}$) and one a likely deleterious missense variant (ARMC2$_{2}$), (Table 1, Table S1). In addition, we identified a fifth subject with an *ARMC2* homozygous loss of function variant (ARMC2$_{5}$) (Table S2). This individual was of Chinese origin, and consulted for primary infertility at the First Affiliated Hospital of Anhui Medical University (Hefei City, Anhui Province) in 2012. He was born to first cousin parents, presented a typical MMAF phenotype, and WES analysis for this subject was performed as described in the previous report [10]. Informed consent was obtained from all tested individuals and the study approved by the local institutional board.

The five *ARMC2* variants were found in five unrelated individuals and all were absent from control sequence databases (dbSNP, 1000 Genomes Project, NHLBI Exome Variant Server, gnomAD and our in-house control database). The variant identified in individual ARMC2$_{1}$ is a splice variant c.1023+1G>A, altering a consensus splice donor site of *ARMC2* exon 8 (Figure 1I). To evaluate the deleterious effect of this splicing variant, we performed RT-PCR with RNA extracted from sperm cells from ARMC2$_{1}$. Amplification of a sequence ranging from cDNA exon 7-10 yielded a normal band of 608 bp in control whereas a single smaller band was obtained from ARMC2$_{1}$ cDNA (Figure S1A). Sequence analysis of the amplified products demonstrated that the c.1023+1G>A variant results in exon 8 skipping (Figure S1B, C) causing a shift in the translational reading frame and introducing a premature termination codon (p.Glu283AlafsTer2). Primers sequences and RT-PCR conditions are indicated in Table S3. The variant identified in individual ARMC2$_{2}$ is a missense variation c.2279T>A (p.Ile760Asn) located in exon 16 (Figure 1I). No mRNA analysis or immunostaining could be performed on sperm cells from this individual due to the lack of biological samples. However, using prediction software for non-synonymous SNPs we found that this missense change is predicted to be deleterious by SIFT (score of 0) and probably damaging by PolyPhen (score of 0.987). Moreover the concerned amino acid (Ile760) was

found conserved in *ARMC2* orthologs (Figure S2). The two others variants identified in ARMC2_3 and ARMC2_4, were small frameshift indels, c.2353_2354delTT (p.Leu785MetfsTer5) and c.1284_1288delACAAA (p.Lys428AsnfsTer3), located in exon 17 and 10, respectively (Figure 1I) inducing a premature translation termination. The last variant identified (ARMC2_5) is a stop-gain variant c.421C>T (p.Gln141Ter) located in exon 4 (Figure 1I). Familial study confirmed the presence of the homozygous loss of function variant in ARMC2_5 and indicated that his parents were both heterozygous and his non-affected brother was homozygous wild-type (Figure S3A,B). The last three variants (in subjects ARMC2_3-5) introduce a premature stop codon and are therefore expected to induce non-sense mediated mRNA decay expected to prevent protein synthesis. All *ARMC2* variants are deposited in ClinVar under reference SUB4929442.

Overall a total of 168 exomes from MMAF subjects were analyzed and five unrelated affected individuals were shown to carry a homozygous deleterious mutation in *ARMC2*. No other candidate variants reported to be associated with cilia, flagella or male fertility were present in any of the five *ARMC2* mutated subjects. We also note that none of the subjects analyzed here carried a homozygous deleterious variant in any two (or more) of the pathological MMAF-associated genes (*DNAH1*, *CFAP43*, *CFAP44*, *CFAP69*, *FSIP2*, *WDR66* and *ARMC2)*, *ie* the 54 subjects with an established causal mutation did not carry another candidate variant. All mutations identified by WES were validated by Sanger sequencing as previously described[9,10] and illustrated in Figure 1H. PCR primers and protocols used for each individual are listed in Table S4.

ARMC2 is preferentially expressed in the testis according to data from ENCODE, FANTOM and GTEx[16–18] and described to be associated with cilia and flagella[19]. We confirmed these data by RT-qPCR experiments in a panel of human tissues, which indicate that expression

of *ARMC2* mRNA in testis is significantly higher than in other tested tissues (Figure S4). Primers sequences and RT-qPCR conditions are indicated in Table S5. According to the Uniprot server, ARMC2 is an armadillo protein composed of 12 armadillo repeats (ARM-repeat) flanked by unique C-terminal and N-terminal domains[20] (Figure 1I).

Sperm analysis was carried out in the source laboratories during routine biological examination of the individuals according to World Health Organization (WHO) guidelines[21]. The morphology of the sperm cells was assessed with Papanicolaou staining (Figure 1A-D). Detailed semen parameters of the five *ARMC2* mutated individuals are presented in Table 1, and the average semen parameters of the studied MMAF cohort separated by genotypes are described in the Table S6. Among the different parameters studied, only viability, total motility and "lack of flagellum" presented statistical difference between the different groups, according to their genotype (one way ANOVA test). For parameters with a positive ANOVA test, a pair-wise statistical Fisher's LSD tests was employed to identify significant differences between subjects with a different genotype (Figure S5). Concerning vitality, sperm from *WDR66* mutated subjects and non-characterized subjects (unknown) presented a significant increase in comparison with *CFAP69*, *CFAP44*, *FSIP2*, *ARMC2* mutated individuals. For motility *CFAP43* and *CFAP44* showed the most pronounced alteration, due to a significant increase in sperm without flagellum and *CFAP44* subjects displayed a significant increase in "no tail", sperm (Figure S5).

Sperm samples for additional phenotypic characterization could only be obtained from ARMC2_1. We studied the ultrastructure of sperm cells from subject ARMC2_1 by transmission electron microscopy (TEM) (Figure 1E-G), according to the protocol previously described[9]. For details, see Supplemental Material and Methods. Due to the low number of sperm cells

available, only a few cross-sections (<10) presented a sufficient quality to be analyzed. Among these sections, all were abnormal and the main defect observed was the absence of the CPC (9+0 conformation) (Figure 1F). In some sections we observed a dramatic axonemal disorganization associated with peri-axonemal structural defects such as unassembled outer dense fibers (ODFs) (Figure 1G), a defect already observed in sperm from MMAF-affected individuals carrying mutations in other genes. Observation of rare longitudinal sections showed severe flagellum abnormalities which appeared completely disorganized resulting in truncated flagella or a cytoplasmic mass encompassing unassembled axonemal components (not shown).

In order to confirm these results, we assessed the impact of the absence ARMC2 on mouse spermatogenesis by generating mutant animals using the CRISPR-Cas9 technology as previously described[9,22]. For all experiments involving mice, animals were handled and euthanized in accordance with methods approved by the Animal Ethic Committees of Grenoble and Geneva. All mice used were adult (6 weeks or older) mice. We generated a strain with a 1 nucleotide duplication in exon 4 (DupT) inducing a translational frameshift expected to lead to the complete absence of the protein or the production of a truncated protein. mRNA was extracted from *Armc2* homozygous mutant mice (*Armc2*^{mutant}) and amplified by RT-PCR. The level of *Armc2*^{mutant} mRNA amplification was much lower in mutant animals than in controls (Figure S6). Sanger sequencing of mRNA from *Armc2* homozygous mutant mice confirmed the production of abnormal transcripts with a premature stop codon 12 nucleotides after the first modified codon at position 135 (NM_001034858.3:c.403dupT, NP_001030030.2:p.Tyr135LeufsTer12) (Figure S6). The guide RNA sequence for CRISPR/Cas9 mice generation and the list of primers used for mice genotyping and RT-PCR are indicated in Table S7 and S8. We studied sperm morphology and observed that in contrast to what is observed in WT animals (A), epididymal sperm from *Armc2*^{mutant} males displayed a

phenotype identical to the typical human MMAF phenotype with 100% of spermatozoa with short, thick and/or coiled flagella, while sperm heads conserved an overall typical hooked shape (Figure 2A-D). As could be expected homozygous *Armc2*^{mutant} males exhibited complete infertility when mated with WT females (Figure 2E) whereas homozygous *Armc2*^{mutant} females were fully fertile and gave litters of normal size compared to WT females (8.33 ± 1.11 versus 9.67 ± 1.01 pups/litter (mean ± SE, n=6 versus n=3). There was no obvious testicular anomaly, with no difference in testis weight, 94 ± 20 and 92 ± 10.67 mg per testis (mean ± SE, n=6 and n=3) for *Armc2*^{mutant} and WT respectively (Figure 2F). Structural defects were observed in close to 100% of spermatozoa from *Armc2*^{mutant} males (Figure 2G) and were associated with a complete motility deficiency (Figure 2H). Sperm production was also affected as *Armc2*^{mutant} sperm concentrations were 4.47 ± 1.29 x10^6 sperm/ml versus 20.75 ± 8.33 x10^6 sperm/ml (mean ± SE, n=6 and n=3), observed in WT littermates (Figure 2I).

To define the ultrastructural defects evidenced by TEM and to characterize the molecular defects induced by *ARMC2* mutation in human sperm, we studied by immunofluorescence (IF) the presence and localization of several proteins belonging to different sub-structures of the axoneme. The presence of the following proteins was investigated: SPAG6 as a marker of the CPC, DNAI2 and DNALI1 as markers of outer (ODA) and inner dynein arms (IDA) respectively, RSPH1 as a marker of the radial spokes (RS), GAS8 as a marker of the nexin-dynein regulatory complex (N-DRC) and AKAP4 as marker of the fibrous sheath (FS). We observed that in sperm from individual ARMC2_1, staining of SPAG6, an axoneme central pair complex protein[23], was totally absent from the flagellum (Figure 3A). In contrast, immunostaining for AKAP4, DNALI1, DNAH5, RSPH1 and GAS8 were all similar to controls, suggesting that FS, IDAs, ODAs, RS and the N-DRC respectively were not directly affected by mutations in *ARMC2* (Figure S7). Due to limited sample availability, these

analyses could not be repeated on sperm from other *ARMC2* mutated individuals. Also, we regret that we could not obtain any specific antibodies against ARMC2 allowing the localization of the protein in human and mouse sperm. In addition, IF experiments performed in *Armc2*^{mutant} animals showed that SPEF2 staining, another marker of the CPC[24], was totally absent supporting the CPC defects observed in sperm samples from the MMAF individual ARMC2_1 (Figure 3-B).

We showed that the presence of bi-allelic *ARMC2* mutations induce a typical MMAF phenotype in both humans and mice, indicating that this gene is necessary for spermatogenesis and in particular for sperm flagellum structure and motility. Bioinformatic analysis suggests that *ARMC2* encodes a protein belonging to the ARM-repeat containing protein family[20] (Figure 1I). ARM-repeats are typically characterized by a 42 amino acid motif composed of three α-helices[25]. Tandem ARM-repeat units fold together as "superhelixes" serving as platforms for protein-protein interactions[26]. ARM-repeat proteins are involved in a wide range of cellular functions including intracellular signaling, cytoskeletal regulation, and protein degradation or folding[27]. ARM-repeat proteins are also involved in different functions in cilia and flagella such as intraflagellar transport and assembly/stability of different axonemal components[26,27]. Moreover, several ARM-repeat proteins have been described to be involved in ciliopathies or male infertility [23,28–31]. In particular, in cattle and mouse, mutations in *ARMC3* and *SPAG6* (a protein with eight contiguous armadillo repeats), are involved in male infertility due to sperm flagellum malformations [23,32,33]. More recently, we also demonstrated that the absence of *CFAP69*, another ARM-repeat flagellar associated protein, leads to a MMAF phenotype in human and mouse[11]. These data clearly demonstrate that Armadillo-domain repeat proteins are critical for correct spermatogenesis and flagellum formation, but the precise function of ARMC2 and of the other ARM-repeat flagellar associated protein remains to be elucidated.

Immunostaining experiments in sperm cells from individual ARMC2_1 harboring a splicing mutation evidenced an absence of the SPAG6 protein, which normally locates on the C1 singlet[34,35], strongly suggesting defects on the CPC structure (Figure 3A). These findings are in concordance with TEM analysis, which revealed the absence of the central pair in the few cross-sections observed in sperm from ARMC2_1 (Figure 1F). Similar observations were described in the *Armc2*^{mutant} mouse with abnormal staining of the SPEF2 protein (Figure 3B). IF experiments showed that other axonemal or peri-axonemal structures did not appear to be affected by the *ARMC2* mutation (Figure S7), suggesting that in absence of ARMC2, these proteins were correctly addressed and were able to maintain their normal localization within the axoneme. These observations suggest that ARMC2 may be specifically involved in the CPC assembly or stability. Interestingly, the two other Armadillo-repeat protein SPAG6 and CFAP69 associated with MMAF phenotype are also linked to the CPC[11,24]. The CPC complex consists of two singlets of microtubules, named C1 and C2, which are structurally and biochemically distinct, and surrounded by complex protein structures, known as projections that are unique for each microtubule[24]. It is worth to note that cilia without CPC exist in different tissues (9+0 organization), indicating that the CPC is not necessary for the axoneme growing and integrity; interestingly those 9+0 cilia are most often non-motile cilia. In contrast to cilia, the CPC seems to play a key role in maintaining the global organization of the sperm flagellum throughout spermiogenesis. Moreover, inactivation of different genes encoding proteins directly or indirectly associated to the CPC was described to lead to MMAF phenotype in human and several animal models[4] confirming the importance of CPC in sperm flagellum assembly/structure. Further experiments, including the development of functional antibodies should now be performed to precisely determine the axonemal localization and to confirm the putative role of ARMC2 in the CPC assembly or stability.

We present here a cohort of 167 individuals. Overall we identified 50 individuals (29.9%) with mutation in known MMAF-related genes including the 35 individuals previously reported (Table S1). We identified 10 different mutations in *DNAH1* in 10 individuals (6%), 12 mutations in *CFAP43* in 12 individuals (7.2%), 2 mutations in *WDR66* in 11 individuals (6.6%), 6 mutations in *CFAP44* in 8 individuals (4.8%), 6 mutations in *FSIP2* in 7 individuals (4.2%) and 2 mutations in *CFAP69* in 2 individuals (1.2%) (Table S1). When adding the four *ARMC2* mutated subjects of the cohort, we obtained a diagnostic efficiency of 32.3% (54/167). We note that for previously published genes, the newly described variants only include the homozygous loss of function variants and we did not report the unpublished variants of questionable significance (missense variants, in frame deletions or variants located in splice region other than the consensus splice sites) which we feel need additional confirmatory work. From these results, *CFAP43, WDR66* and *DNAH1* appear as the most frequently mutated genes in MMAF-affected individuals. Ten of the eleven *WDR66* subjects however harbored the same founder deletion[13] found only in Tunisian subjects (Table S1). The contribution of this gene in other populations could therefore be limited, although *WDR66* has been reported to induce MMAF by others[36]. WES sequencing in MMAF continues to permit the identification of new genes and pathological variants demonstrating the efficiency of WES to investigate the genetic causes of MMAF syndrome. We however observe here that despite regular new gene identification, about two thirds of MMAF-affected individuals remain with unknown genetics causes highlighting the high genetic heterogeneity of the phenotype consistent with the large number of genes involved in the spermatogenesis[37]. As discussed above, some variants of unknown significance in identified MMAF-associated genes could also be responsible for the infertility of some of the investigated subjects. Reliable techniques to validate the pathogenicity of these variations are currently cruelly lacking. These results also suggest that the WES approach cannot be expected to provide 100% positive diagnoses. This could also be explained in part by the fact

that some variants are not detected by the technique used (non-sequenced deep intronic variants or some exonic variants as only approximately 90% of coding nucleotides were covered) or by the current bio-informatic pipeline used for the analysis (e.g., small duplications, structural variations rearrangements). To improve this diagnosis rate and to detect new variants in yet uncharacterized genes, more powerful techniques such as whole genome sequencing (WGS) may now be envisaged for MMAF subjects for whom no variants were identified[38]. Moreover, while the use of new sequencing technologies now permits to easily identify new candidate genes, the resulting high flow of data can easily lead to erroneous diagnosis. Therefore, human genetic studies pertaining to new candidate genes and missense variants should be confirmed by performing functional work and/or phenotypic characterization using animal models. Such studies are also essential to improve our understanding of the function of the studied genes and their role in spermatogenesis[22].

Like all the other MMAF subjects included in this study, the five individuals with *ARMC2* mutations presented only with primary infertility without any other clinical features, thus excluding a phenotype of Primary Cilia Dyskinesia (PCD). This observation clearly supports that sperm flagellum biogenesis requires a pool of proteins, which is different from the pool necessary for cilia biogenesis. In addition, this also suggests that some assembly mechanisms may be specific to the sperm flagellum. We did not observe any significant differences between the semen parameters of the five individuals carrying mutations in *ARMC2* compared to the parameters of individuals with mutations in other MMAF-related genes. This analysis however revealed that sperm from *CFAP43* and *CFAP44* individuals present the highest level of defects, whereas sperm form the unknown group seems the least altered (Figure S5). Last, some questions remain concerning the prognosis of ICSI using sperm cells from *ARMC2*-mutated subjects. Several studies previously demonstrated that MMAF affected

individuals had a good prognosis using ICSI[15,31,39]. It would however be interesting to compare the success rates after ICSI taking into account all MMAF genotypes. Altogether, these data demonstrate that ARMC2 is essential for sperm tail biogenesis in human and mouse and that mutations in this gene lead to drastic flagellum malformations resulting in severe astheno-teratozoospermia and primary male infertility.

Supplemental Data

Supplemental Data include materials and methods, 7 figures and 8 tables.

Declaration of Interest

The authors declare no competing interests.

Acknowledgments

We thank all individuals for their participation. This work was mainly supported by the following grants: The “MAS-Flagella” project financed by French ANR and the DGOS for the program PRTS 2014 (P.F.R.), the “Whole genome sequencing of subjects with Flagellar Growth Defects (FGD)” financed by the fondation maladies rares (FMR) for the program Séquençage à haut débit 2012 (P.F.R.).

Web Resources

Online Mendelian Inheritance in Man, https://www.omim.org
1000 Genomes, http://www.internationalgenome.org
dbSNP, https://www.ncbi.nlm.nih.gov/projects/SNP
ENCODE, https://www.encodeproject.org
ExAC Browser, http://exac.broadinstitute.org/
FANTOM, http://fantom.gsc.riken.jp
gnomAD Browser, http://gnomad.broadinstitute.orgGTEx, http://www.gtexportal.org
NHLBI Exome Sequencing Project (ESP) Exome Variant Server,
http://evs.gs.washington.edu/EVS/
PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/bgi.shtml
SIFT, http://sift.jcvi.org
ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/

Figure Legends

Figure 1. Morphology of normal and *ARMC2* mutant spermatozoa, and the mutations identified in *ARMC2*-mutant individuals. Light microscopy analysis of spermatozoa from fertile control (A) and individual ARMC$_{_1}$ (B-D). Most spermatozoa from *ARMC2* individual have flagella that are coiled (B), short (C) and/or of irregular caliber (D). Transmission electron microscopy analyses of sperm cells from a control subject (E) and individual ARMC2$_{_1}$ (F-G). (E) Cross-sections of the principal piece from a fertile control. The axoneme is composed of nine doublets of microtubules (DMTs) circularly arranged around a central pair complex of microtubules (CPC) ('9+ 2' organization). The axoneme is surrounded by seven outer dense fibers (ODFs) and by the fibrous sheath (FS) composed of two longitudinal columns (LC) connected by circumferential ribs (CR). (F) Cross-section of sperm flagellum from ARMC2$_{_1}$ showing a 9+0 axoneme lacking the CPC. (G) Cross-section of sperm flagellum from ARMC2$_{_1}$ showing a severe axonemal disorganization with unassembled ODFs and DMTs. Scale bars, 10 µm (A–D) and 200 nm (E–G). (H) Electrophoregrams of Sanger sequencing for the five ARMC2-mutated individuals compared to reference sequence. (I) Location and nature of *ARMC2* mutations in *ARMC2* and protein. Colored squares stand for armadillo repeats as predicted by the Uniprot server. Mutations are annotated in accordance to the HGVS's recommendations.

Figure 2. Reproductive phenotype of heterozygous and homozygous *Armc2* male mice. (A) A spermatozoon with normal morphology from a wild-type male mouse. (B–D) Sperm from *Armc2*-deficient mice show severe morphological defects. All spermatozoa had flagellar abnormalities: short, coiled, absent, and/or of irregular caliber consistent with the MMAF phenotype reported in the *ARMC2*-mutated individuals. Scale bars: 10 µm. (E) Fertility of wild-

type (WT), *Armc2 heterozygous mutant (Armc2*$^{+/-}$) and *Armc homozygous mutant (Armc2*$^{-/-}$) males. (F) Testis weight from WT, *Armc2*$^{+/-}$ and *Armc2*$^{-/-}$ males. (G) Total motility of sperm extracted from the cauda epididymis of WT, *Armc2*$^{+/-}$ and *Armc2*$^{-/-}$ males. (H) Concentrations of sperm from the cauda epididymis from wild-type, heterozygous and homozygous *Armc2* males. (I) Concentration of morphologically abnormal sperm in WT, heterozygous and homozygous *Armc2* males, evaluated after Harris-Schorr coloration (expressed in million/ml (M/mL)). Data represent mean ± SE; statistical differences were assessed with t test; a probability value of less than 0.05 was considered to be statistically significant.

Figure 3. Immunostaining of SPAG6 and SPEF2 revealed that the central pair complex is affected by mutations in *ARMC2* in human and mouse. (A) Sperm cells from a fertile control and ARMC2_1-stained with anti-SPAG6 (rabbit polyclonal, HPA38440, Sigma-Aldrich, 1:500, green), which detects a protein located in the C1 microtubule, and anti-acetylated tubulin (monoclonal mouse T7451, Sigma-Aldrich, 1:2000, red) antibodies. SPAG6 staining uniformly decorates the full-length flagellum in the fertile control whereas it is absent from the flagellum of sperm from individual ARMC2_1. (B) Mouse sperm cells from a WT and *Armc2*$^{-/-}$ males stained with anti-SPEF2 (rabbit polyclonal, HPA040343, Sigma-Aldrich, 1:1000, green), a marker of the projection 1b of singlet C1, and anti-acetylated tubulin (monoclonal mouse, T7451, Sigma-Aldrich, 1:500, red) antibodies. DNA was counterstained with DAPI. Contrary to the WT, the SPEF2 immunostaining is not detectable in the sperm flagellum from the *Armc2*$^{-/-}$ male. Scale bars: 10 µm**.**

TABLES

Table 1. Detailed semen parameters for the five MMAF individuals harboring a *ARMC2* mutation.

***ARMC2* mutated individuals**		**Semen parameters**																
Individuals	*ARMC2* mutation	Sperm volume (ml)	Sperm conc.(10^6/ml)	Total motility 1h	Vitality	Normal spermatozoa	Absent flagella	Short Flagella	Coiled Flagella	Bent Flagella	Flagella of irregular caliber	Tapered head	Thin head	Micro-cephalic	Macro-cephalic	Multiple heads	Abnormal base	Abnormal acrosomal region
ARMC2_1	c.1023+1G>A	4	35	6	NA	0	10	50	16	NA	70	18	28	32	0	2	28	98
ARMC2_2	c.2279T>A	4.5	4.6	5	64	0	14	30	6	16	78	2	2	42	0	4	6	88
ARMC2_3	c.2353_2354delTT	1	58.5	0	16	3	31	37	23	NA	39	60	9	0	0	0	31	52
ARMC2_4	c.1284_1288delACAAA	4.2	2	0	35	2	0	83	0	0	0	11	0	2	0	0	0	0
ARMC2_5	c.421 C>T	9.2	10.1	1.1	NA	0	20	38	36	0	0	NA	NA	NA	NA	NA	NA	NA
Reference limits[a]		1.5 (1.4-1.7)	15 (12-16)	40 (38-42)	58 (55-63)	23 (20-26)	5 (4-6)	1 (0-2)	17 (15-19)	13 (11-15)	2 (1-3)	3 (2-4)	14 (12-16)	7 (5-9)	1 (0-2)	2 (1-3)	42 (39-45)	60 (57-63)

Values are percentages unless specified otherwise.[a] Reference limits (5th centiles and their 95% confidence intervals) according to the World Health Organization.

REFERENCES

1. Jan, S.Z., Vormer, T.L., Jongejan, A., Röling, M.D., Silber, S.J., de Rooij, D.G., Hamer, G., Repping, S., and van Pelt, A.M.M. (2017). Unraveling transcriptome dynamics in human spermatogenesis. Dev. Camb. Engl. *144*, 3659–3673.

2. Tüttelmann, F., Ruckert, C., and Röpke, A. (2018). Disorders of spermatogenesis: Perspectives for novel genetic diagnostics after 20 years of unchanged routine. Med. Genet. Mitteilungsblatt Berufsverb. Med. Genet. EV *30*, 12–20.

3. Krausz, C., and Riera-Escamilla, A. (2018). Genetics of male infertility. Nat. Rev. Urol. *15*, 369–384.

4. Coutton, C., Escoffier, J., Martinez, G., Arnoult, C., and Ray, P.F. (2015). Teratozoospermia: spotlight on the main genetic actors in the human. Hum. Reprod. Update *21*, 455–485.

5. Ray, P.F., Toure, A., Metzler-Guillemain, C., Mitchell, M.J., Arnoult, C., and Coutton, C. (2017). Genetic abnormalities leading to qualitative defects of sperm morphology or function. Clin. Genet. *91*, 217–232.

6. Ben Khelifa, M., Coutton, C., Zouari, R., Karaouzène, T., Rendu, J., Bidart, M., Yassine, S., Pierre, V., Delaroche, J., Hennebicq, S., et al. (2014). 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. *94*, 95–104.

7. Amiri-Yekta, A., Coutton, C., Kherraf, Z.-E., Karaouzène, T., Le Tanno, P., Sanati, M.H., Sabbaghian, M., Almadani, N., Sadighi Gilani, M.A., Hosseini, S.H., et al. (2016). Whole-exome sequencing of familial cases of multiple morphological abnormalities of the sperm flagella (MMAF) reveals new DNAH1 mutations. Hum. Reprod. Oxf. Engl. *31*, 2872–2880.

8. Wang, X., Jin, H., Han, F., Cui, Y., Chen, J., Yang, C., Zhu, P., Wang, W., Jiao, G., Wang, W., et al. (2016). Homozygous DNAH1 frameshift mutation causes multiple morphological anomalies of the sperm flagella in Chinese. Clin. Genet.

9. Coutton, C., Vargas, A.S., Amiri-Yekta, A., Kherraf, Z.-E., Ben Mustapha, S.F., Le Tanno, P., Wambergue-Legrand, C., Karaouzène, T., Martinez, G., Crouzy, S., et al. (2018). Mutations in CFAP43 and CFAP44 cause male infertility and flagellum defects in Trypanosoma and human. Nat. Commun. *9*, 686.

10. Tang, S., Wang, X., Li, W., Yang, X., Li, Z., Liu, W., Li, C., Zhu, Z., Wang, L., Wang, J., et al. (2017). Biallelic Mutations in CFAP43 and CFAP44 Cause Male Infertility with Multiple Morphological Abnormalities of the Sperm Flagella. Am. J. Hum. Genet. *100*, 854–864.

11. Dong, F.N., Amiri-Yekta, A., Martinez, G., Saut, A., Tek, J., Stouvenel, L., Lorès, P., Karaouzène, T., Thierry-Mieg, N., Satre, V., et al. (2018). Absence of CFAP69 Causes Male Infertility due to Multiple Morphological Abnormalities of the Flagella in Human and Mouse. Am. J. Hum. Genet. *102*, 636–648.

12. Martinez, G., Kherraf, Z.-E., Zouari, R., Fourati Ben Mustapha, S., Saut, A., Pernet-Gallay, K., Bertrand, A., Bidart, M., Hograindleur, J.P., Amiri-Yekta, A., et al. (2018). Whole-exome sequencing identifies mutations in FSIP2 as a recurrent cause of multiple morphological abnormalities of the sperm flagella. Hum. Reprod. Oxf. Engl.

13. Kherraf, Z.-E., Amiri-Yekta, A., Dacheux, D., Karaouzène, T., Coutton, C., Christou-Kent, M., Martinez, G., Landrein, N., Le Tanno, P., Fourati Ben Mustapha, S., et al. (2018). A Homozygous

Ancestral SVA-Insertion-Mediated Deletion in WDR66 Induces Multiple Morphological Abnormalities of the Sperm Flagellum and Male Infertility. Am. J. Hum. Genet.

14. Lorès, P., Coutton, C., El Khouri, E., Stouvenel, L., Givelet, M., Thomas, L., Rode, B., Schmitt, A., Louis, B., Sakheli, Z., et al. (2018). Homozygous missense mutation L673P in adenylate kinase 7 (AK7) leads to primary male infertility and multiple morphological anomalies of the flagella but not to primary ciliary dyskinesia. Hum. Mol. Genet. *27*, 1196–1211.

15. Sha, Y.-W., Xu, X., Mei, L.-B., Li, P., Su, Z.-Y., He, X.-Q., and Li, L. (2017). A homozygous CEP135 mutation is associated with multiple morphological abnormalities of the sperm flagella (MMAF). Gene *633*, 48–53.

16. Gerstein, M.B., Kundaje, A., Hariharan, M., Landt, S.G., Yan, K.-K., Cheng, C., Mu, X.J., Khurana, E., Rozowsky, J., Alexander, R., et al. (2012). Architecture of the human regulatory network derived from ENCODE data. Nature *489*, 91–100.

17. Lizio, M., Harshbarger, J., Abugessaisa, I., Noguchi, S., Kondo, A., Severin, J., Mungall, C., Arenillas, D., Mathelier, A., Medvedeva, Y.A., et al. (2017). Update of the FANTOM web resource: high resolution transcriptome of diverse cell types in mammals. Nucleic Acids Res. *45*, D737–D743.

18. GTEx Consortium (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science *348*, 648–660.

19. Ivliev, A.E., 't Hoen, P.A.C., van Roon-Mom, W.M.C., Peters, D.J.M., and Sergeeva, M.G. (2012). Exploring the transcriptome of ciliated cells using in silico dissection of human tissues. PloS One *7*, e35618.

20. The UniProt Consortium (2017). UniProt: the universal protein knowledgebase. Nucleic Acids Res. *45*, D158–D169.

21. Wang, Y., Yang, J., Jia, Y., Xiong, C., Meng, T., Guan, H., Xia, W., Ding, M., and Yuchi, M. (2014). Variability in the morphologic assessment of human sperm: use of the strict criteria recommended by the World Health Organization in 2010. Fertil. Steril.

22. Kherraf, Z.-E., Conne, B., Amiri-Yekta, A., Kent, M.C., Coutton, C., Escoffier, J., Nef, S., Arnoult, C., and Ray, P.F. (2018). Creation of knock out and knock in mice by CRISPR/Cas9 to validate candidate genes for human male infertility, interest, difficulties and feasibility. Mol. Cell. Endocrinol. *468*, 70–80.

23. Sapiro, R., Tarantino, L.M., Velazquez, F., Kiriakidou, M., Hecht, N.B., Bucan, M., and Strauss, J.F. (2000). Sperm antigen 6 is the murine homologue of the Chlamydomonas reinhardtii central apparatus protein encoded by the PF16 locus. Biol. Reprod. *62*, 511–518.

24. Teves, M.E., Nagarkatti-Gude, D.R., Zhang, Z., and Strauss, J.F. (2016). Mammalian axoneme central pair complex proteins: Broader roles revealed by gene knockout phenotypes. Cytoskelet. Hoboken NJ *73*, 3–22.

25. Peifer, M., Berg, S., and Reynolds, A.B. (1994). A repeating amino acid motif shared by proteins with diverse cellular roles. Cell *76*, 789–791.

26. Coates, J.C. (2003). Armadillo repeat proteins: beyond the animal kingdom. Trends Cell Biol. *13*, 463–471.

27. Tewari, R., Bailes, E., Bunting, K.A., and Coates, J.C. (2010). Armadillo-repeat protein functions: questions for little creatures. Trends Cell Biol. *20*, 470–481.

28. Sapiro, R., Kostetskii, I., Olds-Clarke, P., Gerton, G.L., Radice, G.L., and Strauss III, J.F. (2002). Male infertility, impaired sperm motility, and hydrocephalus in mice deficient in sperm-associated antigen 6. Mol. Cell. Biol. *22*, 6298–6305.

29. Onoufriadis, A., Shoemark, A., Munye, M.M., James, C.T., Schmidts, M., Patel, M., Rosser, E.M., Bacchelli, C., Beales, P.L., Scambler, P.J., et al. (2014). 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. *51*, 61–67.

30. Hjeij, R., Lindstrand, A., Francis, R., Zariwala, M.A., Liu, X., Li, Y., Damerla, R., Dougherty, G.W., Abouhamed, M., Olbrich, H., et al. (2013). ARMC4 mutations cause primary ciliary dyskinesia with randomization of left/right body asymmetry. Am. J. Hum. Genet. *93*, 357–367.

31. Coutton, C., Escoffier, J., Martinez, G., Arnoult, C., and Ray, P.F. (2015). Teratozoospermia: spotlight on the main genetic actors in the human. Hum. Reprod. Update *21*, 455–485.

32. Pausch, H., Venhoranta, H., Wurmser, C., Hakala, K., Iso-Touru, T., Sironen, A., Vingborg, R.K., Lohi, H., Söderquist, L., Fries, R., et al. (2016). A frameshift mutation in ARMC3 is associated with a tail stump sperm defect in Swedish Red (Bos taurus) cattle. BMC Genet. *17*, 49.

33. Zhang, Z., Sapiro, R., Kapfhamer, D., Bucan, M., Bray, J., Chennathukuzhi, V., McNamara, P., Curtis, A., Zhang, M., Blanchette-Mackie, E.J., et al. (2002). A sperm-associated WD repeat protein orthologous to Chlamydomonas PF20 associates with Spag6, the mammalian orthologue of Chlamydomonas PF16. Mol. Cell. Biol. *22*, 7993–8004.

34. Sapiro, R., Tarantino, L.M., Velazquez, F., Kiriakidou, M., Hecht, N.B., Bucan, M., and Strauss, J.F. (2000). Sperm antigen 6 is the murine homologue of the Chlamydomonas reinhardtii central apparatus protein encoded by the PF16 locus. Biol. Reprod. *62*, 511–518.

35. Sapiro, R., Kostetskii, I., Olds-Clarke, P., Gerton, G.L., Radice, G.L., and Strauss III, J.F. (2002). Male infertility, impaired sperm motility, and hydrocephalus in mice deficient in sperm-associated antigen 6. Mol. Cell. Biol. *22*, 6298–6305.

36. Auguste, Y., Delague, V., Desvignes, J.-P., Longepied, G., Gnisci, A., Besnier, P., Levy, N., Beroud, C., Megarbane, A., Metzler-Guillemain, C., et al. (2018). Loss of Calmodulin- and Radial-Spoke-Associated Complex Protein CFAP251 Leads to Immotile Spermatozoa Lacking Mitochondria and Infertility in Men. Am. J. Hum. Genet. *103*, 413–420.

37. Matzuk, M.M., and Lamb, D.J. (2008). The biology of infertility: research advances and clinical challenges. Nat. Med. *14*, 1197–1213.

38. Meienberg, J., Bruggmann, R., Oexle, K., and Matyas, G. (2016). Clinical sequencing: is WGS the better WES? Hum. Genet. *135*, 359–362.

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

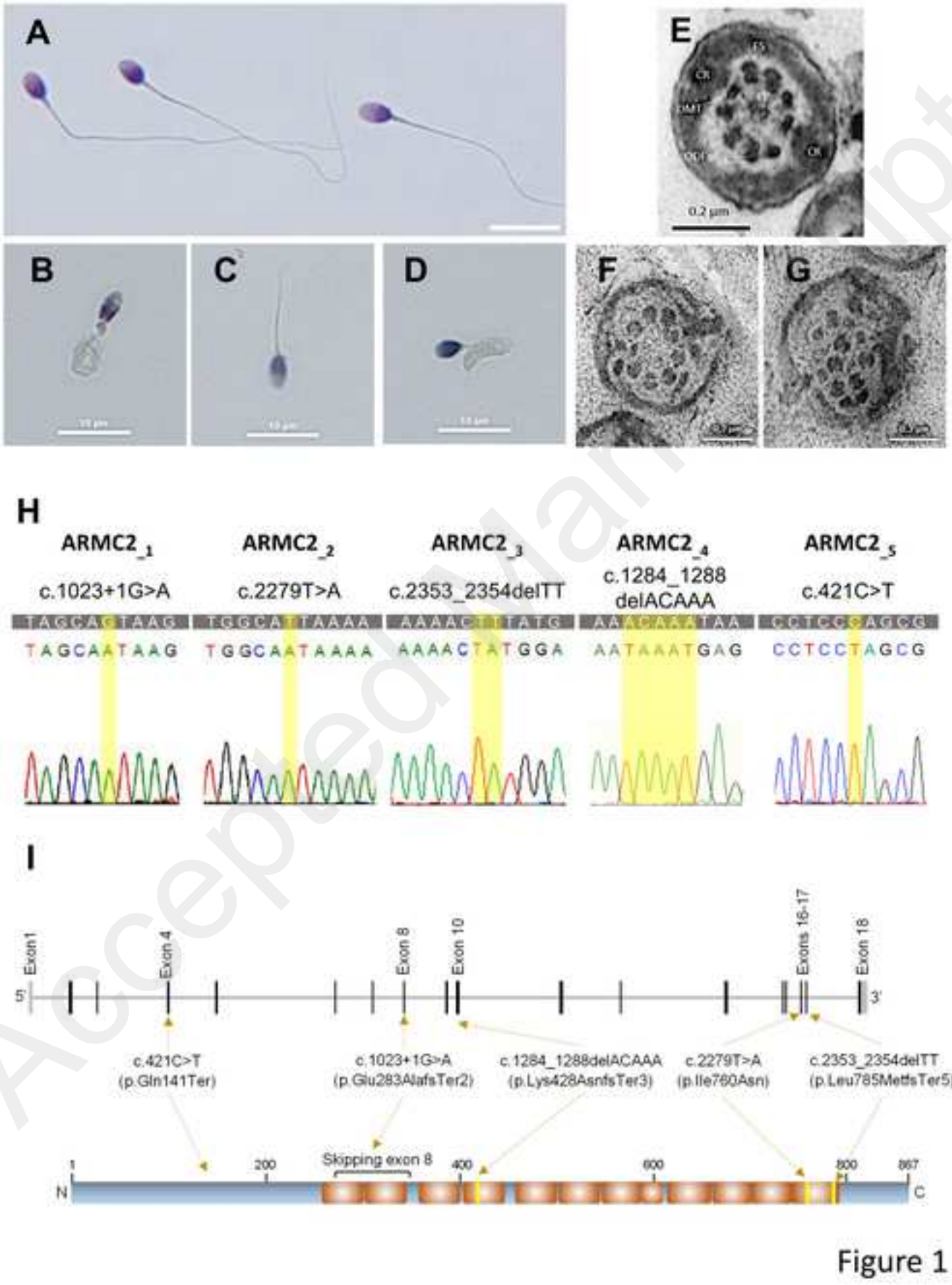

Figure 1

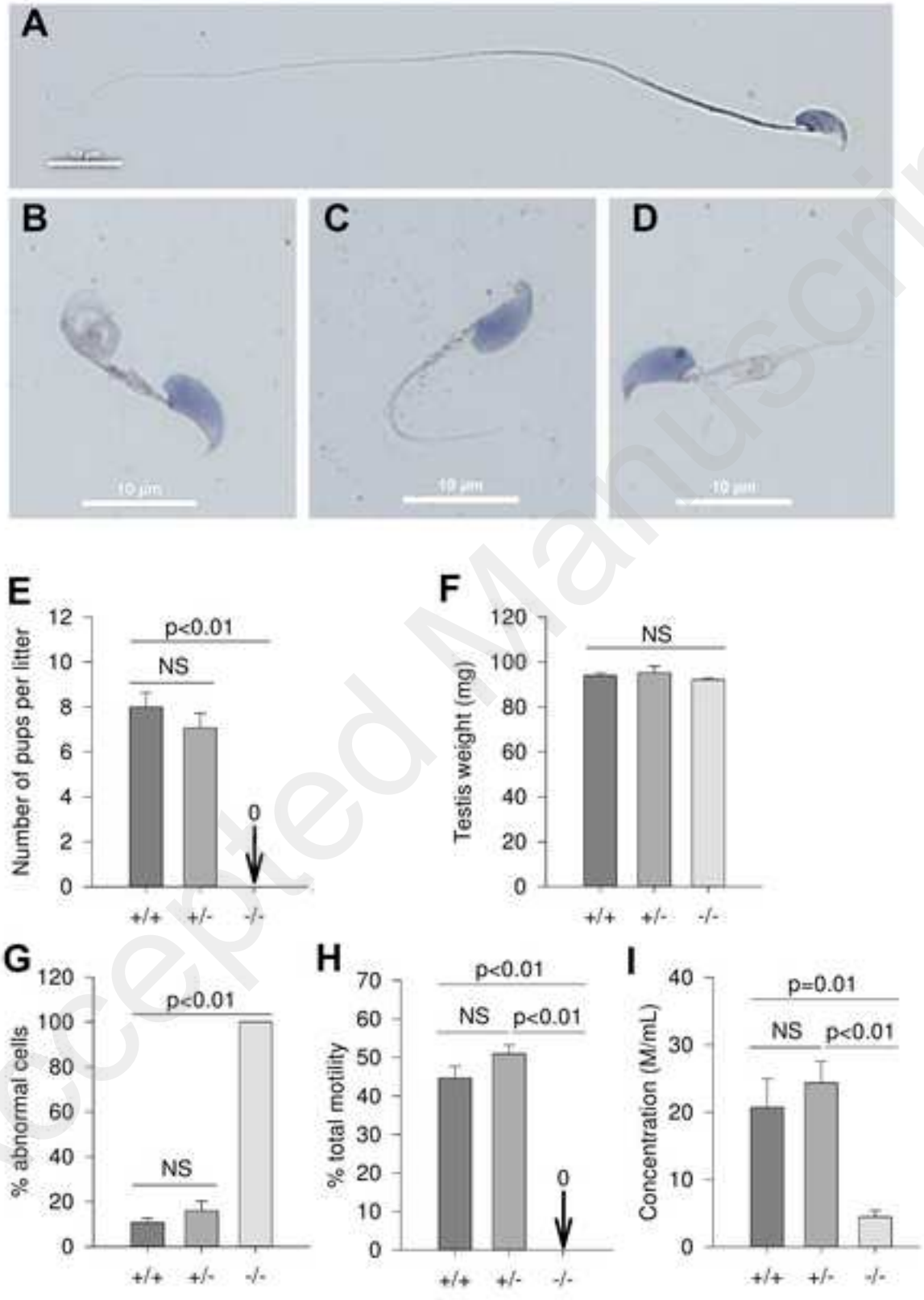

Figure 2

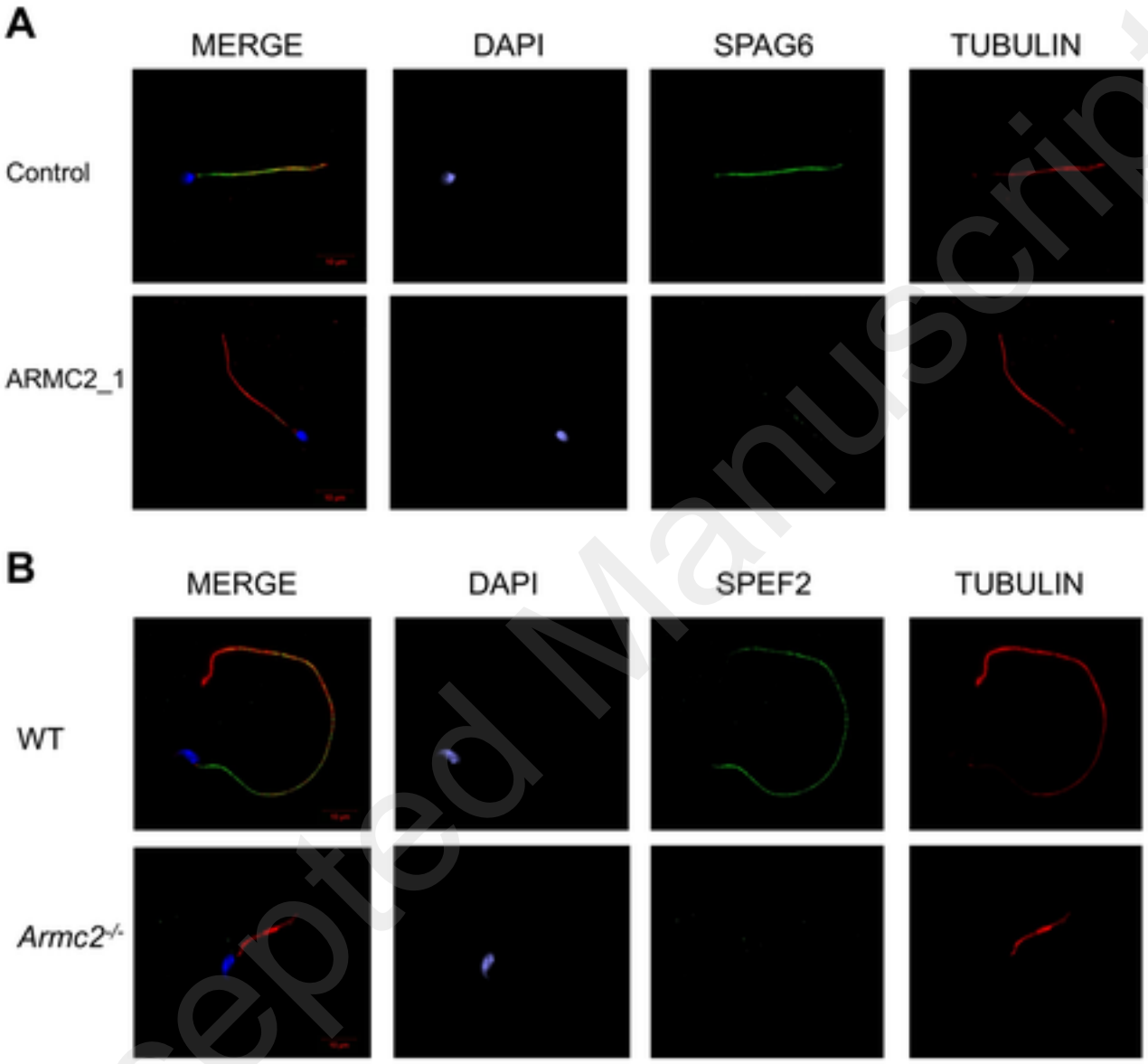

Figure 3