Mutations in CFAP43 and CFAP44 cause male infertility and flagellum defects in *Trypanosoma* and human

Charles Coutton et al. 

Spermatogenesis defects concern millions of men worldwide, yet the vast majority remains undiagnosed. Here we report men with primary infertility due to multiple morphological abnormalities of the sperm flagella with severe disorganization of the sperm axoneme, a microtubule-based structure highly conserved throughout evolution. Whole-exome sequencing was performed on 78 patients allowing the identification of 22 men with bi-allelic mutations in DNAH1 (n = 6), CFAP43 (n = 10), and CFAP44 (n = 6). CRISPR/Cas9 created homozygous CFAP43/44 male mice that were infertile and presented severe flagellar defects confirming the human genetic results. Immunoelectron and stimulated-emission-depletion microscopy performed on CFAP43 and CFAP44 orthologs in *Trypanosoma brucei* evidenced that both proteins are located between the doublet microtubules 5 and 6 and the paraflagellar rod. Overall, we demonstrate that CFAP43 and CFAP44 have a similar structure with a unique axonemal localization and are necessary to produce functional flagella in species ranging from *Trypanosoma* to human.

Correspondence and requests for materials should be addressed to P.F.R. (email: pray@chu-grenoble.fr)

*A full list of authors and their affiliations appears at the end of the paper.*
Medical treatment of infertility has rapidly evolved over the past four decades, but much remains to be accomplished. Despite recent success in identifying infertility genes, most genetic causes of male infertility are currently uncharacterized and additional efforts should be pursued to better characterize male infertility. We demonstrated previously that mutations in the DNAH1 gene are responsible for multiple morphological abnormalities of the flagella (MAF), an infertility phenotype characterized by severe asthenozoospermia due to a combination of flagellar defects including short, curled, abnormal width, rolled, or absent flagella. DNAH1 encodes an axonemal inner dynein arm heavy chain, the lack of which leads to a strong disorganization of the axoneme. DNAH1 mutations were identified in approximately one-third of the studied patients, indicating that MMAF is genetically heterogeneous and that other genes are likely to be involved in this syndrome.

In the present study, we analyzed 78 MMAF patients using whole-exome sequencing (WES) and showed that in addition to mutations in DNAH1, mutations in CFAP43 and CFAP44, two genes encoding for WD repeat domains (WDR) containing proteins, are responsible for MMAF syndrome and account for 20.5% of this cohort. Most importantly, we investigated the role of these novel genes by performing gene invalidation and silencing in two evolutionary distant models, Trypanosoma and mouse, yet sharing an extremely conserved flagellar structure. Using this original approach, we demonstrate the importance of WDR proteins for axonemal structure of flagella and male fertility in humans.

Results
Identification of CFAP43 and CFAP44 mutations in MMAF patients. In the present study, we analyzed a cohort of 78 individuals presenting with a MMAF phenotype defined by the presence in the ejaculate of immotile spermatozoa with several abnormalities of the sperm flagellum including short, coiled, absent, and flagella of irregular caliber (Fig. 1a). A majority of patients originated from North Africa, 46 were recruited in Tunisia, 10 in Iran, and 22 in France. The average semen parameters of all 78 MMAF patients included in the cohort are

![Fig. 1 Morphology of spermatozoa from the patients P43-8 and P44-3 and position of CFAP43 and CFAP44 mutations in the intron-exon structure and in the protein.](image-url)

- All spermatozoa have a much shorter flagellum than those of controls. Additional features of MMAF spermatozoa are absent (*), thick (&), and rolled flagella (@).
- Mutations identified in the CFAP43 gene.
- Mutations identified in the CFAP44 gene. Blue squares stand for WD repeat domains and green squares for coiled-coiled domains as described by Uniprot server. Mutations are annotated in accordance to the HGVS’s recommendations.
Table 1 Average semen parameters in the different genotype groups for the 78 MMAF subjects

| semen parameters                  | MMAF CFA43 patients, n = 10 | MMAF CFA44 patients, n = 6 | MMAF DNAH1 patients, n = 6 | MMAF all mutations, n = 22 | MMAF with unknown causes, n = 56 | Overall MMAF, n = 78 |
|-----------------------------------|-----------------------------|----------------------------|---------------------------|-----------------------------|---------------------------------|----------------------|
| Mean age (years)                  | 37.7 ± 9.6 (n = 9)          | 41.3 ± 4.3 (n = 6)          | 41.5 ± 4.4 (n = 6)        | 39.7 ± 7 (n = 21)           | 42.3 ± 7.9 (n = 56)            | 41.6 ± 7.7 (n = 77) |
| Sperm volume (ml)                 | 3.5 ± 1.4 (n = 8)           | 3.2 ± 0.87 (n = 6)         | 3.5 ± 1.2 (n = 6)         | 3.4 ± 1.2 (n = 20)          | 3.5 ± 1.5 (n = 55)             | 3.5 ± 1.4 (n = 75)  |
| Sperm concentration (10^6/ml)     | 27.2 ± 23.4 (n = 8)         | 7.9 ± 8.4 (n = 6)          | 22.9 ± 15.2 (n = 6)       | 20.1 ± 18.8 (n = 20)        | 27.6 ± 35.7 (n = 55)           | 25.6 ± 32.1 (n = 75) |
| Motility (a + b + c) / h           | 0 ± 0 (n = 9)               | 0 ± 0 (n = 6)              | 2.6 ± 0.6 (n = 6)         | 0.7 ± 2.4* (n = 21)         | 5 ± 6.1 (n = 55)               | 3.9 ± 5.6 (n = 76)  |
| Vitality                          | 55.5 ± 24 (n = 8)           | 43.3 ± 22.6 (n = 6)        | 51.2 ± 23.1 (n = 5)       | 50.5 ± 22.7 (n = 19)        | 53.4 ± 20 (n = 50)             | 57.2 ± 20 (n = 72)  |
| Normal spermatozoa                | 1.25 ± 3.5 (n = 8)          | 0 ± 0 (n = 6)              | 0.5 ± 2.3 (n = 20)        | 0.5 ± 2.3 (n = 20)          | 21 ± 41 (n = 54)               | 1.6 ± 2.7 (n = 61)  |
| Absent flagella                   | 21.8 ± 17.6 (n = 6)         | 36.8 ± 43.5 (n = 4)        | 25.6 ± 15.9 (n = 5)       | 28.1 ± 14.4* (n = 15)       | 18.5 ± 15.3 (n = 51)           | 20.7 ± 15.7 (n = 66) |
| Short flagella                    | 65.3 ± 31.7 (n = 8)         | 52.2 ± 27 (n = 6)          | 49.8 ± 24.3 (n = 5)       | 57.1 ± 27.9* (n = 19)       | 38.9 ± 25.7 (n = 53)           | 43.7 ± 27.3 (n = 72) |
| Coiled flagella                   | 8.2 ± 6 (n = 6)             | 9 ± 3.5 (n = 5)            | 2 ± 1.3 (n = 2)           | 10.4 ± 6.6 (n = 16)         | 13.5 ± 10 (n = 53)             | 12.8 ± 9.4 (n = 69)  |
| Bent flagella                     | 10.3 ± 6 (n = 3)            | 9 ± 3 (n = 3)              | 14 ± 0.9 (n = 2)          | 7.8 ± 5.8 (n = 3)           | 12 ± 9.9 (n = 20)              | 14 ± 8.4 (n = 26)   |
| Flagella of irregular caliber     | 20.2 ± 19.3 (n = 5)         | 28.4 ± 16.9 (n = 5)        | 35 ± 22.7 (n = 5)         | 27.9 ± 19.4 (n = 15)        | 32.8 ± 26.7 (n = 52)           | 31.7 ± 25.1 (n = 67) |
| Multiple anomalies index          | 2.3 ± 0.2 (n = 4)           | 3.4 ± 0.4 (n = 5)          | 0 ± 0 (n = 5)             | 0 ± 0 (n = 6)               | 15.5 ± 15.7 (n = 66)           | 0 ± 0 (n = 6)       |

Values are expressed in percent, unless specified otherwise.

We compared statistical differences between MMAF due to CFA43, CFA44, and DNAH1 mutations versus MMAF due to uncharacterized genetic cause.

CFA4 cilia and flagella associated protein; MMAF, multiple morphological abnormalities of the flagella.

*A significant difference P < 0.05.

described in Table 1. Nearly no spermatozoa with normal morphology could be observed in the ejaculate of MMAF individuals (1.6%); an average of 20.7 and 43.7% of spermatozoa had no flagella and short flagella, respectively, and 31.7% of the spermatozoa had flagella with an irregular caliber. As a result, total sperm motility was dramatically reduced to 3.9% (normal value >40%), which prevented natural conception for all individuals. Given the notion of consanguinity for most individuals from the cohort, we postulated that infertility was likely transmitted through recessive inheritance and probably often resulted from homozygous mutations. After exclusion of frequent variants and applying stringent filters, a limited list of homozygous variants was identified for each proband. First, we identified six patients (7.7%) with mutations in the DNAH1 gene (Table 2), previously identified as the main cause of MMAF phenotype3,4. We subsequently identified 10 subjects with variants in CFA43 (12.8%), eight of which had a homozygous loss-of-function variant and two with two likely deleterious variants (Table 2). In addition, six subjects (7.7%) had a homozygous loss-of-function variant in CFA44 (Table 2). These two CFA genes (for cilia and flagella associated protein) were reported in public expression databases to be strongly expressed in the testis and to be connected with cilia and flagella structure and/or functions5. Quantitative real-time reverse transcription PCR (RT-qPCR) experiments performed in human and mouse tissue panels confirmed that CFA43 and CFA44 mRNA in testis was predominant and very significantly higher than in the other tested tissues (Supplementary Fig. 1). Taking into account the high number of mutated patients and the specific expression pattern of the two genes, we focused on these two genes, which appeared as
The best candidates to explain the primary infertility observed for these individuals. All individuals with CFAP43 or CFAP44 mutations were unrelated; to our knowledge, none carried rare variants in genes previously reported to be associated with male infertility. CFAP43 and CFAP44 encode proteins that belong to the WDR protein family and are both composed of nine WD repeats.

CFAP43 (also known as WDR96, NM_025145) is localized on chromosome 10 and contains 38 exons encoding a predicted 1,665-amino-acid protein (Q8N67M). We identified 9 different pathogenic variants in CFAP43 in 10 unrelated individuals from the cohort. The splicing variant c.3541+2A>C was identified in two unrelated individuals and affects a consensus splice acceptor site of CFAP43 intron 27 (Table 2, Fig. 1b). This variant was absent in the 60,706 unrelated individuals aggregated in the ExAC database (http://exac.broadinstitute.org), consistent with a deleterious effect on protein function. Three other variants corresponding to stop-gain mutations identified in three subjects: c.2658G>T, c.2680C>T, and c.3175C>T in the exon 21. The last mutation, c.1890+1G>T, affects the consensus splice donor site. All these mutations are located in exon 12 and 23, respectively. The frameshift mutations, c.4767delT, with a single-nucleotide deletion, are located in exon 8 of the CFAP43 gene. Although prediction software classified this last variant as likely "benign," this mutation affects a conserved residue located in the N-terminal part of the protein within a WDR known to be important for protein/protein interactions. Moreover, this missense variant is found at a very low prevalence in the general population estimated to 7.45e-06 to 9.89e-05. Two others mutations were small frameshift indels not listed in ExAC: c.1240_1241delGT (found in two unrelated patients) and c.3882delA (Table 2, Fig. 1b). All these mutations generate premature stop codons and are predicted to produce no protein or truncated non-functional proteins. Apart from these obvious harmful mutations found in eight patients, we also identified two patients harboring likely deleterious mutations in CFAP43. The first patient harbors a homozygous splicing mutation c.2141+5G>A, not listed in ExAC which, according to the splice site prediction algorithm Human Splicing Finder (http://www.umd.be/HSF3), likely alters the consensus splice donor site of CFAP43 exon 16. Unfortunately, we could not obtain any additional biological material from this patient and could not verify the effect of this variant on mRNA expression. The second patient is a compound heterozygous carrying one single-nucleotide duplication, c.1320dupT, located in the exon 11 and an additional missense mutation, Val347Ala, located in exon 8 of the CFAP43 gene. Although prediction software classified this last variant as likely "benign," this mutation affects a conserved residue located in the N-terminal part of the protein within a WDR known to be important for protein/protein interactions. Moreover, this missense variant is found at a very low prevalence in the general population estimated to 7.45e-05 to 9.89e-05. Altogether the above-identified mutations clearly indicate that in humans, loss of function of CFAP43 is associated with asthenozoospermia and sperm flagellum defects.

CFAP44 (also known as WDR52, NM_001164496) is localized on chromosome 3 and contains 35 exons encoding a predicted 1,854-amino-acid protein (Q96MT7). Five different homozygous variants were identified in six unrelated patients (Table 2, Fig. 1c). Stop-gain mutations c.1387G>T and c.3175C>T are located in exons 12 and 23, respectively. The frameshift mutations, c.4767delT, with a single-nucleotide deletion, are located in exon 31, and the variant c.2818dupG, with a single-nucleotide duplication, is located in exon 21. The last mutation, c.1890+1C>T, affects exon 15 consensus splice donor site. All these mutations are predicted to generate a premature stop codon. None of these variants were listed in the ExAC database.

The presence of all variants was verified by bidirectional Sanger sequencing (Supplementary Fig. 2). The detailed sperm parameters of the patients carrying DNAH1, CFAP43, or CFAP44 mutations were compared with each other. There was no significant difference between the three groups (Table 1) as illustrated by the similar morphologies of sperm from patient P43-8 and P44-3, mutated for CFAP43 and CFAP44, respectively (Fig. 1a). Data from all these groups were therefore pooled and compared with data from the patients with no identified mutation. Patients with a mutation had a significantly higher rate of spermatozoa with short or absent flagella and presented a significantly lower motility rate compared to patients with no identified mutations. There was no difference in the other sperm parameters (Table 1).

**CFAP43 and CFAP44 mutations induce severe axonemal disorganization.** The internal skeleton of motile cilia and flagella, named the axoneme, is a highly evolutionarily conserved structure, which consists of nine doublets of microtubules (DMS) circularly arranged around a central pair complex (CPC) of microtubules (9+2 structure). All DMS are not identical and are organized in a highly reproducible fashion with substructures specific to certain doublets. The convention adopted in the numbering of the DMS was based on the relative position of each DMT with regard to the plane of the CPC. The beating of cilia and flagella is orchestrated by multiprotein-ATPase complexes, located on the peripheral doublets, which provide the sliding force for sperm motility. In addition, the sperm flagellum harbors specific peri-axonemal structures, which are not found in other motile cilia, a helical mitochondrial sheath (MS) in the midpiece, the fibrous sheath (FS) in the principal piece (PP), and outer dense fibers (ODFs) in the midpiece and the proximal part of the PP.

We studied the ultrastructure of sperm cells from mutated patients in CFAP43 and CFAP44 by transmission electron microscopy (TEM) (Fig. 2). Due to insufficient amount of sperm cells collected from most CFAP44 and CFAP43 patients, only one patient for each gene was studied by TEM. For each patient, we could observe longitudinal sections and >20 cross-sections presenting a sufficient quality to observe the ultrastructure of the axonemal components. Longitudinal sections showed severe axonemal and peri-axonemal defects affecting the ODF, the FS, and the MS, which appeared completely disorganized resulting in aborted flagella or their replacement by a cytoplasmic mass englobing unassembled axonemal components (Fig. 2). In CFAP43-mutated and CFAP44-mutated patients, approximately 95% of the cross-sections were abnormal and the main defect observed was the absence of the CPC (9+0 conformation) observed in 81.8, and 66.7% of the cases, respectively, compared to 0% in control fertile subject (Supplementary Table 1). In the residual fraction (~5%) of axoneme with normal (9+2) structure, all DMTs are not identical and are organized in a highly reproducible fashion with substructures specific to certain doublets. The convention adopted in the numbering of the DMTs was based on the relative position of each DMT with regard to the plane of the CPC. The beating of cilia and flagella is orchestrated by multiprotein-ATPase complexes, located on the peripheral doublets, which provide the sliding force for sperm motility. In addition, the sperm flagellum harbors specific peri-axonemal structures, which are not found in other motile cilia, a helical mitochondrial sheath (MS) in the midpiece, the fibrous sheath (FS) in the principal piece (PP), and outer dense fibers (ODFs) in the midpiece and the proximal part of the PP.

To define the ultrastructural defects evidenced by TEM, we performed immunofluorescence (IF) experiments using antibodies targeting different axonemal proteins. In CFAP43-mutated and CFAP44-mutated patients, staining of SPAG6, a protein located in the CPC, was abnormal and atypical compared to control staining (Fig. 3a–c). SPAG6 staining was absent in CFAP43 patient sperm cells, whereas in CFAP44 patients' sperm cells, SPAG6 immunostaining was present but displayed an abnormal and diffuse pattern concentrated in the midpiece of the spermatozoa, quickly diminishing along the flagellum (Fig. 3d–i).
Additionally, staining of the radial spoke head protein RSPH1 presented significantly different patterns compared to controls. In CFAP43-mutated patients, the RSPH1 staining was completely abnormal with a marked diffuse staining, concentrated in the midpiece, whereas the tubulin staining remained restricted to the axoneme (Fig. 3m–o). In CFAP44-mutated patients’ sperm cells, RSPH1 staining was significantly reduced (Fig. 3p–r). For CFAP43-mutated and CFAP44-mutated patients, immunostaining for AKAP4, DNALI1, DNAI2, and GAS8 were all comparable with controls (Fig. 3j–l), suggesting that FS, outer dynein arms, inner dynein arms, and the nexin-dynein regulatory complex, respectively, were not directly affected by mutations in CFAP43 or CFAP44 (Supplementary Table 2—Human). Unfortunately, we could not obtain any specific antibodies against CFAP43 and CFAP44 that would allow the localization of these proteins in mouse and human flagella.

Analysis of Cfap43 and Cfap44 knockout mice. We assessed the impact of Cfap43 and Cfap44 absence on mouse spermatogenesis by generating knockout (KO) animals using the CRISPR-Cas9 technology. We obtained four independent mutational events for Cfap43 and three for Cfap44; all were insertions/deletions of a few nucleotides inducing a translational frameshift expected to lead to the complete absence of the protein or the production of a truncated protein. Because all KO strains for Cfap43 presented the same reproductive phenotype (identical sperm morphology associated with complete infertility), we randomly chose one and restricted our study to a strain with a 4 bp deletion in exon 21 (delAAGG). Similarly, the Cfap44 lines presented the same reproductive phenotype, and we focused on a strain with a 7 bp insertion in exon 3 (InsTCAGATA). RT-PCR was performed on testis RNA from Cfap43−/− and Cfap44−/− mice, which confirmed the production of abnormal transcripts in both mutants leading to a premature stop codon (Supplementary Fig. 3).

The reproductive phenotype was studied for both KO mice models. Homozygous KO females were fully fertile and gave litters of normal size (7.8 ± 1.8 and 7.3 ± 3.5 versus 6.7 ± 0.5 pups/litter (mean ± SD, n = 7) for Cfap43−/−, Cfap44−/−, and wild-type (WT), respectively, contrary to homozygous KO males, which exhibited complete infertility when mated with WT females (Fig. 4a). Epididymal sperm concentrations were 11.4 ± 0.9, 11.2 ± 0.4, and 14.4 ± 3.1×10^6 sperm/ml (mean ± SD, n = 3), for Cfap43−/−, Cfap44−/−, and WT, respectively, and fall

---

Fig. 2 Transmission electron microscopy analysis of sperm cells from CFAP43 and CFAP44 patients shows a severe axonemal disorganization. (Left panel) Longitudinal sections (scale bar = 500 nm) and cross-sections (scale bars, 100 nm) of sperm flagellum from control. (Central panel) Longitudinal sections (scale bars, 500 nm) and cross-sections (scale bars, 100 nm) of sperm flagellum from CFAP43-mutated patient (P43-8). We can notice a short short tail corresponding to a cytoplasmic mass containing the different components of the flagellum, all unorganized. In CFAP43 upper cross-section, the CPC is not aligned with DMTs 3 and 8 (red line) and is rotated by 90°. We can observe the absence of central pair of microtubules in other cross-sections. (Right panel) Ultrastructure of CFAP44-mutated sperm (P44-3) longitudinal sections (scale bar = 500 nm) show similar ultrastructure of short tail (cytoplasmic mass). In CFAP44 upper cross-section, the central pair is disassembled and displaced (red arrow). We can observe the absence of central pair of microtubules in other cross-sections. Scales bars for cross-sections = 100 nm.
within the normal values for the mouse. We next studied sperm morphology and observed that for Cfap43−/− males, 100% of sperm displayed a typical human MMAF phenotype with short, thick, and coiled flagella (Fig. 4b); although slightly deformed, sperm head exhibited an overall normal hooked form (Fig. 4e). In contrast, sperm from Cfap44−/− males had normal flagellum length, but most of them showed abnormal forms and irregular caliber of the midpiece (Fig. 4c–f). In sperm from Cfap44−/− mice, staining of Mpc1l, a sperm mitochondrial protein, was discontinuous or punctiform, strongly suggesting that the distribution of the mitochondria surrounding the axoneme was irregular, consistent with the observed midpiece abnormalities (Fig. 4b–m). For both Cfap43−/− and Cfap44−/− sperm, the observed structural flagellum defects were associated with severe motility deficiencies (Fig. 4g) with sperm from Cfap43−/− males showing complete immotility whereas those from Cfap44−/− sometime presented some small vibrations (Supplementary Movie 1), in contrast to the motile sperm from Cfap44+/- heterozygotes (Supplementary Movie 2). Interestingly, we observed that in heterozygous animals, Cfap43+/- mice had an impaired progressive sperm motility and Cfap44+/- mice had an increased proportion or morphologically abnormal spermatozoa (Fig. 4g).

To better characterize the molecular defects induced by the absence of Cfap43 and Cfap44 in mouse sperm, we studied by IF the presence and localization of several proteins belonging to different substructures of the axoneme in sperm from both KO mouse models. The presence of the following proteins was investigated: Dnah5 and Dnah1 as markers of dynein arms, Rsph1 and Rsph4a as markers of radial spokes, Gas8 as a marker of nexin links, and Spef2 as a marker of the central pair (Supplementary Table 2—Mouse). In Cfap43−/−-mutated animals, Spef2, Rsph1, and Rsph4a were clearly missing (Fig. 5d–f, j–l) as compared to controls (Fig. 5a–c, g–i, Supplementary Table 2, and Supplementary Fig. 4), indicating that the center of the axoneme, including the head of the radial spoke interacting with the central pair, was absent. In contrast, no obvious defects were observed in sperm from Cfap44−/−, suggesting that all substructures were present (Supplementary Table 2—Mouse).

The impact of the absence of these proteins on the flagellum ultrastructure was analyzed on sperm from Cfap43−/− and Cfap44−/− by TEM. In the midpiece of WT mouse flagellum, the 9+2 axoneme is surrounded by nine ODFs and the mitochondria sheath. The ODFs 1, 5, 6, and 9 are parallel and aligned with the central pair, was absent. In contrast, no obvious defects were observed in sperm from Cfap44−/−, suggesting that all substructures were present (Supplementary Table 2—Mouse).

Fig. 3 Immunofluorescence staining in Cfap43 and Cfap44 patients reveals an abnormal axonemal organization. a–c sperm cells from a fertile control stained with anti SPAG6 (green), a protein located in the CPC, and anti-acetylated tubulin (red) antibodies. DNA was counterstained with Hoechst 33342. c Corresponds to a and b overlay and shows that in control sperm, SPAG6 and tubulin staining superimpose. Scale bars = 5 μm. d–f SPAG6 staining is absent in sperm from the patient P44-5 homozygous for the c.3175C>T variant in Cfap43. d–i Similar IF experiments performed with sperm cells from the patient P44-2 homozygous for the c.3175C>T variant in Cfap44. Scale bar = 5 μm. Contrary to the control, the SPAG6 immunostaining (green) is abnormal with a diffuse pattern concentrated in the midpiece of the spermatozoa and is not detectable in the principle piece. j–l Sperm cells from a fertile control stained with anti RSPH1 (green), a protein of the radial spoke head, and anti-acetylated tubulin (red) antibodies. DNA was counterstained with Hoechst 33342. l corresponds to j and k overlay and shows that RSPH1 and tubulin staining superimpose in control sperm. Scale bar = 5 μm. m–o In sperm from the patient P44-2, RSPH1 staining (green) is significantly different from control (m) with a marked diffuse staining. p–r In sperm from the patient P45-2 the intensity of the RSPH1 staining is strongly reduced.
distribution of the DMTs with CPC misorientation (Fig. 6d–f). 

*Cfap44*−/− sperm also exhibited important defects of the ODFs and of the FS (Supplementary Fig. 6b). In the midpiece, the number of dense fibers was significantly higher and their localization and orientation were defective (Fig. 6b, c). In the PP, ODFs 3 and 8 were abnormally retained, preventing normal anchoring of the FS’s stalks of on DMTs 3 and 8 (Fig. 6g–i). The longitudinal columns also presented several defects: they were misaligned and not facing the 3-central-8 complex, leading to notable asymmetry of the structure (Fig. 6k–n), which was

![Graph showing number of pups per litter and distribution of abnormal sperm](image)

![Overlay images showing Mpc11 and Tubulin](image)
reinforced in numerous sperm by the presence of a third longitudinal column (Fig. 6m, n). Finally, the defects detected in Cfap44<sup>−/−</sup> MS by optic and fluorescence microscopy were specified by TEM analyses of sperm midpieces, evidencing uneven distribution and fragmentation of the mitochondria (Supplementary Fig. 8).

**TbCFAP43 and TbCFAP44 are necessary axonemal proteins.** To better characterize CFAP43 and CFAP44 localization and function and to overcome the absence of reliable antibodies in human and murine cells, we decided to use *Trypanosoma brucei* (T. brucei) as an additional model organism because this flagellated protozoan has largely contributed to elucidating the molecular pathology of human ciliopathies<sup>15</sup>. The *T. brucei* axoneme is similar to that of mammalian flagella, but its peri-axonemal structure, although comparable in function, is different. Instead of the FS and the ODFs found in mammals’ flagella, it contains the paraflagellar rod (PFR), a complex structure connecting with the axonemal doublets 4–7<sup>16,17</sup> which plays a role in flagellum motility<sup>18,19</sup> and serves as a platform for metabolic and signaling enzymes<sup>20–22</sup>. In *T. brucei* the DMTs are numbered from the doublet opposite the PFR (no. 1), which is normally parallel to the central singlets. The doublets are then numbered clockwise, doublets 5 and 6 facing the PFR and doublets 4 and 7 being positioned on each side of the PFR (Supplementary Fig. 5b for DMT annotation and relative PFR localization).

BLASTp analysis on *T. brucei* genome database<sup>23</sup> using human CFAP43 and CFAP44 sequences identified *T. brucei* orthologs Tb927.7.3560 (named TbCFAP43 in this study), and Tb927.4.5380 (named TbCFAP43 in this study), respectively. The sequence identity, calculated with clustal omega<sup>24</sup>, between human CFAP43 and CFAP44 orthologs Tb927.4.5380 (named TbCFAP43 in this study), respectively. The sequence identity, calculated with clustal omega<sup>24</sup>, between human CFAP43 and CFAP44 orthologs Tb927.4.5380 (named TbCFAP43 in this study), respectively. The sequence identity, calculated with clustal omega<sup>24</sup>, between human CFAP43 and CFAP44 orthologs Tb927.4.5380 (named TbCFAP43 in this study), respectively. The sequence identity, calculated with clustal omega<sup>24</sup>, between human CFAP43 and CFAP44 orthologs Tb927.4.5380 (named TbCFAP43 in this study), respectively. The sequence identity, calculated with clustal omega<sup>24</sup>, between human CFAP43 and CFAP44 orthologs Tb927.4.5380 (named TbCFAP43 in this study), respectively. The sequence identity, calculated with clustal omega<sup>24</sup>, between human CFAP43 and CFAP44 orthologs Tb927.4.5380 (named TbCFAP43 in this study), respectively.

To further investigate this localization we used the 10TY1 tag which produces a strong IF staining<sup>22</sup> necessary for stimulated-emission-depletion (STED) confocal microscopy<sup>34</sup> and created two double-labeled cell lines 10TY1TbCFAP44/TbCFAP43<sup>myc</sup> and 10TY1TbCFAP43/TbCFAP44<sup>myc</sup>. Triple labeling STED confocal microscopy of 10TY1-tagged proteins, tubulin, and PFR2 on permeabilized whole cells demonstrated that 10TY1TbCFAP44 and 10TY1TbCFAP43 are closely associated to the axoneme along the PFR (Fig. 7a–c, respectively); similar results were obtained on detergent-extracted cytoskeletons. To confirm this remarkable localization, TEM immunogold labeling was performed using differently tagged proteins. We used the myc tag, a small tag of 10 amino acids which is known to have minimum effect on localization. Concordant with the STED microscopy results, all immunogold beads were located only along the axoneme and facing the PFR, therefore indicating that TbCFAP44 and TbCFAP43 are in direct contact only with the doublets facing the PFR, which correspond to DMTs 5 and 6<sup>35</sup> (Fig. 7g–j). The intervals between two adjacent gold particles were measured and the minimal distance between two particles was around 25 nm for both proteins (Fig. 7g, h, inset). Finally, 10TY1-tagged proteins presented an identical subcellular localization in TEM (Supplementary Fig. 11).

To evaluate the impact of the absence of TbCFAP44 and TbCFAP43 on the structure of the trypanosome flagellum and its beating, we knocked down protein expression by inducible RNA interference (RNAi), either in the parental cell line (cell lines TbCFAP44<sup>RNAi</sup>, TbCFAP43<sup>RNAi</sup>) or in the cell lines expressing the myc-tagged proteins (cell lines TbCFAP44<sup>myc</sup><sup>RNAi</sup>, TbCFAP43<sup>myc</sup><sup>RNAi</sup>) (Fig. 8). Efficiency and specificity of RNAi knockdown of TbCFAP44 and TbCFAP43 were validated by quantitative real-time PCR (qPCR) and Western blot (WB) (see Supplementary Fig. 1).
RT-PCR (Supplementary Fig. 12) and by IF, showing a clear decrease of myc labeling in the new flagellum of TbCFAP44myc RNAi-induced and TbCFAP43myc RNAi-induced cells (Fig. 8e–l, respectively). Cell proliferation was assessed in parental, non-induced, and induced TbCFAP44 RNAi and TbCFAP43 RNAi cells (and in TbCFAP44myc and TbCFAP43myc cell lines, showing the same results). In all cases, both TbCFAP44 RNAi-induced and TbCFAP43 RNAi-induced cells showed abnormal beating (Supplementary Movies 2–5) and stopped proliferating after 24 h and eventually died (Fig. 8m). These growth defects were accompanied by a defect in cytokinesis producing multi-flagellated cells (Supplementary Fig. 13), a phenotype previously described when proteins directly or indirectly involved in flagellar motility are knocked down in BSF 36. We observed that the absence of TbCFAP44myc and TbCFAP43myc did not impact flagella length (Fig. 8n). Overall, these results indicate that TbCFAP44 and TbCFAP43 are necessary for flagellar function and thus cell proliferation in *T. brucei* BSF.

To assess if the observed cellular phenotypes are caused by each protein independently or if the absence of one induces the absence of the other, we took advantage of the double-labeled cell lines 10TY1 TbCFAP44/TbCFAP43myc and 10TY1 TbCFAP43/TbCFAP44myc. Inactivation of TbCFAP43 by RNAi did not alter the amount and localization of TbCFAP44; conversely, inactivation of TbCFAP44 had no impact on TbCFAP43 protein (Supplementary Fig. 14).

The impact of TbCFAP44 and TbCFAP43 knockdown was also investigated by TEM. In control longitudinal sections of parental cells, the flagellum exits the cell through the flagellar pocket (Fig. 8o), and, in cross-section, the canonical ultrastructure of the
**Fig. 6** Electron microscopy cross-sections of sperm from *Cfap44*Δ/Δ males reveal multiple structural axonemal defects. 

**a** Cross-section of the midpiece of a WT sperm, showing the arrangement of the ODFs around the axoneme. 

**b** Presence of extra ODFs in midpieces sections of sperm from *Cfap44*Δ/Δ males. The orientation of ODFs is also defective, leading to an increase of the midpiece diameter. 

**c** Graph showing the increased number of ODF in the mutant. Data represent mean ± SD; the statistical difference was assessed with t test, *P* value as indicated. Twenty-one cross-sections in the midpiece region were analyzed for WT and 21 for *Cfap44*Δ/Δ males. 

**d** Four DMTs (4–7) were missing. The central pair is shifted at the periphery (red arrowhead). 

**e** DMTs 5–7 were missing. Note the presence of a third longitudinal column (LC). 

**f** Irregular distribution of the DMTs associated with a rotation of the central pair (straight red line). Scale bars = 196 nm. 

**g** In WT sperm, the fibrous sheath is linked to the 3-central-8 complex by stalks emerging from the longitudinal columns (white arrowheads). 

**h** The presence of extra ODFs facing DMTs 3 and 8 (red arrowheads) prevents a normal anchoring of the fibrous sheath's stalks on DMTs 3 and 8. Scale bars = 270 nm. 

**j–o** LCs are not aligned with 3-8 CPC axis. In contrast to WT, where the 3-central-8 complex is aligned with LC to form almost a straight line (j, red line), LC are misaligned in sperm from *Cfap44*Δ/Δ males, leading to notable asymmetry of the structure (k, m, red lines). 

**l** The presence of a third LC increases asymmetry. Scale bars = 184 nm. 

**o** Bar graph showing the % of defects observed in the principal piece as described in g–n. One hundred cross-sections in the principal piece region were analyzed for WT and 50 for *Cfap44*Δ/Δ.
axoneme, composed of nine DMTs and the CPC, is observed (Fig. 8r). In control cell line only a small proportion (0.6%) of axonemes ($n = 172$) were abnormal. In contrast, TbCFAP44-RNAi-induced and TbCFAP43-RNAi-induced cells were abnormal and displayed more than two flagella in one abnormally enlarged flagellar pocket (Fig. 8p, q). In cross-section, flagella exhibited numerous abnormal axonemes with $90^\circ$ rotated CPC, a defect never observed in control trypanosome cell lines (Fig. 8p) and mainly displaced CPC and DMTs (Fig. 8s, t). Overall 27.3% of TbCFAP44 ($n = 139$) and 27.7% of the TbCFAP43-RNAi ($n = 150$) showed axonemal defects. We could not exclude some minor modifications in PFR; however, the disruption of the axoneme was so extreme after TbCFAP44 or TbCFAP44-RNAi knockdown that it was difficult to assess them. Altogether, these results confirm the essential role of TbCFAP43 and TbCFAP44 in the organization of the axoneme ultrastructure.

Discussion

Our work illustrates the efficiency of the combination of WES with an original workflow for the validation of the candidate genes. We identified two new genes responsible for MMAF syndrome and several new DNAH1 mutations, reinforcing the importance of this gene in MMAF syndrome. The prevalence of DNAH1 mutations was lower than what was reported previously, maybe due to a wider geographic recruitment of patients in the current study. Altogether, DNAH1, CFAP43, and CFAP44 mutations were identified in 28.2% of the analyzed subjects ($n = 78$) originating from North Africa, Europe, and the Middle East. These results underline the importance of these three genes in MMAF syndrome and will permit to improve the diagnosis efficiency of infertile MMAF patients. To investigate the possibility of a potential genotype–phenotype correlation, we examined the semen parameters of patients carrying mutations in DNAH1, CFAP43, and CFAP44 and MMAF patients with yet...
unidentified mutations. There were not differences between the first three groups except for a higher rate of spermatozoa with short and absent flagella and a lower motility rate evidenced in the mutated subjects compared to patients with no identified mutation, highlighting the severity of mutations in these three genes. The identified mutations are distributed throughout the whole CFAP43 and 44 genes indicating that the entirety of both proteins is necessary for preserving their functionality. Interestingly among CFAP43-mutated patients, only one (P43-9) carried a missense variant, with potentially a milder effect (Val347Ala). This patient has the lowest percentage of sperm with short flagella (22%) with 10% of morphologically normal spermatozoa (versus 0% for the other CFAP43-mutated subjects) and therefore presents a milder phenotype compared to the others with
homzygous truncating mutations. These data support the existence of a genotype–phenotype correlation as had been demonstrated for Dnah1-mutated subjects. We have previously demonstrated that MMAF patients with Dnah1 mutations had a good prognosis using ICSI, but it remains difficult to predict the success rate for the other MMAF genotypes. Additional correlation studies should now be performed to take into account the individual genotypes in the counseling of MMAF patients. Exome sequencing of MMAF patients permitted a diagnosis efficiency of 27% for the three validated genes. Subsequent analysis of exome data from the remaining 56 subjects permitted to identify additional candidate genes with bi-allelic mutations found in at least two individuals in a total of 23 subjects. The identified mutations and these additional genes are currently being investigated. If all these variants are confirmed to be deleterious, a diagnosis will be obtained for 45/78 patients (58%), confirming the interest of WES as a diagnostic tool for MMAF syndrome.

Identification of disease-causing genes has recently been catalyzed by high-throughput genome-wide sequencing technologies. New challenges are now arising, in particular, the validation of candidate genes in the investigated phenotypes. CRISPR/Cas9 is a very efficient and fast technique to create KO models, compatible with the increased number of mutations/gene found by high-throughput sequencing. To confirm the pathogenicity of the identified mutations, we produced two knockout mice lines using CRISPR/Cas9 technology. Cfp43+/− males were infertile and 100% of their spermatozoa were immotile and morphologically abnormal with short and coiled tail, a phenotype very similar to human MMAF (Fig. 4e). Cfp44+/− male mice were also infertile due to flagellar immotility, yet presenting subtler flagellar defects than those observed in patients with Cfp44 mutations (Fig. 4f). Such a phenotypic discrepancy between mouse and human is not uncommon. For example, Dnah1 KO mice display asthenozoospermia without morphological defects of the flagellum, whereas in humans Dnah1 truncating mutations induce a MMAF phenotype. Nevertheless, TEM analysis of Cfp44+/− sperm evidenced ultrastructural defects similar to those found in Cfp44 patients, confirming that Cfp44 is necessary for axonemal organization and function both in mouse and human. The reason why this axonemal disorganization does not impact the overall flagellum morphology and length in mice is not known. It is however highly unlikely that a functional truncated protein could persist and have a rescue effect in KO mice since they are predicted at best to produce a truncated peptide of 49 out of the 1,854 amino acids of the full-length protein (Supplementary Fig. 3). Overall, the tandem use of WES and CRISPR/Cas9 technology in mouse was a very efficient strategy to identify and validate the mutations responsible for infertility phenotypes.

The structure and organization of motile cilia or flagellum was mainly deciphered from models including the green alga Chlamydomonas reinhardtii, the protozoa T. brucei and Tetrahymena, sea urchins, zebrashift, Xenopus, and mouse. Such a wide selection of distant models is possible because motile cilia or flagella are built on a canonical 9+2 axoneme which forms a highly organized protein network remarkably conserved during evolution. Here we chose to use T. brucei as this model has two advantages, first, in contrast to Chlamydomonas and other models, it shares some specific structural axonemal characteristics with mammalian flagellum such as a full RS3 and a fixed orientation of the central pair during flagellum beating, and second, it allows forward and reverse genetics for an easier characterization of gene function. BlastP enabled us to identify Cfp43 and Cfp44 likely orthologs in T. brucei with good sequence identity. Interestingly, in-depth structural analyses showed that these proteins share some unique structural similarity, in particular, a common general motif with β-strand domains in the N-terminal region followed by α-helical domains in the C-terminal region. These results confirm the relatedness of the identified orthologs and suggest that Cfp43 and Cfp44 proteins may have a similar function as reported for others proteins with similar structures. Consistent with this, we showed that in T. brucei, inactivation of one protein by RNAi did not alter the amount and localization of the other protein, indicating that (i) the phenotype observed when invalidating each protein is independent of the other protein, (ii) the presence of one cannot compensate for the absence of the other, and (iii) if the two proteins interact with one another, they do not depend on each other for adequate positioning and docking.

In all studied models (i.e. humans, mouse, and T. brucei), the invalidation of each protein induced important axonemal disorganization. It is important to underline that the destabilization of the axoneme can be triggered by the absence of proteins located either in axonemal or peri-axonemal structures, making it difficult to assess the function from the observed defects. Only the localization at the ultrastructural level would provide a clue on protein function. For this purpose, we generated several T. brucei cell lines expressing endogenously tagged proteins, and studied their localization by STED and immuno-EM. Using two different tags, we showed that both proteins have a restricted location within the flagellum, closely associated with the part of the axoneme facing the PFR (Fig. 7) and most probably corresponding to DMTs 5 and 6. This localization rules out a potential role of both proteins in the intraflagellar transport.

Fig. 8 T. brucei Cfp44 and Cfp43 are necessary for cell survival and proper axonemal organization. a-d Immunofluorescence on detergent-extracted cells of parental T. brucei stained with anti-PFR (red) and anti-myc (green) antibodies. e-h Immunofluorescence on detergent-extracted cells expressing TbCfp44myc and RNAi induced (24 h) for TbCfp44 and TbCfp44myc RNAi. Cells showed no or weak myc labeling (green) on the new flagellum (NF), while the old flagellum (OF) remained labeled. PFR is labeled in red. Note: cells with a maximum of two flagella were chosen for clear imaging. i-l Similar RNAi experiments as performed in e-h for TbCfp43 RNAi in cells expressing TbCfp43myc. Cells showed no or weak myc labeling (green) on the new flagellum (NF), while the old flagellum (OF) remained labeled. PFR is labeled in red. Nuclei and kinetoplasts (mitochondrial genome) are labeled with DAPI (blue). Scale bars=5 μm. Insets are enlargements of images of flagella taken from the main panels and display areas indicated by white arrows (scale bar=1 μm). m Growth curves for parental cells, and TbCfp44myc and TbCfp43myc cells, non-induced or induced with tetracycline. The graph represents the cumulative number of cells per ml. Error bars represent the standard error from three independent experiments. n Flagellum length measurement of the new flagellum in cells bearing two flagella from parental cells, and from cells expressing TbCfp44myc and TbCfp44myc induced (24 h) for TbCfp44 RNAi and TbCfp43 RNAi, respectively. Flagellum length was measured in cells showing a clear decrease in myc labeling in the new flagellum. o-t Electron micrographs of stained thin sections of parental cells (o, r) bearing one flagellar pocket (FP) and one flagellum (′), and of 24 h induced TbCfp44myc (p, s) and TbCfp43myc (q, t) cells. Cells and t are enlargements of p and q, respectively. Scale bars=100 nm in e–q, 200 nm in r–t. Note: in TbCfp44myc and TbCfp43myc cells, the flagellar pocket is enlarged and bears more than two flagella (p, q). Some of these flagella present axonemal defects including displaced and rotated CPC and shifted DMTs (p, white arrows, s, t)
(IFT) since IFT in *T. brucei* flagella was never described between the PFR and the axoneme and is circumscribed to two sets of doublet microtubules 3–4 and 7–8, located on each side of the PFR8. Moreover, the absence of cytoplasmic material accumulated at the basis of the flagellum and in the flagellar pocket and the normal length of the flagellum in both RNAi mutants (Fig. 8) do not support the IFT hypothesis8. Interestingly, it has been reported that two specific structures are present around DMTs 5 and 6: the first one is proposed to consolidate the interaction between the axoneme and the para-axonemal components11,47 and the second one to specifically strengthen DMTs 5–6 interaction, known as the 5–6 bridge11, made of inner and outer subunits. Both structures, the 5–6 bridge and the connecting proteins linking the axoneme to the PFR, are known to present an intercalary pattern11,47. This subcellular distribution is in agreement with our super-resolution and immuno-EM results which showed a minimum interval between two immunogold particles of around 25 nm for both TbCFAP44 and TbCFAP43, a distance similar to the interval between two outer 5–6 bridge subunits observed in the flagellum of sea urchin sperm11. All these connecting structures, located in the same area, could share physical interactions and the absence of one of the CFAP proteins could destabilize the whole complex, leading to both para-axonemal and axonemal defects. This echoes the anomalies evidenced in IF and TEM experiments in both mouse models and in man. This was particularly noticeable in *Csfap44*−/− sperm, where DMTs 5–6 were preferentially missing and para-axonemal structures such as dense fibers and FSs were defective (Fig. 6). Altogether, these results suggest that TbCFAP44 and TbCFAP43 may be the first two proteins specifically located next to DMTs 5–6, and therefore opening the way for a fine characterization of the protein complexes located in this area. Apart from infertility, the studied patients did not present any obvious PCD-associated symptoms, such as cough, rhinitis, sinusitis, and rhinorrhea and chronic bronchitis. It is not very surprising because flagellar and cilia defects are not always associated38. We already have shown that it is also true for *DNAH1*-mutated patients5,6,49,50. These observations suggest that mutations in MMAF genes are only responsible for primary infertility without other PCD features and reinforce the presumption that the sperm flagellum is assembled and organized through a specific way, different from other cilia. These differences are notably illustrated by a different beating pattern and the presence of specific peri-axonemal structures, mainly the FS and the ODFs31. Importantly, we demonstrated here that in the trypanosome, CFAP43 and CFAP44 are associated to a subset of microtubule doublets facing the PFR. The PFR does not exist in mammalian sperm but its function is likely covered by the FS and the ODF. The fact that the absence of CFAP43/44 has no visible impact on motile cilia could therefore be explained by the functional specificity of the proteins: to interact/ link the axoneme with flagella specific extra-axonemal structures, not present in motile cilia in respiratory tissues, brain, or in the oviduct. Further work should now be performed to elucidate the differential composition and organization of motile cilia and flagella. Last, analysis of KO heterozygous mice showed a small deterioration of sperm motility and morphology, suggesting that heterozygous mutations in key spermatogenesis genes might, alone or more likely through cumulative genetic and/or environmental factors, contribute to the less severe but much more frequent phenotype of mild to intermediate olisthogaenozoospemoria.

**Methods**

**Subjects and controls.** We included 78 subjects presenting with asthenozoospermia due to a combination of morphological defects of the sperm flagella including: absent, short, bent, coiled flagella, and of irregular width without any of the additional symptoms associated with primary ciliary dyskinesia (PCD). The morphology of patients’ sperm was assessed with Papanicolaou staining. Small variations in protocol might occur between the different laboratories. Subjects were recruited on the basis of the identification of >5% of at least three of the aforementioned flagellar morphological abnormalities (absent, short, coiled, bent, and irregular flagella).

The global average of all semen parameters are presented in Table 1 and were compared between the different genotype groups using a two-tailed test. Forty-six patients were of North African origin and consulted for primary infertility at the Clinique des Jazmin in Tunis. Ten individuals originated from the Iranian (Iranians) and were treated in Tehran at the Royan Institute (Reproductive Biomedicine Research Center) for primary infertility and 22 patients were recruited in France: 21 at the Cochin Institute and 1 in Lille. Participants were recruited between 2008 and 2016. All subjects had normal somatic karyotypes. Approximately half of the patients declared to be born from related parents. 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) guidelines35. Saliva and/or peripheral blood was obtained for all participants. During their medical consultation, all subjects answered a health questionnaire focused on PCD manifestations for infertility. Informed consent was obtained from all the subjects participating in the study according to the local protocols and the principles of the Declaration of Helsinki. In addition, the study was approved by local ethic committees. The samples were then stored in the CRB Gemethque (certified under ISO-9001 and NF 9004) following a standardized procedure.

**Subjects from fertile individuals with normal spermograms were obtained from CRB Biomedicine Research Center) for primary infertility and 22 patients were recruited on the basis of the following inclusion criteria: male partners of fertile women with normal semen analysis and no history of PCD, or related infertility. No use of fertility drugs.**

**WES and bioinformatics analysis.** Genomic DNA was isolated from saliva using Oragen DNA Extraction Kit (DNAgenotech®, Ottawa, ON, Canada). Coding regions and intron/exon boundaries were enriched using the all Exon V5 Kit (Agilent Technologies, Santa Clara, CA, USA) according to the manufacturer’s protocol. DNA sequencing was undertaken at the Genoscope, Evry, France, on the HiSeq 2000 from Illumina®. Sequence reads were aligned to the reference genome (hg19) using MAGIC3. MAGIC produces quality-adjusted variant and reference read counts on each strand at every covered position in the genome. Duplicate reads and reads that mapped to multiple locations in the genome were excluded from further analysis. Positions whose sequence coverage was below 10 on either the forward or reverse strand were marked as low confidence, and positions whose coverage was below 10 on both strands were excluded. Single-nucleotide variations and small insertions/deletions (indels) were identified and quality-filtered using in-house scripts. Briefly, for each variant, independent calls are made on each strand and the two calls agree are retained. Candidate variants were identified using an in-house bioinformatics pipeline, as follows. Variants with a minor allele frequency >5% in the NHLBI ESP6500 (Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA) or in 1,000 Genomes Project phase 1 datasets5,6,49,51. These observations suggest that mutations in MGAF genes are only responsible for primary infertility without other PCD features and reinforce the presumption that the sperm flagellum is assembled and organized through a specific way, different from other cilia. These differences are notably illustrated by a different beating pattern and the presence of specific peri-axonemal structures, mainly the FS and the ODFs31. Importantly, we demonstrated here that in the trypanosome, CFAP43 and CFAP44 are associated to a subset of microtubule doublets facing the PFR. The PFR does not exist in mammalian sperm but its function is likely covered by the FS and the ODF. The fact that the absence of CFAP43/44 has no visible impact on motile cilia could therefore be explained by the functional specificity of the proteins: to interact/ link the axoneme with flagella specific extra-axonemal structures, not present in motile cilia in respiratory tissues, brain, or in the oviduct. Further work should now be performed to elucidate the differential composition and organization of motile cilia and flagella. Last, analysis of KO heterozygous mice showed a small deterioration of sperm motility and morphology, suggesting that heterozygous mutations in key spermatogenesis genes might, alone or more likely through cumulative genetic and/or environmental factors, contribute to the less severe but much more frequent phenotype of mild to intermediate olisthogaenozoospemoria.

**RT-qPCR analysis.** RT-qPCR was performed with cDNAs from various tissues of human and mouse including testes. A panel of eight organs was used for mouse tissue RT-qPCR analysis. RT-qPCR was performed with cDNAs from various tissues of human and mouse including testes. A panel of eight organs was used for mouse tissue RT-qPCR analysis. RT-qPCR was carried out on ABI 3130XL (Applied Biosystems). Sequences were analyzed using SeqScape software (Applied Biosoft, Applied Biosystems).

**ARTICLE NATURE COMMUNICATIONS | DOI: 10.1038/s41467-017-02792-7**

**DOI: 10.1038/s41467-017-02792-7**
Immunostaining in human and mouse sperm cells. We performed IF staining on human and mouse spermatozoa as described by our laboratory25. In human, immunostaining was carried out on sperm cells collected from two of the DPBS with a stop-gain mutation, one mutated in each gene. IF experiments were performed using sperm cells from control individuals, from the patient P43-5 homozygous for the c.2658C>T variant in CFAP43 and from the patient P44-2 homozygous for the c.3175G>T variant in CFAP44. Sperm cells were fixed in phosphate-buffered saline (PBS) (pH 7.4) or paraformaldehyde for 1 h at room temperature. Slides were then blocked in 5% corresponding normal serum for 1 h, washed with PBS, and incubated overnight with the primary antibodies. The IF experiments on human spermatozoaREFERENCES and mouse sperm cells were performed. The primary antibodies used are listed in Supplementary Table S3. Western blot analyses were performed with 0.1% (v/v) Triton X-100–PBS, followed by 1 h incubation at room temperature with secondary antibodies. Highly cross-adsorbed secondary antibodies (Dylight 488 and Dylight 549) were from Jackson Immunoresearch. 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). Fluorescence images were captured with a confocal microscope (Zeiss LSM 710).

Electron microscopy of Trypanosoma cells. Cells were fixed in 1% glutaraldehyde to a final concentration of 2.5%, for 60 min. They were pelleted (1,000 g, 10 min), resuspended in fixation buffer (2.5% glutaraldehyde, 2% PFA, 100 mM phosphate buffer, pH 7.4, 50 mM sucrose for 1 h). Fixed cells were washed in water for 10 min, post-fixed in 1% OsO4 for 1 h, washed three times in water, and then samples were stained in 2% uranyl acetate in water at 4 °C overnight. Samples were next washed in water, dehydrated in ethanol, embedded in Spurr resin, and polymerized overnight at 60 °C. Sections (60–80 nm) were stained in 2% uranyl acetate, then lead citrate, and visualized on a Tecnai 12 TEM.

Immunogold labeling on detergent-extracted cytoskeleton from Trypanosoma cells. Cells were cultured in pVPBS (sucrose 15.7 g/l, glucose 1.8 g/l) and extracted with PIPES 100 mM, NP–40 (1% Igepal), 1% mmonium. Cytoskeletons were pelleted by centrifugation 10 min at 2,500 g and washed twice in PIPES 100 mM, 1% MgCl2, 1 mM, benzocyan. Two hundred and fifty microliters of droplets of cytoskeletons were placed on parafilm. Butvar covered, charged, carbon-coated nickel grids were floated on each drop and incubated for 10 min to let the cytoskeletons adhere. The grids were then moved through two blocking droplets for 5 min each (blocking buffer—1% fish skin gelatin, 0.01% Tween-20 in PBS) and then to droplets containing the primary antibody in blocking buffer (anti-TY1 1:10, anti-myc, anti-PFR mouse monoclonal LiCoA) for 2 h at room temperature. The grids were washed four times in blocking buffer and incubated with gold-conjugated secondary antibodies—EM GAR10 1:20 (anti-mouse, 6 nm), in blocking buffer for 2 h at room temperature. After incubation, grids were washed 3x for 5 min in blocking buffer, 2x 5 min in 0.1% fish skin gelatin, 0.001% Tween-20 in PBS, 2x 5 min in PBS, and then fixed in 2.5% glutaraldehyde in milliQ H2O for 5 min. Samples were negatively stained with 10 μl auro-thioglucose for 20 s. Samples were visualized on a FEI Tecnai 12 electron microscope, camera ORIUS 1000 11M Pixel (resolution 3–5 nm). Images were acquired with DigitalMicrograph and processed with ImageJ.

Western blotting on detergent-extracted cytoskeleton from Trypanosoma cells. Cells were cultured in pVPBS and extracted as described above. The cytoskeletons were pelleted by centrifugation 10 min at 55 °C and resuspended in 3% glutaraldehyde in PBS. Blots were revealed using the Clarity Western ECL Substrate Kit (Bio-Rad) with the ImageQuant LAS4000. After membrane stripping, protein loading was controlled by probing tubulin (TAT1, 1:500) as described above.

CRISPR/Cas9 KO mice. All animal procedures were run according to the Swiss and French guidelines on the use of animals in scientific investigations with the approval of the Swiss Federal Office of Veterinary Services, the Fédération Suisse des Laboratoires d’Excellence and the Direction Générale de la Santé (DGS) for the State of Geneva. All the procedures were done in Geneva and Lausanne. Animals were housed in the animal facilities of the Geneva University Hospitals (Coolsnap HQ2 camera, with Zeiss ×100 or ×63 objectives (NA 1.4) using the Metamorph software (Molecular Devices), and processed with ImageJ. Signals were quantified by measuring the intensity of bands.
Phenotypic analysis of mouse mutant

RNA expression analysis used for the subsequent PCR and Sanger sequencing. Primers sequences and RT-PCR was used to compare the different genotypes. Statistical tests with a two-tailed

 Approximately 25% of the offspring were homozygous for the mutated allele, eliminating any possible mosaicism. From the second generation of mice, genotyping was resolved on 1% agarose running gel with BET and visualized by UV light.

Phenotypic analysis of mouse mutant. To test the fertility, pubescent Cjap45 embryos were mated with C57Bl/6J females. To determine the sperm concentration, sperm samples were collected from the cauda epididymis of 8-week-old Cjap45 and Cjap44 sperm and number was determined using a hemocytometer under a light microscope. For sperm morphology, sperm was displayed over a slide, dried at room temperature, and then fixed in 75% ethanol. Harris–Schroeder staining was performed according to the WHO protocol. Schorr staining solution was obtained from Merck and at least 100 sperm per animal were analyzed. Mobility of sperm was assessed with computer-assisted motility analysis (CEOS I, Hamilton Thorn Research, Beverly, MA, USA) in an analysis chamber (100 μm depth, 30 μl volume, Leja Products B.V., Netherlands) at 37 °C. The settings employed for the analysis were as follows: acquisition rate, 60 Hz; number of frames, 100; minimum contrast, 25; minimum cell size, 10; low static size gate, 2.4; high static size gate, 2.4; low static intensity gate, 1.02; high static intensity gate, 1.37; minimum elongation gate, 12; maximum elongation gate, 100; magnification factor, 0.70. The motility parameters measured were curvilinear velocity, average path velocity (VAP), and straight line velocity (VSL). At least 100 motile sperm were analyzed for each assay while sperm and progressive sperm were characterized by VAP >1 μm/s, by average path velocity >30 μm/s and straightness (VSL/VAP) >70%, respectively. For each genotype, five mice were used.

Statistical analyses were performed with the SigmaPlot software. Unpaired t test was used to compare the different genotypes. Statistical tests with a two-tailed P values ≤0.05 were considered significant.

RNA expression analysis. Total RNA was isolated from 105 cells of parental, non-induced, and tetracycline-induced RNAi cells using the TRIzol reagent according to the manufacturer’s instructions (Life Technologies). The constitutively expressed housekeeper telomerase reverse transcriptase was used as an internal control35. RT-PCR was carried out with a SuperScript III One-step RT-PCR System with Platinum Taq-high-fidelity polymerase (Life Technologies Ltd, Paisley, UK) following the manufacturer’s protocol. Briefly, 100 ng of total RNA were mixed with primers and reverse transcriptase solution in 50 μl total volume and using the following cycle protocol: 30 min at 55 °C (reverse transcription); 2 min at 94 °C (inactivation of reverse transcriptase and activation Taq polymerase); followed by 25 cycles of regular PCR (denaturation: 94 °C for 15 s; annealing: 55 °C for 30 s; extension: 68 °C for 40 s); and finalized with a hold for 5 min at 68 °C. The RT-PCR products were resolved on 1% agarose running gel with BET and visualized by UV light. Primer sequence is provided in Supplementary Table 9.

References

1. Coutton, C., Fissore, R. A., Palermo, G. D., Stouffs, K. & Touré, A. Male infertility: genetics, mechanism, and therapies. Biomed. Res. Int. 2016, 7372362 (2016).
2. Coutton, C., Escoffier, J., Martinez, G., Arnoult, C. & Ray, P. F. Teratozoospermia: spotlight on the main genetic actors in the human. Hum. Reprod. Update 21, 455–485 (2015).
3. Ray, P. F. et al. Genetic abnormalities leading to qualitative defects of sperm morphology or function. Clin. Genet. 91, 217–232 (2017).
4. Kherra, Z.-E. et al. SPINK2 deficiency causes infertility by inducing sperm defects in heterozygotes and azoospermia in homozygotes. EMBO Mol. Med. 9, 1132–1149 (2017).
5. Ben Khelifa, M. 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. 94, 95–104 (2014).
6. Amiri-Yekta, A. et al. 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 (2016).
7. Ilyver, A. E., ’t Hoorn, P. A., van Roon-Mom, W., Peters, D. J. & Sergeeva, M. G. Exploring the transcriptome of cultured cells using in silico dissection of human tissues. PLoS ONE 7, e35618 (2012).
8. Smith, T. F. Diversity of WD-repeat proteins. Subcell. Biochem. 48, 20–30 (2008).
9. Smith, T. F. Diversity of WD-Repeat Proteins (Landes Bioscience, Austin, 2003).
10. Afzelius, B. Electron microscopy of the sperm tail; results obtained with a new fixative. J. Biophys. Biochem. Cytol. 5, 269–278 (1959).
11. Lin, J., Heuser, T., Song, K., Fu, X. & Nicastro, D. One of the nine doublet microtubules of eukaryotic flagella exhibits unique and partially conserved structures. PLoS ONE 7, e46494 (2012).
12. Vandeperre, B. et al. MCI-like is a placental mammal-specific mitochondrial preys carrier subunit expressed in postmeiotic male germ cells. J. Biol. Chem. 291, 16448–16461 (2016).
13. Fawcett, D. W. The mammalian spermatozoon. Dev. Biol. 44, 394–436 (1975).
14. Linde mann, C. B., Orlando, A. & Kanous, K. S. The flagellar beat of rat sperm is organized by the interaction of two functionally distinct populations of dynein bridges with a stable central axonemal junction. J. Cell Sci. 102, 249–260 (1992).
15. Vincensini, L., Blisnick, T. & Bastin, P. 1001 model organisms to study cilia and flagella. Biol. Cell. Annu. Eur. Cell Biol. Organ. 103, 109–130 (2011).
16. de Souza, W. & Souto-Padrón, T. The paraxial structure of the flagellum of Trypanosoma brucei. Mol. Cell. Proteom. 145, 117–120 (2006).
17. Sherwin, T. & Gull, K. The cell division cycle of Trypanosoma brucei brucei: timing of event markers and cytoskeletal modulations. Philos. Trans. R. Soc. Lond. Ser. B 323, 573–588 (1989).
18. Santrich, C. et al. A motility function for the rod of Leishmania parasites revealed by PFR-2 gene knockouts. Mol. Biochem. Parasitol. 90, 95–109 (1997).
19. Bastian, S., Sherwin, T. & Gull, K. Parafflagellar rod is vital for trypanosome motility. Nature 391, 548 (1998).
20. Oberholzer, M., Morand, S., Kunz, S. & Séebeck, T. A vector series for rapid PCR-mediated C-terminal in situ tagging of Trypanosoma brucei genes. Mol. Biochem. Parasitol. 145, 117–120 (2006).
21. Pullen, T. J., Ginger, M. L., Gaskell, S. J. & Gull, K. Protein targeting of an unusual, evolutionarily conserved adenylate kinase to a eukaryotic flagellum. Mol. Biol. Cell 15, 3257–3265 (2004).
22. Ridgley, E., Webster, P., Patton, C. & Ruben, L. Calmodulin-binding properties of the parafflagellar rod complex from Trypanosoma brucei. Mol. Biochem. Parasitol. 109, 195–201 (2000).
23. Aslett, M. et al. TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic Acids Res. 38, D457–D462 (2010).
24. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539 (2011).
25. Subota, I. et al. Proteomic analysis of intact flagella of procyclic Trypanosoma brucei cells identifies novel flagellar proteins with unique sub-localization and dynamics. Mol. Cell. Proteom. 13, 1769–1786 (2014).
26. Broadhead, R. et al. Flagellar motility is required for the viability of the bloodstream trypanosome. Nature 440, 224–227 (2006).
27. Baron, D. M., Ralston, K. S., Kabatutu, Z. P. & Hill, K. L. Functional genomics in Trypanosoma brucei identifies evolutionarily conserved components of motile flagella. J. Cell Sci. 120, 478–491 (2007).
28. Mirabelli, C. & Pollastri, G. Porter, PaleAle 4.0: high-accuracy prediction of protein secondary structure and relative solvent accessibility. Bioinformatics (Oxf., Engl.) 29, 2056–2058 (2013).
29. Wilson, D. et al. SUPERFAMILY—sophisticated comparative genomics, data mining, visualization and phylogeny. Nucleic Acids Res. 37, D388–D389 (2009).
30.ough, J., Karplus, V., Hughes, R. & Chothia, C. Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. J. Mol. Biol. 313, 903–919 (2001).

Received: 1 June 2017 Accepted: 28 December 2017
Published online: 15 February 2018
31. Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J. Mol. Biol. 247, 536–540 (1995).

32. Bastin, P., Bagherzadeh, Z., Matthews, K. R. & Gull, K. A novel epito tag system to study protein targeting and organelle biogenesis in Trypanosoma brucei. Mol. Biochem. Parasitol. 77, 235–239 (1996).

33. Albisetti, A. et al. Interaction between the flagellar pocket collar and the hook complex via a novel microtubule-binding protein in Trypanosoma brucei. PLoS Pathog. 13, e1006710 (2017).

34. Dean, S. et al. A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. Open Biol. 5, 140197 (2015).

35. Langousis, G. & Hill, K. L. Motility and more: the flagellum of Trypanosoma brucei. Nat. Rev. Microbiol. 12, 505–518 (2014).

36. Ralston, K. S., Kisalu, N. K. & Hill, K. L. Structure-function analysis of dynein light chain 1 identifies viable motility mutants in bloodstream Trypanosoma brucei. Eukaryot. Cell 10, 884–894 (2011).

37. Wamberger, C. 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. (Oxf., Engl.) 31, 1164–1172 (2016).

38. Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8, 2281–2308 (2013).

39. Neesen, J. et al. Disruption of an inner arm dynein heavy chain gene results in intracytoplasmic sperm injection. Hum. Reprod. (Oxf., Engl.) 31, 1164–1172 (2016).

40. Ostrowski, E. L., Dutcher, S. K. & Lo, C. W. Cilia and models for studying flagellum. Cilia 2, 2281–2308 (2013).

41. Silflow, C. D. & Lefebvre, P. A. Assembly and motility of eukaryotic cilia and flagella. Lessons from Olsiamonomas reinhartii. Plant Physiol. 127, 1300–1307 (2001).

42. Alberts, B. et al. Molecular Biology of the Cell (Garland Science, New York, 2002).

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

44. Yan, W. Male infertility caused by spermiogenic defects: lessons from gene knockouts. Mol. Cell. Endocrinol. 306, 24–32 (2009).

45. Morga, B. & Bastin, P. Getting to the heart of intraglagger transport using Trypanosoma and Olsiamonomas models: the strength is in their differences. Cilia 2, 16 (2013).

46. Huet, D., Blasnick, T., Perrot, S. & Bastin, P. The GTPass IFT27 is involved in both anterograde and retrograde intraglagger transport. eLife 3, e02419 (2014).

47. Koyfman, A. Y. et al. Structure of Trypanosoma brucei flagellum accounts for its biblical motion. Proc. Natl. Acad. Sci. USA 108, 11105–11108 (2011).

48. Praveen, K., Davis, E. E. & Katsani, N. Unique among ciliopathies: primary ciliary dyskinesia, a motile cilia disorder. F1000Prime Rep. 7, 36 (2015).

49. Wang, X. et al. Homozygous DNAH1 frameshift mutation causes multiple morphological abnormalities of the sperm flagella in Chinese. Clin. Genet. https://doi.org/10.1111/ctg.12857 (2016).

50. Sha, Y. et al. DNAH1 gene mutations and their potential association with dysplasia of the sperm fibrous sheath and infertility in the Han Chinese population. Fertil. Steril. 107, 1312–1318.e2 (2017).

51. Oberholzer, M. et al. Trypanosomes and mammalian sperm: one of a kind? Trends Parasitol. 23, 71–77 (2007).

52. World Health Organization. WHO Laboratory Manual for the Examination and Processing of Human Sperm, 5th ed. (World Health Organization, 2010).

53. SEQC/MAQC-III Consortium. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. Nat. Biotechnol. 32, 903–914 (2014).

54. 1000 Genomes Project Consortium et al. An integrated map of genetic variation from 1,092 human genomes. Nature 491, 56–65 (2012).

55. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285–291 (2016).

56. McLaren, W. et al. The Ensembl Variant Effect Predictor. Genome Biol. 17, 122 (2016).

57. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. 4, 1073–1081 (2009).

58. Adzhubei, I. A. et al. A method and server for predicting damaging missense mutations. Nat. Methods 7, 248–249 (2010).

59. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2−(ΔΔCT) method. Methods 25, 402–408 (2001).

60. Escoffier, J. et al. Subcellular localization of phospholipase Cδ in human sperm and its absence in DPY19L2-deficient sperm are consistent with its role in oocyte activation. Mol. Hum. Reprod. 21, 157–168 (2015).

61. Wirtz, E., Leal, S., Ochatt, C. & Cross, G. A. A tightly regulated inducible expression system for conditional gene knockouts and dominant-negative genetics in Trypanosoma brucei. Mol. Biochem. Parasitol. 99, 89–101 (1999).

62. Dauchy, F.-A. et al. Trypanosoma brucei CYP51: essentiality and targeting therapy in an experimental model. PLoS Negl. Trop. Dis. 10, e0005125 (2016).

63. LaCount, D. J., Barrett, B. & Donelson, J. E. Trypanosoma brucei FL1A is required for flagellum attachment and cytokinesis. J. Biol. Chem. 277, 17580–17588 (2002).

64. Woods, A. et al. Definition of individual components within the cytoskeleton of Trypanosoma brucei by a library of monoclonal antibodies. J. Cell Sci. 93 (Part 3), 491–500 (1989).

65. Brenndorfer, M. & Boshart, M. Selection of reference genes for mRNA quantification in Trypanosoma brucei. Mol. Biochem. Parasitol. 172, 52–55 (2010).

Acknowledgements
We thank Anne Bertrand from the GIN electron microscopy platform and the IAB microscopy platform and Alexis Gruchin and Jacques Mazega for their technical help. We thank Klaus Ersfeld (Bayreuth University) for the anti-myc antibody, Nicolas Biteau (Bordeaux University) for the anti-PFR2 antibody, Keith Gull (Oxford University) for the TATI antibody, Philippe Bastin (Institut Pasteur, Paris) for the anti-TY1 antibody, and Vanderpere Benoit and Martinou Jean Claude (UNIGE university, Geneva, CH) for the MPcl1l antibody.

We thank Denise Escaler and Gérard Gascron for discussions and expertise on the topic. The STED microscopy was done in the Bordeaux Imaging Centre a service unit of the CNRS-INSERM and Bordeaux University, member of the national infrastructure France Bioimaging supported by the French National Research Agency (ANR-10-INBS-04). The help of Christel Poupol and Patrice Mascalchi is acknowledged. Some of the electron microscopy was done at the Bordeaux Imaging Centre. We thank our patients and control 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 and the "The Whole genome sequencing of patients with Flagellar Growth Defects (FGD)" financed by the fondation maladies rares (FMR) for the program Squeuence à haut débit 2012.

Author contributions
C.C., M.B., A.T., C.A., and P.F.R. analyzed the data and wrote the manuscript; Z.-E.K., A. A.-Y., L.S., P.L., E.EK., I.F., and C.C. performed molecular work; T.K., J.-F.D., A.B., and S.-T.M. analyzed genetic data; C.C., A.S.V., P.L.T., C.W.-L., and S.P.B. performed IF experiments; S.C. performed fluorescent analysis in silico, K.P.-G., A.S.V. A.S., and A.T. performed EM experiments, Z.-E.K., A.A.-Y., A.S.V., B.C., S.F.B.M., S.N., and J.E. performed mouse work; D.D., N.L., D.R.R., and M.B. performed Trypanosoma work, N.L. and M.B. performed STED analysis. S.F.B.M., G.M., A.D., S.H.H., V.M., L.H., O.M., M. M., H.L., M.K., S.H., V.S., P.-S.I., J.-P.W., H.G., E.D., and R.Z. provided clinical samples and data; M.B., A.T., C.A., and P.F.R. designed the study, supervised all molecular laboratory work, had full access to all of the data in the study, and takes responsibility for the integrity of the data and its accuracy. All authors contributed to the report.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-017-02792-7.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Charles Coutton1,2, Alexandre S. Vargas1, Amir Amiri-Yekta1,3,4, Zine-Eddine Kherraf1,3, Selima Fourati Ben Mustapha5, Pauline Le Tanno1,2, Clémentine Wamberge-Legrand1,3, Thomas Karaouzène1,3, Guillaume Martinez1,2, Serge Crouzy6, Abbas Daneshpouri4, Seyyedeh Hanieh Hosseini7, Valérie Mitchell8, Lazhar Halouani5, Ouafi Marrakchi5, Mounir Makni5, Habib Latrous5, Mahmoud Kharouf5, Jean-François Deleuze9, Anne Boland9, Sylviane Hennebicq1,10, Véronique Satre1,2, Pierre-Simon Jouk11, Nicolas Thierry-Mieg12, Beatrice Conne13, Denis Dacheux14,15, Nicolas Landrein14, Alain Schmitt16,17,18, Laurence Stouvenel16,17,18, Patrick Lorès16,17,18, Elma El Khouri16,17,18, Serge P. Bottari1, Julien Fauré19,20, Jean-Philippe Wolf18,21, Karin Pernet-Gallay20, Jessica Escoffier1, Hamid Gourabi4, Derrick R. Robinson14, Serge Nef13, Emmanuel Dulioust18,21, Raoudha Zouari5, Mélanie Bonhivers14, Aminata Touré16,17,18, Christophe Arnoult1 & Pierre F. Ray1,3

1Genetic Epigenetic and Therapies of Infertility, Institute for Advanced Biosciences, Inserm U1209, CNRS UMR 5309, Université Grenoble Alpes, 38000 Grenoble, France. 2CHU de Grenoble, UM de Génétique Chromosomique, 38000 Grenoble, France. 3CHU de Grenoble, UM GI-DPI, 38000 Grenoble, France. 4Department of Genetics, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACER, Tehran, 16635-148, Iran. 5Polyclinic les Jasmins, Centre d’Aide Médicale à la Procréation, Centre Urbain Nord, 1003 Tunis, Tunisia. 6Laboratoire de Chimie et Biologie des Métaux, Institut de Recherche en Technologie et Sciences pour le Vivant, CEA IRTSV/LCBM/GMCT, CNRS UMR 5249, Université Grenoble Alpes, 38054 Grenoble, France. 7Department of Andrology, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR, Tehran, 16635-148, Iran. 8EA 4308, Department of Reproductive Biology and Spermiology-CECOS Lille, University Medical Center, Lille, 59037, France. 9Centre National de Génotypage, Institut de Génomique, CEA, 91000Evry, France. 10CHU de Grenoble,UF de Biologie de la procréation, 38000 Grenoble, France. 11CHU de Grenoble, UF de Génétique Médicale, 38000 Grenoble, France. 12Univ. Grenoble Alpes/CNRS TIMC-IMAG, 38000 Grenoble, France. 13Department of Genetic Medicine and Development, University of Geneva Medical School, 1211 Geneva, Switzerland. 14Microbiologie Fondamentale et Pathogénicité CNRS UMR 5234, University Bordeaux, 33000 Bordeaux, France. 15Bordeaux-INP, Microbiologie Fondamentale et Pathogénicité, UMR-CNRS 5234, 33000 Bordeaux, France. 16Institut National de la Santé et de la Recherche Médicale, INSERM U1016, Institut Cochin, 75014 Paris, France. 17Centre National de la Recherche Scientifique, CNRS UMR8104, 75014 Paris, France. 18Sorbonne Paris Cité, Faculté de Médecine, Université Paris Descartes, 75014 Paris, France. 19CHU de Grenoble UF de Biochimie Génétique et Moléculaire, 38000 Grenoble, France. 20Grenoble Neuroscience Institute, INSERM 1216, 38000 Grenoble, France. 21Laboratoire d’Histologie Embryologie - Biologie de la Reproduction, GH Cochin Broca Hôtel Dieu, Assistance Publique-Hôpitaux de Paris, 75014 Paris, France. Mélanie Bonhivers, Aminata Touré, Christophe Arnoult, and Pierre F. Ray contributed equally to this work.