Gametes

ORIGINAL RESEARCH

Biallelic variants in MAATS1 encoding CFAP91, a calmodulin-associated and spoke-associated complex protein, cause severe astheno-teratozoospermia and male infertility

Guillaume Martinez,1,2 Julie Beurois,1 Denis Dacheux,3,4 Caroline Cazin,1 Marie Bidart,1,5 Zine-Eddine Kherraf,1,6 Derrick R Robinson,4 Véronique Satre,1,2 Gerald Le Gac,7,8 Chandran Ka,7 Isabelle Gourlaouen,7 Yann Fichou,7 Graciane Petre,9 Emmanuel Dulioust,10,11 Raoudha Zouari,12 Nicolas Thierry-Mieg,13 Aminata Touré,11,14,15 Christophe Arnoult,1 Mélanie Bonhivers,3 Pierre Ray,1,6 Charles Coutton1,2

ABSTRACT

Background Multiple morphological abnormalities of the flagella (MMAF) consistently lead to male infertility due to a reduced or absent sperm motility defined as asthenozoospermia. Despite numerous genes recently described to be recurrently associated with MMAF, more than half of the cases analysed remain unresolved, suggesting that many yet uncharacterised gene defects account for this phenotype.

Methods Exome sequencing was performed on 167 infertile men with an MMAF phenotype. Immunostaining and transmission electron microscopy (TEM) in sperm cells from affected individuals were performed to characterise the ultrastructural sperm defects. Gene inactivation using RNA interference (RNAi) was subsequently performed in Trypanosoma.

Results We identified six unrelated affected patients carrying a homozygous deleterious variants in MAATS1, a gene encoding CFAP91, a calmodulin-associated and spoke-associated complex (CSC) protein. TEM and immunostaining experiments in sperm cells showed severe central pair complex (CPC) and radial spokes defects. Moreover, we confirmed that the WDR66 protein is a physical and functional partner of CFAP91 into the CSC. Study of Trypanosoma MAATS1’s orthologue (TbcFAP91) highlighted high sequence and structural analogies with the human protein and confirmed the axonemal localisation of the protein. Knockdown of TbcFAP91 using RNAi impaired flagellar movement led to CPC defects in Trypanosoma as observed in humans.

Conclusions We showed that CFAP91 is essential for normal sperm flagellum structure and function in human and Trypanosoma and that biallelic variants in this gene lead to severe flagellum malformations resulting in astheno-teratozoospermia and primary male infertility.

INTRODUCTION

Spermatozoa are highly differentiated haploid cells resulting from spermatogenesis, a complex developmental process in which germline stem cells undergo a series of divisions and morphological modifications. At the simplest level, mature spermatozoa are polarised cells composed of a head and a flagellum joined by the connecting piece. Each of these parts contains highly specialised structural features reflecting the unique functions required for fertilisation. The flagellum contains the energy sources and machinery to generate the motility necessary for the spermatozoa to reach the egg. Flagellum integrity is thus critical for normal sperm function and sperm flagellum defects consistently lead to male infertility due to a reduced or absent sperm motility defined as asthenozoospermia. Asthenozoospermia is a frequent sperm phenotype found in approximately 19% of infertile men1 and can be combined with morphological sperm defects. Since 1984, morphological abnormalities of the sperm flagella leading to asthenozoospermia have been regularly reported. This rare phenotype has been subsequently reported as ‘dysplasia of the fibrous sheath’, ‘short tails’ or ‘stump tails’. In 2014, this heterogeneous group of flagellar defects was called ‘MMAF’ for multiple morphological anomalies of the flagella to standardise all these terms.2 Exome sequencing allowed us to partially elucidate genetic and physiopathological mechanisms leading to sperm flagellum defects. To date, variants in at least 16 genes have been found to be associated with MMAF phenotype.3–17 We however observe that despite regular new gene identification, about two-thirds of MMAF individuals remain unresolved demonstrating the high genetic heterogeneity of this phenotype.4 Several other genes are therefore likely to be involved in this phenotype and further efforts are needed to improve our genetic understanding of the molecular pathogenesis of asthenozoospermia.

MATERIALS AND METHODS

Patients

With the aim of identifying further genetic causes associated with human asthenozoospermia due to MMAF, we analysed whole exome sequencing data from a cohort of 167 patients with MMAF...
previously established by our team. All patients presented with a typical MMAF phenotype characterised 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. All patients had a normal somatic karyotype (46,XY) with normal bilateral testicular size, normal hormone levels (FSH, testosterone and prolactin) and secondary sexual characteristics. All these patients presented only with isolated infertility without any other clinical features. Sperm analysis was carried out in the source laboratories during routine biological examination of the patients according to WHO guidelines. The morphology of the patients’ sperm was assessed with Papanicolaou staining (figure 1A–C). Detailed semen parameters of the six MAATS1 affected individuals are presented in table 1. Sperm samples for additional phenotypic characterisation could only be obtained from MAATS1_6 and MAATS1_7. The six patients were (MAATS1_1 to MAATS1_6) were initially recruited at the Clinique des Jasmin in Tunis (Tunisia) and were originated from North Africa. All these patients were born to first-cousin parents.

Exome sequencing and bioinformatic analysis
Data processing of the whole cohort of 167 patients with MMAF were performed according to our previously described protocol. For details, see online supplementary methods.

Sanger sequencing
MAATS1 variants identified by exome sequencing were validated by Sanger sequencing as previously described. PCR primers and protocols used for each patient are listed in the online supplementary table S1.

Real-time RT-qPCR analysis
RT-qPCR was performed with cDNAs from a panel of 10 human tissues purchased from Life Technologies as previously described. RT-qPCR was performed as described in the online supplementary methods. The primer sequences are presented in the online supplementary table S2.

Transmission electron microscopy analysis of human sperm cells
Transmission electron microscopy (TEM) experiments were performed using sperm cells from control individuals and from MAATS1_1 and MAATS_1, both carrying the missense variant c.124G>C; p.(Asp42His). TEM was performed as described in the online supplementary methods.

Minigene assays
All recombinant plasmid constructs were produced as described previously and details about MAATS1 minigene assays are presented in the online supplementary methods. Briefly, the wild-type MAATS1 plasmid construct (pSplicePOLR2G–MAATS1_WT) was generated by cloning sequence of exon 1, 5' untranslated region and the intronic flanking regions of MAATS1 into the pSplicePOLR2G vector. A c.124G>C variant was also introduced in the pSplicePOLR2G–MAATS1__ c.124G>C vector by using the QuickChange Site-Directed mutagenesis kit according to the manufacturer’s instructions (Agilent Technologies, Santa Clara, California, USA). Sequencing analyses were performed to check the integrity of all of the plasmid constructs.

Immunostaining in human sperm cells
Immunofluorescence (IF) experiments were performed using sperm cells from control individuals and from MAATS1_1 and MAATS_1, both carrying the missense variant c.124G>C; p.(Asp42His). For each studied patient with MMAF, 200 sperm cells were analysed by two different experienced operators and the IF staining intensity and pattern were compared with a fertile control. IF experiments were performed as described in the online supplementary methods. The list of primary antibodies used in the present study is detailed in the online supplementary table S3.

Trypanosoma brucei culture and transfection
The trypanosome cell lines used in this study derived from the procyclic form parental Trypanosoma brucei SmOxP427 strain co-expressing the T7 RNA polymerase and the tetracycline repressor and were grown and transfected as described. When required, the culture medium was supplemented with puromycin (1 µg/ml), hygromycin (25 µg/ml), phleomycin (5 µg/ml). RNA interference (RNAi) was induced with tetracycline (10 µg/ml).

T. brucei cell lines
We generated in SmOxP427 cells a TbCFAP91 transgene at its endogenous locus encoding N-terminal 10 HA epitope-tagged TbCFAP91 (1HA–TbCFAP91) using a pPOTv7-blast-10HA vector as described. For TbCFAP91 RNAi, Bp 1-864 were cloned between the XhoI and XbaI sites of p2T7tiB and transfected into the 1HA–TbCFAP91-expressing background cell line. Primers are described in the online supplementary table S4.

Immunofluorescence of Trypanosoma cells
Cells were collected, washed and processed for immunolabeling on methanol fixed detergent-extracted cells (cytoskeleton (CSK)) as described by Albisetti et al. The antibodies used are described in the online supplementary table S5. Images were acquired on a Zeiss Imager Z1 microscope, using a Photometrics Coolscape HQ2 camera, with a 100x Zeiss objective (NA 1.4) using Metamorph software (Molecular Devices), and processed with ImageJ.

Sedimentation assays and video-microscopy
Sedimentation assays and videos were done as described. Briefly, the cells were placed in cuvettes and incubated 24 hours without shaking. The optical density (ODmax) was measured before mixing (ODb, the cell density reflects the ‘swimming’ cells) and after mixing (ODa, the cell density reflects ‘swimming’ and ‘sedimenting’ cells). The graphs presented in figure 1F represent the percentage of swimming cells calculated as 100−(ODb/ODa)×100 and, for the RNAi cells, normalised with the 1HA–TbCFAP91 parental cell line. Video-microscopy was carried out on a Zeiss AxiosImager Z1, 40x lens (NA 1.3). 25 s of digital video from separate regions were captured and analysed using Metamorph software (Molecular Devices) (speed x4).

Electron microscopy of Trypanosoma cells
Electron microscopy blocks of cells and thin sections were prepared as described. Thin sections were visualised on a FEI Tecnai 12 electron microscope, camera ORIUS 1000 11M Pixel (resolution 3 ± 3 nm). Images were acquired with DigitalMicrograph and processed with ImageJ.

Western blot analysis
Proteins from whole cell extracts or detergent-extracted cells and flagella were separated on sodium dodecyl sulfate-polyacrylamide gels and transferred to nitrocellulose membranes. Blots were incubated with the appropriate primary antibodies overnight at 4°C. After washing, the blots were incubated with alkaline-phosphatase-conjugated secondary antibodies. The reaction was developed with 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) and nitroblue tetrazolium (NBT).
Figure 1  Morphology of normal and MAATS1 mutant spermatozoa and the variants identified in MAATS1-mutant patients. (A–C) Light microscopy analysis of spermatozoa from fertile control individuals (A), patient MAATS1_3 (B) and patient MAATS1_4 (C). Most spermatozoa from MAATS1 patients have flagella that are short (#), absent (*), coiled (@) or of irregular calibre (&). Head malformations were also observed (△). (D) Electropherograms from Sanger sequencing indicating the homozygous state of the two identified variants: the c.682+1G>A variant (NM_033364.3) affecting the splice donor site following exon 6 and the c.124G>C:p.(Asp42His) variant (NM_033364.3) affecting the last nucleotide of exon 1 in patients MAATS1_1-2 and MAATS1_3-6, respectively. Variants are annotated following the Human Genome Variation Society (HGVS) recommendations. (E) Structure of the canonical transcript of MAATS1 showing the position of the observed variants. The functional structure of the encoded protein is shown in the lower panel. Coiled-coil domains are indicated by black boxes. CTAP91-domain (IPR032840) is highlighted in green. (F) Transmission electron microscopy analyses of sperm cells from a control individual and the two patients MAATS1_3 and MAATS1_4. In longitudinal sections, we can notice a short tail corresponding to a cytoplasmic mass containing the different components of the flagellum, all disorganised. In cross-sections of sperm flagellum from MAATS1 patients, we can observe some sections showing a 9+0 axoneme lacking the CPC (blue arrow) with an abnormal number of outer dense fibres (orange arrow) (MAATS1_4) as well as some sections with a normal axonemal 9+2 conformation (MAATS1_3). Scale bars: 200 nm.
RESULTS

Exome sequencing identified homozygous variants in MAATS1 in patients with MMAF

In the whole cohort of 167 patients with MMAF, we previously identified 60 patients (36 %) with harmful variants in known MMAF-related genes. After reanalysis of the remaining exomes, we identified six additional unrelated patients originated from North Africa (MAATS1_1) with a homozygous variant in MAATS1 (MIM: 609910), a gene not previously associated with any pathology, accounting for 3.6% of our cohort. Two patients (MAATS1_1) had a splice-site variant c.682+1G>A (NM_033364.3, GRCh38 g.119715744G>A, rs147597066) and four (MAATS1_2-6) had a missense variant c.124G>C; p.(Asp42His) (NM_033364.3, GRCh38 g.119703222G>C, rs149348782) located in a splice-site region (figure 1, table 1 and online supplementary table S6). All variants identified by exome sequencing were validated by Sanger sequencing as illustrated in figure 1D. All MAATS1 variants are deposited in ClinVar under reference SUB6200231. No other candidate variants reported to be associated with cilia, flagella or male fertility were present in any of the six MAATS1 affected individuals. No heterozygous MAATS1 variants were found in our full cohort.

MAATS1 (also known as CFAP91, AAT-1 or csof15) is located on chromosome 3 and contains 18 exons encoding cilia-associated and flagella-associated protein 91 (CFAP91), a predicted 767-amino acid protein (Q7Z4T9). MAATS1 is mostly expressed in the pituitary and the testis according to data from GTEx and described to be associated with cilia and flagella. Moreover, quantitative single-cell RNA sequencing datasets from human adult testis (ReproGenomics Viewer) indicate an abundant expression in the germ cells from zygotene spermatocyte to late spermatid stages suggesting a role in sperm cell differentiation and/or function. In addition, the encoded CFAP91 protein was significantly detected in human sperm proteome, whereas it was found at a very low level in human airway cilia. RT-qPCR experiments performed with a panel of 10 human tissues including other ciliated tissues such as trachea confirmed these results, showing that MAATS1 transcripts are largely overexpressed in the testis compared with all the other tested tissues (online supplementary figure S1).

The first variant c.682+1G>A identified in patients MAATS1_1 and MAATS1_2 is a splice variant which changes the invariable +1 guanine and, thus, abolishes the splicing donor site according to the splice-site prediction algorithm Human Splicing Finder (http://www.umd.be/HSF3). The variant is present in the Genome Aggregation Database (gnomAD, http://gnomad.broadinstitute.org/variant/3-119434591-G-A) with a minor allele frequency of 3.98×10^−6 (online supplementary table S6). The second variant identified in four patient (MAATS1_3-6) is a missense variant c.124G>C; p.(Asp42His) located in exon 1 (figure 1E). This variant is present at a very low frequency in the general population (8.42×10^−6 in gnomAD database/3-1194242069-G-C) (online supplementary table S6) and is absent from our in-house database of 94 control exomes obtained from other infertile patients from North Africa (n=60) and the Middle East (n=34) (corresponding to the geographical origin of most patients in this study) presenting with a different sperm phenotype (azoospermia). Using prediction software for non-synonymous SNPs, we found that this missense change...
is predicted to be deleterious by the Sorting Intolerant from Tolerant (SIFT) algorithm (score of 0) and probably damaging by PolyPhen (score of 0.999) (online supplementary table S6). Moreover, the concerned amino acid (Asp42) was found conserved in CFAP91 orthologues (online supplementary figure S2). The in silico splicing prediction module in the Alamut software V2.10 (Interactive Biosoftware) was used to predict the effect of the variant c.124G>C on the effectiveness of splicing. This software regroups five splice-site prediction tools: SpliceSiteFinder-like; MaxEntScan; NNSPLICE; GeneSplicer and Human Splicing Finder. Interestingly, the c.124G>C variant which concerns the last nucleotide of MAATS1 exon 1 was predicted to weaken the consensus splice donor site of MAATS1 exon 1 (online supplementary figure S3). Another tool for predicting variant spliceogenicity was also used: SPICE, accessible at https://sourceforge.net/projects/spicev2-1/. SPICE uses logistic regression by running different combinations of in silico tools and reveals an unprecedented sensitivity and specificity of 99.5% and 95.2%, respectively. The variant was predicted to alter splicing with a SPICE probability of 0.99196 (online supplementary table S6).

In vitro splicing analysis of the missense variant c.124G>C; p.(Asp42His)

In order to confirm the bioinformatic predictions and because mRNA from affected individuals was not available, we tested the consequences of the c.124G>C; p.(Asp42His) variant using a splicing minigene assay. HEK293T cells were transiently transfected with different pSplicePOLR2G–MAATS1 minigenes (online supplementary figure S4A), and then the splicing pattern of the mutant minigenes was compared with that of the wild type. Transfection of the pSplicePOLR2G–MAATS1_WT minigene in HEK293 cells resulted in the production of one major mRNA product (online supplementary figure S4B). Sequencing analysis confirmed that this product consisted of the succession of exon 1 of MAATS1 and exon 2 of PolR2G (online supplementary figure S4C). RT-PCR analysis of the c.124G>C variant in exon 1 (pSplicePOLR2G–MAATS1_c.124G>C) resulted in the synthesis of a product that differs in size from those generated with the wild-type construct (online supplementary figure S4B). Direct sequencing of the band indicated a 49 bp partial retention of intron in mRNA: r[124G>C; 124_125ins124+1_124+49] (online supplementary figure S4B). A second plasmid construct was generated for the study with the first three exons of MAATS1 as well as the 5’ and 3’ intronic flanking regions and confirmed these results (data not shown). Overall, these data indicate that the c.124G>C variant alters the cellular splicing of a full-length mature transcript, and leads to an aberrant transcript with a 49 bp insertion inducing a stop codon after the first exon. These observations permit to define at the protein level a new nomenclature for this genetic variant: p.Asp42ArgfsTer15.

**MAATS1 variants lead to a severe axonemal disorganisation of the sperm flagellum**

Using TEM, we first studied the ultrastructure of sperm cells from the affected individuals MAATS1_3 and MAATS1_4 both carrying the missense variant c.124G>C; p.(Asp42His) (figure 1F). Due to the low number of available sperm cells, only a few cross-sections (<10) had a sufficient quality to be analysed and therefore no statistical analysis could be done. Among these cross-sections, we observed various defects from normal sections (figure 1F) to abnormal conformations of the axoneme with in particular the absence of the central pair complex (CPC) (9+0 conformation) (figure 1F). Observation of longitudinal sections showed severe flagellum abnormalities, which appeared completely disorganised, resulting in truncated flagella or a cytoplastic mass encompassing unassembled axonemal components (figure 1F). In some sections, we also observed peri-axonemal structural defects such as an abnormal number of outer dense fibres (figure 1F), a defect already observed in sperm from MMAF-affected patients carrying variants in other genes.

To define the ultrastructural defects established by TEM and to characterise the molecular defects induced by the c.124G>C MAATS1 variant in human sperm, we subsequently studied the presence and localisation of several proteins belonging to different axonemal substructures by IF. The presence and localisation of several proteins was investigated: SPAG6 as a marker of the radial spokes (RS). We observed that in sperm from patients MAATS1_3 and MAATS1_4, staining of SPAG6 revealed a heterogeneous staining depending on flagellar morphology. Scale bars: 10 µm. (B) Histogram showing the percentage of analysed spermatozoa with or without SPAG6 staining according to their morphology (number of sperm cells analysed=200). Scale bars: 10 µm.
the full-length flagella, but in patients MAATS1_3 and MAATS1_4, WDR66 staining was weak or completely absent whereas tubulin staining remained detectable (online supplementary figure S7). This result demonstrates that the CSC is strongly disorganised, and is in agreement with the fact that CFAP91 and WDR66 (CFAP251) are two interacting proteins, likely to be closely associated within the same axonemal complex.

**T. brucei CFAP91 is an axoneme-associated protein involved in flagellum motility and central pair positioning**

To validate our candidate gene and to better characterise CFAP91 localisation and function, we decided to use a *Trypanosoma* model, which could be used for forward and reverse genetics and which has largely contributed to the characterisation of the molecular pathology of different human diseases caused by defective cilia and/or flagella. The *T. brucei* axoneme is similar to that of mammalian flagella with a 9+2 structure. *T. brucei* flagellum contains a para-axonemal structure named parlagellur rod (PFR), which plays a role in flagellum motility and serves as a platform for metabolic and signalling enzymes, similar to para-axonemal structures FS and ODFs on mammalian sperm. Using BlastP analysis on the *T. brucei* genome database, we identified the putative *T. brucei* CFAP91 orthologue Tb927.11.2790 (named TbCFAP91 in this study). TbCFAP91 is a 709 amino acid protein annotated as a tetratricopeptide repeat putative protein. The calculated identity and similarity between human CFAP91 and TbCFAP91 is 27.97% and 46%, respectively. We thus generated a *T. brucei* cell line expressing N-terminus-HA tagged TbCFAP91 (*ha-TbCFAP91*) and localised the protein by IF on detergent-extracted cells (CSK). *ha-TbCFAP91* localised specifically to the flagellum, as established by IF with the PFR protein, PFR2 (figure 4A) and by western blot analysis from flagellum fraction (figure 4B). *ha-TbCFAP91* is associated with the axoneme fraction of the flagellum as its labelling is observed in flagellum part lacking PFR labelling. Furthermore, *ha-TbCFAP91* labelling was observed at both the old flagellum and the new flagellum, suggesting that it is present at every stage of the cell cycle. To assess the functional role of TbCFAP91 in the parasite, we used a tetracycline-inducible RNAi system to knockdown the expression of TbCFAP91, and generated the *ha-TbCFAP91* RNAi T. brucei cell line. Induction of RNAi by addition of tetracycline in the culture medium did not affect cell growth (figure 4C), but led to a decrease in *ha-TbCFAP91* levels. Indeed, detection of *ha-TbCFAP91* was strongly reduced by western blot analysis (figure 4D) and the protein was not detected by IF in the newly formed flagellum (figure 4E). Knockdown of *TbCFAP91* did not affect the cell morphology and the flagella appeared normal in number, length and overall structure. However, induced cells were sedentary, suggesting a flagellar beating defect (figure 4F). Indeed, sedimentation assays show that after 96 hours of induction, 90% of the induced cells sedimented. This sedimentation reflects a flagellar motility defect that was directly observed by video microscopy (online supplementary videos S1 and S2) and tracking (figure 4G). Remarkably, comparison of TEM images of thin sections of WT and 7 days RNAi-induced cells revealed that knockdown of *TbCFAP91* induced a rotation of the central pair (CP) (figure 4H). Indeed, measurements of the α angle (see scheme in figure 4I) demonstrated that in WT cells, the CP orientation was fixed as previously shown, while in *TbCFAP91* RNAi-induced cells some cross-sections exhibit a rotated CP axis (figure 4H, I).

**Figure 3** Radial spokes are affected in MAATS1-mutant patients. (A) Sperm cells from a fertile control individual and MAATS1_1 and MAATS1_4 patients stained with anti-RSPH1 (HPA017382, Sigma-Aldrich, rabbit, 1:100, green), a protein located at the head of the radial spokes and anti-acytlated tubulin (32–2500, ThermoFisher, mouse, 1:1000, red) antibodies. DNA was counterstained with DAPI II. RSPH1 immunostaining is present throughout the flagellum in control sperm cells but is mainly absent or strongly reduced in MAATS1 patients. In a few sperm cells, the RSPH1 staining is present but displays an abnormal pattern with a dotted and irregular signal. Scale bars: 10 µm. (B) Histogram showing the percentage of analysed spermatozoa with or without RSPH1 staining according to their morphology (number of sperm cells analysed=200). Scale bars: 10 µm.
DISCUSSION

In the present work, we showed that the presence of biallelic MAATS1 variants induce a typical MMAF phenotype in humans, demonstrating that this gene is necessary for sperm flagellum structure. Several ultrastructural components of the axoneme were found to be affected by the absence of the encoded protein. First, CPC anomalies were shown by IF or TEM experiments (figures 1F and 2). We observed a heterogeneous pattern of CPC staining, and the presence of the CPC seems closely associated with the flagellar morphology, with a higher proportion of CPC defects arising in shorter flagellum (figure 2B). This observation supports the hypothesis that the CPC is a key component of the sperm flagellum ensuring a stable construction and/or a global cohesion of the full axonemal structure. The RS proteins seemed more severely and constantly affected by the MAATS1 variants even though few sperm cells showed normal RSPH1 staining (figure 3). RS enhance the stability of the CP and of the entire axoneme. When the RS are defective the CPC can be assembled but tends to be misplaced with CP transposition, shift or rotation. In Trypanosoma, we confirmed that CFAP91 is an axonemal protein and by RNAi experiments we observed that its absence leads to structural defects (figure 4). We indeed showed using TEM that flagella from RNAi-induced parasites exhibit an abnormal axoneme with a rotated CPC, a defect not observed in control trypanosome cell lines (figure 4H, I). All these results are consistent with the known localisation of CFAP91 in the axoneme: CFAP91 is an A-kinase anchoring protein (AKAP)-binding protein and is a major component of the
The CSC is located within the region that interconnects the bases of RS2, RS3 and the N-DRC. It mediates regulatory signals between the radial spokes and the dynein arms and therefore is involved in the regulation of flagellar motility. Based on the localisation of the CSC at the base of the radial spokes, it was expected that the absence of members of this complex would result in radial spoke assembly defects. In *Chlamydomonas*, knockdown of FAP91 by microRNAs results in ultrastructural defects in RS2 and RS3 and in a partial destabilisation of the N-DRC. Interestingly, knockout of each CSC subunit in *Chlamydomonas* or Tetrabrachymena led to a reduced RS density but not to the total absence of the RS indicating heterogeneity among the repeats. Such heterogeneity was also observed in sperm cells from MAATS1 patients, as illustrated by the RSPH1 staining (figure 3).

In *Chlamydomonas*, in addition to FAP91, the CSC consists of CaM and two other conserved proteins: FAP61 (also known as c2orf126) and FAP251 (known as WDR66). The two subunits FAP61 and FAP251 are within the base of the RS3 while the FAP91 subunit contacts the base of RS2 on the side facing RS3.

We and another team previously demonstrated that deleterious variants in CFA2P1/WDR66 (MIM: 618146) lead to a MMAF phenotype with similar flagellar defects to what is observed in MAATS1 affected individuals. Moreover, *Chlamydomonas* mutants lacking RSP3 (pf14) completely failed to assemble the radial spokes. Overall, these data suggest that CFA9P1 is critical for radial spokes assembly and stability through its essential role within the CSC and its interaction with other axonemal partners such as WDR66 and RSP3. In humans, the absence of CFA9P1 in sperm cells likely destabilises the RS structure and as a consequence the stability of the CPC, both of these defects resulting in a severe axonemal disorganisation responsible for a MMAF sperm phenotype.

In our cohort of 167 patients, we had identified 60 patients (36%) with variant in known MMAF-related genes. The diagnostic yield rises to 40% when adding the six new MAATS1-affected individuals, highlighting the performance of the exome sequencing strategy in the genetic exploration of patients with MMAF. MAATS1 is one of the most frequently mutated genes in our cohort with 3.6% of patients. Combined with patients with a WDR66 alteration (11 patients), variants in CSC-related genes are found in 10.2% of our cohort, suggesting that the CSC is the main structure affected in patients with MMAF. *T. brucei* was chosen to validate our candidate gene because this flagellated protozoan has an axonemal organisation similar to human with a 9+2 structure. Moreover, the *T. brucei* flagellum also contains the para-axonemal PFR structure, which plays a role in flagellum motility and serves as a platform for metabolit and signalling enzymes, similar to the mammalian sperm peri-axonemal structures, the fibrous sheath and outer dense fibres. Overall, these results demonstrated that the combined use of exome sequencing and *Trypanosoma* model is an efficient strategy to identify new candidate gene responsible for MMAF phenotype in human.

Altogether, these data demonstrate that CFA9P1 is essential for normal sperm flagellum structure and function in human and the flagellated protist *T. brucei*, and that variants in MAATS1 lead to severe flagellum malformations resulting in asthenoteratozoospermia and primary male infertility. Interestingly, a mouse model inactivated for the *Maats1* gene is also available from the Mouse Genome Informatics (MGI) database (http://www.informatics.jax.org/allele/MGI:6152510). Infertility is reported in male mice but no additional details about the sperm phenotype are currently provided.

**Author affiliations**

1. Univ. Grenoble Alpes, INSERM U1209, CNRS UMR 5209, Institute for Advanced Biosciences, Team Genetics Epigenetics and Therapies of Infertility, Grenoble, France
2. CHU Grenoble Alpes, UM de Génétique Chimique, Grenoble, France
3. Univ. de Bordeaux, Microbiologie Fondamentale et Pathogénicité, CNRS UMR 5224, Bordeaux, France
4. Institut Polytechnique de Bordeaux, Microbiologie Fondateille et Pathogénicité, CNRS UMR 5224, Bordeaux, France
5. CHU Grenoble Alpes, Unité de Génétique Moléculaire : Maladies Héréditaires et Oncologie, Pôle Biologie, Institut de Biologie et de Pathologie, Grenoble, France
6. CHU Grenoble Alpes, UM GI-DPI, Grenoble, France
7. INSERM UMR 1078, Université Bretagne Loire – Université de Brest, Etablissement Français du Sang – Bretagne, Institut de Biologie de Brest, Brest, France
8. Service de Génétique Médicale et Biologie de la Reproduction, Laboratoire de Génétique Moléculaire et Histocompatibilité, CHRU de Brest, Hôpital Morvan, Brest, France
9. INSERM U1205, UFR Chimie Biologie, Univ. Grenoble Alpes, Grenoble, France
10. Laboratoire d’Histoélogie Embryologique - Biologie de la Reproduction, GH Cochin Broca Hôtel Dieu, Assistance Publique-Hôpitaux de Paris, Paris, France
11. Université Paris Descartes, Sorbonne Paris Cité, Faculté de Médecine, Paris, France
12. Polyclinique les Jasmins, Centre d’Aide Médicale à la Procréation, Centre Urbain Nord, Tunis, Tunisia
13. Univ. Grenoble Alpes, CNRS UMR 5525, TICM-IMAG / BCM, Grenoble, France
14. INSERM U1016, Institut Cochin, Paris, France
15. Centre National de la Recherche Scientifique UMR8104, Paris, France

**Acknowledgements**

The authors would like to thank all individuals for their participation. The authors would like to thank Samuel Dean (Oxford University) for the pPOT vector, Nicolas Biteau (Bordeaux University) for the anti-PFR2 antibody and Frédéric Brincaud (Bordeaux University) for the anti-enolase antibody.

**Contributors**

All authors of this manuscript fulfil the criteria of authorship. VS, ED and RZ recruited patients and collected clinical information. CCa, MB, Z-EK, GLK, CG, IG, YF and GP performed molecular analyses. GM and JB performed IF experiments. NT-M performed the bioinformatic analysis. DD, DRB and MB performed the Trypanosoma study. AT and ED performed the TEM experiments. GM, JB, AT, CA, MB, PPR and CC performed data analysis and interpretation. GM, JB, AT, CA, MB, PPR and CC designed the study and wrote the manuscript.

**Funding**

This work was mainly supported by the following grants by the Agence Nationale de la Recherche (ANR) (MUCOFERTIL 12-BSFV1–0011 and NASFLAGELLA 14-CE15), the LabEx ParaFrap (ANR-11-LABX-0024), the ‘Whole genome sequencing of subjects with Flagellar Growth Defects (FGD)’ financed by the fondation maladies rares (FMR) for the programme Sénquévac à haut débit 2012. The electron microscopy analysis was done in the Bordeaux Imaging Centre, a service unit of the CNRS/INSERM and Bordeaux University, member of the national infrastructure France Bionanaging supported by the ANR (ANR-19-INSB-04).

**Competing interests**

None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available on reasonable request.

**ORCID iD**

Charles Coutton http://orcid.org/0000-0002-8873-8098

**REFERENCES**


SUPPLEMENTARY MATERIALS AND METHODS

Exome Sequencing and bioinformatic analysis
Genomic DNA was isolated from saliva using the Oragene DNA extraction kit (DNAgenotech®, Ottawa, Canada). Coding regions and intron/exon boundaries were sequenced after enrichment with Agilent kits (Agilent Technologies, Wokingham, UK) on four distinct platforms: the IGBMC sequencing platform, Strasbourg, France (agilent 50M, HiSeq2000), the Genoscope, Evry, France (agilent v5, HiSeq2000); IntegraGen, Evry, France (Agilent CR, HiSeq4000), and Novogen, Hong Kong, China (agilent v6, HiSeqX).

All datasets were reanalyzed using our updated bioinformatics pipeline, following the GATK 3.7 best practice recommendations. Briefly: sequencing reads were aligned against the GRCh38 reference genome using BWA-MEM 0.7.17, resulting BAM files were cleaned and sorted and duplicates were marked using Picard 2.7.1, and variants were called using GATK HaplotypeCaller to produce GVCF files. The GVCF files were then merged with a custom script (https://github.com/bcm-uga/mergeGVCFs) to obtain a single GVCF, and a final VCF file was produced with GATK GenotypeGVCFs. The most promising candidate variants were identified using an in-house bioinformatics pipeline (Perl scripts available upon request), as follows. Low-quality variant calls (DP<5, GQ<20, or less than 15% of reads supporting the ALT allele) were removed. We used Variant Effect Predictor (VEP version 92)¹ to annotate the variants and predict their impact. Variants with a minor allele frequency greater than 5% in the NHLBI ESP6500 [Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA], greater than 3% in 1000 Genomes Project phase 3 datasets, or greater than 1% in gnomAD v2.0 (http://gnomad.broadinstitute.org/) were discarded. We also compared these variants to an in-house database of 193 control exomes enriched for North African subjects (n=153) corresponding to the geographical origin of the majority of subjects.
in this study and which are under-represented in public SNP databases. All variants present in homozygous state in this control database were excluded. We then only retained variants impacting splice donor or acceptor sites or causing frameshifts, in-frame insertions or deletions, stop gain, stop loss or missense variants except those scored as “tolerated” by SIFT (sift.jcvi.org) and as “benign” by Polyphen-2 (genetics.bwh.harvard.edu/pph2). Gene expression data from the Genotype-Tissue Expression (GTEx) Project were added and used to strengthen the likely implication of the selected candidate variants.

**Quantitative real-time RT-PCR (RT-qPCR) analysis**

A panel of 10 organs was used for experiments: placenta, lung, pancreas, testis, trachea, kidney, whole brain, skeletal muscle, liver and heart. Each sample was assayed in triplicate for each gene on a StepOnePlus (LifeTechnologies®) with Power SYBR®Green PCR Master Mix (Life Technologies®). The PCR cycle was as follows: 10 min at 95°C, 1 cycle for enzyme activation; 15 s at 95°C, 60 s at 60°C with fluorescence acquisition, 40 cycles for the PCR. RT-qPCR data were normalized using the two reference housekeeping genes *RPL6* and *RPL27* for human with the \( \Delta \Delta C_t \) method\(^2\). The \( 2^{-\Delta \Delta C_t} \) value was set at 1 in brain cells, resulting in an arbitrary expression of 1. The efficacy of primers was checked using a standard curve. Melting curve analysis was used to confirm the presence of a single PCR product. Statistics were performed using a two-tailed \( t \)-test on Prism 4.0 software (GraphPad, San Diego, CA) to compare the relative expression of *MAATS1* transcripts in several organs. ***\( P<0.001 \).***

**Minigene assays**

All recombinant plasmid constructs were produced as described previously\(^3\). The wild-type *MAATS1* plasmid construct (pSplicePOLR2G–MAATS1_WT) was generated by cloning sequence of exon 1, 5’UTR, and the intronic flanking regions of *MAATS1* into the
pSplicePOLR2G vector. Generation of the pSplicePOLR2G vector was described previously⁴. The c.124G>C; p.(Asp42His) variant was also introduced in the pSplicePOLR2G–MAATS1_c.124G>C vector by using the QuickChange Site-Directed mutagenesis kit according to the manufacturer’s instructions (Agilent Technologies, Santa Clara, CA). Sequencing analyses were performed to check the integrity of all of the plasmid constructs.

The pSplicePOLR2G vector was digested with EcoRI (New England Biolabs) at 37°C overnight, gel-purified using the NucleoSpin Extract II system (Macherey-Nagel) and stored at 20°C until use. MAATS1 inserts were amplified with the proofreading HotStar Hi-Fi DNA polymerase (Qiagen). Thirty-five cycles of PCR were performed as follows: denaturation at 94°C for 15 s, annealing at 58°C for 60 s, and extension at 72°C for 60 s. In-Fusion cloning reactions were done following the manufacturer's recommendations (Clontech).

HEK293T cells from the American Type Culture Collection were incubated at 37°C in a 5% CO2 humidified atmosphere and propagated in Dulbecco’s Modified Eagle Medium (DMEM; Lonza, Newington, NH) supplemented with 10% fetal bovine serum and 5 µg/mL plasmocin (InvivoGen, San Diego, CA). Cells were transiently transfected using Lipofectamine 2000 (Invitrogen) according to the manufacturer’s recommendations and using a 3:1 transfection reagent (µL)/plasmid DNA ratio (µg).

Forty eight hours after transfection, total RNA was extracted and purified with a RNA extraction kit (RNaseasyMini Kit, Qiagen) and eluted with RNasefree water. One microgram of total RNA served as a template for a subsequent RT-PCR. RT-PCR products were loaded on an agarose gel for sizing and gel-purified for direct sequencing.

Transmission Electron Microscopy analysis of human sperm cells

Human sperm cells were fixed by incubation in 0.1 M phosphate buffer pH 7 supplemented with 3% glutaraldehyde (Grade I; Sigma-Aldrich Co. Saint-Louis, MO, USA) for 2 h at room
temperature. The samples were washed twice in PBS and resuspended in 0.2 M sodium cacodylate buffer. The samples were then post-fixed by incubation with 1% osmium tetroxide (Electron Microscopy Sciences, Hatfield, UK), after which they were dehydrated by immersion in a graded series of alcohol solutions and embedded in Epon resin (Polysciences Inc., Warrington, USA). Semi-thin sections were cut and stained with toluidine blue-Azur II. Ultra-thin sections (90 nm) were cut with a Reichert Ultracut S ultramicrotome (Reichert-Jung AG, Wien, Austria) and were then stained with uranyl acetate and lead citrate. Sections were analyzed with a JEOL 1011 microscope and digital images were acquired with a Gatan Erlangshen CCD camera and Digital Micrograph software.

Immunostaining in human sperm cells

Sperm cells were fixed in phosphate-buffered saline (PBS)/4% paraformaldehyde for 1 min at room temperature. After washing in 1 ml PBS, the sperm suspension was spotted onto 0.1% poly L-lysine pre-coated slides (Thermo Scientific). After attachment, sperm were permeabilized with 0.1% (v/v) Triton X-100 – DPBS (Triton X-100; Sigma-Aldrich) for 5 min at RT. Slides were then blocked in 5% normal serum–DPBS (normal goat or donkey serum; Gibco, Invitrogen) and incubated overnight at 4°C with the primary antibodies. The list of primary antibodies is described in Table S3. Washes were performed with 0.1% (v/v) Tween 20–DPBS, followed by 1 h incubation at room temperature with secondary antibodies. Highly cross-adsorbed secondary antibodies (Dylight 488 and Dylight 549, 1:1000) were from Jackson Immunoresearch®. Appropriate controls were performed, omitting the primary antibodies. Samples were counterstained with DNA was counterstained with DAPI II and mounted with DAKO mounting media (Life Technology). Fluorescence images were captured with a confocal microscope (Zeiss LSM 710).
SUPPLEMENTARY REFERENCES


SUPPLEMENTARY FIGURES

**Figure S1.** Relative mRNA Expression of human *MAATS1* transcripts. RT-qPCR was performed with cDNAs from various human tissues purchased from Life Technologies®. A panel of 10 organs was used for experiments: placenta, lung, pancreas, testis, trachea, kidney, whole brain, skeletal muscle, liver and heart. Each sample was assayed in triplicate for each gene on a StepOnePlus (LifeTechnologies®) with Power SYBR®Green PCR Master Mix (Life Technologies®). The PCR cycle was as follows: 10 min at 95°C, 1 cycle for enzyme activation; 15 s at 95°C, 60 s at 60°C with fluorescence acquisition, 40 cycles for the PCR. RT-qPCR data were normalized using the two reference housekeeping genes *RPL6* and *RPL27* for human with the $-\Delta\Delta C_t$ method. The 2$^{-\Delta\Delta C_t}$ value was set at 1 in brain cells, resulting in an arbitrary expression of 1. The efficacy of primers was checked using a standard curve. Melting curve analysis was used to confirm the presence of a single PCR product. Statistics were performed using a two-tailed *t*-test on Prism 4.0 software (GraphPad, San Diego, CA) to compare the relative expression of *MAATS1* transcripts in several organs. ***$P<0.001$.**
**Figure S2** Alignment of CFAP91 sequences from several orthologs around human amino-acid 42 (marked by a blue square).
Figure S3 In silico prediction of 5' splice sites in MAATS1 exon 1 by Alamut v2.10. Wild-type sequence (top) and c.124G>C; p.(Asp42His) (NM_033364.3) (bottom). Black arrows indicate the preferential 5' donor splice site, respectively.
**Figure S4** Functional analysis of the c.124G>C; p.(Asp42His) (NM_033364.3) variant in a splicing minigene assay. (A) Plasmid constructs used in the study. The POLR2G gene region consisting of exons 1 to 3 (black boxes) was cloned into the pcDNA3.1 vector. The wild-type MAATS1 plasmid construct (pSplicePOLR2G–MAATS1_WT) was generated by cloning sequence of exon 1, 5’UTR, and the intronic flanking regions of MAATS1 (grey boxes) into the pSplicePOLR2G vector. The c.124G>C; p.(Asp42His) variant was introduced in the plasmid pSplicePOLR2G–MAATS1_c.124G>C. (B) Agarose gel analysis of the RT-PCR products obtained from transcripts generated from the pSplicePOLR2G–MAATS1_WT and pSplicePOLR2G–MAATS1_c.124G>C minigene constructs. (C) Sequencing of the purified RT-PCR products. c.124G>T causes partial intronic retention of 49 bp (D) Schematic representation of the alternative splicing event.
Figure S5 Outer dynein arms are not affected by MAATS1 variants. Sperm cells from a fertile control individual and MAATS1_3 and MAATS1_4 patients stained with anti-DNAI2 (H00064446-M01, Abnova, mouse, 1:400, red), a protein located in the outer dynein arm, and anti-acetylated tubulin (PA5-19489, ThermoFisher, rabbit, 1:500, green) antibodies. DNA was counterstained with DAPI II. DNAI2 immunostaining is present throughout the flagellum with no differences between control and patients. Scale bars: 10 µm.
**Figure S6** Inner dynein arms are not affected by *MAATS1* variants. Sperm cells from a fertile control individual and MAATS1<sub>3</sub> and MAATS1<sub>4</sub> patients stained with anti-DNALI1 (HPA028305, Sigma-Aldrich, rabbit, 1:100, green), a protein located in the inner dynein arm, and anti-acetylated tubulin (32-2500, ThermoFisher, mouse, 1:1000, red) antibodies. DNA was counterstained with DAPI II. DNALI1 immunostaining is present throughout the flagellum with no differences between control and *MAATS1* patients. Scale bars: 10 µm.
Figure S7 WDR66 immunostaining is absent from MAATS1 patients. Sperm cells from a fertile control individual and MAATS1\textsubscript{3} and MAATS1\textsubscript{4} patients stained with anti-WDR66 (HPA040005, Sigma-Aldrich, rabbit, 1:50, green) and anti-acetylated tubulin (32-2500, ThermoFisher, mouse, 1: 1000, red) antibodies. DNA was counterstained with DAPI II. Contrary to the control, the WDR66 immunostaining is not detectable in the sperm flagellum from MAATS1 patients. Scale bars: 10 µm.
**SUPPLEMENTARY TABLES**

**Table S1** Primer sequences used for Sanger sequencing verification of *MAATS1* variants in human.

<table>
<thead>
<tr>
<th>Primer names</th>
<th>Primer sequences (5’-3’)</th>
<th>Tm</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAATS1_F_ex1</td>
<td>CAGTAACCATCGAGGAGCCC</td>
<td>60°C</td>
</tr>
<tr>
<td>MAATS1_R_ex1</td>
<td>CCAGAGGGTCTGTTTCGTT</td>
<td></td>
</tr>
<tr>
<td>MAATS1_F_ex6</td>
<td>CTACTGTGGGCACCTGACTG</td>
<td>60°C</td>
</tr>
<tr>
<td>MAATS1_R_ex6</td>
<td>ACATACATAGGAATGTGCATTTCAGT</td>
<td></td>
</tr>
</tbody>
</table>

Tm: melting temperature

**Table S2** Primers used for RT-qPCR of *MAATS1* in human.

<table>
<thead>
<tr>
<th>Primer names</th>
<th>Primer sequences (5’-3’)</th>
<th>Tm</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAATS1_RT_F</td>
<td>AAGCTACCCCTGATTCGCAGC</td>
<td>60°C</td>
</tr>
<tr>
<td>MAATS1_RT_R</td>
<td>ATGTCCCTTCCATCCCGAC</td>
<td></td>
</tr>
<tr>
<td>RPL6_RT_F</td>
<td>TCCATTGCAGAGCAAAAAA</td>
<td>56°C</td>
</tr>
<tr>
<td>RPL6_RT_R</td>
<td>TACGGAGCAGCGCAAGAT</td>
<td></td>
</tr>
<tr>
<td>RPL27_RT_F</td>
<td>AGAGTACCTTGTTGGGCAT</td>
<td>56°C</td>
</tr>
<tr>
<td>RPL27_RT_R</td>
<td>TGATGGCACCTCAGATCGC</td>
<td></td>
</tr>
</tbody>
</table>

Tm: melting temperature
**Table S3** List of primary antibodies used in immunofluorescence experiments in human.

<table>
<thead>
<tr>
<th>Primary antibodies</th>
<th>Reference</th>
<th>Species</th>
<th>Protein localization</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Human experiments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNAI2</td>
<td>Abnova® H00064446-M01</td>
<td>mouse</td>
<td>ODA</td>
<td>1/400</td>
</tr>
<tr>
<td>DNALI1</td>
<td>Sigma-Aldrich® HPA028305</td>
<td>rabbit</td>
<td>IDA</td>
<td>1/100</td>
</tr>
<tr>
<td>RSPH1</td>
<td>Sigma-Aldrich® HPA017382</td>
<td>rabbit</td>
<td>RS</td>
<td>1/100</td>
</tr>
<tr>
<td>WDR66</td>
<td>Sigma-Aldrich® HPA040005</td>
<td>rabbit</td>
<td>CSC</td>
<td>1/50</td>
</tr>
<tr>
<td>SPAG6</td>
<td>Sigma-Aldrich® HPA038440</td>
<td>rabbit</td>
<td>CPC</td>
<td>1/500</td>
</tr>
<tr>
<td>Acetylated tubulin</td>
<td>ThermoFisher® 32-2500</td>
<td>mouse</td>
<td>Microtubules</td>
<td>1/1000</td>
</tr>
<tr>
<td>Acetylated tubulin</td>
<td>ThermoFisher® PA5-19489</td>
<td>rabbit</td>
<td>Microtubules</td>
<td>1/500</td>
</tr>
</tbody>
</table>

Abbreviations are as follows: ODA, Outer Dynein Arms; IDA, Inner Dynein Arms; RS, Radial Spokes; FS, Fibrous sheath; CPC, Central Pair Complex; CSC, calmodulin- and spoke-associated complex
**Table S4** Endogenous tagging and RNAi primer sequences for T. brucei experiments.

<table>
<thead>
<tr>
<th>Endotag</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1340</td>
<td>GCGTGTGCAAGGCGTGTGTTGCCCTTCAACGTACTTTATTTTTTACCTTACTTCTCGTACCATAAAGTCGCT</td>
</tr>
<tr>
<td></td>
<td>ACCTCTTACgtataatgcagacctgctgc</td>
</tr>
<tr>
<td>1387</td>
<td>CGGTCACGCCCTCCACTTCCGCCGTGAAACCATTGCAACCCAACCGACGGCTGGTTTCTGCTCGTTTC</td>
</tr>
<tr>
<td></td>
<td>GGTAGTACATactacccgatcctgatcc</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RNAi</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1342</td>
<td>GTGGCGGCCGCGCGCTCTAGAATGTACTACCGAACGCCAGCAG</td>
</tr>
<tr>
<td>1343</td>
<td>AGCCCGGCCGCCGCCCTCGAGacgtcccatccccacaacgcttttg</td>
</tr>
</tbody>
</table>
Table S5 List of primary antibodies used in immunofluorescence and Western-blot experiments in *Trypanosoma*.

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Source</th>
<th>dilution IF</th>
<th>dilution WB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-HA.11</td>
<td>Biolegend</td>
<td>1:1,000</td>
<td>1,1000</td>
</tr>
<tr>
<td>Anti-PFR2</td>
<td>Ref1</td>
<td>1:2,000</td>
<td></td>
</tr>
<tr>
<td>Anti-TbSAXO</td>
<td>Ref2</td>
<td>1:1,000</td>
<td></td>
</tr>
<tr>
<td>Anti-enolase</td>
<td>Ref3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>anti-mouse HRP-</td>
<td>Jackson 115-035-062</td>
<td></td>
<td></td>
</tr>
<tr>
<td>conjugated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anti-rabbit HRP-</td>
<td>Sigma A91-69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>conjugated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anti-mouse FITC-</td>
<td>Sigma F-2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>conjugated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anti-rabbit Alexa594-</td>
<td>Thermofischer A11012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>conjugated</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Table S6 MAATS1 variants identified by exome sequencing in the cohort of 167 MMAF individuals.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variant genomic coordinates (GRCh38)</th>
<th>Transcript</th>
<th>cDNA Variation</th>
<th>Amino acid variation</th>
<th>Individuals</th>
<th>Origin</th>
<th>Allelic status</th>
<th>Minor allele frequency (gnomAD)</th>
<th>dbSNP rs</th>
<th>SIFT</th>
<th>PolyPhen</th>
<th>SPICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAATS1</td>
<td>chr3:119715744 NM_033364.3</td>
<td>NM_033364.3</td>
<td>c.682+1G&gt;A</td>
<td>Splice donor variant</td>
<td>MAATS1_1</td>
<td>North Africa</td>
<td>Homozygous</td>
<td>3.98×10⁻⁶</td>
<td>rs147597066</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MAATS1</td>
<td>chr3:119715744 NM_033364.3</td>
<td>NM_033364.3</td>
<td>c.682+1G&gt;A</td>
<td>Splice donor variant</td>
<td>MAATS1_2</td>
<td>North Africa</td>
<td>Homozygous</td>
<td>3.98×10⁻⁶</td>
<td>rs147597066</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MAATS1</td>
<td>chr3:119703222 NM_033364.3</td>
<td>NM_033364.3</td>
<td>c.124G&gt;C</td>
<td>p.Asp42His</td>
<td>MAATS1_3</td>
<td>North Africa</td>
<td>Homozygous</td>
<td>8.42×10⁻⁵</td>
<td>rs149348782</td>
<td>0</td>
<td>0.9999</td>
<td>0.99196</td>
</tr>
<tr>
<td>MAATS1</td>
<td>chr3:119703222 NM_033364.3</td>
<td>NM_033364.3</td>
<td>c.124G&gt;C</td>
<td>p.Asp42His</td>
<td>MAATS1_4</td>
<td>North Africa</td>
<td>Homozygous</td>
<td>8.42×10⁻⁵</td>
<td>rs149348782</td>
<td>0</td>
<td>0.9999</td>
<td>0.99196</td>
</tr>
<tr>
<td>MAATS1</td>
<td>chr3:119703222 NM_033364.3</td>
<td>NM_033364.3</td>
<td>c.124G&gt;C</td>
<td>p.Asp42His</td>
<td>MAATS1_5</td>
<td>North Africa</td>
<td>Homozygous</td>
<td>8.42×10⁻⁵</td>
<td>rs149348782</td>
<td>0</td>
<td>0.9999</td>
<td>0.99196</td>
</tr>
</tbody>
</table>


SUPPLEMENTARY VIDEOS

**Video S1** Illustration of cell mobility of WT cells.

**Video S2** Illustration of cell mobility of *TbMAATS1*-RNAi cells induced for 8 days.