REPORT

Bi-allelic Mutations in *M1AP* Are a Frequent Cause of Meiotic Arrest and Severely Impaired Spermatogenesis Leading to Male Infertility

Margot J. Wyrwoll,[1,25] Şehime G. Temel,[2,25] Liina Nagirnaja,[3] Manon S. Oud,[4] Alexandra M. Lopes,[5,6] Godfried W. van der Heijden,[4,7] James S. Heald,[8] Nadja Rotte,[1,9] Joachim Wistuba,[9] Marius Wöste,[10] Susanne Ledig,[1] Henrike Krenz,[10] Roos M. Smits,[7] Filipa Carvalho,[6,11] João Gonçalves,[12,13] Daniela Fietz,[14] Burcu Türkgenç,[15] Mahmut C. Ergören,[16] Murat Çetinkaya,[17] Murad Başar,[18] Semra Kahraman,[19] Kevin McEleny,[20] Miguel J. Xavier,[8] Helen Turner,[21] Adrian Pilatz,[22] Albrecht Röpke,[1] Martin Dugas,[10] Sabine Kliesch,[23] Nina Neuhaus,[9] GEMINI Consortium, Kenneth I. Aston,[24] Donald F. Conrad,[3] Joris A. Veltman,[4,8] Corinna Friedrich,[1,25] and Frank Tüttelmann[1,25,*]

Male infertility affects ~7% of men, but its causes remain poorly understood. The most severe form is non-obstructive azoospermia (NOA), which is, in part, caused by an arrest at meiosis. So far, only a few validated disease-associated genes have been reported. To address this gap, we performed whole-exome sequencing in 58 men with unexplained meiotic arrest and identified the same homozygous frameshift variant c.676dup (p.Trp226LeufsTer4) in *M1AP,* encoding meiosis 1 associated protein, in three unrelated men. This variant most likely results in a truncated protein as shown *in vitro* by heterologous expression of mutant *M1AP*. Next, we screened four large cohorts of infertile men and identified three additional individuals carrying homozygous c.676dup and three carrying combinations of this and other likely causal variants in *M1AP*. Moreover, a homozygous missense variant, c.1166C>T (p.Pro389Leu), segregated with infertility in five men from a consanguineous Turkish family. The common phenotype between all affected men was NOA, but occasionally spermatids and rarely a few spermatozoa in the semen were observed. A similar phenotype has been described for mice with disruption of *M1ap*. Collectively, these findings demonstrate that mutations in *M1AP* are a relatively frequent cause of autosomal recessive severe spermatogenic failure and male infertility with strong clinical validity.

Around 7% of all men in Western societies experience infertility,[1] which is primarily diagnosed by semen analysis comprising sperm concentration and count as the most relevant parameters. More than 10% of all infertile men exhibit azoospermia[2]—the absence of spermatozoa in the ejaculate. Azoospermia constitutes the most challenging and clinically severe form of male infertility and is further classified into obstructive azoospermia (OA) with normal spermatogenesis and non-obstructive azoospermia (NOA) due to impaired spermatogenesis. In some men, a few spermatozoa can be identified after centrifugation of the semen, which is denoted as cryptozoospermia. From the biological point of view, NOA and cryptozoospermia are closely related, just as they are in their clinical implications, i.e., virtually no chance of natural conception.[3]

The variable spermatogenic impairment in these men correlates to a diverse spectrum of testicular histological phenotypes. This spectrum includes Sertoli cell-only (SCO), over maturation arrest, and hypospermatogenesis, and all of these can be complete, focal, or mixed. Maturation arrest most frequently presents as meiotic arrest in which spermatocytes are the most advanced germ cell

[1]Institute of Human Genetics, University of Münster, 48149 Münster, Germany; [2]Bursa Uludag University, Faculty of Medicine, Department of Medical Genetics & Department of Histology & Embryology & Health Sciences Institute, Department of Translational Medicine, 16059 Bursa, Turkey; [3]Division of Genetics, Oregon National Primate Research Center, Oregon Health & Science University, Beaverton, OR 97006, USA; [4]Department of Human Genetics, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, 6525 Nijmegen, the Netherlands; [5]Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), 4200-804 Porto, Portugal; [6]Instituto de Investigação e Inovação em Saúde (i3s), Universidade do Porto, 4099-002 Porto, Portugal; [7]Department of Obstetrics and Gynecology, Radboud University Medical Center, 6525 Nijmegen, the Netherlands; [8]Biosciences Institute, Faculty of Medical Sciences, Newcastle University, NE2 4HH Newcastle upon Tyne, UK; [9]Centre of Reproductive Medicine and Andrology, Institute of Reproductive Medicine, University of Münster, 48149 Münster, Germany; [10]Institute of Medical Informatics, University of Münster, 48149 Münster, Germany; [11]Serviço de Genética, Departamento de Patologia, Faculdade de Medicina da Universidade do Porto, 4099-002 Porto, Portugal; [12]Departmento de Genética Humana, Instituto Nacional de Saúde Dr. Ricardo Jorge, 1649-016 Lisboa, Portugal; [13]ToxOmics - Centro de Toxicogenómica e Saúde Humana, Nova Medical School, 1169-056 Lisboa, Portugal; [14]Institute of Veterinary Anatomy, Histology and Embryology, Justus Liebig University, 35392 Gießen, Germany; [15]University of Acibadem, Acibadem Genetic Diagnostic Centre, 34662 Istanbul, Turkey; [16]Near East University, Faculty of Medicine, Department of Medical Biology, 99138 Nicosia, Cyprus; [17]Istanbul Memorial Hospital, Assisted Reproductive Technologies and Reproductive Genetics Centre, 34385 Istanbul, Turkey; [18]Istanbul Memorial Hospital, Department of Urology & Andrology, 34385 Istanbul, Turkey; [19]Istanbul Memorial Hospital, Assisted Reproductive Technologies and Reproductive Genetics, 34385 Istanbul, Turkey; [20]Newcastle Fertility Centre, The Newcastle upon Tyne Hospitals NHS Foundation Trust, NE1 4EP Newcastle upon Tyne, UK; [21]Department of Cellular Pathology, The Newcastle upon Tyne Hospitals NHS Foundation Trust, NE1 4LP Newcastle upon Tyne, UK; [22]Clinic for Urology, Pediatric Urology and Andrology, Justus Liebig University, 35392 Gießen, Germany; [23]Centre of Reproductive Medicine and Andrology, Department of Clinical and Surgical Andrology, University Hospital Münster, 48149 Münster, Germany; [24]Andrology and IVF Laboratories, Department of Surgery, University of Utah School of Medicine, Salt Lake City, UT 84132, USA
[25]These authors contributed equally to this work
*Correspondence: frank.tuettelmann@ukmuenster.de
https://doi.org/10.1016/j.ajhg.2020.06.010.

types in the testes. If germ cell arrest is complete, no mature spermatozoa develop; testicular biopsy and sperm extraction (TESE) will not be successful and assisted reproductive technology (ART) will not be possible.

In a large fraction of severely disturbed spermatogenesis, a genetic origin is assumed,[4] and affected men are routinely screened for chromosomal aberrations and Y chromosome azoospermia factor (AZF) microdeletions. Yet these diagnostic tests only establish a causal diagnosis in 15%–20% of azoospermia-affected individuals.[2] Recently, monogenic alterations associated with germ cell arrest in human males have been described.[5,6] However, according to a standardized clinical validity assessment, the X chromosome gene *TEX11* (MIM: 300311) currently remains the only one in which variants are associated with male infertility with strong evidence.[7] Given the large number of genes in which variants are known to cause meiotic arrest in mice, the vast majority of mutations causing this phenotype in humans are yet to be identified.

To this end, we first screened the exomes of well-characterized men with complete bilateral meiotic arrest and identified bi-allelic loss-of-function (LoF) variants in the gene encoding meiosis 1 associated protein (*M1AP*) in three unrelated men. In mice, *M1ap* is primarily expressed in male germ cells throughout spermatogenesis, and its knockout leads to infertility due to meiotic arrest and severe oligozoospermia.[8,9] Our subsequent analyses in four independent cohorts and a consanguineous Turkish family, as well as *in vitro* analyses of a recurring *M1AP* frameshift variant, corroborated that disruption of M1AP is associated with a variable spectrum of severely impaired spermatogenesis, mostly at meiosis and resulting in azoospermia, but also compatible with sparse postmeiotic germ cell development and retrieval of sperm in some instances.

We originally selected 64 azoospermic but otherwise healthy males who attended the Centre of Reproductive Medicine and Andrology (CeRA), University Hospital Münster (n = 51) or the Clinic for Urology, Pediatric Urology and Andrology, Gießen (n = 13) for couple infertility. All men were diagnosed with complete bilateral germ cell arrest at the spermatocyte stage after the evaluation of at least 100 seminiferous tubules in tissue sections of both testes accompanied by a negative TESE outcome, i.e., no sperm could be recovered. This is a subset of all individuals included in our large-scale Male Reproductive Genomics (MERGE) study, which currently comprises 735 men with lacking or severe quantitatively impaired spermatogenesis and 53 individuals with normal spermatogenesis (OA and controls) (Figure S1). Specifically, we performed whole-exome sequencing (WES; for details, see Supplemental Methods) in 569 men with NOA, 116 with cryptozoospermia, and 50 with severe oligozoospermia (sperm concentration < 5 M/mL). Chromosomal aberrations and AZF deletions were excluded in this and all other cohorts and subjects (detailed below). All participants gave written informed consent and the study protocol was approved by the respective ethics committees and institutional review boards (details in Supplemental Information).

We identified likely causal variants in the three genes *TEX11*, *STAG3*, and *SYCP2* in six of the men with complete meiotic arrest.[5,10,11] The WES data of the remaining 58 men were filtered for rare (minor allele frequency [MAF] < 0.01 according to the Genome Aggregation Database[12] [gnomAD]) bi-allelic LoF variants. Affected genes were prioritized with regard to the level of expression in the testes and previous evidence for an association with infertility in either human or model species (Figure S1). The highest-ranked gene was *M1AP* because three unrelated men (M330, M864, and M1792, Table 1) carried the same homozygous LoF variant (c.676dup, MAF = 0.0021, no homozygotes in gnomAD[12]), the *M1AP* mRNA displayed the highest expression in the testis (according to both the Genotype-Tissue Expression [GTEx] project[13] and the Human Protein Atlas [HPA]), and it was shown to play a crucial role in spermatogenesis in mice.[8,9] The c.676dup variant in *M1AP* was confirmed by Sanger sequencing (GenBank: NM_138804.4, the longest isoform with highest testis expression; for primer sequences see Table S1) in all affected men from the MERGE cohort. Testis biopsies of all three individuals were collected for TESE and research use. These were fixed in Bouin's solution and embedded in paraffin, and subsequently, sections were stained with periodic acid-Schiff (PAS) as previously described.[14] These were re-analyzed to confirm complete meiotic arrest (Figure 1). Because DNA from the three individuals' parents was not available, and to exclude a hemizygous deletion on the other allele, quantitative PCR (qPCR) of *M1AP*'s exon 5 was performed on gDNA (primers and conditions in Table S1 and the Supplemental Methods). This excluded an intragenic deletion in all individuals (Table S2). No regions of homozygosity (ROHs) involving *M1AP* were detected for any of the affected men, rendering consanguinity of their parents unlikely. We also did not notice evidence for consanguinity between the men (analysis by H3M2 and vcftools algorithms,[15] data not shown).

By screening the complete MERGE cohort, we identified an additional man (M2062) carrying the same homozygous LoF variant, c.676dup (Figure S2), and another man (M1943) carrying c.676dup in combination with a second rare missense variant, c.797G>A (p.Arg266Gln). This missense variant is consistently predicted as pathogenic by all *in silico* prediction programs (PolyPhen-2, SIFT, MutationTaster, and HOPE[17]). Both men had varying azoospermia and cryptozoospermia in repeated semen analyses and did not undergo a testicular biopsy. In contrast, no individuals with two rare *M1AP* variants predicted as pathogenic via *in silico* programs were observed in the remaining MERGE cohort or in the individuals with spermatogenic impairment and other testicular phenotypes, such as SCO (n = 213), or normal spermatogenesis (OA or controls, n = 53).

Next, through collaborations established within the International Male Infertility Genomics Consortium

Table 1. Genetic and Clinical Data of Infertile Men with *M1AP* Variants

Individual	Age, Origin	*M1AP* Variant	Fertility Parameters	Testicular Phenotype, TESE Outcome
M330	38 years, Germany	c.[676dup];[676dup], p.[Trp226LeufsTer4];[Trp226LeufsTer4]	FSH, 9; LH, 5.3; T, 14.6; TV, 17/23; azoospermia	meiotic arrest (0/0% tubules with ES, 0/0% RS, 91/99% SC, 6/1% SG, 3/0% SCO, 0/0% TS), TESE negative
M864	41 years, Germany	c.[676dup];[676dup], p.[Trp226LeufsTer4];[Trp226LeufsTer4]	FSH, 4.7; LH, 1.5; T, 9.6; TV, 19/26; azoospermia	meiotic arrest (0/0% tubules with ES, 0/0% RS, 71/91% SC, 10/4% SG, 17/1% SCO, 2/4% TS), TESE negative
M1792	36 years, Germany	c.[676dup];[676dup], p.[Trp226LeufsTer4];[Trp226LeufsTer4]	FSH, 7.8; LH, 5.1; T, 10.1; TV, 15/15; azoospermia	meiotic arrest (0/0% tubules with ES, 0/0% RS, 96/97% SC, 0/2% SG, 3/1% SCO, 1/0% TS), TESE negative
M1943	43 years, Croatia	c.[676dup];[797G>A], p.[Trp226LeufsTer4];[Arg266Gln]	FSH, 5.5; LH, 3.2; T, 43.6; TV, 28/20; cryptozoospermia or azoospermia[a]	N/A
M2062	26 years, Poland	c.[676dup];[676dup], p.[Trp226LeufsTer4];[Trp226LeufsTer4]	FSH, 3.5; LH, 3.6; T, 18.6; TV, 26/23; cryptozoospermia or azoospermia[a]	N/A
Y126	34 years, Portugal	c.[676dup];[949G>A], p.[Trp226LeufsTer4];[Gly317Arg]	FSH, 6.7; LH, 3.2; T, N/A; TV, N/A; azoospermia	maturation arrest at round spermatid stage, (quantification N/A), TESE negative
P86	44 years, Portugal	c.[148T>C];[1289T>C], p.[Ser50Pro];[Leu430Pro]	FSH, N/A; LH, N/A; T, N/A; TV, N/A; azoospermia	dispersed Sertoli cells, some tubules contained only spermatogonia, (quantification N/A), TESE negative
RU01691	41 years, the Netherlands	c.[676dup];[676dup], p.[Trp226LeufsTer4];[Trp226LeufsTer4]	FSH, 5; LH, 2.0; T, 11.3; TV, N/A; azoospermia	predominant meiotic arrest with occasional spermatids, (unilateral TESE: 4% tubules with ES, 5% RS, 88% SC, 2% SG, 0% SCO, 0% TS), TESE positive
MI-0006-P	33 years, UK	c.[676dup];[676dup], p.[Trp226LeufsTer4];[Trp226LeufsTer4]	FSH, 10.4; LH, N/A; T, 15.4; TV, 20/20; azoospermia	predominant meiotic arrest with occasional postmeiotic germ cells (quantification N/A), TESE negative
T1024	28 years, Turkey	c.[1166C>T];[1166C>T], p.[Pro389Leu];[Pro389Leu]	FSH, 8.3; LH, 4.4; T, 8.8; TV, 15/15; azoospermia	maturation arrest at round spermatid stage, (quantification N/A), TESE negative
F1: II-1[16] (Tu et al.)	34 years, China	c.[1435−1G>A];[1435−1G>A], p.?	FSH, 3.65; LH, 3.17; T, 18.24; TV, 12/12; severe oligozoospermia	N/A

Abbreviations are as follows: FSH, follicle-stimulating hormone (IU/L); LH, luteinizing hormone (IU/L); T, testosterone (nmol/L); TV, testicular volume right/left (mL); ES, elongating spermatids; RS, round spermatids; SC, spermatocytes; SG, spermatogonia; SCO, Sertoli cell-only; TS, tubular shadows; N/A, not available. Reference values: FSH 1–7 IU/L, LH 2–10 IU/L, T > 12 nmol/L, TV > 15 mL per testis.
[a]Semen contained none or below 10 spermatozoa/sample on repeated analyses.

(IMIGC), three cohorts of infertile men in whom WES was performed in independent studies (details are provided in the Supplemental Methods) were screened for bi-allelic variants in *M1AP*: 930 men with unexplained NOA from the Genetics of Male Infertility Initiative (GEMINI) study, 283 men with unexplained azoospermia (n = 214) or oligozoospermia (n = 69) who presented at Radboud University Medical Center (Radboudumc, Nijmegen), and 48 men with unexplained azoospermia (n = 36) or oligozoospermia (n = 12) recruited at The Newcastle upon Tyne Hospitals NHS Foundation Trust (Newcastle, UK). From these 1,261 individuals in total, we identified four additional infertile men with likely bi-allelic variants in *M1AP*. Clinical data of all individuals carrying *M1AP* variants is shown in Table 1. The variants, gnomAD frequencies, *in silico* predictions, and classification according to the guidelines of the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP)[18] are given in Table 2 (for detailed variant description and interpretation, see Supplemental Note). On the basis of RNA analyses from public sources (GTEx and HPA), *M1AP* might also be expressed in the bone marrow and other tissues, such as the pituitary. However, no blood-system-related or hormonal abnormalities were noted in any of the affected men reported herein.

Two individuals of Portuguese origin (Y126 and P86) analyzed within the GEMINI study each carried two different variants in *M1AP* (Figure S2). Individual Y126 carried the missense variant c. 949G>A (p.Gly317Arg) and the recurrent frameshift variant c.676dup and had germ cell arrest at the round spermatid stage. Individual P86 had the two missense variants c.148T>C (p.Ser50Pro) and c.1289T>C (p.Leu430Pro), suggesting compound heterozygosity. Testicular histology showed severely disturbed spermatogenesis. Individuals RU01691 from the Netherlands and MI-0006-P from the UK were also homozygous for the frameshift variant c.676dup. RU01691's parents were both heterozygous carriers. His

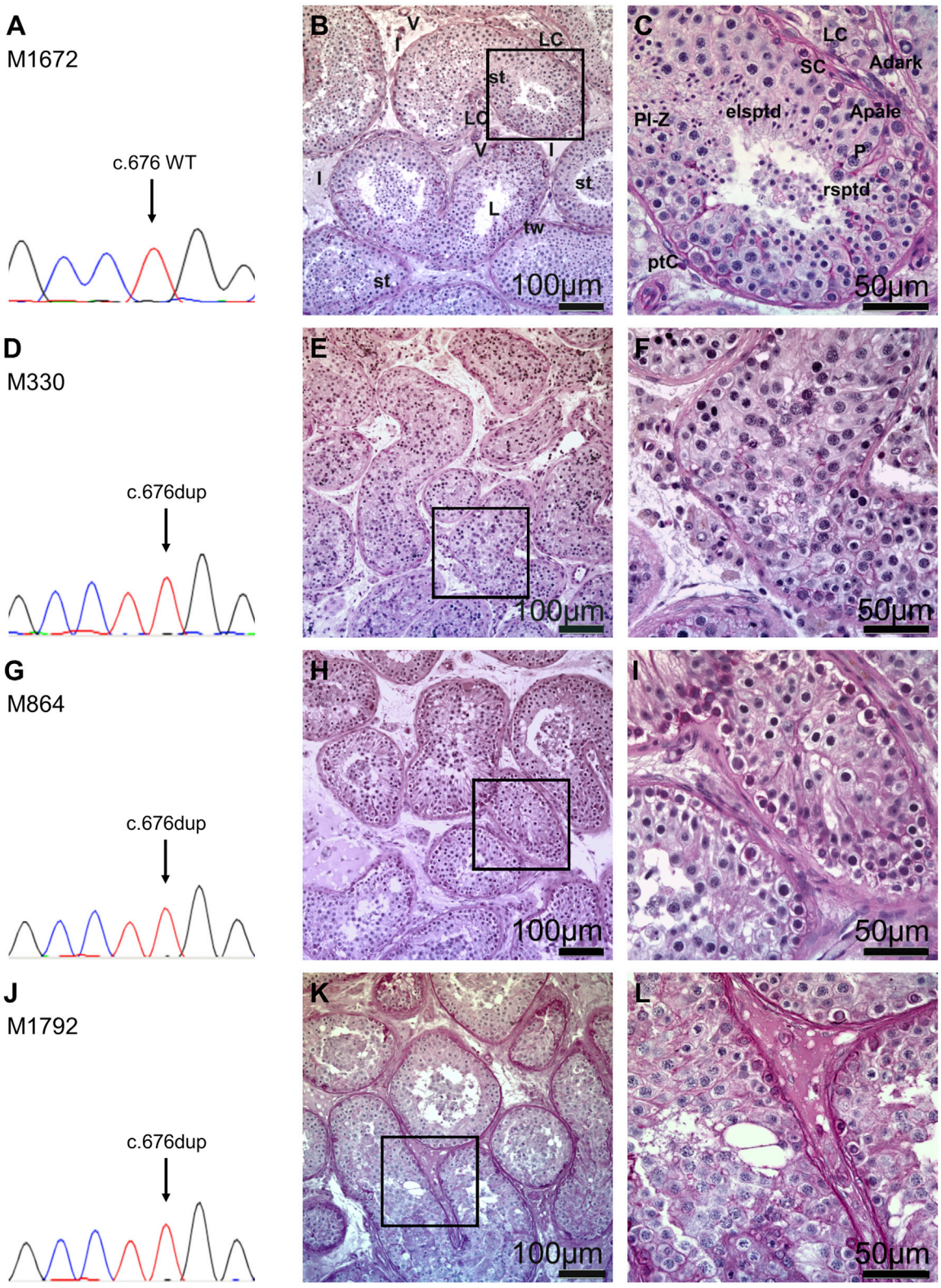


Figure 1. Recurrent Homozygous Variant c.676dup in *M1AP* Leading to Complete Bilateral Meiotic Arrest in Three Individuals from the MERGE Study

(A) Electropherogram with the wild-type sequence of *M1AP* exon 5 (M1672 with obstructive azoospermia).

(B and C) Testicular tissue showing complete spermatogenesis, PAS staining.

(B) Testicular tissues are composed of seminiferous tubules and interstitium. The seminiferous tubules are separated from the interstitial space (I) by tubular walls (tws) formed by myoid peritubular cells and the lamina propria. Inside, the seminiferous epithelium and the lumen (L) are localized. In the interstitium, groups of steroidogenic Leydig cells (LCs) and blood vessels (Vs) are observed. Tubular cross-section showed the regular appearance of a functioning testis exhibiting complete germ cell differentiation.

(legend continued on next page)

testicular biopsy showed bilateral severe hypospermatogenesis with predominantly meiotic arrest; sporadically, spermatids were present, and material was cryopreserved for intracytoplasmic sperm injection (ICSI) but hitherto not used. Individual MI-0006-P had azoospermia and predominant meiotic arrest with rare postmeiotic germ cells (Figure S2).

In parallel, WES was performed in two infertile, azoospermic brothers from a consanguineous Turkish family as well as in their fertile brother. The index T1024 (V.2; Figure 2) presented at Istanbul Memorial Hospital, Assisted Reproductive Technologies and Reproductive Genetics Centre and Bursa Uludag University Faculty of Medicine Hospital, Turkey because of couple infertility. Testicular histology demonstrated maturation arrest at round spermatid stage, and no sperm could be recovered by TESE. The WES data were analyzed focusing on rare homozygous variants shared between both infertile brothers but not found in the fertile brother. The two affected men carried rare homozygous missense variants in the autosomal genes *AMPD2*, *CELSR2*, *CEP164*, and *M1AP* as well as rare hemizygous variants in the X chromosome genes *ATG4A* and *ENOX2*. Of these genes, only *M1AP* has been described in the context of infertility. Both infertile men carried the homozygous missense variant c.1166C>T (p.Pro389Leu) (MAF = 0.00001), which was also found in a heterozygous state in the fertile brother, in *M1AP*. No homozygous individuals with this variant have been described in any public databases, whereas it was found in a homozygous state in three additional infertile males from this family: two third cousins once removed from the maternal side and one second cousin once removed from the paternal side. We did identify both a fertile man and a fertile woman (IV.13 and V.6, respectively, in Figure 2A) as heterozygous carriers of the same variant (example result of Sanger sequencing for subject V.6 shown in Figure 2B).

In a complementary approach, an updated version of the population sampling probability (PSAP) pipeline[19] was used to prioritize potentially causative variants. PSAP models the significance of observing a single person's genotype in comparison to genotype frequencies in unaffected populations. This enabled us to rank all variants per individual by following the prioritization criteria MAF $\leq$ 0.01, CADD $\geq$ 20, and PopScore $\leq$ 0.005.[20] The bi-allelic *M1AP* LoF variants were ranked in the first position for M1792 and in the third position for M330, M864, and M2062 in the discovery cohort (Table S3). The missense variant c.797G>A and the duplication c.676dup of M1943 were ranked in the ninth position under a compound heterozygous recessive disease model. The five men identified in the follow-up analyses exhibited highly ranked *M1AP* variants as well: position seven in RU01691, position four in Y126, position seven in P86, position two in MI-0006-P, and position three in T1024.

The rare but recurring *M1AP* variant c.676dup warranted further analyses. It is located in exon 5 of 11 and causes a frameshift and premature stop codon (p.Trp226LeufsTer4) as confirmed by testicular cDNA sequencing of exon 5 of individual M864 (Figure S3). This results in a truncated protein as shown by heterologous expression of mutated M1AP in HEK293T and subsequent immunoblot analysis displaying a protein band of almost 23 kD (Supplemental Methods and Figure 3), indicating a protein lacking 57% of its normal length. This is in line with the analysis of individual M864's testicular RNA that resulted in an equal band compared to control testis RNA, excluding elimination of the mRNA through nonsense-mediated decay (Figure S3). Still, because of the non-native, ectopic expression of the protein in HEK293T cells, it remains possible that no product is translated by the mutant mRNA *in vivo*. The relevance of the homozygous frameshift variant c.676dup in *M1AP* is further supported by the exceptionally low PSAP-PopScore (9.7×10^{-7}) and the high prioritization (Table S3). Moreover, the expected mode of inheritance for *M1AP* is autosomal recessive according to a general prediction,[21] fitting our observations of bi-allelic variants in the affected men. According to the ACMG-AMP guidelines,[18] this variant is categorized as pathogenic (Table S4).

The fact that the same frameshift variant, c.676dup, was also found in individuals from Croatia, Poland, the Netherlands, UK, and Portugal, suggests that it is relatively prevalent in European populations most likely originating from a founder mutation. In total, we screened 1,996 infertile males across the four cohorts (735, 930, 283, and 48 from the MERGE, GEMINI, Nijmegen, and Newcastle studies, respectively), six of whom were homozygous for c.676dup. In contrast, c.676dup is rarely described in global large databases: gnomAD (v2.1.1: 141,421 individuals corresponding to 282,842 alleles) does not contain any homozygous individuals. Hence, subjects homozygous for c.676dup in *M1AP* are highly significantly overrepresented in our cohort (Fisher's exact $p = 7.2 \times 10^{-12}$; only for the 76,685 males, $p = 2.6 \times 10^{-10}$). To establish the allele frequency of *M1AP* c.676dup in an ancestry-matched control group, we performed Sanger sequencing of exon 5 of *M1AP* in an additional 285 normozoospermic men recruited at the CeRA, Münster. Indeed, five

(C) Detail of B; the tubules are surrounded by the lamina propria and the peritubular cells (ptCs), forming the wall. Within the seminiferous epithelium, somatic Sertoli cells (SCs) are supporting the germ cells differentiating from A spermatogonia (Apale/Adark) via premeiotic spermatocytes (preleptotene to zygotene stage; Pl-Z) into the meiotic pachytene spermatocytes (Ps). After meiosis is completed, haploid round spermatids (rsptds), which mature further into elongated spermatids (elsptds), are formed.
(D–L) Identification of a recurrent homozygous variant in *M1AP* c.676dup (p.Trp226LeufsTer4). Sanger sequencing verified the variant in M330 (D), M864 (G), and M1792 (J), leading to complete bilateral meiotic arrest as indicated by histological examination of testis biopsies (M330 [E, magnified in F], M864 [H, magnified in I], M1792 [K, magnified in L]), which show spermatocytes as the most advanced germ cells in all tubules. Magnified areas and scale bars are indicated.

Table 2. Assessment of *M1AP* Variants

cDNA Change	Protein Change	*In Silico* Prediction for Missense Variants (PolyPhen-2/SIFT/ MutationTaster)	MAF[a] (gnomAD)	MAF (Local Controls)	Conservation	Classification According to ACMG-AMP Guidelines[18]
c.676dup	p.Trp226LeufsTer4	N/A	0.0021[a]	0.0088	N/A	pathogenic
c.1435−1G>A[16]	p.?	N/A	0	ND	N/A	pathogenic
c.148T>C	p.Ser50Pro	T/P/D	0	ND	platypus	uncertain significance
c.797G>A	p.Arg266Gln	D/D/D	0.0002	ND	zebrafish	uncertain significance
c.949G>A	p.Gly317Arg	D/D/D	0.00007	ND	platypus	uncertain significance
c.1166C>T	p.Pro389Leu	D/D/D	0.00001	ND	tetraodon	uncertain significance
c.1289T>C	p.Leu430Pro	D/D/D	0.000008	ND	platypus	uncertain significance

Abbreviations are as follows: D, damaging, deleterious, or disease-causing; B, benign; T, tolerated; P, polymorphism; MAF, minor allele frequency; N/A, not applicable; ND, not determined.
[a]Overall MAF is presented. This is, for example, slightly higher for the recurring variant c.676dup in non-Finnish Europeans with an MAF of 0.0038. By contrast, this variant has not been reported in East and South Asian populations.

heterozygous but no homozygous subjects were detected. Subsequent sequencing of the complete coding region of *M1AP* (exon 2 to 11, primers in Table S1) in all five heterozygous carriers ruled out the presence of a second relevant variant in all individuals. Additionally, we queried a previously established database of 5,784 Dutch fertile men and 5,803 fertile women who had conceived at least one child. WES had been performed as part of clinical diagnostic workup of a child with severe development delay (trio-WES, Nijmegen, the Netherlands). Again, 27 heterozygous male and 21 heterozygous female carriers but no homozygous individuals were detected. Also, there were no homozygous subjects of other LoF variants in the whole coding region of *M1AP* among either fathers or mothers. In summary, statistical evidence strongly supports that homozygous c.676dup in *M1AP* is associated with severe male infertility, while the different allele frequencies can most likely be explained by population stratification.

To gain insight into the function of M1AP and better assess the relevance of identified variants, we pursued two strategies. First, we attempted to model M1AP's 3D structure. However, because of the lack of information on M1AP and comparable 3D structures, it was not possible to achieve a reliable prediction (BLAST results for sequence of UniProt: Q8TC57 are below 30% sequence identity to known protein structures, details in Supplemental Methods). Second, we tried to establish immunohistochemistry as well as immunoblot analyses with the two most promising (based on available data from the manufacturers and the HPA) commercially available M1AP antibodies (#PA5-31627, ThermoFisher Scientific and #HPA045420, Sigma-Aldrich) (see Supplemental Methods for details). After optimization, both antibodies did result in a specific signal in immunohistochemistry (Figures S4 and S5) of testicular control sections. However, one of them (#PA5-31627) showed poor results (high amounts of background and antibody precipitates, Figure S4), whereas #HPA045420 seems not to bind M1AP but detect a different target instead (Figure S5). Most importantly, the presumed epitope resides downstream of the variant p.Trp226LeufsTer4 and, therefore, is disrupted in homozygous subjects (further details in Figures S5 and S6). Unfortunately, we were not successful in contacting the colleagues who published the *M1ap* knockout mice and immunoblot staining with a self-raised antibody.[8] In conclusion, the structure of M1AP is currently unknown and it is impossible to predict functional domains of M1AP with sufficient reliability. Moreover, the specific molecular function of M1AP remains to be elucidated in subsequent studies, which would also open the possibility of functional assessment of the missense variants.

The assessment of the other detected *M1AP* variants has, for the time being, to rely on established *in silico* tools, and we followed the strict clinical ACMG-AMP criteria (Table S4).[18] In addition to the recurrent frameshift variant c.676dup, individuals M1943 and Y126 each carry a missense variant (p.Arg266Gln and p.Gly317Arg, respectively), while individual P86 carries two assumed compound-heterozygous missense variants (p.Ser50Pro and p.Leu430Pro). Although *in silico* assessment supports the relevance of identified missense variants, all of these are categorized as being of uncertain significance. However, the co-segregation in the Turkish family especially is highly suggestive of the pathogenicity of (at least some) missense variants in *M1AP* (LOD score = 3.28). Of note, the group of Şehime G. Temel identified *M1AP* as a candidate independently from the initial identification and the clinical and testicular phenotype of the index individual T1024 fits the spectrum of the other affected men very well. In contrast, all investigated fertile family members had at least one wild-type allele, which further supports the impact of the bi-allelic variant.

Overall, we identified ten unrelated men with likely causal bi-allelic variants in *M1AP*, nine stemming from four independent study cohorts and one from a Turkish family (Table 1), as well as four additional infertile men

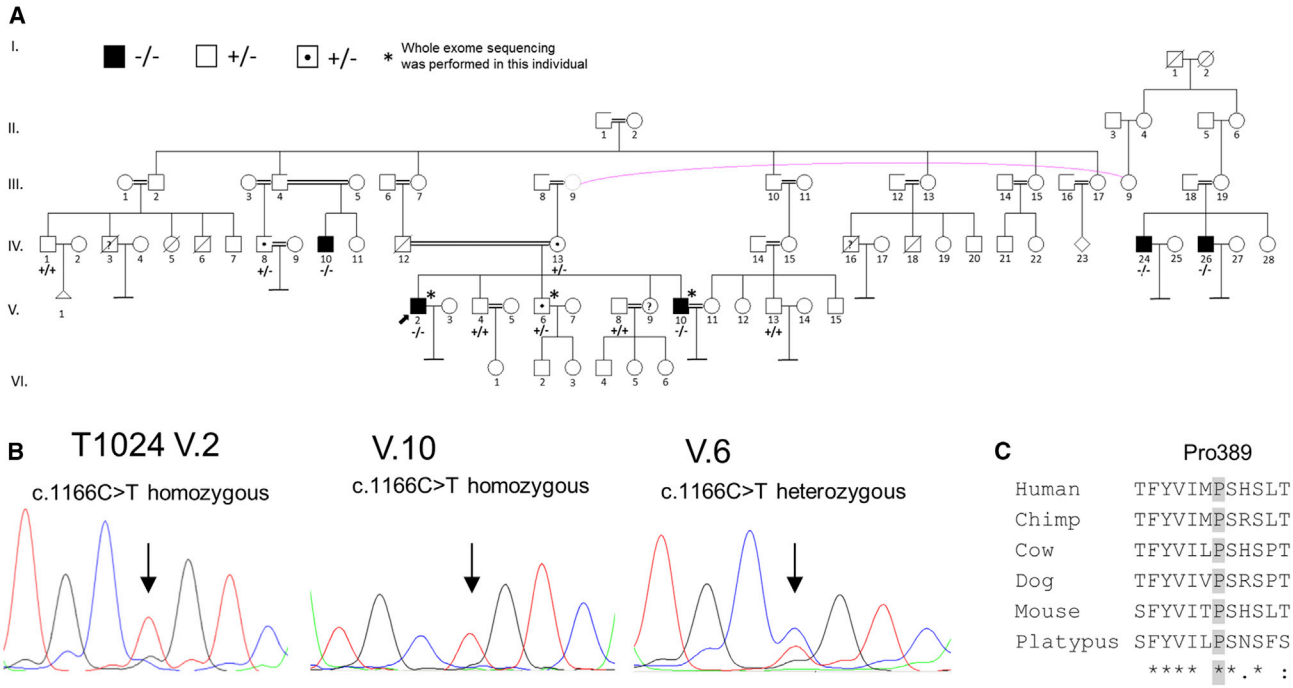


Figure 2. Turkish Consanguineous Family with Infertile, Azoospermic Men Homozygous for *M1AP* Missense Variant and Fertile Heterozygous Carriers

(A) Pedigree of the Turkish family with five infertile azoospermic men carrying the homozygous *M1AP* variant c.1166C>T (p.Pro389Leu) indicated with black boxes and −/−. The index individual T1024, who presented at Uludag University Faculty of Medicine Hospital, is marked with an arrow (V.2). Heterozygous carriers of the *M1AP* variant are marked with a point and +/-. Examined family members with the homozygous *M1AP* wild-type allele are marked with +/+. Homozygous men are infertile, whereas heterozygous carriers are fertile.
(B) Representative electropherograms of the index affected individual (V.2), his infertile brother (V.10), and his fertile brother (V.6), who is a heterozygous carrier.
(C) The missense variant affects a highly conserved amino acid.

from this family. Out of these, eight men underwent testicular biopsy, and histology showed arrested spermatogenesis in all of them. Arrest occurred at meiosis, i.e., the spermatocyte stage, in the majority and was either complete (n = 3, M330, M864, and M1792) or predominant (n = 2, RU1691 and MI-0006-P). Fewer men had a later arrest at the round spermatid stage (n = 2, Y126 and T1024) or earlier with only spermatogonia present (n = 1, P86). Thus, the common phenotype was NOA, but sporadically spermatids and rarely spermatozoa in the semen (below 10 per sample) were observed in two individuals constituting a continuum at the very severe end of spermatogenic failure. A similar phenotype was described for mice with disruption of *M1ap* that sporadically had some spermatozoa in their semen.[8]

Very recently, a homozygous splice-site variant (c.1435−1G>A) in *M1AP* has been published as cause for severe oligozoospermia in a single male from a consanguineous Han Chinese family.[16] Of note, this variant resides in the last exon and, on the basis of our in-depth characterization of the same antibody, we put into question Tu et al.'s analyses using the same antibody. Likewise, the proposed localization of M1AP in the sperm midpiece should be critically assessed in light of the clear evidence we provide that the antibody does not seem to stain M1AP specifically (see above and Figures S5 and S6). Still, this variant likely impairs M1APs function, and the finding in this infertile male from yet another population (1) further broadens the phenotypical spectrum of *M1AP*-associated spermatogenic disturbances and (2) provides more evidence for *M1AP*'s relevance for male infertility. We therefore included this individual and the variant in Tables 1 and 2 to provide a comprehensive overview of all presumably relevant *M1AP* variants, affected individuals, and clinical characteristics.

So far, only very few other genes, such as *TEX11* and *STAG3*, with mutations leading to germ cell arrest in both men and mice and validated in independent cohorts have been published.[5,7,10,22] From our initial cohort of 64 men with complete bilateral meiotic arrest, we also identified likely causal variants in three other genes, namely *TEX11*, *STAG3*, and *SYCP2*.[5,10,11] Among the remaining group of 58 men analyzed herein, we identified three unrelated men with the same homozygous frameshift variant, c.676dup, in *M1AP*. Thus, disruptive *M1AP* variants are highly enriched in men affected by male infertility, NOA, and meiotic arrest (5%, 3 out of 64), a frequency comparable to the currently best-characterized *TEX11* mutations with 6% (4 out of 64) in this selected group. Concerning the full spectrum of phenotypes associated with *M1AP*, variants in this gene most likely contribute to less than 1% of the highly heterogeneous individuals with severe spermatogenic defects, i.e., NOA, cryptozoospermia, and

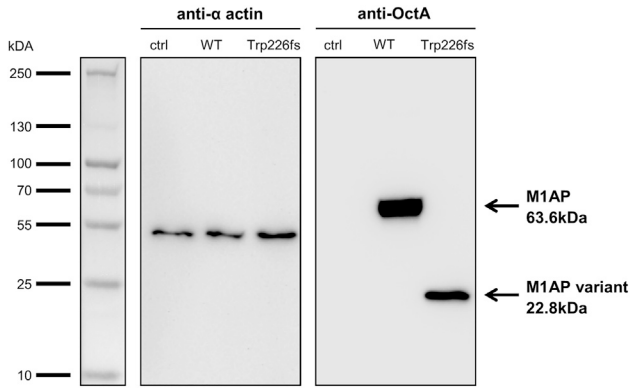


Figure 3. Heterologous Expression of M1AP in HEK293T Cells
Non-transfected control (ctrl), wild-type (WT), and mutated c.676dup, p.[Trp226LeufsTer4];[Trp226fs] *M1AP* cDNA (N-terminal located DYK-tag, in pcDNA3.1) plasmids were transfected into HEK293T cells. 24 h post-transfection, whole-cell lysates were prepared and separated on an SDS-PAGE. A PAGE Ruler Plus Prestained Protein Ladder (Thermo Scientific) was used for validation of protein sizes, and Biorad ImageLab Software was used to calculate protein sizes. Anti-alpha-actin antibody (1:5000, ab5694, abcam) served as a loading control, and M1AP was detected by a monoclonal anti-OctA antibody (1:1000, sc-166355, Santa Cruz) against the integrated DYK-tag (1:1000) (n = 3).

oligozoospermia (0.5%, 9 out of 1,996 individuals). Still, *M1AP* is one of the most commonly mutated genes associated with severe male infertility to date and ranges directly behind the well-established genetic causes Klinefelter syndrome (47,XXY, almost exclusively found in NOA) and Y chromosome AZF deletions (found in azoospermia and severe oligozoospermia).[2]

Assessing the clinical validity of gene-disease associations is now recognized as being of the utmost importance. Collectively, the presented data, accumulated from several independent groups and populations as well as a consanguineous family, together with previously described evidence from murine studies and the in parallel published second family from China, result in *M1AP* immediately reaching a strong clinical validity based on a structured assessment according to Smith et al.[23] (Table S5). This is an unprecedented example of any gene in which variants are associated with male infertility and clearly shows the power of large-scale collaborative efforts.

The process of meiosis is in part orchestrated similarly in both sexes but initiated at different times in life. Fittingly, a few genes, such as *STAG3*, have been reported in which variants impair male as well as female meiosis and result in infertility in both sexes.[10,22,24] *M1AP* is predominantly expressed in the adult testis in both men and mice, but it is also reported to be expressed in the fetal mouse ovary.[9] However, the structure of ovaries in female *M1ap* knockout mice appeared normal and fertility was preserved.[8] To shed some light onto a potential role of M1AP in human female meiosis, we screened 101 women diagnosed with unexplained premature ovarian insufficiency (POI) (62 with isolated POI and 39 with ovarian dysgenesis) by Sanger sequencing of *M1AP*'s full coding region. Details of a part of this cohort (n = 25) have been published previously.[25] The additional 76 individuals followed the same inclusion and exclusion criteria. No women carrying two rare variants in *M1AP* were identified. Furthermore, we did not identify a fertile woman carrying a homozygous LoF variant in the large Dutch trio cohort or a woman homozygous for the missense variant in the Turkish family. Thus, we cannot exclude that variants in *M1AP* might also be a rare cause for POI, but current evidence suggests that M1AP is only required for male meiosis.

In conclusion, the presented data from four independent cohorts strongly supports that both homozygous LoF and deleterious missense variants in *M1AP* as well as compound heterozygosity for either variant type results in severe spermatogenic failure and male infertility. Cosegregation in multiple affected men from a Turkish family provides further strong and independent evidence. *M1AP* disruption is associated primarily with germ cell arrest but might also be compatible with the production of a few spermatozoa, similar to the phenotypic spectrum observed in men with Klinefelter's syndrome or AZFc deletions. On the basis of our data, we cannot reliably predict the probability of successful sperm retrieval by testicular biopsy and TESE, but if we extrapolate the findings from the men reported here, TESE success is quite low (1 of 8, <15%). Our findings provide further evidence that germ cell arrest is often of monogenic origin. According to the commonly applied structured gene assessment,[23] *M1AP* has strong clinical validity for causing NOA, cryptozoospermia, or severe oligozoospermia and should be used as a screening marker before testicular biopsy to estimate the chances of successful TESE. Finally, identifying mutations in *M1AP* in unexplained infertile men with severe spermatogenic failure provides them with a causal diagnosis for their infertility.

Data and Code Availability

All variants have been submitted to ClinVar (VCV000805830–VCV000805834) and can also be accessed in the Male Fertility Gene Atlas (MFGA), a public platform for collecting and integrating datasets about genetic causes of male infertility produced in a subproject of the Clinical Research Unit "Male Germ Cells: from Genes to Function." Overall, the published article includes all datasets generated or analyzed during this study. The exceptions are the primary WES data that have not been deposited in a public repository because the individual's consent did not include this. However, additional information is available from the corresponding author on request.

Supplemental Data

Supplemental Data can be found online at https://doi.org/10.1016/j.ajhg.2020.06.010.

Acknowledgments

We are indebted to all individuals consenting to research evaluation of their data and donating their DNA as well as the physicians who took care of them. We thank Christian Ruckert for his bioinformatic support, Martin Bergmann for many years of skillfully evaluating testicular histologies, Nurten Akarsu for encouraging us to start a project with this Turkish Family, Zeliha Görmez for bioinformatics analyses of WES, Şeref Gül and Laurens van de Wiel for their attempts to model M1AP, Joachim Kremerskothen and Verena Höffken for attempting to establish a immunoblot analysis from testicular samples, as well as Christina Burhöi, Nicole Terwort, and Katja Hagen for their excellent technical assistance. We thank Celeste Brennecka for language editing of the manuscript. This work was carried out within the frame of the German Research Foundation Clinical Research Unit "Male Germ Cells: from Genes to Function" (DFG CRU326). Funding for sequencing of the GEMINI cohort was provided by the National Institutes of Health, United States (R01HD078641). The analyses in the Turkish family were supported by a grant from the Bursa University of Uludag Project Unit [KUAP(T)-2014/36].

Declaration of Interests

The authors declare no competing interests.

Received: October 12, 2019
Accepted: June 12, 2020
Published: July 15, 2020

Web Resources

Clinical Research Unit "Male Germ Cells," https://www.male-germ-cells.de
ClinVar, https://www.ncbi.nlm.nih.gov/clinvar
GEMINI, https://gemini.conradlab.org
gnomAD, https://gnomad.broadinstitute.org
GTEx, https://gtexportal.org
HOPE, https://www3.cmbi.umcn.nl/hope/input
Human Protein Atlas, https://www.proteinatlas.org
International Male Infertility Genomics Consortium, http://www.imigc.org
Male Fertility Gene Atlas, https://mfga.uni-muenster.de
MutationTaster, http://www.mutationtaster.org
OMIM, https://www.omim.org
PolyPhen 2, http://genetics.bwh.harvard.edu/pph2
PSAP, https://github.com/conradlab/PSAP
SIFT, https://sift.bii.a-star.edu.sg

References

1. Juul, S., Karmaus, W., and Olsen, J. (1999). Regional differences in waiting time to pregnancy: pregnancy-based surveys from Denmark, France, Germany, Italy and Sweden. The European Infertility and Subfecundity Study Group. Hum. Reprod. *14*, 1250–1254.
2. Tüttelmann, F., Ruckert, C., and Röpke, A. (2018). Disorders of spermatogenesis: Perspectives for novel genetic diagnostics after 20 years of unchanged routine. Med. Genetik *30*, 12–20.
3. World Health Organization (2010). WHO laboratory manual for the examination and processing of human semen (World Health Organization).
4. Lee, J.Y., Dada, R., Sabanegh, E., Carpi, A., and Agarwal, A. (2011). Role of genetics in azoospermia. Urology *77*, 598–601.
5. Yatsenko, A.N., Georgiadis, A.P., Röpke, A., Berman, A.J., Jaffe, T., Olszewska, M., Westernströer, B., Sanfilippo, J., Kurpisz, M., Rajkovic, A., et al. (2015). X-linked *TEX11* mutations, meiotic arrest, and azoospermia in infertile men. N. Engl. J. Med. *372*, 2097–2107.
6. Gershoni, M., Hauser, R., Yogev, L., Lehavi, O., Azem, F., Yavetz, H., Pietrokovski, S., and Kleiman, S.E. (2017). A familial study of azoospermic men identifies three novel causative mutations in three new human azoospermia genes. Genet. Med. *19*, 998–1006.
7. Oud, M.S., Volozonoka, L., Smits, R.M., Visser, L.E., Ramos, L., and Veltman, J.A. (2018). A systematic review and standardized clinical validity assessment of male infertility genes. bioRxiv. https://doi.org/10.1101/425553.
8. Arango, N.A., Li, L., Dabir, D., Nicolau, F., Pieretti-Vanmarcke, R., Koehler, C., McCarrey, J.R., Lu, N., and Donahoe, P.K. (2013). Meiosis I arrest abnormalities lead to severe oligozoospermia in meiosis 1 arresting protein (M1ap)-deficient mice. Biol. Reprod. *88*, 76.
9. Arango, N.A., Huang, T.T., Fujino, A., Pieretti-Vanmarcke, R., and Donahoe, P.K. (2006). Expression analysis and evolutionary conservation of the mouse germ cell-specific D6Mm5e gene. Dev. Dyn. *235*, 2613–2619.
10. van der Bijl, N., Röpke, A., Biswas, U., Wöste, M., Jessberger, R., Kliesch, S., Friedrich, C., and Tüttelmann, F. (2019). Mutations in the stromal antigen 3 (STAG3) gene cause male infertility due to meiotic arrest. Hum. Reprod. *34*, 2112–2119.
11. Schilit, S.L.P., Menon, S., Friedrich, C., Kammin, T., Wilch, E., Hanscom, C., Jiang, S., Kliesch, S., Talkowski, M.E., Tüttelmann, F., et al. (2020). SYCP2 Translocation-Mediated Dysregulation and Frameshift Variants Cause Human Male Infertility. Am. J. Hum. Genet. *106*, 41–57.
12. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. bioRxiv. https://doi.org/10.1101/531210.
13. Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., et al.; GTEx Consortium (2013). The Genotype-Tissue Expression (GTEx) project. Nat. Genet. *45*, 580–585.
14. Nieschlag, E., Behre, H.M., and Nieschlag, S. (2010). Andrology: Male reproductive health and dysfunction, Third Edition (Springer-Verlag Berlin Heidelberg).
15. Magi, A., Tattini, L., Palombo, F., Benelli, M., Gialluisi, A., Giusti, B., Abbate, R., Seri, M., Gensini, G.F., Romeo, G., and Pippucci, T. (2014). H3M2: detection of runs of homozygosity from whole-exome sequencing data. Bioinformatics *30*, 2852–2859.
16. Tu, C., Wang, Y., Nie, H., Meng, L., Wang, W., Li, Y., Li, D., Zhang, H., Lu, G., Lin, G., et al. (2020). An M1AP homozygous splice-site mutation associated with severe oligozoospermia in a consanguineous family. Clin. Genet. *97*, 741–746.
17. Venselaar, H., Te Beek, T.A.H., Kuipers, R.K.P., Hekkelman, M.L., and Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics *11*, 548.
18. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al.; ACMG Laboratory Quality Assurance Committee (2015). Standards

and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet. Med. *17*, 405–424.

19. Wilfert, A.B., Chao, K.R., Kaushal, M., Jain, S., Zöllner, S., Adams, D.R., and Conrad, D.F. (2016). Genome-wide significance testing of variation from single case exomes. Nat. Genet. *48*, 1455–1461.
20. Kasak, L., Punab, M., Nagirnaja, L., Grigorova, M., Minajeva, A., Lopes, A.M., Punab, A.M., Aston, K.I., Carvalho, F., Laasik, E., et al.; GEMINI Consortium (2018). Bi-allelic Recessive Loss-of-Function Variants in FANCM Cause Non-obstructive Azoospermia. Am. J. Hum. Genet. *103*, 200–212.
21. Cassa, C.A., Weghorn, D., Balick, D.J., Jordan, D.M., Nusinow, D., Samocha, K.E., O'Donnell-Luria, A., MacArthur, D.G., Daly, M.J., Beier, D.R., and Sunyaev, S.R. (2017). Estimating the selective effects of heterozygous protein-truncating variants from human exome data. Nat. Genet. *49*, 806–810.
22. Riera-Escamilla, A., Enguita-Marruedo, A., Moreno-Mendoza, D., Chianese, C., Sleddens-Linkels, E., Contini, E., Benelli, M., Natali, A., Colpi, G.M., Ruiz-Castañé, E., et al. (2019). Sequencing of a 'mouse azoospermia' gene panel in azoospermic men: identification of RNF212 and STAG3 mutations as novel genetic causes of meiotic arrest. Hum. Reprod. *34*, 978–988.
23. Smith, E.D., Radtke, K., Rossi, M., Shinde, D.N., Darabi, S., El-Khechen, D., Powis, Z., Helbig, K., Waller, K., Grange, D.K., et al. (2017). Classification of Genes: Standardized Clinical Validity Assessment of Gene-Disease Associations Aids Diagnostic Exome Analysis and Reclassifications. Hum. Mutat. *38*, 600–608.
24. Caburet, S., and Vilain, É. (2015). [STAG3 in premature ovarian failure]. Med. Sci. (Paris) *31*, 129–131.
25. Ledig, S., Röpke, A., and Wieacker, P. (2010). Copy number variants in premature ovarian failure and ovarian dysgenesis. Sex Dev. *4*, 225–232.

The American Journal of Human Genetics, Volume *107*

Supplemental Data

Bi-allelic Mutations in *M1AP* Are a Frequent Cause of Meiotic Arrest and Severely Impaired Spermatogenesis Leading to Male Infertility

Margot J. Wyrwoll, Şehime G. Temel, Liina Nagirnaja, Manon S. Oud, Alexandra M. Lopes, Godfried W. van der Heijden, James S. Heald, Nadja Rotte, Joachim Wistuba, Marius Wöste, Susanne Ledig, Henrike Krenz, Roos M. Smits, Filipa Carvalho, João Gonçalves, Daniela Fietz, Burcu Türkgenç, Mahmut C. Ergören, Murat Çetinkaya, Murad Başar, Semra Kahraman, Kevin McEleny, Miguel J. Xavier, Helen Turner, Adrian Pilatz, Albrecht Röpke, Martin Dugas, Sabine Kliesch, Nina Neuhaus, GEMINI Consortium, Kenneth I. Aston, Donald F. Conrad, Joris A. Veltman, Corinna Friedrich, and Frank Tüttelmann

Figure S1. Prioritization scheme for WES data.

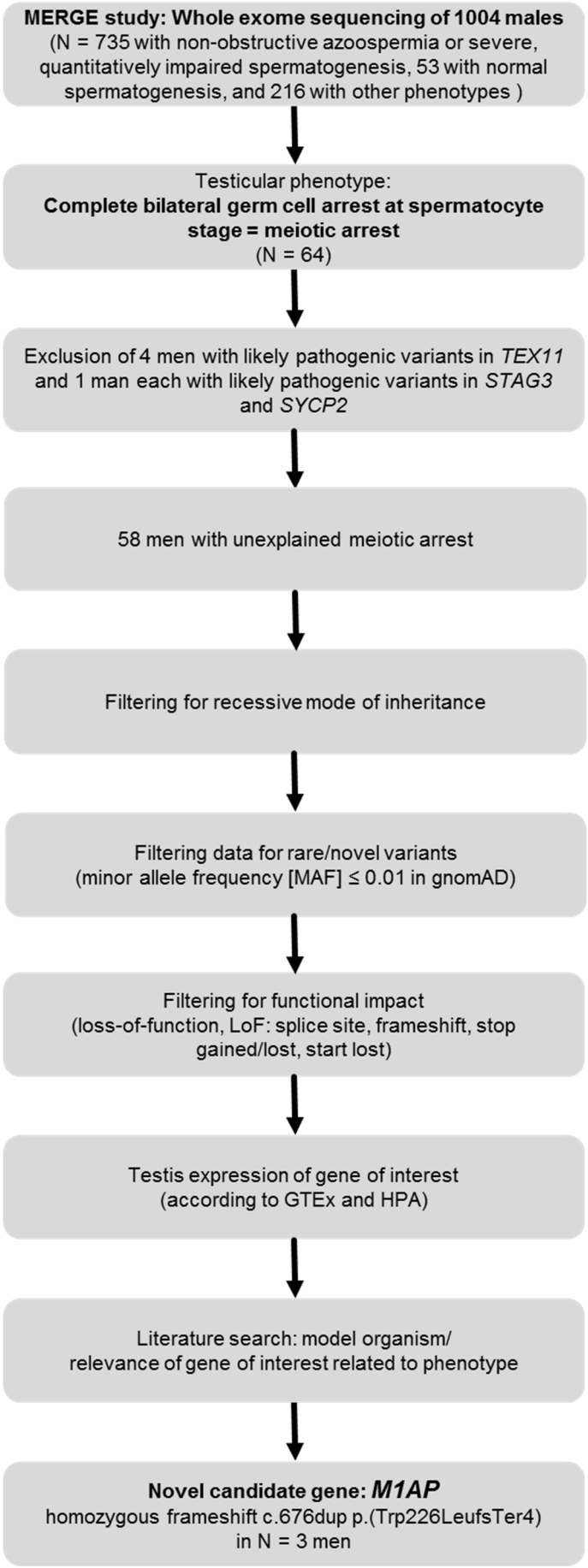
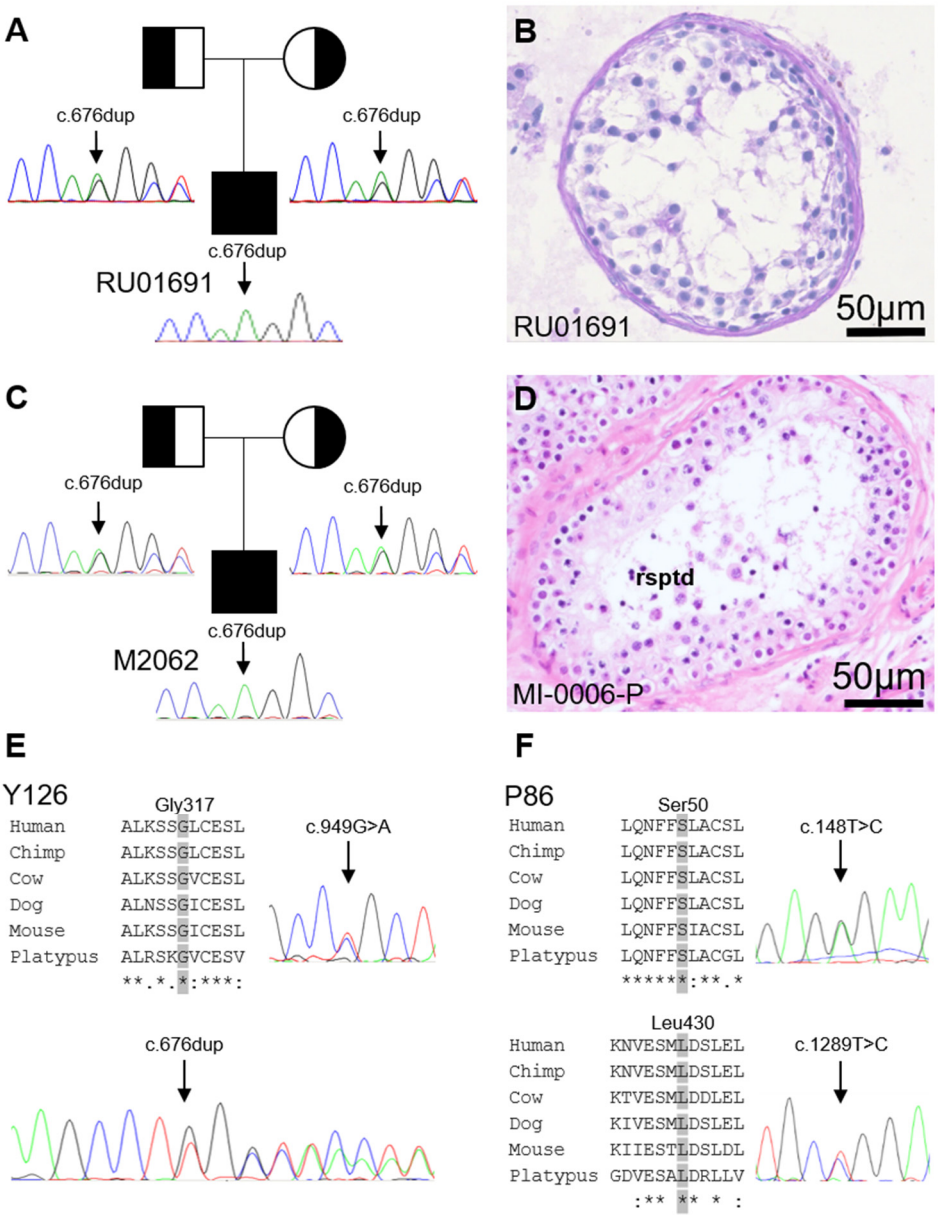
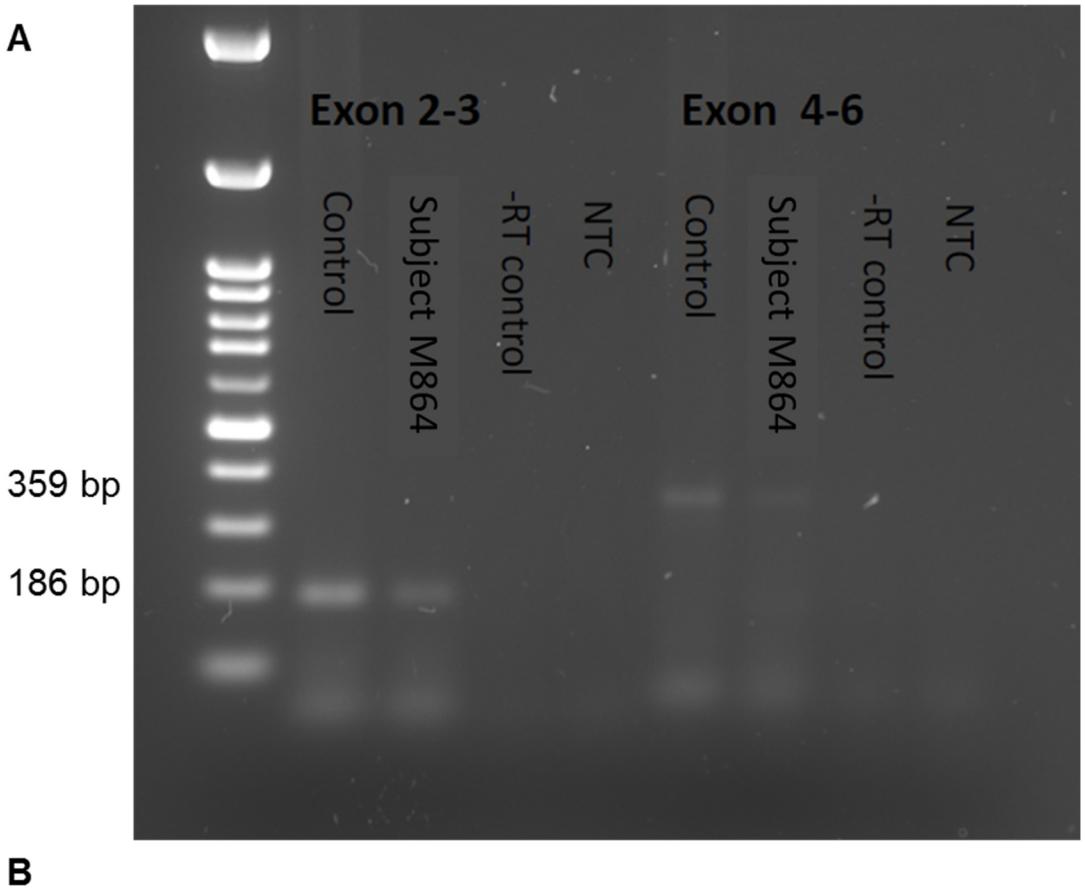


Figure S2. Variants in *M1AP* in follow-up studies.



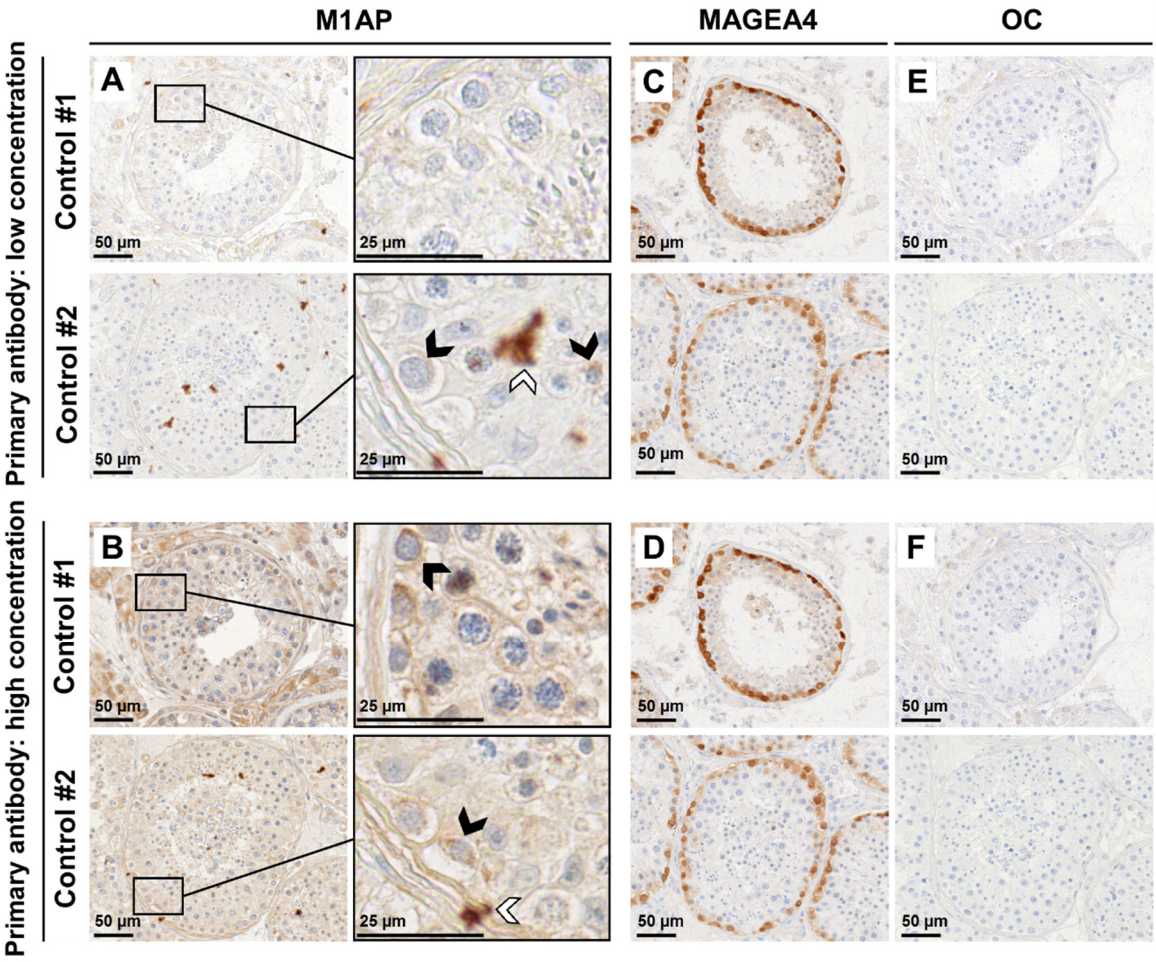
(**A**) Recurrent homozygous duplication c.676dup (p.Trp226LeufsTer4) in *M1AP* in RU01691 from Nijmegen, NL. Both parents are heterozygous for the same variant. (**B**) Testicular histology of RU01691 indicates predominant germ cell arrest at the spermatocyte stage. (**C**) Individual M2062 from Poland with azoo-/cryptozoospermia carries the same homozygous duplication in *M1AP*, also inherited from his heterozygous parents. (**D**) Testicular histology of another homozygous c.676dup subject, MI-0006-P from the UK, revealed a predominant germ cell arrest, but sporadically round spermatids (rsptd) were present in some seminiferous tubules. **(E/F)** Identification of potentially biallelic variants in *M1AP* in Portuguese men from the GEMINI study. (**E**) Y126 carries the recurrent LoF variant c.676dup (p.Trp226LeufsTer4) and the missense variant c.949G>A (p.(Gly317Arg)). (**F**) P86 carries two missense variants c.148T>C(;)1289T>C p.(Ser50Pro)(;)(Leu430Pro). All missense variants affect highly conserved amino acids, as seen from multiple sequence alignments (**E**/**F**).

Figure S3. Testis RNA analysis of individual M864.



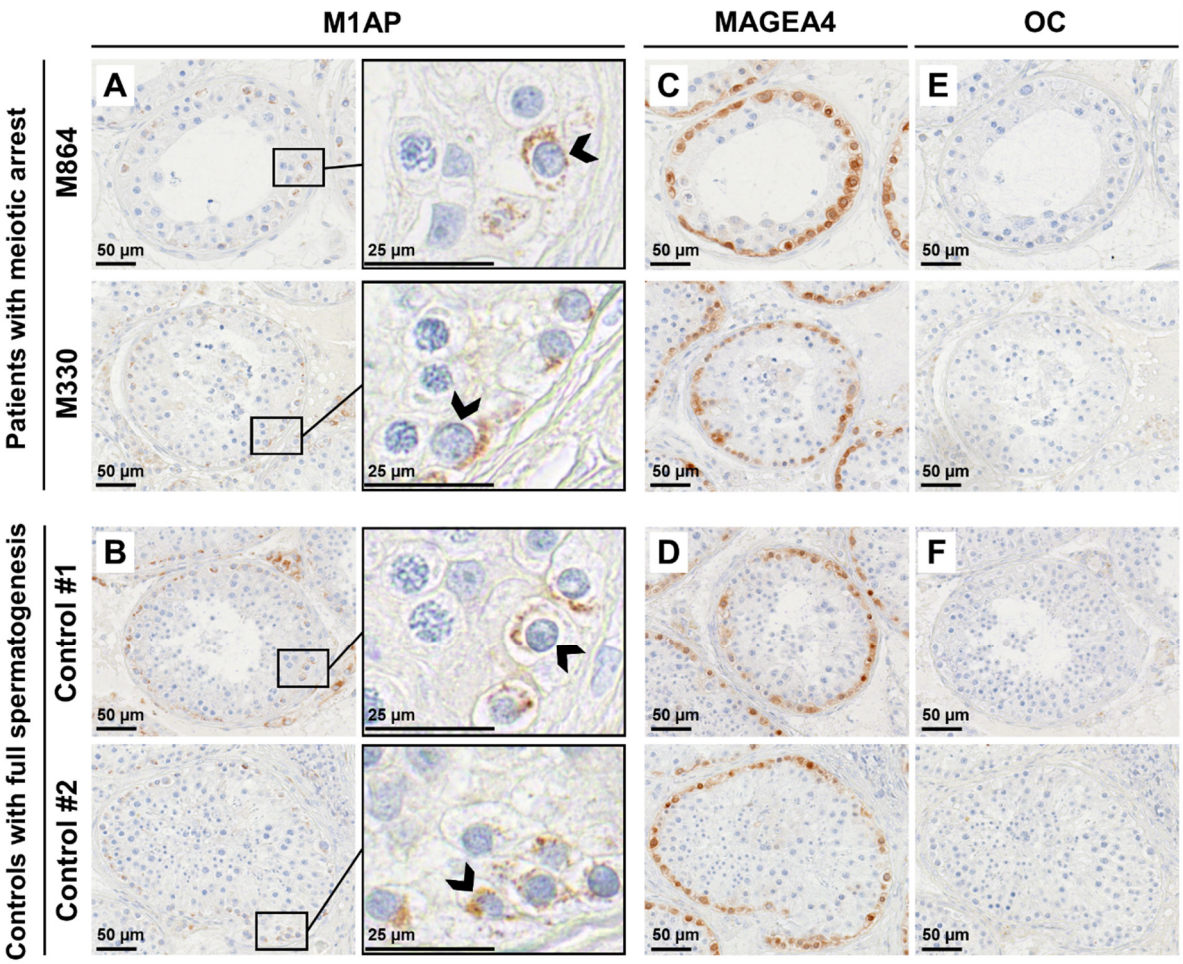
(**A**) PCR of testicular cDNA of exons 2-3 and 4-6 of a control and M864. Exons 2-3 and 4-6 were amplified. Expected band sizes were 186 bp (Exon 2-3) and 359 bp (Exon 4-6). + fertile control DNA, Mutation: M864 DNA, - RT: negative control without reverse transcriptase (RT), NTC: no template control. (**B**) Sanger sequencing of the PCR product amplified from cDNA of M864 was performed according standard procedures and validates the homozygous variant c.676dup in Exon 5.

Figure S4. Analysis of M1AP protein expression by immunohistochemical staining using #PA5-31627.



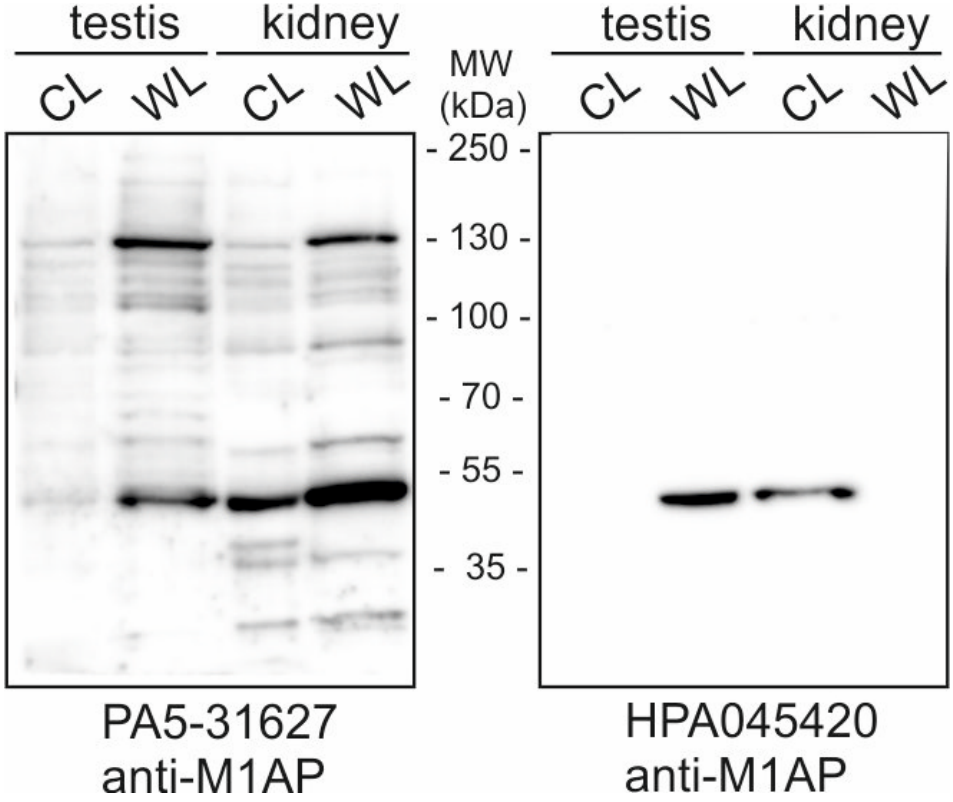
(**A**) Low concentration of M1AP antibody (#PA5-31627, ThermoFisher Scientific, exemplarily depicted for 1:100) was evaluated and black arrow heads indicate spermatogonia, which potentially show weak but specific cytoplasmic protein expression. Additionally, unspecific antibody precipitation in control #2 can be observed (indicated by white arrow head). (**B**) Analysis of higher concentrations of M1AP antibody (exemplarily depicted for 1:50) lead to enhanced unspecific background signal and specific cytoplasmic protein expression primarily in spermatogonia (indicated by black arrow head). (**C**, **D**) MAGEA4 was used as a marker for early germ cells. (**E**, **F**) Omission of primary antibody (OC) is shown as representative technical control. Scale bars are indicated in each micrograph, respectively.

Figure S5. Analysis of M1AP protein expression by immunohistochemical staining using #HPA045420.



(**A**) Testis tissue from two individuals M330 and M864 with a homozygous duplication (c.676dup p.Trp226LeufsTer4) in *M1AP* resulting in bilateral meiotic arrest was stained with a commercially available M1AP antibody (#HPA045420, Sigma-Aldrich, exemplarily depicted for 1:100). Representative micrographs show cytoplasmic protein expression specifically in spermatogonia (indicated by black arrow head). (**B**) Sections from men with full spermatogenesis were used as controls. The detected signal did not differ from *M1AP* mutated individuals. Therefore, the antibody does not seem to bind M1AP but pick up a different target instead: both men carry the homozygous frameshift variant p.Trp226LeufsTer4, probably resulting in a truncated protein and consequently leading to a disruption of the supposed epitope of the antibody, which was presumed to reside downstream of the variant and should be specific to M1AP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). By database search (Ensembl, GTEx), we ruled out that *M1AP* transcript isoforms exist lacking exon 5 and thus none of the isoforms should be escaping the truncation. (**C**, **D**) MAGEA4 was used as a marker for early germ cells. (**E**, **F**) Omission of primary antibody (OC) is shown as representative technical control. Scale bars are indicated in each micrograph, respectively.

Figure S6. Western blot analysis from tissue lysates.



Cytosolic (CL) and whole cell (WL) lysates from testis and kidney were obtained from fertile healthy donors. Western blots were stained with two different M1AP antibodies as indicated. Predicted molecular weight of M1AP is 59 kDa. The observed band size of almost 55 kDa is similar to those expected for PA5-31627 antibody by the manufacturer's instructions. In contrast, validation experiments of the Human Protein Atlas of HPA045420 antibody detect protein bands of ~30 kDa in different tissues and cell lines. Wether different band sizes may result from different post-translational modifications remains elusive due to the lack of knowledge about M1AP protein. Furthermore, the Western blot did not only result in detection of a band in testis lysate, but also in kidney lysate where M1AP is putatively not expressed (GTEx, HPA). Taken the results from IHC staining and Western blot analysis together, the specificity of both antibodies is questionable.

Table S1. Primer sequences.

Primer	Sequence Forward	Sequence Reverse
Validation of c.676dup (MERGE study)	TGGGTCTGGAAATGTTGCTGA	GATTGCTAGAGCCCAGGCAT
qPCR of Exon 5 in *M1AP*	TCTGGGAACTGACATTGACCTTC	TGGGTCTGGAAATGTTGCTGA
Sanger-Sequencing of all exons of *M1AP* in heterozygous control men		
Exon 2	TGGATTTTCTCTTCAACAGTACACA	TGTGCACCTGTAGTCCTAGC
Exon 3	CAGTTTTTCCTCATAATTCACTTCAGT	TCCTTTCATGTTTCTGGTAACTCT
Exon 4	CCTCAGTGAAATCTGCTGGC	GCCTGATTGGAAAGGTCCTGT
Exon 5	CTAACTGGCCCTTGCTGGTT	TGCTAGAGCCCAGGCATTTG
Exon 6	CACCATCTGCACATTTGGCC	AACCAGTCAGGCTTTCCTCT
Exon 7	ACAGAATATATATCTAGGGCTTGACAC	GAGTCTGCTTCAACTCTTCCCA
Exon 8	GCCGAAGTTAAATGGCTCTG	TGGCATAATTGCCTATCCTT
Exon 9+10	GGGGGACAGCATCTATTTCA	TTCCCTCTTCAACCCCAACT
Exon 11	CCTTGAGGCTGTCACTCCA	CTTGCTGGAGAAAGGACAGG
RNA analyses		
M1AP cDNA Ex2-3	ACATTGCTCTACCGTCCTGG	TGCAACCTAGCAAAGTTCCCT
M1AP cDNA Ex4-6	CCTAGCCAGAGTCAGGAGGT	ATTCTCAAGGAGCCGTCAGC
Mutagenesis Primer		
M1AP mut p.W226L dup	GATTTCTTCAAAGCCTTGGCTACATAA CAGTGG	CCACTGTTATGTAGCCAAGGCTT TGAAGAAATC

Table S2. qPCR results to exclude the hemizygosity of LoF variant c.676dup in exon 5 of *M1AP* of men from the MERGE study.

Individual	Repeat	Crossing Point Exon 5 *M1AP*	Concentration Exon 5 *M1AP*	Mean concentration Exon 5 *M1AP*	Crossing Point *Albumin*	Concentration *Albumin*	Mean concentration *Albumin*	Ratio mean concentration Exon 5 *M1AP* / mean concentration *Albumin*
M330	1	21.58	10.2		22.64	9.42		
M330	2	21.57	10.3	10.03	22.65	9.35	9.40	1.10
M330	3	21.57	10.3		22.64	9.43		
M864	1	22.62	4.96		23.46	5.34		
M864	2	22.58	5.12	5.04	23.46	5.34	5.35	0.94
M864	3	22.60	5.03		23.45	5.37		
M1792	1	21.96	7.86		23.21	6.37		
M1792	2	21.97	7.79	7.80	23.12	6.77	6.78	1.16
M1792	3	21.98	7.73		23.03	7.20		
M2062	1	24.49	5.76		22.61	11.40		
M2062	2	23.84	9.02	8.56	22.68	10.80	11.1	0.77
M2062	3	23.57	10.89		22.64	11.10		
Control	1	21.62	10.0	10.0	22.54	10.0	10.0	1.0
Control	2	21.61	10.0	10.0	22.55	10.0	10.0	1.0
Control	3	21.60	10.0	10.0	22.57	10.0	10.0	10.0
Control	4	23.6	10.0	10.0	22.36	10.0	10.0	10.0

Table S3. Top 50-List of PSAP results of individuals with *M1AP* variants.

(see respective Excel-file)

Table S4. Classification of *M1AP* variants.

cDNA change	Protein change	Categories	Classification according to ACMG-AMP guidelines[1]
c.676dup	p.Trp226LeufsTer4	PVS1, PS3, PM2, PP4	Pathogenic
c.148T>C	p.(Ser50Pro)	PM1, PP3, PP4	Uncertain significance
c.797G>A	p.(Arg266Gln)	PP3, PP4	Uncertain significance
c.949G>A	p.(Gly317Arg)	PM2, PP3, PP4	Uncertain significance
c.1166C>T	p.(Pro389Leu)	PM2, PP1, PP3, PP4	Uncertain significance
c.1289T>C	p.(Leu430Pro)	PM1, PP3, PP4	Uncertain significance
c.1435-1G>A[2]	?	PVS1, PS3, PM2, PP4	Pathogenic

Table S5. Structured clinical validity assessment of *M1AP* according to Smith et al. 2018[3].

	Points	Comments/Reference
Number of unrelated individuals (1-2 → 1 pt, 3-4 → 2 pt, 5-9 → 3 pt, 10-24 → 4 pt; > 25 → classification "definitive")	4	11 individuals (M330, M864, M1792, M1943, M2062, Y126, P86, RU01691, MI-0006-P, T1024, F1:II-1[2])
Other statistical evidence (AD disease with significant excess of de novos OR AR disease with e.g. LOD score > 3 → 1 pt)	1	Consanguineous Turkish family (homozygous variant segregates with infertility, heterozygous carriers fertile, LOD score 3.28)
Number of publications reporting independent probands (per publication 1 pt, max. 3)	1	This publication and Tu *et al.* 2020[2]
Number of pathogenic variants (per VLP or mutation 1 pt, max. 4)	2	c.676dup (p.Trp226LeufsTer4) c.1435-1G>A (p.?)[2]
Gene function (function/expression consistent with disease → 1 pt	1	Testicular expression in mice, expressed in last stages of spermatogenesis,[4] Human Protein Atlas, GTEx
…and/or physically interacts with gene characterized for same disease → 1 pt)	0	
Gene disruption (relevant pathology *in vitro* after similar genetic modification → 1 pt	1	*in vitro* expression in HEK293T cells demonstrates putatively truncated protein
...and/or determination of mutational mechanism → 1 pt)	1	Frameshift variant → truncated protein → LoF (this study: analysis of testis RNA, heterologous expression)
Model organism (gene function *in vivo* related to pathology of human disease → 1 pt	1	Male KO-mouse is infertile[5]
…and/or phenotype and genotype match human disease → 1 pt)	1	Male knockout mouse exhibits meiotic arrest and severe oligozoospermia[5]
TOTAL POINTS	13	
Classification no evidence (0-4 pts) limited (2-9 pts) moderate (8-12 pts) strong (13+ pts) definitive (canonical)	strong	

Table S6. Members of the GEMINI consortium.

Group Leaders	Institution
Donald F. Conrad, Liina Nagirnaja	Department of Genetics, Oregon National Primate Research Center, Oregon Health & Science University, Beaverton, OR, USA
Kenneth I. Aston, Douglas T. Carrell, James M. Hotaling, Timothy G. Jenkins	Andrology and IVF Laboratory, Department of Surgery (Urology), University of Utah School of Medicine, Salt Lake City, UT, USA
Rob McLachlan[1,2]	1) Hudson Institute of Medical Research and the Department of Obstetrics and Gynaecology, Monash University, Clayton, Victoria, Australia 2) Monash IVF and the Hudson Institute of Medical Research, Clayton, Victoria, Australia
Moira K. O'Bryan	School of Biological Sciences, Monash University, Clayton, Victoria, Australia
Peter N. Schlegel	Department of Urology, Weill Cornell Medicine, New York, NY, USA
Michael L. Eisenberg	Department of Urology, Stanford University School of Medicine, Stanford, CA 94305, USA
Jay I. Sandlow	Department of Urology, Medical College of Wisconsin, Milwaukee, WI, 53226, USA
Emily S. Jungheim, Kenan R. Omurtag	Washington University in St Louis, School of Medicine, St Louis, MO, USA
Alexandra M. Lopes[1,2], Susana Seixas[1,2], Filipa Carvalho[1,3], Susana Fernandes[1,3], Alberto Barros[1,3]	1) i3S - Instituto de Investigação e Inovação em Saúde , Universidade do University of Porto 2) IPATIMUP - Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Porto, Portugal 3) Serviço de Genética, Departamento de Patologia, Faculdade de Medicina da Universidade do Porto, Porto, Portugal
João Gonçalves[1,2], Iris Caetano[1], Graça Pinto[3], Sónia Correia[3]	1) Departamento de Genética Humana, Instituto Nacional de Saúde Dr Ricardo Jorge, Lisboa, Portugal 2) ToxOmics, Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Portugal 3) Centro de Medicina Reprodutiva, Maternidade Dr. Alfredo da Costa, Lisboa, Portugal
Maris Laan	Institute of Biomedicine and Translational Medicine, University of Tartu, 51010 Tartu, Estonia
Margus Punab	Andrology Center, Tartu University Hospital, 50406 Tartu, Estonia
Ewa Rajpert-De Meyts, Niels Jørgensen, Kristian Almstrup	Department of Growth and Reproduction, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark
Csilla G. Krausz[1,2]	1) Department of Experimental and Clinical Biomedical Sciences, University of Florence, Florence, Italy 2) Andrology Department, Fundacio Puigvert, Instituto de Investigaciones Biomédicas Sant Pau (IIB-Sant Pau), Barcelona, Spain
Keith A. Jarvi	Division of Urology, Department of Surgery, Mount Sinai Hospital and; Institute of Medical Sciences, University of Toronto, Toronto, Ontario, Canada

Supplemental methods

Individual's consent and study approval

All participants gave written informed consent for the evaluation of their clinical data and analysis of their DNA samples. The study protocol was approved by the respective Ethics Committees/Institutional Review Boards (Ref. No. Münster: 2010-578-f-S, Gießen: 26/11, GEMINI consortium: 201502059, Porto: PTDC/SAU-GMG/101229/2008, Nijmegen: NL50495.091.14 version 4, Newcastle: REC Ref: 18/NE/0089, Bursa: 05.01.2015/04).

Whole exome sequencing (WES) and bioinformatics analysis (MERGE study)

Genomic DNA was extracted from peripheral blood leukocytes via standard methods. WES sample preparation and enrichment were carried out in accordance with the protocols of either Agilent's SureSelect^{QXT} Target Enrichment kit or Twist Bioscience's Twist Human Core Exome kit. Agilent's SureSelect^{XT} Human All Exon Kits V4, V5 and V6 or Twist Bioscience's Human Core Exome plus RefSeq spike-in's were used to capture libraries. For multiplexed sequencing, the libraries were index tagged using appropriate pairs of index primers. Quantity and quality of the libraries were assessed with the ThermoFisher Qubit and Agilent's TapeStation 2200, respectively. Sequencing was conducted on the Illumina HiScan®SQ, NextSeq®500, or HiSeqX® systems using the TruSeq SBS Kit v3 - HS (200 cycles), the NextSeq 500 V2 High-Output Kit (300 cycles) or the HiSeq Rapid SBS Kit V2 (300 cycles), respectively.

After trimming, Cutadapt v1.15 was used to remove the remaining adapter sequences and primers.[6] Sequence reads were aligned against the reference genome GRCh37.p13 using BWA Mem v0.7.17.[7] We excluded duplicate reads and reads that mapped to multiple locations in the genome from further analysis. Small insertions/deletions (indels) and single nucleotide variations were identified and quality-filtered by GATK toolkit v3.8 with HaplotypeCaller, in accordance with the best practice recommendations.[8] Ensembl Variant Effect Predictor was used to annotate called variants.[9]

WES and bioinformatic analysis in individual RU01691 and MI-0006P

RU01691 came from a cohort of 171 offspring-parent trios and 112 singleton cases (N = 283) of men with severe oligozoospermia (<5 million sperm/ml; N = 69) or non-obstructive azoospermia (N = 214). MI-0006P belongs to a cohort of 16 offspring-parent trios and 32 singleton cases (N = 48) of men with unexplained azoospermia (N = 36) or oligozoospermia (N = 12). WES samples were prepared and enriched following the manufacturer's the protocols

of either Illumina's Illumina's Nextera DNA Exome Capture kit or Twist Bioscience's Twist Human Core Exome Kit. All sequencing was performed on the NovaSeq 6000 Sequencing System (Illumina) at an average depth of 72x (Illumina's Nextera Kit) and 99x (Twist Bioscience's Kit). Sequenced reads were aligned to GRCh37.p5 using BWA Mem v0.7.17, Picard and GATK v4.1.4.1. Following best practice recommendations single nucleotide variantions and small indels were identified and quality-filtered using GATK's HaplotypeCaller. Ensembl's Variant Effect Predictor (VEP) was used to fully annotate detected variants.

WES and bioinformatic analysis in individual Y126 and P86

Genetics of Male Infertility Initiative (GEMINI; https://gemini.conradlab.org/) is a multi-center consortium dedicated to identifying and describing the underlying genetic causes of male infertility. To date, N = 930 men with non-obstructive azoospermia have been studied using the whole-exome sequencing (WES) approach. In short, WES was performed at the McDonnell Genome Institute of Washington University (genome.wustl.edu) on Illumina HiSeq 4000 and using an in-house exome targeting reagent which captures 39.1 Mb of exome at an average coverage of 80x. The sequence reads were aligned to hg38 using bwa-mem, Picard and Genome Analysis Toolkit (GATK; https://software.broadinstitute.org/gatk) in an alternate contig-aware manner. Genotype calling was performed jointly for all samples using GATK tools and following their procedures of best practices. The resulting genotype callset was subjected to thorough quality control procedures, including but not limited to removing positions with high missingness rate (>15%) and removing samples with low coverage (<30x), high contamination (VerifyBamID freemix >5%) or low call rate (<85%). Individual genotypes with read depth (DP) <10x and genotype quality (GQ) <30 were further excluded from the data. All sequenced cases were screened for the known infertility causes such as Klinefelter syndrome, deleterious CFTR mutations, Y-chromosome microdeletions and large structural variation on sex chromosomes utilizing the WES genotype dataset.

In order to prioritize the deleterious lesions most likely disrupting the function of the respective genes and potentially leading to the disease phenotype, a modified version of the population sampling probability (PSAP) software (https://github.com/conradlab/PSAP) was applied to the WES genotype callset. The genomic coordinates of identified variation were lifted over to hg19 for the PSAP analysis. The list of prioritized variants was subsequently filtered by only including mutations with PSAP popScore values less than 10^{-4} and minor allele frequency <1% across all populations in the gnomAD database (v2.1.1, https://gnomad.broadinstitute.org/). As the study aims to identify rare DNA lesions most likely observed in few, if not singleton, cases, genes enriched for and positions found to be commonly affected by rare deleterious variation among the cases were excluded from the study.

WES and bioinformatic analysis in individual T1024

Genomic DNA was extracted from peripheral venous blood using the QIAamp® DNA Mini Kit (QIAGEN, Ankara, Turkey). SureSelectXT Library Prep Kit was used for target enrichment. All procedures were carried out according to the manufacturer's protocols. Paired-end sequencing was performed on an Illumina NovaSeq system with a read length of 151. Base calling and image analysis were conducted using Illumina's Real-Time Analysis software. The BCL (base calls) binary is converted into FASTQ utilizing Illumina package bcl2fastq. All bioinformatics analysis performed on Sophia DDMTM platform which includes algorithms for alignment, calling SNPs and small indels (Pepper), calling copy number variations (Muskat) and functional annotation (Moka). Raw reads were aligned to the human reference genome (GRCh37/hg19). Variant filtering and interpretation performed on Sophia DDMTM. Integrative Genomics Viewer (IGV)16 was used to bam file visualization. In families with consanguineous marriages, the homozygosity mapping was carried out with HomSI.

Attempt at 3D modelling of M1AP protein

The longest coding transcript of *M1AP* (GENCODE: ENST00000290536.5, RefSeq: NM_001281296.1, NM_138804.4) translates to the protein sequence with UniProt identifier Q8TC57. This sequence was used to perform a basic local alignment search tool (BLAST) to the sequences of known protein structures in the protein data bank (PDB). Four structures were obtained that partially match the M1AP sequence.

1. PDB ID: 5CK3 (Chain B, D, and F) with sequence identity: 29%
2. PDB ID: 5CK4 (Chain A, and B) with sequence identity: 29%
3. PDB ID: 5CK5 (Chain A, B, C, and D) with sequence identity: 29%
4. PDB ID: 6EGC (Chain A) with sequence identity: 24%

The highest sequence identity of these was 29%. Using these template structures will create unreliable results for a homology model, as templates with below 30% identity are likely to lead to serious mispredictions.

Quantitative PCR analysis

Quantitative PCR (qPCR) was carried out in 96-well plates on the LightCycler 480 using the manufacturer's default settings. The ALB (albumin) gene was used for normalization. The reactions were performed in triplicates using the SensiMix Real-Time PCR Kit (Bioline). The

PCR consisted of an initial incubation step at 95°C for 10 min followed by 40 cycles of 95°C for 15 sec, 60°C for 30 sec and 72°C for 15 sec. Baseline and threshold values were automatically detected using the LightCycler software. Primers are provided in Table S1.

RNA analysis

For RNA isolation, snap frozen testicular material of a individual M864 carrying the *M1AP* variant and a control proband with full spermatogenesis was used. RNA isolation was conducted with the miR-Neasy Micro kit (217084, Qiagen, Hamburg, Germany) according to the manufacturer's protocol. 500 ng of RNA were used as starting material for cDNA synthesis employing the iScript cDNA Synthesis Kit (Biorad). Subsequently, 4 µl of cDNA were amplified via PCR reaction using the Qiagen Taq Polymerase Kit. Reactions were performed in 20 µl of total volume with primer concentrations of 20 pmol (sequences in Table S1) and 1 mM dNTPs. The PCR included an initial incubation step at 94°C for 2 min followed by 35 cycles at 94°C for 30 sec, 56°C for 45 sec and 72°C for 1 min as well as a final step at 72°C for 10 min. PCR products were evaluated using a 2% agarose gel.

Attempt at M1AP immunohistochemical localization

Testicular tissue sections of 3 µm thickness were prepared (SM2010R sliding microtome, Leica Biosystems, Nussloch, Germany) from two human control samples with complete spermatogenesis or men with a homozygous duplication (c.676dup p.Trp226LeufsTer4) in *M1AP*, resulting in bilateral meiotic arrest. Sections were deparaffinized, rehydrated in a descending ethanol row, and rinsed with tap water. Heat-induced antigen retrieval was performed in citrate buffer (pH 6). After cooling to room temperature (RT), sections were washed with 1X Tris-buffered saline (TBS) prior to incubation with 3% H_2O_2 for 15 min at RT to inactivate endogenous peroxidases. Washing steps with distilled water and TBS followed. Nonspecific binding sites were blocked by applying 25% goat serum (#G6767-100ML, Sigma-Aldrich, Munich, Germany) in TBS containing 0.5% bovine serum albumin (BSA) for 30 min at RT in a humid chamber. Primary antibody incubation was performed overnight at 4°C in a humid chamber and different concentrations were evaluated for both M1AP antibodies (1:20, 1:50, 1:100, 1:200, 1:500, 1:1000 diluted in blocking solution, respectively). MAGE-A4, a marker for early human germ cells, was used as a positive control (kindly provided by Prof. G. C. Spagnoli, University Hospital of Basel, Switzerland). Respective IgG (#I5006, Sigma-Aldrich, Munich, Germany) and omission of primary antibody controls were included for each staining. On the following day, sections were washed again in TBS and incubated with corresponding secondary antibodies (goat anti-rabbit Biotin, #ab6012, Abcam, Cambridge,

USA - 1:100, 1:200) for 1 h at RT in a humid chamber. After further washing with TBS, Streptavidin conjugated with horse-radish peroxidase (#S5512, Sigma-Aldrich, Munich, Germany) was diluted in blocking solution (1:500) and sections were incubated for 45 min at RT in a wet chamber. Subsequently, sections were washed with TBS and incubated with 3,3'-Diaminobenzidine tetrahydrochloride (DAB, #D3757, AppliChem, Darmstadt, Germany) for visualization of antibody binding. Staining was validated by microscopical acquisition and distilled water was used to stop the reaction. Counterstains were conducted using Mayer's hematoxylin (#109249, Merck Millipore, Darmstadt, Germany). Sections were rinsed with tap water, dehydrated, and mounted using Merckoglas® mounting medium (#103973, Merck Millipore, Darmstadt, Germany). Images were captured using the PreciPoint M8 Scanning Microscopy System (PreciPoint, Freising, Germany).

Tissue preparation and Western blotting analysis

Cytosolic lysates were prepared by mincing biopsy samples from human testis or kidney in IP buffer (1% Triton-X 100, 20 mM Tris-HCl [pH 7.5], 25 mM NaCl, 50 mM NaF, 15 mM Na4P2O7, and 1.5 mM EDTA) containing protease inhibitors (Complete; Roche, Mannheim, Germany). Samples were centrifuged at 10,000×g for 45 min at 4°C to pellet cell debris and nuclei. Cytosolic supernatants were removed and stored at −80°C until further use. For total cell lysates, tissue was minced in Laemmli buffer ((20% Glycerin, 125 mM Tris-HCl pH 6.8, 10% SDS, 0.2% Bromphenol blue, 5% β-Mercaptoethanol), boiled for 5 min at 95°C and then stored at -20°C. Sodium dodecyl sulfate polyacrylamid gel electrophoresis (SDS-PAGE) and semidry Western blotting analysis was performed using standard techniques. Primary anti-M1AP antibodies were from ThermoFisher (PA5-31627) and Sigma (HPA045420), respectively. Signals were detected using secondary peroxidase-coupled antibodies and enhanced chemiluminescence (ECL).

Generation of M1AP variant-specific plasmid DNA

The LoF variant (c.676dup) was introduced in wildtype (WT) *M1AP* cDNA in the mammalian expression vector pcDNA3.1 (+) (Genscript, Leiden, NL) using site-directed-mutagenesis (SDM) according to manufacturer's instructions (QuikChange II XL Site-Directed Mutagenesis Kit, Agilent Technologies, Waldbronn, Germany). Clones carrying the respective mutations were verified by Sanger sequencing. For follow-up validation, a DYK-tag was integrated at the N-terminus of M1AP. Primer sequences can be found in Table S1.

Heterologous expression of M1AP in HEK293T cells

Human embryonic kidney 293T (HEK293T) cells were cultivated in 6 well plates (500.000 cells/well) in DMEM medium (Dulbecco's modified Eagle's medium, Sigma-Aldrich, Munich, Germany) supplemented with 1% PenStrep and 10% FCS at 37 °C, 5% CO_2. After 24 h post seeding, heterologous expression was induced by transfecting 2 µg plasmid DNA with XTR9 transfection reagent (Sigma-Aldrich, Munich, Germany) according to manufacturer's instructions. Transfection and subsequent analysis were performed in independent triplicates.

24 h post-transfection whole cell lysates were generated as following: cells were washed with 1X PBS and collected in a microcentrifuge tube. After centrifugation (4°C, 5 min, 200 g) and additional washing, the supernatant was discarded, and cells were re-suspended with lysis buffer (120 mM NaCl, 25 mM HEPES, 1% Triton X-100, 2 mM EDTA, 25 mM NaF, 1 mM $NaVO_3$, 0.2% SDS, 1x cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich, Munich, Germany), 1 mM PMSF). Cells were incubated on ice for 15 min and collected by centrifugation (4°C, 15 min, full speed). The protein-containing supernatant was transferred into a fresh tube and stored at -20°C until further processing.

For Western blot analysis, protein samples were separated on Mini-PROTEAN® TGX Stain-Free™ Precast gels (Bio-Rad, Feldkirchen, Germany) and transferred to a PVDF membrane using a Trans-Blot Turbo Mini Transfer Pack kit (Bio-Rad, Feldkirchen, Germany) according to manufacturer's instructions. Membranes were blocked with 5% milk powder TBST solution overnight at 4 °C prior to primary antibody (anti-OctA FLAG®, 1:1000, sc-166355, Santa Cruz and anti-α smooth muscle actin, 1:5000, ab5694, abcam) incubation for 2 h at RT. After washing, secondary antibody incubation (1 h, RT, anti-mouse IgG HRP, 1:1000, sc-2357, Santa Cruz)) followed. Chemiluminescence was detected with the Clarity™ Western ECL Substrate kit (Bio-Rad, Feldkirchen, Germany) and the ChemiDoc MP Imaging System (Bio-Rad, Feldkirchen, Germany).

Supplemental Note: M1AP missense variant description and interpretation

We identified another missense variant, c.797G>A, in individual M1943 which results in the substitution of arginine, a positively charged amino acid, with a neutrally charged glycine at position 266 (p.(Arg266Gly)).[10] The substituted amino acid is the only residue found at this position and can therefore be classified as highly conserved (up to zebrafish; Table 2). A mutation at this position is usually damaging for the protein, especially if being exchanged with such a different amino acid like glycine. As a smaller, more flexible and hydrophobic amino acid, glycine could for example disturb the required rigidity of the resulting protein or disturb correct protein folding.[10] The effect of this mutation is predicted to be damaging, deleterious or disease causing by all used *in silico* programs.

In addition to the heterozygous frameshift c.676dup, individual Y126 carries the substitution c.949G>A. The missense variant replaces the highly conserved (up to platypus; Figure 3C) neutral and nonpolar amino acid glycine with the larger, positively charged amino acid arginine (p.(Gly317Arg)). Based on conservation information, the variation in this position is highly likely to impair M1AP protein function.[10] The introduction of a charge may cause the repulsion of interaction partners or of other positively charged residues. In addition, the altered torsion angles may have an influence on the correct conformation and disturb the local structure of the protein. Correspondingly, the amino acid change p.(Gly317Arg) is predicted to be deleterious by all *in silico* algorithms.

Individual P86 carries two missense variants in *M1AP*. The variant c.1289T>C leads to a substitution of the hydrophobic amino acid leucine, which is predicted to be located in an alpha helix, with the less hydrophobic amino acid proline (p.(Leu430Pro)).[10] Because proline is an alpha helix breaker, the alteration likely has severe effects on protein structure. This change is predicted to be disease causing by all *in silico* programs. Furthermore, leucine at position 430 is a highly conserved amino acid (up to platypus; Figure 3D). The substitution c.148T>C replaces the polar amino acid serine at position 50 with the nonpolar amino acid proline (p.(Ser50Pro)). Again, the wildtype residue is predicted to be located in an alpha helix, and, therefore, its substitution by a proline likely has severe effects on protein structure and function.[10] Although the *in silico* algorithms SIFT and MutationTaster predict this change as being tolerated and as a polymorphism, the variant has not previously been described in any population, which supports a possible pathogenic impact. Moreover, this amino acid is likewise highly conserved (up to platypus; Figure 3D).

The substitution c.1166C>T identified in a consaguinous Turkish family leads to a replacement of the highly conserved and weakly hydrophobic proline with the more hydrophobic leucine at position 389 (p.(Pro389Leu), Figure 4C). Proline is known to have a rigid structure, giving a protein a specific conformation, which could be disrupted by substitution with leucine.[10]

Accordingly, the amino acid exchange is predicted to be pathogenic by all *in silico* programs and clinical variant interpretation classifies it as variant of uncertain significance.

Supplemental references

1. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424. doi:10.1038/gim.2015.30
2. Tu C, Wang Y, Nie H, et al. An M1AP homozygous splice-site mutation associated with severe oligozoospermia in a consanguineous family. *Clin Genet*. 2020;(December 2019):741-746. doi:10.1111/cge.13712
3. Smith ED, Radtke K, Rossi M, et al. Classification of Genes: Standardized Clinical Validity Assessment of Gene–Disease Associations Aids Diagnostic Exome Analysis and Reclassifications. *Hum Mutat*. 2017;38(5):600-608. doi:10.1002/humu.23183
4. Arango NA, Huang TT, Fujino A, Pieretti-Vanmarcke R, Donahoe PK. Expression analysis and evolutionary conservation of the mouse germ cell–specificD6Mm5e gene. *Dev Dyn*. 2006;235(9):2613-2619. doi:10.1002/dvdy.20907
5. Arango NA, Li L, Dabir D, et al. Meiosis I Arrest Abnormalities Lead to Severe Oligozoospermia in Meiosis 1 Arresting Protein (M1ap)-Deficient Mice1. *Biol Reprod*. 2013;88(3):1-6. doi:10.1095/biolreprod.111.098673
6. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. 2011;17(1):10. doi:10.14806/ej.17.1.200
7. Li H, Durbin R. Fast and accurate long-read alignment with Burrows – Wheeler transform. *Bioinformatics*. 2010;26(5):589-595. doi:10.1093/bioinformatics/btp698
8. McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010. doi:10.1101/gr.107524.110
9. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122. doi:10.1186/s13059-016-0974-4
10. Venselaar H, te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics*. 2010;11(1):548. doi:10.1186/1471-2105-11-548