bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1 **Biallelic mutations in *M1AP* are a frequent cause of meiotic arrest**
2 **leading to male infertility**

3
4

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

12
13

14 [1]Institute of Human Genetics, University of Münster, Münster, Germany
15 [2]Bursa Uludag University, Faculty of Medicine, Department of Medical Genetics &
16 Department of Histology & Embryology & Health Sciences Institute, Department of
17 Translational Medicine, Bursa, Turkey
18 [3]Division of Genetics, Oregon National Primate Research Center, Oregon Health & Science
19 University, Portland, USA
20 [4]Department of Human Genetics, Donders Institute for Brain, Cognition and Behavior,
21 Radboud University Medical Center, Nijmegen, The Netherlands
22 [5]Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), Porto,
23 Portugal
24 [6]Instituto de Investigação e Inovação em Saúde (i3s), Universidade do Porto, Porto, Portugal
25 [7]Department of Obstetrics and Gynecology, Radboud University Medical Center, Nijmegen,
26 The Netherlands
27 [8]Centre of Reproductive Medicine and Andrology, Institute of Reproductive Medicine,
28 University of Münster, Münster, Germany
29 [9]Institute of Medical Informatics, University of Münster, Münster, Germany
30 [10]Serviço de Genética, Departamento de Patologia, Faculdade de Medicina da Universidade
31 do Porto, Porto, Portugal
32 [11]Dep. de Genética Humana, Instituto Nacional de Saúde Dr. Ricardo Jorge, Lisboa, Portugal
33 [12]Institute of Veterinary Anatomy, Histology and Embryology, Justus Liebig University,
34 Gießen, Germany
35 [13]University of Acibadem, Acibadem Genetic Diagnostic Centre, Istanbul, Turkey
36 [14]Near East University, Faculty of Medicine, Department of Medical Biology, Nicosia, Cyprus

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

37 [15]Istanbul Memorial Hospital, Assisted Reproductive Technologies and Reproductive
38 Genetics Centre, Istanbul, Turkey
39 [16]Istanbul Memorial Hospital, Department of Urology & Andrology, Istanbul, Turkey
40 [17]Istanbul Memorial Hospital, Assisted Reproductive Technologies and Reproductive
41 Genetics, Istanbul, Turkey
42 [18]Clinic for Urology, Pediatric Urology and Andrology, Justus Liebig University, Gießen,
43 Germany
44 [19]Centre of Reproductive Medicine and Andrology, Department of Clinical and Surgical
45 Andrology, University Hospital Münster, Münster, Germany
46 [20]Andrology and IVF Laboratories, Department of Surgery, University of Utah School of
47 Medicine, Salt Lake City, Utah, USA
48 [21]Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK
49
50
51 [#]These authors contributed equally to this work and should be considered joint first/last
52 authors.
53
54
55 ***Correspondence:** Institute of Human Genetics, University of Münster, Vesaliusweg 12-14,
56 48149 Münster, Germany, frank.tuettelmann@ukmuenster.de
57
58
59 **Word count:** 5162 (text), 250 (abstract)
60

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

61 **Abstract**

62 Male infertility affects ~7% of men in Western societies, but its causes remain poorly
63 understood. The most clinically severe form of male infertility is non-obstructive azoospermia
64 (NOA), which is, in part, caused by an arrest at meiosis, but so far only few genes have been
65 reported to cause germ cell arrest in males. To address this gap, whole exome sequencing
66 was performed in 60 German men with complete meiotic arrest, and we identified in three
67 unrelated men the same homozygous frameshift variant c.676dup (p.Trp226LeufsTer4) in
68 *M1AP,* encoding meiosis 1 arresting protein. Then, with collaborators from the International
69 Male Infertility Genomics Consortium (IMIGC), we screened a Dutch cohort comprising 99
70 infertile men and detected the same homozygous variant c.676dup in a man with
71 hypospermatogenesis predominantly displaying meiotic arrest. We also identified two
72 Portuguese men with NOA carrying likely biallelic loss-of-function (LoF) and missense
73 variants in *M1AP* among men screened by the Genetics of Male Infertility Initiative (GEMINI).
74 Moreover, we discovered a homozygous missense variant p.(Pro389Leu) in *M1AP* in a
75 consanguineous Turkish family comprising five infertile men. M1AP is predominantly
76 expressed in human and mouse spermatogonia up to secondary spermatocytes and
77 previous studies have shown that knockout male mice are infertile due to meiotic arrest.
78 Collectively, these findings demonstrate that both LoF and missense *M1AP* variants that
79 impair its protein cause autosomal-recessive meiotic arrest, non-obstructive azoospermia
80 and male infertility. In view of the evidence from several independent groups and
81 populations, *M1AP* should be included in the growing list of validated NOA genes.

82

83 **Key words:** infertility, non-obstructive azoospermia, spermatogenesis, meiotic arrest, M1AP

84

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

85 **Introduction**

86 Around 7% of all men in Western societies are infertile,[1] which is primarily diagnosed by
87 semen analysis, comprising as the most relevant parameters sperm concentration and count.
88 More than 10% of all infertile men exhibit azoospermia,[2] which is defined as the absence of
89 spermatozoa in the ejaculate and constitutes the most clinically severe form of male
90 infertility, resulting in zero chance of natural conception.[3] Azoospermia is further classified
91 into obstructive and non-obstructive azoospermia (OA and NOA, respectively). In a large
92 fraction of NOA cases a genetic origin is assumed,[4] such that patients with azoospermia are
93 routinely screened for chromosomal aberrations and Y-chromosomal azoospermia factor
94 (AZF) microdeletions. However, these diagnostic tests only establish a reason for the
95 azoospermia in 15-20% of cases.[2]

96 NOA can be stratified into four groups based on histological analysis of the seminiferous
97 tubules: normal spermatogenesis (a finding which indicates OA and post-testicular defects),
98 hypospermatogenesis, Sertoli-cell only (SCO), and maturation arrest. Maturation arrest most
99 frequently presents as meiotic arrest, in which spermatocytes are the most advanced germ
100 cell types in the testes.[5] If germ cell arrest is only partial, some mature spermatozoa will
101 develop, meaning that men with partial germ cell arrest can become parents by undergoing
102 testicular biopsy and sperm extraction (TESE) and then using the extracted sperm in
103 assisted reproductive technology (ART). By contrast, if germ cell arrest is complete, no
104 mature spermatozoa will develop and TESE cannot be successful. However, thus far, germ
105 cell arrest can only be reliably diagnosed by testicular biopsy, i.e., after the surgery,
106 emphasizing the urgent need for better diagnostic tests before biopsy to avoid unnecessary
107 and unsuccessful surgical procedures.

108 Better diagnostic workup could be developed upon identifying causal genes or mutations that
109 promote germ cell arrest. Recently, the first monogenic alterations associated with germ cell
110 arrest in human males have been described. However, according to a standardized
111 assessment of clinical validity, the X-linked gene *TEX11*[6] (MIM: 300311) currently remains
112 the only gene with strong evidence.[7] Given the large number of genes known to cause

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

113 meiotic arrest in mice, the vast majority of causal mutations causing this phenotype in
114 humans are yet to be identified.

115 To this end, we first screened the exomes of well-characterized patients with complete,
116 bilateral germ cell arrest at the spermatocyte stage for variants in testis-expressed genes. As
117 a result, we identified biallelic loss-of-function (LoF) variants in three unrelated infertile men
118 in the gene encoding meiosis 1 arresting protein (*M1AP*). Furthermore, by screening two
119 independent cohorts, we found the same homozygous LoF variant in a patient with
120 hypospermatogenesis predominantly displaying meiotic arrest, and we found likely
121 pathogenic missense variants in *M1AP* in two NOA patients. Additionally, we identified a
122 homozygous missense variant in *M1AP* segregating with azoospermia in five infertile men in
123 a consanguineous Turkish family. *M1ap* is primarily expressed in male germ cells during
124 spermatogenesis, and its knockout in male mice leads to infertility due to meiotic arrest.[8,9]

125 Our results, together with previously published findings, provide sufficient evidence that
126 M1AP plays an essential role during spermatogenesis and its loss causes NOA in a
127 considerable proportion of infertile men. This allows for a better understanding of the
128 molecular basis of meiotic arrest and improved counseling and treatment of infertile couples.

129

130 **Subjects and Methods**

131 *Study cohorts*

132 We originally selected 64 azoospermic but otherwise healthy male patients who attended the
133 Centre of Reproductive Medicine and Andrology (CeRA), University Hospital Münster
134 (N = 51) or the Clinic for Urology, Pediatric Urology and Andrology, Gießen (N = 13), for
135 couple infertility. This is a subset of all patients included in our large-scale Male Reproductive
136 Genomics (MERGE) study, which currently comprises >800 men including 514 with NOA
137 (Figure S1). All of the 64 patients were diagnosed with complete bilateral germ cell arrest at
138 the spermatocyte stage after evaluating at least 100 seminiferous tubules in tissue sections
139 of both testes accompanied by a negative TESE outcome, i.e., no sperm could be recovered.

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

140 Chromosomal aberrations and AZF deletions were excluded in advance. Four out of 64
141 patients were diagnosed with a LoF variant in *TEX11* previously (Yatsenko *et al.* 2015 and
142 unpublished data).[2,6] In addition, 27 men, also attending the CeRA for couple infertility and
143 with normal semen parameters (normozoospermia according to WHO[3]), were included as
144 controls. All patients gave written informed consent for the evaluation of their clinical data
145 and analysis of their DNA samples. The study protocol was approved by the respective
146 Ethics Committees/Institutional Review Boards (Ref. No. Münster: 2010-578-f-S, Gießen:
147 26/11, Nijmegen: NL50495.091.14 version 4, GEMINI consortium: 201502059, Porto:
148 PTDC/SAU-GMG/101229/2008, Bursa: 05.01.2015/04) in accordance with the Helsinki
149 Declaration of 1975.

150 As a next step, a study cohort of 99 men with unexplained azoospermia (N = 55) or severe
151 oligozoospermia (N = 44) who presented at Radboud University Medical Center
152 (Radboudumc, Nijmegen) was screened for biallelic variants in *M1AP*. In parallel, we
153 screened the whole exome sequencing data produced within the GEMINI study
154 (https://gemini.conradlab.org/) of 979 unrelated men with unexplained NOA for biallelic
155 variants in *M1AP*. In both cohorts, chromosomal aberrations, AZF deletions and *CFTR*-
156 mutations had been excluded.

157 Additionally, we performed whole exome sequencing in two brothers with unexplained
158 infertility from a consanguineous Turkish family and one fertile brother. The index patient
159 T1024 (V.2; Figure 3A) and his wife presented at Uludag University Faculty of Medicine
160 Hospital because of couple infertility. Semen analysis revealed azoospermia, and
161 chromosomal aberrations as well as AZF deletions were excluded. The patient reported that
162 he had an infertile brother and three further infertile male relatives (Figure 3A). We focused
163 on rare homozygous variants shared by both affected brothers. Subsequently, seven male
164 family members and the mother of T1024 were screened for the *M1AP* variant detected in
165 the infertile brothers.

166

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

167 *Whole exome sequencing (WES) (MERGE study)*
168 Genomic DNA was extracted from peripheral blood leukocytes via standard methods.[10] WES
169 sample preparation and enrichment were carried out in accordance with the protocols of
170 either Agilent's SureSelect^{QXT} Target Enrichment kit or Twist Bioscience's Twist Human Core
171 Exome kit. Agilent's SureSelect^{XT} Human All Exon Kits V4, V5 and V6 or Twist Bioscience's
172 Human Core Exome plus RefSeq spike-in's were used to capture libraries. For multiplexed
173 sequencing, the libraries were index tagged using appropriate pairs of index primers.
174 Quantity and quality of the libraries were assessed with the ThermoFisher Qubit and Agilent's
175 TapeStation 2200, respectively. Sequencing was conducted on the Illumina HiScan®SQ,
176 NextSeq®500, or HiSeqX® systems using the TruSeq SBS Kit v3 - HS (200 cycles), the
177 NextSeq 500 V2 High-Output Kit (300 cycles) or the HiSeq Rapid SBS Kit V2 (300 cycles),
178 respectively.

179 After trimming, Cutadapt v1.15 was used to remove the remaining adapter sequences and
180 primers[11]. Sequence reads were aligned against the reference genome GRCh37.p13 using
181 BWA Mem v0.7.17[12]. We excluded duplicate reads and reads that mapped to multiple
182 locations in the genome from further analysis. Small insertions/deletions (indels) and single
183 nucleotide variations were identified and quality-filtered by GATK toolkit v3.8 with
184 HaplotypeCaller, in accordance with the best practice recommendations.[13] Ensembl Variant
185 Effect Predictor was used to annotate called variants.[14]

186 DNA extraction, WES and variant calling in patients from the other groups (RU01691, Y126,
187 P86 and T1024) were carried out according to the standard local procedures. The respective
188 details are provided in the Supplemental Methods.

189

190 *Data analysis and variant prioritization*
191 In the MERGE study, variation categories, transcript and functional consequences,
192 population frequencies, and *in silico* predicted relevance were annotated to each variant
193 utilizing the in-house pipeline Sciobase©. Likely causative variants were identified by filtering

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

194 the data according to (i) the recessive mode of inheritance, (ii) the population frequency in
195 the Genome Aggregation Database[15] (gnomAD, minor allele frequency [MAF] < 0.01), and
196 (iii) the functional impact of the variant (loss of function: splice site, frameshift, stop
197 gained/lost, start lost) (Figure S1). Finally, the relevance for the phenotype was assessed
198 using comprehensive expression data (Genotype-Tissue Expression [GTEx] project[16]) and
199 model organism data from the literature. In a complementary approach, an updated version
200 of the population sampling probability (PSAP) pipeline[17] was used to prioritize potentially
201 causative variants. PSAP models the significance of observing a single subject's genotype in
202 comparison to genotype frequencies in unaffected populations (commonly referred to as the
203 'n-of-one' problem). The resulting variant lists were filtered as described by Kasak *et al.*[18]
204 (MAF ≤ 0.01, PopScore ≤ 0.005, and CADD ≥ 20).

205 To assess the pathogenicity of detected missense variants, we used common *in silico*
206 prediction programs (PolyPhen-2, SIFT, MutationTaster, HOPE[19]). We attempted to model
207 the 3D structure of the M1AP protein. However, due to the lack of previous information on
208 M1AP and comparable 3D structures, it was not possible to achieve a reliable prediction
209 (BLAST for sequence of UniProt identifier Q8TC57 all below 30% sequence identity to known
210 protein structures, details in Suppl. Methods).

211

212 *Sanger sequencing for variant validation and screening of normozoospermic controls*
213 All relevant variants identified in azoospermic men were confirmed by direct Sanger
214 sequencing of the respective exons of *M1AP* (NM_001321739.1) according to standard
215 procedures.[10] To establish the carrier frequency of the recurring variant c.676dup in exon 5 in
216 an ancestry-matched control group, 285 normozoospermic men (from couples attending the
217 CeRA) were analyzed. In the five heterozygous carriers of this variant, the whole coding
218 region of *M1AP* (exons 2-11) was subsequently sequenced, to exclude a second pathogenic
219 variant in *M1AP*. Primer sequences are provided in Table S1.

220

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

221 *Quantitative PCR analysis*

222 To exclude a hemizygous deletion on the other allele, quantitative PCR (qPCR) of exon 5
223 was performed on gDNA of the three patients from the MERGE study (M330, M864, M1792)
224 carrying the variant c.676dup. qPCR was carried out in 96-well plates on the LightCycler 480
225 using the manufacturer's default settings. The *ALB* (albumin) gene was used for
226 normalization. The reactions were performed in triplicates using the SensiMix Real-Time
227 PCR Kit (Bioline). The PCR consisted of an initial incubation step at 95°C for 10 min followed
228 by 40 cycles of 95°C for 15 sec, 60°C for 30 sec and 72°C for 15 sec. Baseline and threshold
229 values were automatically detected using the LightCycler software. Primers are provided in
230 Table S1.

231

232 *Histological evaluation and testing of M1AP antibodies*

233 Testis biopsies of patients M330, M864 and M1792 were collected for TESE and research
234 use. Biopsies were fixed in Bouin's solution overnight at 4°C, washed with 70% ethanol and
235 embedded in paraffin using an automatic ethanol and paraffin row (Bavimed Laborgeräte
236 GmbH, Birkenau, Germany). Subsequently, 5 µm sections were stained with periodic acid-
237 Schiff (PAS) as previously described.[20]

238 We attempted to establish immunohistochemistry as well as Western blot analyses with
239 commercially available M1AP antibodies (#PA5-31627, ThermoFisher Scientific,
240 Langenselbold, Germany and #HPA045420, Sigma-Aldrich, Munich, Germany). These were
241 selected based on available data from the manufacturers and the human protein atlas (see
242 Suppl. Data for details).

243

244 *Search for M1AP variants in women with premature ovarian insufficiency*

245 Because some genes have been reported to be associated with both male and female
246 infertility, we screened 101 women diagnosed with unexplained premature ovarian
247 insufficiency (POI) (62 with isolated POI and 39 with ovarian dysgenesis) by direct Sanger

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

248 sequencing of the full coding region. Details of a part of this cohort (N = 25) have been
249 published previously.[21] The additional 76 patients followed the same in- and exclusion
250 criteria.

251

252 **Results**

253 *Identification of M1AP as a candidate gene and follow-up study in three independent groups*

254 In the initial MERGE study, the WES data of 60 highly selected, azoospermic infertile
255 patients with unexplained, complete, bilateral germ cell arrest at the spermatocyte stage
256 were screened for rare (MAF < 0.01 according to gnomAD-database[22]), biallelic LoF variants.
257 Two patients were subsequently excluded from this study, because likely causal variants in
258 other genes had been identified in parallel: patient M870 had compound heterozygous
259 variants in *STAG3,*[23] and patient M1401 had a heterozygous LoF variant in *SYCP2*.[24] The
260 prioritized genes in the remaining 58 patients were analyzed with regard to the level of
261 expression in the testes. A literature search was performed to identify genes with previous
262 evidence for an association with infertility in either human or model species. The highest
263 prioritized gene was *M1AP* because three unrelated men (M330, M864, M1792, Figure 1)
264 carried the same potentially homozygous LoF variant (c.676dup, MAF = 0.0021, no
265 homozygotes in gnomAD[22]), the M1AP protein displays the highest expression in the testis,
266 and it has been shown to play a crucial role in spermatogenesis in mice.[8,9] The variant
267 c.676dup is located in exon 5 of 11 and causes a frameshift and premature stop codon
268 (p.Trp226LeufsTer4), as confirmed by testicular cDNA sequencing of exons 5 of patient
269 M864 (Figure S2B). Quantitative PCR analysis of *M1AP* exon 5 excluded an intragenic
270 deletion and, thus, confirmed homozygosity for c.676dup in all patients (Table S2). No
271 regions of homozygosity (ROH) involving *M1AP* were detected for any of the three patients
272 homozygous for c.676dup, rendering consanguinity of their parents unlikely. Neither did we
273 notice evidence for consanguinity between the patients (analysis by H3M2 algorithm,[25] data
274 not shown). No patient carrying two rare variants in *M1AP* (MAF ≤ 0.01 from gnomAD) was
275 identified by WES in the remaining MERGE cohort of almost 750 patients with other infertility

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

276 phenotypes (mostly azoospermia due to other testicular phenotypes such as SCO or without
277 biopsy, N = 548, or severe oligozoospermia, N = 120).

278 Next, and through a collaboration established within the International Male Infertility
279 Genomics Consortium (IMIGC, imigc.org), we identified three additional infertile patients with
280 likely biallelic mutations in *M1AP* from two independent study groups. All identified *M1AP*
281 variants in all six patients were confirmed by Sanger sequencing. The variants, gnomAD
282 frequencies, *in silico* predictions, and clinical data of all patients carrying *M1AP* variants are
283 shown in Table 1.

284 Patient RU01691 from the Netherlands was also homozygous for the frameshift variant
285 c.676dup (p.Trp226LeufsTer4) in *M1AP* (Figure 2A). The patient's parents were both
286 heterozygous carriers. Testicular biopsy in this patient showed bilateral severe
287 hypospermatogenesis with predominantly meiotic arrest (Figure 2B); occasionally spermatids
288 were present.

289 Two patients analyzed within the GEMINI study and of Portuguese origin (P86, Y126) each
290 carried two different variants in *M1AP* (Figure 2C and 2D). Patient Y126 carried the
291 missense variant c.949G>A (p.(Gly317Arg); MAF = 0.00007) and, in addition, the recurrent
292 frameshift variant c.676dup (p.Trp226LeufsTer4). This patient had germ cell arrest at the
293 round spermatid stage as diagnosed by testicular biopsy and histological evaluation. The
294 other patient P86 had two missense variants c.148T>C (p.(Ser50Pro); not described in
295 gnomAD) and c.1289T>C (p.(Leu430Pro); MAF = 0.000008), suggesting compound-
296 heterozygosity. Testicular histology of this patient showed thickened seminiferous tubules
297 and dispersed Sertoli cells with some tubules containing only spermatogonia.

298 In parallel, and independent of the identification of *M1AP* in the MERGE study, WES was
299 performed in two infertile, azoospermic brothers from a consanguineous Turkish family as
300 well as their fertile brother (Figure 3A). The data were analyzed focusing on rare
301 homozygous variants, which were shared between the infertile brothers but not found in the
302 fertile brother. The two affected men carried rare, homozygous missense variants in the

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

autosomal genes *AMPD2*, *CELSR2*, *CEP164*, and *M1AP* as well as rare hemizygous variants in the X-linked genes *ATG4A* and *ENOX2*. Of these genes, *M1AP* is the only one that has been described in the context of infertility. Both infertile men carried the homozygous missense variant c.1166C>T (p.(Pro389Leu); MAF = 0.00001) in *M1AP*, which was found in a heterozygous state in the fertile brother. No homozygous carriers of this variant have been described so far in any public databases. Subsequently, this variant was also found in a homozygous state in three additional infertile males from the same family. Additionally, after recruiting and examining nine fertile members of the Turkish family, we found that none of them were homozygous for this variant. We did identify both a fertile man and a fertile woman (IV.13 and V.6 in Figure 3A, respectively) as heterozygous carriers of the same variant (example result of Sanger sequencing for subject V.6 shown in Figure 3B). Clinical data was only available for the index case T1024, and this patient had borderline follicle stimulating hormone (FSH, 8.3 U/L) as well as testicular volume (15 mL both left and right) (Table 1). This combination is often found in azoospermic men with germ cell arrest, while most other infertile NOA patients lacking advanced germ cells have elevated FSH levels because of the pituitary-gonadal feedback loop.

In addition, we analyzed all identified variants from the respective WES analyses in each patient individually utilizing the PSAP pipeline. This pipeline enabled us to rank all variants per patient, following the prioritization criteria (MAF ≤ 0.01, PopScore ≤ 0.005 and CADD ≥ 20). The biallelic *M1AP* LoF variants were ranked in the first position for patient M1792 and in the third position for patients M330 and M864 in the discovery cohort (Table S3). The four patients identified in the follow-up analyses exhibited highly ranked *M1AP* variants as well: position 17 in patient RU01691, position eleven in patient Y126, position 33 in patient P86 and position 12 in T1024, respectively.

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

328 *M1AP immunostaining and Western blot*

329 After optimization, both commercial antibodies resulted in a specific signal in
330 immunohistochemistry (Figure S3/4). However, both antibodies seem not to pick up M1AP
331 but a different target. First of all, similar signals could also be detected in two of the patients
332 carrying the homozygous frameshift variant c.676dup, which lead to a disruption of the
333 supposed epitope of the antibodies presumed to reside in the region of or downstream of the
334 variant. Moreover, the Western blot did not only result in detection of a band in testis lysate
335 but also in a kidney lysate where M1AP is not expressed and, last not least, the bands were
336 not at the expected size (Figure S5). Therefore, we also reached out to the colleagues who
337 published the *M1AP* knockout mice and Western blot staining with a self-raised antibody,[9]
338 but were unsuccessful in getting in contact.

339

340 *Control cohorts and women with POI*

341 Because of the rather surprising finding of a rare but recurring variant c.676dup in the
342 primary MERGE study group, we aimed to establish the carrier frequency in an ancestry-
343 matched control population. To this end, we performed Sanger sequencing of exon 5 of
344 *M1AP* in 285 fertile men. We indeed detected five fertile men carrying the same frameshift
345 variant c.676dup (p.Trp226LeufsTer4) in *M1AP*, but these were, in contrast to the three
346 affected patients (M330, M864, M1792), in a heterozygous state. No homozygous carriers
347 were detected, resulting in an allele frequency of 0.0088. We next performed Sanger
348 sequencing for the complete coding region of the *M1AP* gene (exon 2 to 11) in the five
349 heterozygous carriers to rule out the presence of a second variant. No coding alterations
350 were detected.

351 Additionally, we queried a previously established database of 3347 Dutch fertile couples who
352 had conceived at least one child. WES had been performed in these subjects as part of
353 clinical diagnostics and workup of a child with severe development delay (trio-WES). Again,
354 20 heterozygous male as well as 10 heterozygous female carriers were detected but no

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

355 homozygous subjects were found, resulting in an allele frequency of *M1AP* c.676dup in the
356 Dutch cohort of 0.0022. Also, there were no homozygous carriers of other LoF variants
357 among either fathers or mothers. No women carrying two rare variants in *M1AP* were
358 identified in the analyzed 101 women affected by POI.

359

360 **Discussion**

361 We have identified a total of eleven patients from four independent cohorts and provide
362 convincing evidence that biallelic variants in *M1AP* are associated with predominantly germ
363 cell arrest in otherwise healthy men. Previously, disruption of *M1ap* has been shown to
364 cause a highly similar testicular phenotype in mice. The *M1ap* knockout mice showed severe
365 oligozoospermia due to predominantly tubular defects with no germ cells beyond the
366 spermatocyte stage, consequently resulting in infertility.[9] So far, only very few other genes
367 with mutations leading to germ cell arrest in both men and mice and validated in independent
368 cohorts have been published. The first was the X-chromosomal gene *TEX11*,[6,26] which is,
369 according to a current structured assessment, one of only a few genes with strong clinical
370 validity for an association with NOA.[7] Another example is the autosomal gene *STAG3*, which
371 has only recently been described in publications by us and others in parallel,[23,27] and its
372 clinical validity is currently 'moderate'. Most of the proteins involved in DNA recombination,
373 including *M1AP*, are highly evolutionarily conserved.[8] This suggests that these genes are not
374 tolerant to variation likely due to an infertility phenotype. Because of the previously available
375 evidence from mice and because we replicated our primary finding in several independent
376 groups as well as a consanguineous family, *M1AP* immediately reaches a 'moderate' clinical
377 validity (Table S4). This clearly underlines the strength of such collaborative efforts, which
378 have only recently been established in the context of the previously slowly progressing field
379 of male infertility genetics.

380 From our initial cohort of 64 men with complete bilateral meiotic arrest, we also identified
381 likely causal variants in three other genes, namely *TEX11*, *STAG3* and *SYCP2*. These

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

382 results have been reported elsewhere.[6,23,24] Among the remaining group of 58 men, we
383 identified three unrelated patients with the same homozygous frameshift variant c.676dup
384 (p.Trp226Leufs*4) in *M1AP*. Thus, up to 5% of men affected by male infertility, non-
385 obstructive azoospermia and meiotic arrest (3 out of 64) carry causal mutations in *M1AP*,
386 making this a comparably common reason for germ cell arrest at the spermatocyte stage like
387 *TEX11* mutations with 6% (4 out of 64) in this selected patient population. Overall, *M1AP*
388 contributes to less than 1% of the highly heterogeneous NOA cases (0.4%, 6 out of 1548
389 patients in total across the three screened cohorts: 514 and 979 from the MERGE and
390 GEMINI studies, respectively, and 55 from the Nijmegen trio cohort).

391 The fact that the same frameshift variant c.676dup was also found to be homozygous in
392 RU01691 from the Netherlands and heterozygous in patient Y126 from Portugal suggests
393 that this variant is relatively prevalent in European populations. The likely explanation for its
394 relative commonness is that this mutation is a founder mutation present in men of European
395 ancestry. Yet, although we found this frameshift variant four times in a homozygous state in
396 unrelated men with germ cell arrest, it is a rarely described variant in global large databases.
397 The gnomAD database does not list any homozygous men, and the maximum allele
398 frequency is 0.0038 in the European (non-Finnish) population. In our control cohort of 285
399 German men with normal sperm production, we detected five heterozygous carriers of
400 c.676dup, corresponding to an allele frequency of 0.0088. This difference in allele frequency
401 can be explained either by the relatively small size of the control cohort or by an enrichment
402 in the population attending the CeRA in Münster, i.e., of Westphalian origin. In the larger
403 Dutch cohort of 3347 fertile couples, 30 heterozygous carriers were found, corresponding to
404 an allele frequency of 0.0022.

405 The relevance of the homozygous frameshift variant c.676dup in *M1AP* found in four patients
406 is underlined by the very low PopScore ($9.7x10^{-7}$) obtained by PSAP and the high
407 prioritization of the *M1AP* variants in all analyzed patients who carried the variants
408 (Table S3). Moreover, the expected mode of inheritance for this gene is autosomal recessive

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

409 according to a general prediction,[28] fitting our observations of biallelic variants in the affected
410 patients.

411 This frameshift variant c.676dup very likely causes a premature stop codon four amino acids
412 downstream (p.Trp226Leufs*4). This could result either in an altered and severely truncated
413 protein with 230 amino acids (normal protein length: 530 amino acids), or in nonsense-
414 mediated decay (NMD) of the mRNA. Analysis of patient M864's testis RNA resulted in an
415 equal band compared to control testis RNA, and we therefore exclude elimination through
416 NMD (Figure S2A). Unfortunately, both protein analysis via Western blot and
417 immunohistological stainings failed because of the lack of suitable antibodies raised against
418 M1AP (Figures S3, S4, and S5).

419 In addition to the heterozygous frameshift c.676dup, patient Y126 carries the substitution
420 c.949G>A. The missense variant replaces the highly conserved (up to platypus; Figure 2C)
421 neutral and nonpolar amino acid glycine with the larger, positively charged amino acid
422 arginine (p.(Gly317Arg)). Based on conservation information, the variation in this position is
423 highly likely to impair M1AP protein function.[19] The introduction of a charge may cause the
424 repulsion of interaction partners or of other positively charged residues. In addition, the
425 altered torsion angles may have an influence on the correct conformation and disturb the
426 local structure of the protein. Correspondingly, the amino acid change p.(Gly317Arg) is
427 predicted to be deleterious by all *in silico* algorithms.

428 Patient P86 carries two missense variants in *M1AP*. The mutation c.1289T>C leads to a
429 substitution of the hydrophobic amino acid leucine, which is predicted to be located in an
430 alpha helix, with the less hydrophobic amino acid proline (p.(Leu430Pro)).[19] Because proline
431 is an alpha helix breaker, the alteration likely has severe effects on protein structure. This
432 change is predicted to be disease causing by all *in silico* programs. Furthermore, leucine at
433 position 430 is a highly conserved amino acid (up to platypus; Figure 2D). The substitution
434 c.148T>C replaces the polar amino acid serine at position 50 with the nonpolar amino acid
435 proline (p.(Ser50Pro)). Again, the wildtype residue is predicted to be located in an alpha

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

436 helix, and, therefore, its substitution by a proline likely has severe effects on protein structure
437 and function.[19] Although the *in silico* algorithms SIFT and MutationTaster predict this change
438 as being tolerated and as a polymorphism, the variant has not previously been described in
439 any population, which supports a pathogenic impact. Moreover, this amino acid is likewise
440 highly conserved (up to platypus; Figure 2D).

441 We add another layer of evidence that *M1AP* missense variants cause NOA with data from a
442 consanguineous family showing segregation of the homozygous *M1AP* missense variant
443 c.1166C>T (p.(Pro380Leu) with affected family members (LOD score = 3.28). Of note, the
444 group of SGM identified *M1AP* as candidate for the affected, azoospermic family members
445 independently from our initial identification. Although the index case T1024 did not undergo a
446 testicular biopsy, his clinical data is indicative of NOA due to germ cell arrest. In contrast, all
447 fertile family members investigated had at least one wildtype allele, which underlines the
448 pathogenicity of this variant when it is biallelic. The substitution c.1166C>T leads to a
449 replacement of the highly conserved and weakly hydrophobic proline with the more
450 hydrophobic leucine at position 389 (p.(Pro389Leu), Figure 3C). Proline is known to have a
451 rigid structure, giving a protein a specific conformation, which could be disrupted by
452 substitution with leucine.[19] Accordingly, the amino acid exchange is predicted to be
453 pathogenic by all *in silico* programs.

454 The possible effects of the detected missense variant on post-translational modifications and
455 folding of the M1AP protein are difficult to assess and, thus, inherently uncertain.
456 Unfortunately, the structure of the M1AP protein is currently unknown and it is impossible to
457 predict functional domains of the M1AP protein with high reliability. Moreover, the specific
458 molecular function of M1AP remains to be elucidated in subsequent studies.

459 The process of meiosis is in part similar between males and females, but orchestrated highly
460 differently concerning its timing. Thus far, only few genes have been reported to impair male
461 as well as female meiosis. As an example, variants in *STAG3* had previously been reported
462 to cause POI[29] and now have recently been shown to also cause NOA and meiotic arrest.[23,27]

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

463 *M1AP* is predominantly expressed in the adult testis (GTEx), but also reported to be
464 expressed in the fetal mouse ovary.[8] However, the structure of ovaries in female *M1ap*
465 knockout mice appeared normal and fertility was preserved.[9] Concordantly, we did not
466 identify any relevant variants in *M1AP* in 101 women affected by POI. Still, we neither found
467 a fertile woman carrying a homozygous LoF variant in the large Dutch trio cohort. From this
468 data, we cannot exclude that variants in *M1AP* may be a rare cause also for POI, but it
469 seems likely that M1AP is only required for male meiosis.

470 The common phenotype among our patients is NOA, and four out of six with testicular
471 histology available had either complete (N = 3) or predominant (N = 1) meiotic arrest, i.e.,
472 germ cell arrest at the spermatocyte stage. A similar phenotype was observed in mice with
473 disruption of *M1ap*,[9] but these mice had some sperm in their semen, i.e., severe
474 oligozoospermia. This fact is not inconsistent with our findings, as the authors of the study
475 described mice exhibiting variable efficiency of *M1ap* disruption: Some mice had no apparent
476 wildtype protein, while others displayed reduced levels of the M1AP protein.[9] If the authors
477 had achieved a complete biallelic knockout of *M1ap* in all mice, representing complete
478 meiotic arrest, one would expect azoospermia to result.

479 We did not identify biallelic *M1AP* variants in any other male infertility phenotypes such as
480 SCO or severe oligozoospermia, and no subjects with homozygous *M1AP* LoF variants are
481 present in gnomAD (>140,000 subjects). In conclusion, the presented data strongly suggests
482 that both homozygous LoF and missense variants in *M1AP* impairing its protein as well as
483 compound heterozygosity for either variant type lead to NOA. *M1AP* disruption is associated
484 primarily with germ cell arrest at meiosis/spermatocyte stage, but it may also be compatible
485 with some rare instances of further progressed spermatogenesis. Based on our data, we
486 cannot reliably predict the probability of successful sperm retrieval by testicular biopsy and
487 TESE among men with disrupted *M1AP*, but if we extrapolate the findings from the men
488 reported here, TESE success is quite low (1 of 6, <20%).

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

489 Our finding that biallelic *M1AP* mutations cause predominantly germ cell arrest at the
490 spermatocyte stage in infertile men provides further evidence that meiotic arrest is often of
491 monogenic origin. According to the structured assessment presented here, *M1AP* has
492 considerable clinical validity for causing NOA and could potentially be used as a screening
493 marker before testicular biopsy to estimate the chances of successful TESE. Last not least,
494 identifying mutations in *M1AP* in infertile NOA men provides them with a causal diagnosis for
495 their infertility.

496

497 **Appendices**

498 **Supplemental Data**

499 Supplemental Data includes five tables, five figures, and supplemental methods.

500

501 **Declaration of Interests**

502 The authors declare no competing interests.

503

504 **Acknowledgements**

505 We are indebted to all patients consenting to research evaluation of their data and donating
506 their DNA as well as the physicians who took care of them. We thank Christian Ruckert for
507 his bioinformatic support, Martin Bergmann for many years of skillfully evaluating testicular
508 histologies, Zeliha Görmez for bioinformatics analyses of WES, Şeref Gül and Laurens van
509 de Wiel for their attempts to model the M1AP protein, Joachim Kremerskothen and Verena
510 Höffken for attempting to establish a Western blot analysis, as well as Christina Burhöi,
511 Nicole Terwort, and Katja Hagen for their excellent technical assistance. We thank Dr.
512 Celeste Brennecka for language editing of the manuscript.

513 This work was carried out within the frame of the German Research Foundation Clinical
514 Research Unit 'Male Germ Cells: from Genes to Function' (DFG CRU326). Funding for

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

515 sequencing of the GEMINI cohort was provided by the National Institutes of Health
516 (R01HD078641).The analyses in the Turkish family were supported by a grant from the
517 Bursa University of Uludag Project Unit (KUAP(T)-2014/36).

518

519 **Web Resources**

520 Clinical Research Unit 'Male Germ Cells', http://www.male-germ-cells.de

521 ClinVar, https://www.ncbi.nlm.nih.gov/clinvar

522 GEMINI, https://gemini.conradlab.org/

523 gnomAD, https://gnomad.broadinstitute.org

524 GTEx, https://gtexportal.org

525 HOPE, https://www3.cmbi.umcn.nl/hope/input/

526 Human Protein Atlas, https://www.proteinatlas.org/

527 Male Fertility Gene Atlas, https://mfga.uni-muenster.de

528 MutationTaster, http://www.mutationtaster.org/

529 International Male Infertility Genomics Consortium, http://www.imigc.org

530 OMIM, http://www.omim.org

531 PolyPhen 2, http://genetics.bwh.harvard.edu/pph2/

532 PSAP, https://github.com/conradlab/PSAP

533 SIFT, https://sift.bii.a-star.edu.sg/

534

535 **Accession Numbers**

536 All variants have been submitted to ClinVar (SUB6396814 [*will be substituted for final*
537 *accession no. when available*]) and can also be accessed in the Male Fertility Gene Atlas
538 (MFGA, https://mfga.uni-muenster.de), a public platform for collecting and integrating data
539 sets about epi-/genetic causes of male infertility produced in a subproject of the Clinical
540 Research Unit 'Male Germ Cells: from Genes to Function'.

541

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

542 **References**

543 1. WHO | World Health Organization (2017). Sexual and reproductive health - Fertility and
544 infertility - Assisting couples and individuals.

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

548 3. WHO | World Health Organization (2010) WHO laboratory manual for the examination and
549 processing of human semen.

550 4. Lee, J.Y., Dada, R., Sabanegh, E., Carpi, A., and Agarwal, A. (2011). Role of Genetics in
551 Azoospermia. URL *77*, 598–601.

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

554 6. Yatsenko, A.N., Georgiadis, A.P., Röpke, A., Berman, A.J., Jaffe, T., Olszewska, M.,
555 Westernströer, B., Sanfilippo, J., Kurpisz, M., Rajkovic, A., et al. (2015). X-Linked *TEX11*
556 Mutations, Meiotic Arrest, and Azoospermia in Infertile Men. N. Engl. J. Med. *372*, 2097–
557 2107.

558 7. Oud, M.S., Volozonoka, L., Smits, R.M., Vissers, L.E.L.M., Ramos, L., and Veltman, J.A.
559 (2019). A systematic review and standardized clinical validity assessment of male infertility
560 genes. Hum. Reprod. *34*, 932–941.

561 8. Arango, N.A., Huang, T.T., Fujino, A., Pieretti-Vanmarcke, R., and Donahoe, P.K. (2006).
562 Expression analysis and evolutionary conservation of the mouse germ cell-specific D6Mm5e
563 gene. Dev. Dyn. *235*, 2613–2619.

564 9. Arango, N.A., Li, L., Dabir, D., Nicolau, F., Pieretti-Vanmarcke, R., Koehler, C., McCarrey,
565 J.R., Lu, N., and Donahoe, P.K. (2013). Meiosis I Arrest Abnormalities Lead to Severe
566 Oligozoospermia in Meiosis 1 Arresting Protein (M1ap)-Deficient Mice1. Biol. Reprod. *88*, 1–
567 11.

568 10. Röpke, A., Tewes, A.-C., Gromoll, J., Kliesch, S., Wieacker, P., and Tüttelmann, F.
569 (2013). Comprehensive sequence analysis of the NR5A1 gene encoding steroidogenic factor
570 1 in a large group of infertile males. Eur. J. Hum. Genet. *21*, 1012–1015.

571 11. Martin, M. (2014). Cutadapt removes adapter sequences from high-throughput
572 sequencing reads. EMBnet.journal *17*, 10.

573 12. Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows –
574 Wheeler transform. Bioinformatics *26*, 589–595.

575 13. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A.,
576 Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The genome analysis toolkit:
577 A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res.
578 *20*, 1297–1303.

579 14. McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., and
580 Cunningham, F. (2016). The Ensembl Variant Effect Predictor. Genome Biol. *17*,.

581 15. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins,
582 R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2019). Variation across 141,456
583 human exomes and genomes reveals the spectrum of loss-of-function intolerance across
584 human protein-coding genes. bioRxiv 531210.

585 16. Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters,
586 G., Garcia, F., Young, N., et al. (2013). The Genotype-Tissue Expression (GTEx) project.
587 Nat. Genet. *45*, 580.

588 17. Wilfert, A.B., Chao, K.R., Kaushal, M., Jain, S., Zöllner, S., Adams, D.R., and Conrad,
589 D.F. (2016). Genome-wide significance testing of variation from single case exomes. Nat.

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

590 Genet. *48*, 1455–1461.

591 18. Kasak, L., Punab, M., Nagirnaja, L., Grigorova, M., Minajeva, A., Lopes, A.M., Punab,
592 A.M., Aston, K.I., Carvalho, F., Laasik, E., et al. (2018). Bi-allelic Recessive Loss-of-Function
593 Variants in FANCM Cause Non-obstructive Azoospermia. Am. J. Hum. Genet. *103*, 200–212.

594 19. Venselaar, H., te Beek, T.A., Kuipers, R.K., Hekkelman, M.L., and Vriend, G. (2010).
595 Protein structure analysis of mutations causing inheritable diseases. An e-Science approach
596 with life scientist friendly interfaces. BMC Bioinformatics *11*, 548.

597 20. Nieschlag, E., Behre, H.M., and Nieschlag, S. (2010). Andrology: Male reproductive
598 health and dysfunction.

599 21. Ledig, S., Röpke, A., and Wieacker, P. (2010). Copy number variants in premature
600 ovarian failure and ovarian dysgenesis. Sex. Dev. *4*, 225–232.

601 22. Lek, M., Karczewski, K.J., Minikel, E. V, Samocha, K.E., Banks, E., Fennell, T.,
602 O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of
603 protein-coding genetic variation in 60,706 humans. Nature *536*, 285.

604 23. van der Bijl, N., Röpke, A., Biswas, U., Wöste, M., Jessberger, R., Kliesch, S., Friedrich,
605 C., and Tüttelmann, F. (2019). Mutations in STAG3 cause male infertility due to meiotic
606 arrest. Hum. Reprod.

607 24. Schilit, S.L.P., Menon, S., Friedrich, C., Kammin, T., Wilch, E., Hanscom, C., Jiang, S.,
608 Kliesch, S., Talkowski, M.E., Tüttelmann, F., et al. (2019). SYCP2 translocation-mediated
609 dysregulation and frameshift variants cause human male infertility. bioRxiv 641928.

610 25. Magi, A., Tattini, L., Palombo, F., Benelli, M., Gialluisi, A., Giusti, B., Abbate, R., Seri, M.,
611 Gensini, G.F. ranc., Romeo, G., et al. (2014). H3M2: detection of runs of homozygosity from
612 whole-exome sequencing data. Bioinformatics *30*, 2852–2859.

613 26. Yang, F., Silber, S., Leu, N.A., Oates, R.D., Marszalek, J.D., Skaletsky, H., Brown, L.G.,
614 Rozen, S., Page, D.C., and Wang, P.J. (2015). TEX11 is mutated in infertile men with
615 azoospermia and regulates genome-wide recombination rates in mouse. EMBO Mol. Med. *7*,
616 1198–1210.

617 27. Riera-Escamilla, A., Enguita-Marruedo, A., Moreno-Mendoza, D., Chianese, C.,
618 Sleddens-Linkels, E., Contini, E., Benelli, M., Natali, A., Colpi, G.M., Ruiz-Castañé, E., et al.
619 (2019). Sequencing of a "mouse azoospermia" gene panel in azoospermic men: identification
620 of RNF212 and STAG3 mutations as novel genetic causes of meiotic arrest. Hum. Reprod.
621 *34*, 978–988.

622 28. Cassa, C.A., Weghorn, D., Balick, D.J., Jordan, D.M., Nusinow, D., Samocha, K.E.,
623 O'Donnell-Luria, A., MacArthur, D.G., Daly, M.J., Beier, D.R., et al. (2017). Estimating the
624 selective effects of heterozygous protein-truncating variants from human exome data. Nat.
625 Genet. *49*, 806–810.

626 29. Caburet, S., and Vilain, É. (2015). STAG3 in premature ovarian failure.
627 Medecine/Sciences *31*, 129–131.

628

629 **Table 1. Genetic and clinical data of infertile patients carrying *M1AP* variants.**

Patient	cDNA change	Protein change *in silico* prediction for missense variants (PolyPhen-2/SIFT/MutationTaster)	MAF (gnomAD)	MAF (local controls)	Fertility parameters	Testicular phenotype, TESE outcome	Geographic origin
M330	c.[676dup]; [676dup]	p.[Trp226LeufsTer4]; [Trp226LeufsTer4]	0.0038	0.0088	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	Germany
M864	c.[676dup]; [676dup]	p.[Trp226LeufsTer4]; [Trp226LeufsTer4]	0.0038	0.0088	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	Germany
M1792	c.[676dup]; [676dup]	p.[Trp226LeufsTer4]; [Trp226LeufsTer4]	0.0038	0.0088	FSH: 7.8 LH: 5.1 T: 10.12 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	Germany
RU01691	c.[676dup]; [676dup]	p.[Trp226LeufsTer4]; [Trp226LeufsTer4]	0.0038	0.0030	FSH: 5.0 LH: 2.0 T: 11.3 TV: NA 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	The Netherlands
Y126	c.676dup(;) 949G>A	p.Trp226LeufsTer4(;) (Gly317Arg) -/-/-;D/D/D	0.0038;0.0001	ND	FSH: 6.7 LH: 3.2 T: NA TV: NA Azoospermia	Maturation arrest at round spermatid stage, (quantification NA), TESE negative	Portugal
P86	c. 148T>C(;) 1289T>C	p. (Ser50Pro)(;) (Leu430Pro) T/P/D;D/D/D	0;0.00003	ND	FSH: NA LH: NA T: NA TV: NA Azoospermia	Dispersed Sertoli cells, some tubules contained only spermatogonia, (quantification NA), TESE negative	Portugal
T1024	c.[1166C>T]; [1166C>T]	p.[(Pro389Leu)]; [(Pro389Leu)] D/D/D	0.00001	ND	FSH: 8.3 LH: 4.4 T: 8.8 TV: 15/15 Azoospermia	NA	Turkey

630 Abbreviations: D: damaging, deleterious or disease causing, T: tolerated, P: polymorphism, MAF: minor allele frequency, FSH: follicle stimulating
631 hormone (IU/L), LH: luteinizing hormone (IU/L), T: testosterone (nmol/L), TV: testicular volume right/left (mL), ES: elongating spermatids, RS:
632 round spermatids, SC: spermatocytes, SG: spermatogonia, SCO: Sertoli-cell only, TS: tubular shadows, ND: not determined, NA: not available.
633 Reference values: FSH 1-7 IU/L, LH 2-10 IU/L, T >12 nmol/L, TV >15 mL per testis.

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

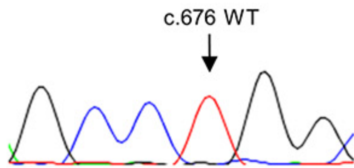
634 **Figure 1. Recurrent homozygous variant c.676dup (duplication of T-nucleotide;
635 p.Trp226LeufsTer4) in *M1AP* leading to complete bilateral meiotic arrest in patients
636 M330 (D-F), M864 (G-I) and M1792 (J-L) from the MERGE study.**

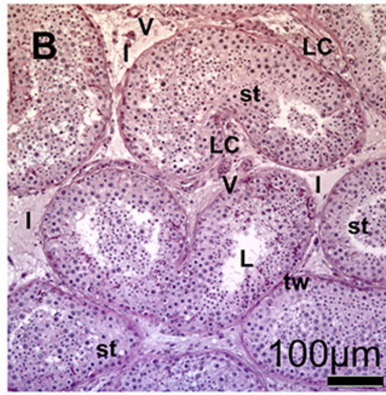
637 (**A**) Electropherogram with the wildtype sequence of *M1AP* exon 5 (patient M1672 with
638 obstructive azoospermia). (**B/C**) Testicular tissue showing complete spermatogenesis, PAS
639 staining. (**B**) Testicular tissues are composed of seminiferous tubules and interstitium. The
640 seminiferous tubules are separated from the interstitial space (I) by tubular walls (tw) formed
641 by myoid peritubular cells and the lamina propria. Inside, the seminiferous epithelium and the
642 lumen (L) are localized. In the interstitium, groups of steroidogenic Leydig cells (LC) and
643 blood vessels (V) are observed. Tubular cross section showed the regular appearance of a
644 functioning testis exhibiting complete germ cell differentiation. (**C**) Detail of B; the tubules are
645 surrounded by the lamina propria and the peritubular cells (ptC), forming the wall. Within the
646 seminiferous epithelium, somatic Sertoli cells (SC) are supporting the germ cells
647 differentiating from A spermatogonia (Apale/Adark) via premeiotic spermatocytes
648 (preleptotene to zygotene stage; Pl-Z) into the meiotic pachytene spermatocytes (P). After
649 meiosis is completed, haploid round spermatids (rsptd) are formed which mature further into
650 elongated spermatids (elsptd). (**D-L**) Identification of a recurrent homozygous variant in
651 *M1AP* (c.676dup, p.Trp226LeufsTer4). Sanger sequencing verified the variant in patients
652 M330 (**D**), M864 (**G**) and M1792 (**J**) leading to complete bilateral meiotic arrest as indicated
653 by histological examination of testis biopsies (**E/F**: M330, **H/I**: M864, **K/L**: M1792), which
654 show spermatocytes as most advanced germ cells in all tubules.

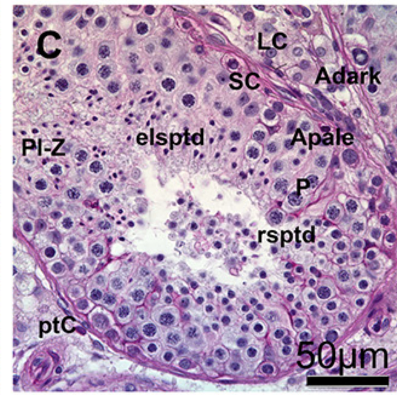
bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A

M1672

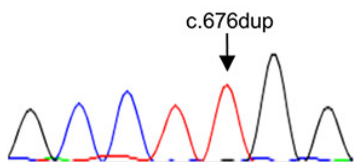


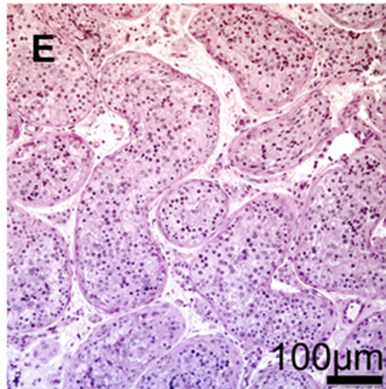


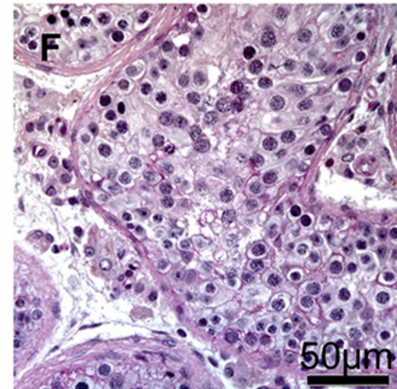


D

M330

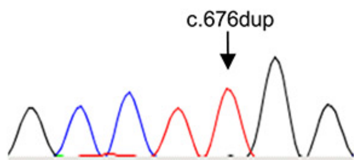


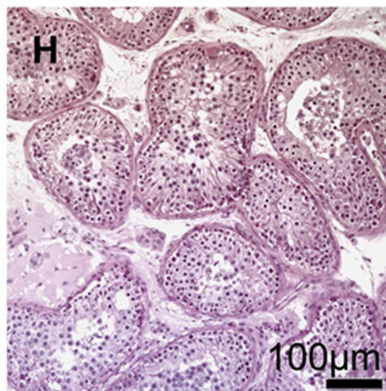


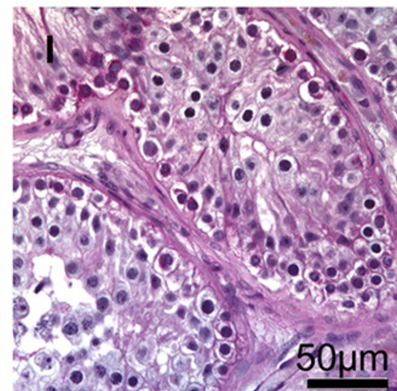


G

M864

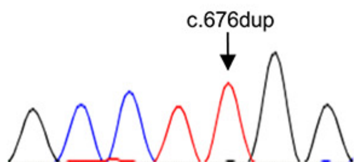


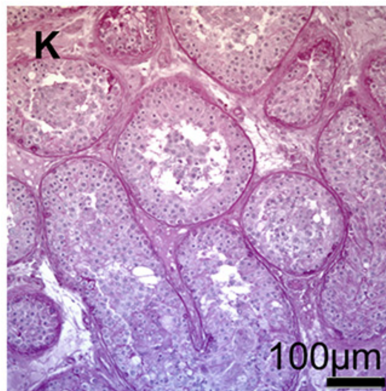


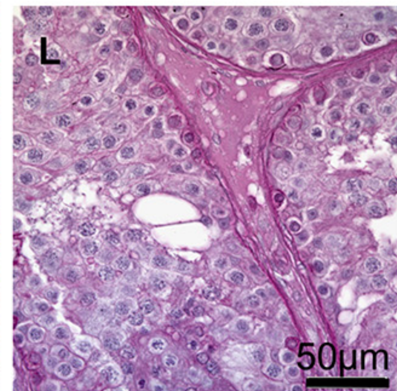


J

M1792







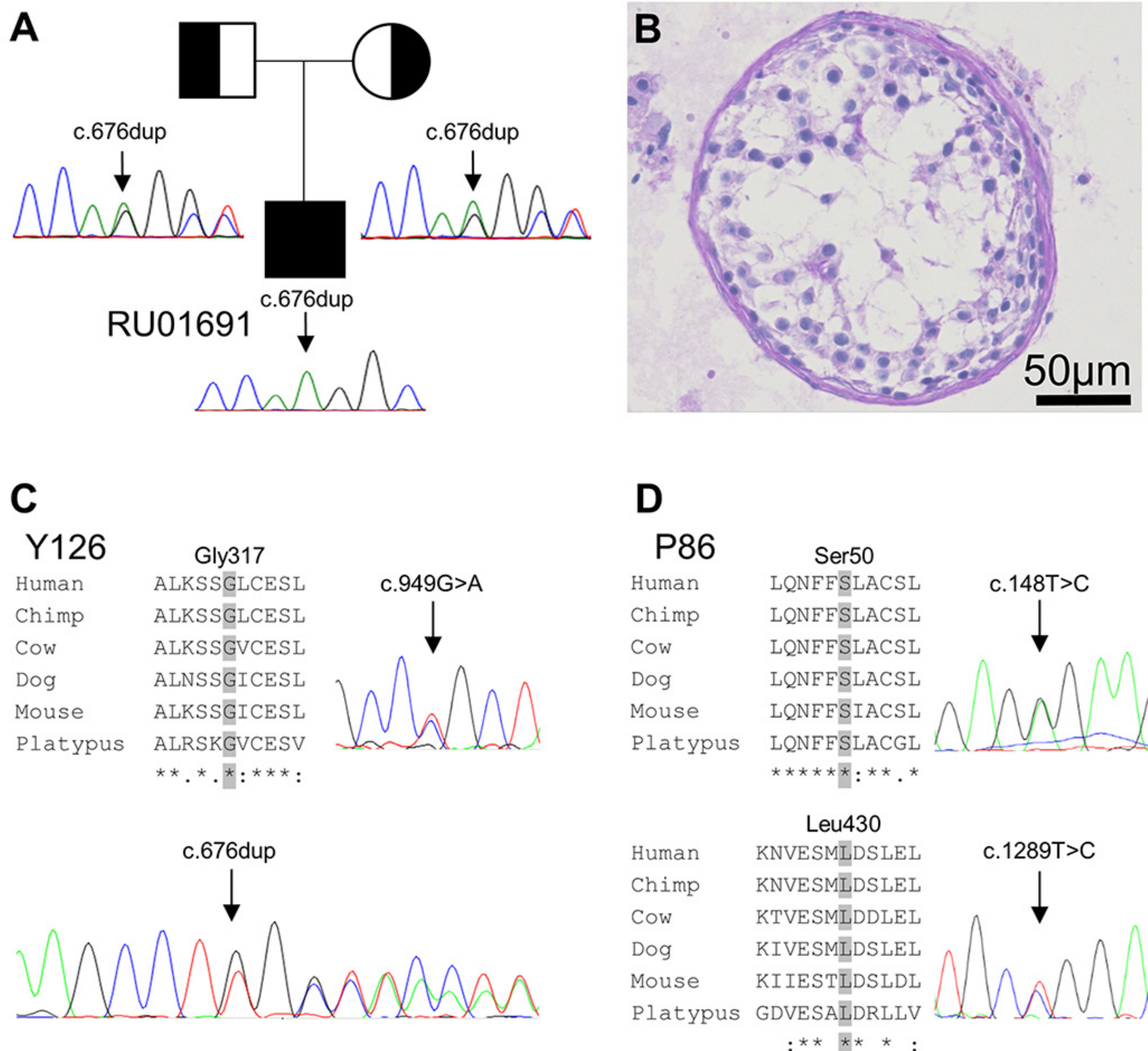
655

656

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

657 **Figure 2. Variants in *M1AP* in two follow-up studies.**

658 Recurrent homozygous duplication c.676dup (p.Trp226LeufsTer4) (**A**) in *M1AP* leading to
659 predominantly meiotic arrest in patient RU01691 from Nijmegen, NL. Both parents are
660 heterozygous for the same variant. (**B**) Histology indicates predominant germ cell arrest at
661 the spermatocyte stage. Identification of potentially biallelic variants in *M1AP* in Portuguese
662 patients from the GEMINI study. (**C**) Patient Y126 carries the recurrent LoF variant c.676dup
663 (p.Trp226LeufsTer4) and the missense variant c.949G>A (p.(Gly317Arg)). (**D**) Patient P86
664 carries two missense variants (c.148T>C(;)1289T>C p.(Ser50Pro)(;)(Leu430Pro)). All
665 missense variants affect highly conserved amino acids, as seen from multiple sequence
666 alignments (**C/D**).



667

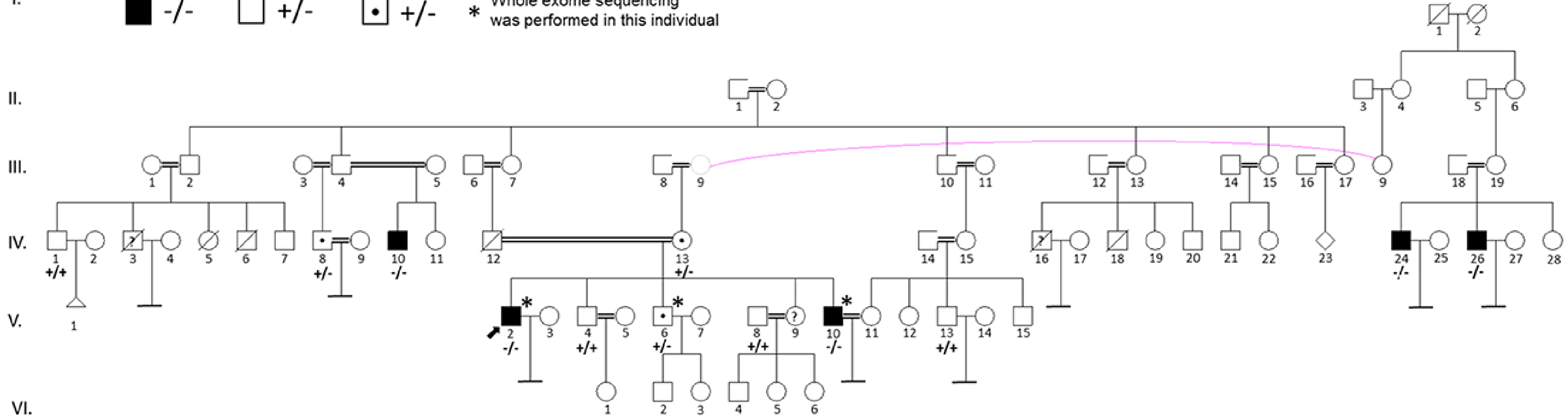
668

669 **Figure 3. Turkish consanguineous family with infertile, azoospermic men homozygous for *M1AP* missense variant and fertile**
670 **heterozygous carriers.**

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

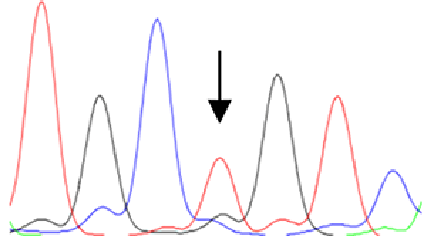
bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

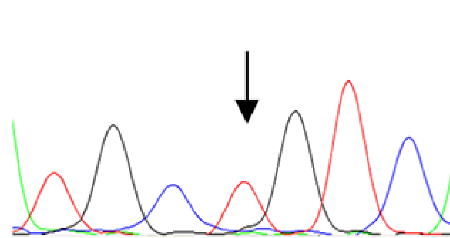
A

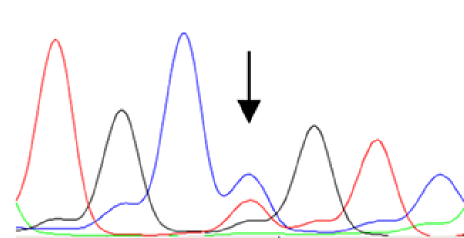


B T1024 V.2 V.10 V.6

c.1166C>T homozygous c.1166C>T homozygous c.1166C>T heterozygous







C

```
                 Pro389
Human      TFYVIMPSHSLT
Chimp      TFYVIMPSRSLT
Cow        TFYVILPSHSPT
Dog        TFYVIVPSRSPT
Mouse      SFYVITPSHSLT
Platypus   SFYVILPSNSFS
            **** **.*  :
```

677

bioRxiv preprint doi: https://doi.org/10.1101/803346; this version posted October 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.