Whole-exome sequencing in patients with maturation arrest: a potential additional diagnostic tool for prevention of recurrent negative testicular sperm extraction outcomes

F Ghieh, A Barbotin, N Swierkowski-Blanchard, C Leroy, J Fortemps, C Gerault, C Hue, H Mambu Mambueni, S Jaillard, M Albert, et al.

Human Reproduction, Vol.37, No.6, pp. 1334–1350, 2022
Advance Access Publication on April 12, 2022 https://doi.org/10.1093/humrep/deac057

human reproduction

ORIGINAL ARTICLE *Reproductive genetics*

Whole-exome sequencing in patients with maturation arrest: a potential additional diagnostic tool for prevention of recurrent negative testicular sperm extraction outcomes

F. Ghieh [1,2], A.L. Barbotin[3], N. Swierkowski-Blanchard[1,2,4], C. Leroy[3], J. Fortemps[5], C. Gerault[6], C. Hue[7], H. Mambu Mambueni[7], S. Jaillard [8,9], M. Albert[1,2], M. Bailly[4], V. Izard[10], D. Molina-Gomes[6], F. Marcelli[3], J. Prasivoravong[3], V. Serazin[1,2,6], M.N. Dieudonne[1,2], M. Delcroix[6], H.J. Garchon [7], A. Louboutin[5], B. Mandon-Pepin [1,2], S. Ferlicot [11], and F. Vialard [1,2,6,*]

[1]Université Paris-Saclay, UVSQ, INRAE, BREED, Jouy-en-Josas, France [2]École Nationale Vétérinaire d'Alfort, BREED, Maisons-Alfort, France [3]Institut de Biologie de la Reproduction-Spermiologie-CECOS, Hôpital Jeanne de Flandre, Centre Hospitalier et Universitaire, Lille, France [4]Département de Gynécologie Obstétrique, CHI de Poissy/Saint-Germain-en-Laye, Poissy, France [5]Service d'Anatomie Pathologique, CHI de Poissy/Saint-Germain-en-Laye, Saint-Germain-en-Laye, France [6]Département de Génétique, Laboratoire de Biologie Médicale, CHI de Poissy/Saint-Germain-en-Laye, Poissy, France [7]Department of Biotechnology and Health, UVSQ, Université Paris-Saclay, Inserm UMR 1173, Montigny-le-Bretonneux, France [8]Service de Cytogénétique, CHU Rennes, Rennes, France [9]INSERM, EHESP, IRSET—UMR_S 1085, Université Rennes 1, Rennes, France [10]Service d'Urologie, AP-HP, Université Paris-Saclay, Hôpital de Bicêtre, Le Kremlin-Bicêtre, France [11]Service d'Anatomie Pathologique, AP-HP, Université Paris-Saclay, Hôpital de Bicêtre, Le Kremlin-Bicêtre, France

*Correspondence address. Tel: +33-139-274-700; E-mail: francois.vialard@uvsq.fr https://orcid.org/0000-0002-5774-2756

Submitted on October 28, 2021; resubmitted on March 07, 2022; editorial decision on March 13, 2022

STUDY QUESTION: Could whole-exome sequencing (WES) be useful in clinical practice for men with maturation arrest (MA) after a first testicular sperm extraction (TESE)?

SUMMARY ANSWER: WES in combination with TESE yields substantial additional information and may potentially be added as a test to predict a negative outcome of a recurrent TESE in patients with MA.

WHAT IS KNOWN ALREADY: At present, the only definitive contraindications for TESE in men with non-obstructive azoospermia (NOA) are a 46,XX karyotype and microdeletions in the azoospermia factor a (AZFa) and/or AZFb regions. After a first negative TESE with MA, no test currently exists to predict a negative outcome of a recurrent TESE.

STUDY DESIGN, SIZE, DURATION: In a cohort study, we retrospectively included 26 patients with idiopathic NOA caused by complete MA diagnosed after a first TESE.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Twenty-six men with MA at the spermatocyte stage in all seminiferous tubules, according to a histopathological analysis performed independently by two expert histologists, and a normal karyotype (i.e. no *AZF* gene microdeletions on the Y chromosome) were included. Single-nucleotide polymorphism comparative genomic hybridization array and WES were carried out. The results were validated with Sanger sequencing. For all the variants thought to influence spermatogenesis, we used immunohistochemical techniques to analyse the level of the altered protein.

MAIN RESULTS AND THE ROLE OF CHANCE: Deleterious homozygous variants were identified in all seven consanguineous patients and in three of the 19 non-consanguineous patients. Compound heterozygous variants were identified in another 5 of the 19 non-consanguineous patients. No recurrent variants were identified. We found new variants in genes known to be involved in azoospermia or

MA [including testis expressed 11 (*TEX11*), meiotic double-stranded break formation protein 1 (*MEI1*), proteasome 26s subunit, ATPase 3 interacting protein (*PSMC3IP*), synaptonemal complex central element protein 1 (*SYCE1*) and Fanconi anaemia complementation group M (*FANCM*) and variants in genes not previously linked to human MA (including CCCTC-binding factor like (*CTCFL*), Mov10 like RISC complex RNA helicase 1 (*MOV10L1*), chromosome 11 open reading frame 80 (*C11ORF80*) and exonuclease 1 (*EXO1*)].

LARGE SCALE DATA: Data available on request

LIMITATIONS, REASONS FOR CAUTION: More data are required before WES screening can be used to avoid recurrent TESE, although screening should be recommended for men with a consanguineous family background. WES is still a complex technology and can generate incidental findings.

WIDER IMPLICATIONS OF THE FINDINGS: Our results confirmed the genetic aetiology of MA in most patients: the proportion of individuals with at least one pathologic variant was 50% in the overall study population and 100% in the consanguineous patients. With the exception of *MEI1* (compound heterozygous variants of which were identified in two cases), each variant corresponded to a specific gene—confirming the high degree of genetic heterogeneity in men with MA. Our results suggest that WES screening could help to avoid recurrent, futile TESE in men with MA in general and in consanguineous individuals in particular, but these results need to be confirmed in future studies before clinical implementation.

STUDY FUNDING/COMPETING INTEREST(S): The study was funded by the Fondation Maladies Rares (Paris, France), Merck (Kenilworth, NJ, USA), IRSF (Montigny le Bretonneux, France) and Agence de la Biomédecine (Saint Denis, France). There are no competing interests.

TRIAL REGISTRATION NUMBER: N/A.

Key words: azoospermia / meiosis / spermatogenic arrest / whole-exome sequencing / consanguinity / testicular sperm extraction / non-obstructive azoospermia

Introduction

Infertility is a major worldwide health issue and concerns more than 15% of couples. The aetiology of infertility is not known in a high proportion of cases. Hence, a greater level of fundamental knowledge is required to improve clinical care.

In over 50% of cases, infertility is caused by sperm abnormalities (Jungwirth *et al.*, 2012). The most severe clinical phenotype (observed in 10% of infertile men) is azoospermia, with the complete absence of spermatozoa in the semen. The condition can be obstructive (i.e. caused by an obstruction in the seminal tract, for example) or non-obstructive (i.e. owing to a defect in sperm production). Testicular sperm extraction (TESE) is the main form of ART that enables men with azoospermia to become fathers (ASRM, 2019; Corona *et al.*, 2019). TESE is successful in over 95% of cases of obstructive azoospermia (OA) and 40–50% of cases of non-obstructive azoospermia (NOA). In NOA, the likelihood of retrieving sperm via TESE is related to the testis' histological phenotype: hypospermatogenesis; Sertoli cell-only (SCO, with the absence of germ cells in the tubules); and germ cell maturation arrest (MA, where development stops at the spermatocyte or spermatid stage). MA can be homogeneous (i.e. all the seminiferous tubules have the same appearance) or heterogeneous (i.e. tubules differ in their appearance) (McLachlan *et al.*, 2007). With a view to subsequent IVF, surgical sperm recovery via TESE is futile in cases of homogeneous MA or SCO.

Many cases of NOA are thought to have a genetic cause. After acquired diseases have been ruled out, genetic screening is always recommended for men with NOA. This screening is typically limited to karyotyping for the identification of chromosomal abnormalities (De Braekeleer and Dao, 1991) and Y chromosome microdeletions (Vogt *et al.*, 1992). Over the last few decades, whole-genome analyses (especially array comparative genomic hybridization (CGH) and whole-exome sequencing (WES)) have been used to discover novel genetic defects associated with spermatogenesis failure or NOA (Krausz *et al.*, 2020; Cannarella *et al.*, 2021). Several lines of evidence indicate that MA can be caused by point mutations in single genes and by copy number variations (CNVs) (Halder *et al.*, 2017); the evidence notably includes: the occurrence of MA in male siblings; the features of monogenic mouse models with MA; and the fact that spermatogenesis is governed by a particular set of genes. In human, the list of gene mutations leading to NOA continues to grow (Houston *et al.*, 2021)—confirming that a large number of genes are involved in spermatogenesis (Matzuk and Lamb, 2008). To date, few mutations in human genes directly involved in meiosis have been linked to MA in particular (Bellil *et al.*, 2021; Houston *et al.*, 2021). The sequencing of these mutated genes has not yet been integrated into the clinical management of men with sperm abnormalities. Some researchers have suggested sequencing a panel of candidate genes prior to TESE, especially for genes that are frequently found to be mutated in NOA (e.g. testis expressed 11 (*TEX11*), stromal antigen 3 (*STAG3*) and synaptonemal complex central element protein 1 (*SYCE1*) (Cannarella *et al.*, 2021)). When at least one of these genes carries a deleterious mutation (e.g. as listed in the ClinVar database (http://www.clinvar.com/)), the likelihood of sperm retrieval is low and so futile TESE could be avoided.

At present, the only definitive contraindications for TESE in men with NOA are a 46,XX karyotype (usually 46,X,der(X)t(X;Y)(p22.3;p11.2) (Ferguson-Smith, 1966)) and microdeletions in the azoospermia factor a (AZFa) and/or AZFb region (Patrat *et al.*, 2010). These abnormalities lead to SCO or MA. Other chromosomal abnormalities do not contraindicate TESE, although genetic counselling is required to evaluate the risk of an unbalanced karyotype in the offspring; this is mainly applied to structural chromosome rearrangements. Thus, karyotyping and screening for Y chromosome microdeletions (De Braekeleer and Dao, 1991; Jungwirth *et al.*, 2012) are still the first-line analyses.

Furthermore, the respective medical benefits of a gene panel approach and WES must be compared; longer sequences will be associated with a greater number of genetic variants, higher costs and a greater probability of incidental findings. Furthermore, incidental findings will: complicate the provision of patient consent; require pre- and post-test genetic counselling; and make the data more difficult to interpret. However, a focus on genes involved in spermatogenesis might help to avoid the incidental discovery of gene defects associated with other pathologies (e.g. cancer).

Here, we first evaluated the incidence of single-nucleotide polymorphism (SNPs) and CNVs in patients with a particular clinical phenotype, in order to refine our clinical practice. Next, we thought about whether new technological approaches, such as WES, should be performed before TESE or after a first negative TESE. In order to address these issues, we performed an SNP-CGH array analysis and WES in a series of 26 men with homogeneous spermatogenetic MA and a successful (sperm-positive: TESE+) or unsuccessful (sperm-negative: TESE−) TESE. We also used immunohistochemical techniques to assess the impact of the identified gene variants on the expression of candidate proteins during spermatogenesis.

Materials and methods

Patients

We included patients with idiopathic NOA consulting at Bicêtre Hospital (Le Kremlin Bicêtre, France), Lille University Hospital (Lille, France) and Poissy General Hospital (Poissy, France) after a TESE procedure. As described previously (Barbotin *et al.*, 2018, 2019), a scrotal incision was made on each testis and a small fragment of the testicular pulp was removed with scissors (one site per testis). The testicular biopsy sample was immersed in culture medium and sent to the reproductive biology laboratory for sperm extraction. The TESE outcome was scored as successful (presence of spermatozoa) or unsuccessful (absence of spermatozoa), and the suspension was frozen if a least >1 live spermatozoon was detected for further use in ICSI cycles (Barbotin *et al.*, 2019): for details see, Supplementary Materials and Methods. NOA was caused by complete MA at the spermatocyte stage in all seminiferous tubules for all patients. The histopathological analysis was performed independently by two expert histologists using an additional fragment fixed in formalin, acetic acid and alcohol, then paraffin-embedded. The histological evaluation encompassed at least 100 visible seminiferous tubule cross-sections (for details see, Supplementary Materials and Methods). In the event of disagreement between the two experts, the patient was excluded. Patients with an abnormal karyotype or Y chromosome microdeletions were excluded. After counselling and the provision of consent, DNA was purified from whole blood or buffy coat samples either automatically (using a QIAsymphony DSP DNA Midi Kit, Qiagen, Venlo, The Netherlands) or manually. Parental DNA was not available for all study participants.

Ethical approval

The study protocol was approved by an independent ethics committee (*CPP Ile de France-Ouest*, Paris, France; reference: 01-132). All participants gave their written, informed consent.

SNP-CGH array analysis and CNV selection

We performed array CGH (SurePrint G3 Human Genome CGH+SNP Microarray Kit, Agilent Technologies, Santa Clara, CA, USA), in order to: identify CNVs associated with MA; identify regions with loss of heterozygosity (LOH, also referred to as runs of homozygosity) for consanguineous patients; and rule out the presence of *TEX11* exonic microdeletions (Yatsenko *et al.*, 2015). The array CGH protocol has been validated by our laboratory (Ghieh *et al.*, 2021b). The experimental protocol and the interpretation procedure have been performed according to the manufacturer. We followed the American College of Medical Genetics and Genomics (ACMG) criteria for the selection of LOH regions and calculation of the inbreeding coefficient F (the homozygosity rate, defined as the fraction of the individual's genome that should be homozygous by descent) (McQuillan *et al.*, 2008) for each patient (Sund *et al.*, 2013): F = total LOH/human genome size (3138 Mb for the hg19 genome) (Wierenga *et al.*, 2013). The observed coefficients were compared with the theoretical coefficients for various degrees of inbreeding (Sund *et al.*, 2013). The identification of regions with LOH also enabled us to focus on regions of interest for the WES analysis in consanguineous patients (for details see, Supplementary Materials and Methods).

WES and variant prioritization

WES was carried out at the genomics facility at the University of Versailles Saint Quentin (Versailles, France) and by Eurofins Genomics (Ebersberg, Germany), using Illumina technology (San Diego, CA, USA). DNA libraries were built with a SureSelect Exome V6+UTR Capture Library Kit (Agilent Technologies), according to the manufacturer's instructions. Only homozygous or compound heterozygous variants were considered. We selected variants causing insertions/deletions, missense, stop-loss, stop-gain or frameshift mutations, or changes to splice acceptor/donor sites. Synonymous variants and variants with untranslated 3′ or 5′ regions were excluded. Missense variants were selected by combining: the allele frequency in the GnomAD and the 1000Genomes databases; the predicted effect on the encoded protein, as judged with the Rare Exome Variant Ensemble Learner, SIFT (http://sift.jcvi.org/), Polyphen2 (http://genetics.bwh.harvard.edu/pph2/) and M-CAP (http://bejerano.stanford.edu/mcap/); and data on predominant or exclusive expression of the protein in the testis or essential status for spermatogenesis and meiosis in the literature or the Gene-Tissue Expression (https://gtexportal.org/home/), Human Protein Atlas (https://www.proteinatlas.org), PubMed and Ensembl databases. Variant calling is detailed in Supplementary Materials and Methods.

Validation of mutations by Sanger sequencing

Variants prioritized in our analysis were experimentally validated using Sanger sequencing with a BigDye Terminator v3.1 sequencing kit A38073 (Thermo Fisher Scientific, Waltham, MA, USA) and a SeqStudio genetic analyser (Thermo Fisher Scientific). After PCR amplification, the products were sequenced in both directions. Primers were chosen with Primer 3 Plus software (http://primer3.ut.ee/), according to the general principles of primer design. The primers for PCR and Sanger sequencing validation are listed in Supplementary Table SI.

The chromatogram files generated by the analyser were read with 4 Peaks software (https://nucleobytes.com/4peaks/index.html).

Immunohistochemical analyses

Protein expression was assessed with immunochemical techniques. Samples from participants were compared with control samples from patients with OA and normal spermatogenesis. All the immunochemical experiments were carried out at the same time, in order to minimize possible inter-session discrepancies in staining. Immunohistochemistry was performed with the Benchmark XT Ventana Roche system and the XT ultraView DAB v3 revelation kit (Roche Life Science, Penzberg, Germany). The primary antibodies are listed in Supplementary Table SII, and the experimental conventional protocol is detailed in Supplementary Materials and Methods.

Results

Clinical characteristics of the study population

Twenty-six patients (including seven with a consanguineous family background) were included in the study. In each case, homogenous MA of spermatogenesis in all seminiferous tubules (according to a testicular biopsy) had been confirmed independently by two experts, independent of the TESE results (Supplementary Fig. S1). The mean ± SD volumes for the left and right testis were, respectively, 10.93 ± 4.88 ml and 12.17 ± 4.93 ml. The mean FSH, LH, oestradiol, inhibin B and testosterone levels were, respectively, 7.58 ± 6.02 IU/l, 5.54 ± 3.52 IU/l, 25.50 ± 11.34 ng/l, 102.52 ± 80.17 ng/l and 5.04 ± 1.81 µg/l. All the data are detailed in Supplementary Table SIII. The measured clinical variables and hormone levels were in line with the literature data on the MA phenotype, i.e. half the normal testis volume and slightly elevated FSH and LH levels.

TESE had enabled sperm retrieval for 5 of the 26 patients, forming the successful TESE (TESE+) group. Hence, the unsuccessful TESE (TESE−) group comprised 21 patients. Although two of the five TESE+ patients lacked a frozen sperm sample, more than 20 spermatozoa were retrieved by TESE for the other three TESE+ patients (i.e. enough for attempting IVF without a further TESE).

CGH analysis

With the exception of patient 24 (P24), all the patients had provided enough DNA for SNP-CGH analysis.

Number of CNVs

The mean ± SD total number of CNVs was 8.56 ± 7.06 (range: 1 to 30). The mean number of nullosomies was 0.68 ± 2.41, with 3.12 ± 3.24 deletions, 3.84 ± 3.04 duplications and 0.92 ± 3.50 amplifications. The nullosomy regions did not contain any genes reported as being essential for spermatogenesis.

TEX11 *CNV screening*

In view of the above results, we considered that the CNVs were not responsible for the patients' phenotype. Furthermore, no *TEX11* exon deletions (Yatsenko *et al.*, 2015) were found in any of the 25 analysed patients.

LOH evaluation

We also identified regions with LOH (in Mb) in each patient's genome. According to the criteria published by the ACMG, the mean ± SD inbreeding ratio was 2.14 ± 2.32 (range: 0.20 to 8.91). As expected, 7 of the 25 analysed MA patients had high inbreeding coefficients (between 2.3% and 8.9%) and 18 had low inbreeding coefficients (between 0.9% and 1.75%). These results enabled us to focus on regions of interest with LOH in the WES analysis of patients with high inbreeding coefficients.

Exome analysis

First, for all the spermatogenesis genes (Supplementary Results) affected by a heterozygous CNV deletion or duplication, we used WES to search for mutations on the other allele; no variant was found in any cases.

Second, we postulated that MA was associated with autosomal recessive or X-linked recessive inheritance. Accordingly, a list of homozygous, compound heterozygous or X-linked variants was drawn up for each patient after the exclusion of frequent variants and the application of stringent filters. We only considered variants in genes reportedly expressed during spermatogenesis in humans or mice. The data for each patient are summarized in Table I.

There were no homozygous or compound heterozygous variants in spermatogenesis-associated or testis-expressed genes for 8 of the 26 patients, all of whom had non-consanguineous family backgrounds. For the remaining 18 patients, we identified at least one candidate variant. By applying additional selection criteria, we only considered variants in 15 different genes for 15 patients (Table II and Supplementary Results).

Sanger sequencing and immunohistochemical analyses

With the exception of P24 (for whom no further material was available), the Sanger sequencing confirmed all the participants' variants (Supplementary Fig. S2). We then used immunohistochemical techniques to semi-quantitatively evaluate the protein levels in testicular biopsies. We performed 15 immunohistochemical analyses, including two analyses for P21. A lack of material prevented us from analysing P4—a compound heterozygote for two *MEI1* (meiotic double-stranded break formation protein 1) variants. The observed protein expression on control samples was always in agreement with the human Protein Atlas database (www.proteinatlas.org/). However, we noticed that the Protein Atlas did not contain data on the testis expression of SPATA22 (spermatogenesis associated 22), EXO1 (exonuclease 1), FANCM (FA complementation group M) or MCMDC2 (minichromosome maintenance domain containing 2).

Next, we compared protein expression in the testis in patient samples versus the control sample (Fig. 1). The results for P18 and P19 differed markedly from the control results. FANCM protein was observed in the cytoplasm only for P18 but in the nucleus only for the control. Expression of MOV10L1 (coding for mov10 like RISC complex RNA helicase 1) was observed in the nucleus only for P19 but in the cytoplasm only for the control. The variants' impact on these

Table I Detailed genetic results and variants obtained for each patient.

Patient	Consanguinity rate	TESE outcome	Gene		Variant			
			Name	NM	Genomic position	Nucleotide	Protein	Type
P1	0.96	Positive +						
P2	6.25*	Negative	*CTCFL*	NM_001269041	chr20: 57498582 chr20: 57515764	exon10:c.T1960C exon4:c.A344G	C654R Y115C	Homozygote Homozygote
P3	7.15*	Positive +	*SPATA22*	NM_001170696	chr17: 3462708	exon3:c.G103A	G35R	Homozygote
P4	1.07	Negative	*MNS1* *MEI1*	NM_018365 NM_152513	chr15: 56444483 chr22: 41705502 chr22: 41729750	exon5:c.T647C exon3:c.299-2A>G exon8:c.C950T	I216T / A317V	Homozygote Heterozygote Heterozygote
P5	0.63	Negative	*MLH1*	NM_001167619	chr3: 37047639 chr3: 37047640	exon15:c.A1129G exon15:c.A1130C	K377E K377T	Heterozygote Heterozygote
P6	3.48*	Negative	*SYCE1*	NM_001143763	chr10: 133557865	exon6:c.A373G	R125G	Homozygote
P7	0.9	Negative	*PSMC3IP*	NM_001256014	chr17: 42573326	exon5:c.333delG	R111fs	Homozygote
P8	1.14	Negative						
P9	2.38*	Negative	*ZNF85* *DNMT3a*	NM_001256173 NM_022552 NM_022552	chr19: 20949441 chr2: 25275056 chr2: 25275066	exon3:c.736dupG exon6:c.G524T exon6:c.G514A	T245fs G175V G172S	Homozygote Heterozygote Heterozygote
P10	2.68*	Negative	*C11orf80*	NM_001302084	chr11: 66788158	exon3:c.26-2A>G	/	Homozygote
P11	1.13	Positive	*MAGEB6* *TOPAZ1*	NM_173523.2 NM_001145030	chrX : 26194134-26194150 chr3: 44242181 chr3: 44244523	exon2:c.288_304del exon1:c.G128C exon2:c.C2017G	T96fs C43S P673A	Heterozygote Heterozygote Heterozygote
P12	1.18	Positive	*EXO1*	NM_003686	chr1: 241853401 chr1: 241885313	exon4:c.G325A exon14:c.2209-1G>C	E109K /	Heterozygote Heterozygote
P13	0.78	Negative	*SLC22A16*	NM_033125	chr6: 110438740 chr6: 110442732	exon5:c.G1291A exon4:c.A695G	V431I E232G	Heterozygote Heterozygote
P14	1.29	Negative						
P15	1.21	Negative	*CCDC36*	NM_001135197	chr3: 49211800 chr3:49256550-49256553	exon2:c.A20G exon8:c.1053_1056del	N7S K351fs	Heterozygote Heterozygote
P16	1.75	Negative						
P17	1.11	Positive +						
P18	1.02	Negative	*LRRCC1* *MDC1* *FANCM*	NM_001349639 NM_014641 NM_001308133	chr8: 85137473 chr6: 30705059 chr6: 30705062 chr14: 45137184 chr14: 45154705	exon11:c.T1202C exon10:c.4124delC exon10:c.4120_4121insA exon2:c.A624G exon6:c.C1114T	L401P P1375fs P1374fs I208M R372W	Homozygote Heterozygote Heterozygote Heterozygote Heterozygote
P19	5.49*	Negative	*MOV10L1* *ATM*	NM_001164104 NM_000051	chr22: 50144223 chr11: 108272729	exon18:c.G2485A exon22:c.C3161G	A829T P1054R	Homozygote Homozygote
P20	0.29	Negative	*TEX11*	NM_031276	chrX: 70651450	exon17:c.G1483A	A495T	Homozygote
P21	8.91*	Negative	*MCMDC2* *ARL2*	/ NM_001199745	chr8: 66905436 chr11: 65018680	intron: T/C exon3:c.C286T	/ R96C	Homozygote Homozygote
P22	0.84	Negative	*MAP7* *MAP7*	NM_001198617	chr6: 136362513 chr6: 136372607	exon9:c.G1181A exon6:c.T488A	R394Q I163N	Heterozygote Heterozygote
P23	1.15	Negative						
P24	not done	Negative	*MEI1*	NM_152513	chr22: 41699654 chr22: 41718236 chr22: 41730527	exon1:c.T116G exon6:c.A695T exon9:c.T986C	V39G D232V L329P	Heterozygote Heterozygote Heterozygote
P25	0.2	Negative						
P26	0.56	Negative						

TESE, testicular sperm extraction.
*Consanguineous patient.

Table II Detailed variant interpretation obtained for each patient.

Patient	Gene	Variant								Decision after WES
		Protein	Type	Frequency GnomaD	Predictive pathogenicity				Classification	
	Name				MCAP	Revel	SIFT	Polyphen		
P2	*CTCFL*	C654R	Homozygote	Unknown	Benign	Benign	Deleterious	Deleterious	VUS	To be confirmed
		Y115C	Homozygote	Unknown	Deleterious	Deleterious	Deleterious	Deleterious	Deleterious	
P3	*SPATA22*	G35R	Homozygote	inf e-04	Unknown	Benign	Benign	Benign	Benign but very rare	To be confirmed
P4	*MNS1*	I216T	Homozygote	0.0178	Unknown	Benign	Deleterious	Deleterious	Deleterious	
	MEI1	intronic	Heterozygote	Unknown	Unknown	Unknown	Unknown	Unknown	VUS	To be confirmed
		A317V	Heterozygote	inf e-04	Deleterious	Benign	Benign	Deleterious	VUS	
P5	*MLH1*	K377E	Heterozygote	0.0034	Deleterious	Deleterious	Deleterious	Deleterious	Cis	
		K377T	Heterozygote	0.0034	Deleterious	Deleterious	Deleterious	Deleterious		
P6	*SYCE1*	R125G	Homozygote	0.0022	Unknown	Benign	Deleterious	Deleterious	Deleterious	To be confirmed
P7	*PSMC3IP*	R111fs	Homozygote	Unknown	Unknown	Unknown	Unknown	Unknown	Frameshift	To be confirmed
P9	*ZNF85*	T245fs	Homozygote	Unknown	Unknown	Unknown	Unknown	Unknown	Frameshift	To be confirmed
	DNMT3a	G175V	Heterozygote	Unknown	Deleterious	Deleterious	Deleterious	Deleterious	Cis	
		G172S	Heterozygote	Unknown	Deleterious	Limite	Deleterious	Limite		
P10	*C11orf80*	intronic	Homozygote	inf e-04	Unknown	Unknown	Unknown	Unknown	VUS	To be confirmed
P11	*MAGEB6*	T96fs	Heterozygote	Unknown	Unknown	Unknown	Unknown	Unknown	Deleterious	To be confirmed
	TOPAZ1	C43S	Heterozygote	inf e-04	Benign	Benign	Benign	Benign	Benign	
		P673A	Heterozygote	0.0182	Unknown	Benign	Deleterious	Deleterious	Likely Deleterious	
P12	*EXO1*	E109K	Heterozygote	0.0002	VUS	Benign	Deleterious	Deleterious	VUS	To be confirmed
		intronic	Heterozygote	0.0016	Unknown	Unknown	Unknown	Unknown	VUS	
P13	*SLC22A16*	V431I	Heterozygote	0.0026	Unknown	Benign	Benign	Benign	Benign	
		E232G	Heterozygote	0.0149	Unknown	Deleterious	Deleterious	Deleterious	Deleterious	
P15	*CCDC36*	N7S	Heterozygote	0.0004	Benign	Benign	Deleterious	Deleterious	VUS	To be confirmed
		K351fs	Heterozygote	inf e-04	Unknown	Unknown	Unknown	Unknown	Deleterious	
P18	*LRRCC1*	L401P	Homozygote	0.0039	Deleterious	Limite	Deleterious	Deleterious	VUS but frequent	
	MDC1	P1375fs	Heterozygote	Unknown	Unknown	Unknown	Unknown	Unknown	Cis	
		P1374fs	Heterozygote	Unknown	Unknown	Unknown	Unknown	Unknown		
	FANCM	I208M	Heterozygote	0.0102	Unknown	Benign	Deleterious	Deleterious	VUS	To be confirmed
		R372W	Heterozygote	inf e-04	Deleterious	Benign	Deleterious	Probablement	Deleterious	
P19	*MOV10L1*	A829T	Homozygote	0	Deleterious	Deleterious	Deleterious	Deleterious	Deleterious	To be confirmed
P20	*TEX11*	A495T	Homozygote	Unknown	Deleterious	Deleterious	Deleterious	Deleterious	Deleterious	To be confirmed
P21	*MCMDC2*	intronic	Homozygote	Unknown	Unknown	Unknown	Unknown	Unknown	VUS	To be confirmed
	ARL2	R96C	Homozygote	inf e-04	Deleterious	Deleterious	Deleterious	Deleterious	Deleterious	To be confirmed
P22	*MAP7*	R394Q	Heterozygote	0.0005	Deleterious	Deleterious	Deleterious	Deleterious	Deleterious	
	MAP7	I163N	Heterozygote	0.0005	Benign	Benign	Benign	Benign	Benign	
P24	*MEI1*	V39G	Heterozygote	0.015	Unknown	Benign	Benign	Benign	Benign	To be confirmed
		D232V	Heterozygote	Unknown	Benign	Benign	Benign	Benign	Benign but unknown	
		L329P	Heterozygote	inf e-04	Deleterious	Limite	Deleterious	Deleterious	Deleterious	

M-CAP, Mendelian Clinically Applicable Pathogenicity: http://bejerano.stanford.edu/mcap/; SIFT, Sorting Intolerant FromTolerant: http://sift.jcvi.org/; VUS, variant of unknown significance; WES, whole-exome sequencing.

patients was therefore uncertain. Relative to the control, P9 showed very low levels of ZNF85 (zinc finger protein 85) expression and P12 showed very low levels of EXO1 expression. Again, the variants' impact on these patients was uncertain.

Protein was absent for the 11 remaining patients. Relative to the control patient, these proteins were either not detected by immunocytochmistry (at spermatogonia stage for ARL2 (ADP ribosylation factor like GTPase 2) (P21)), at the spermatogonia and spermatocyte stages (for CCDC36 (coiled-coil domain-containing protein 36) (P15), MCMDC2 (P21)), predominantly during the spermatocyte stage for (SPATA22 (P3), SYCE1 (P6), PSMC3IP (PSMC3 interacting protein) (P7), C11orf80 (chromosome 11 open reading frame 80) (P10),

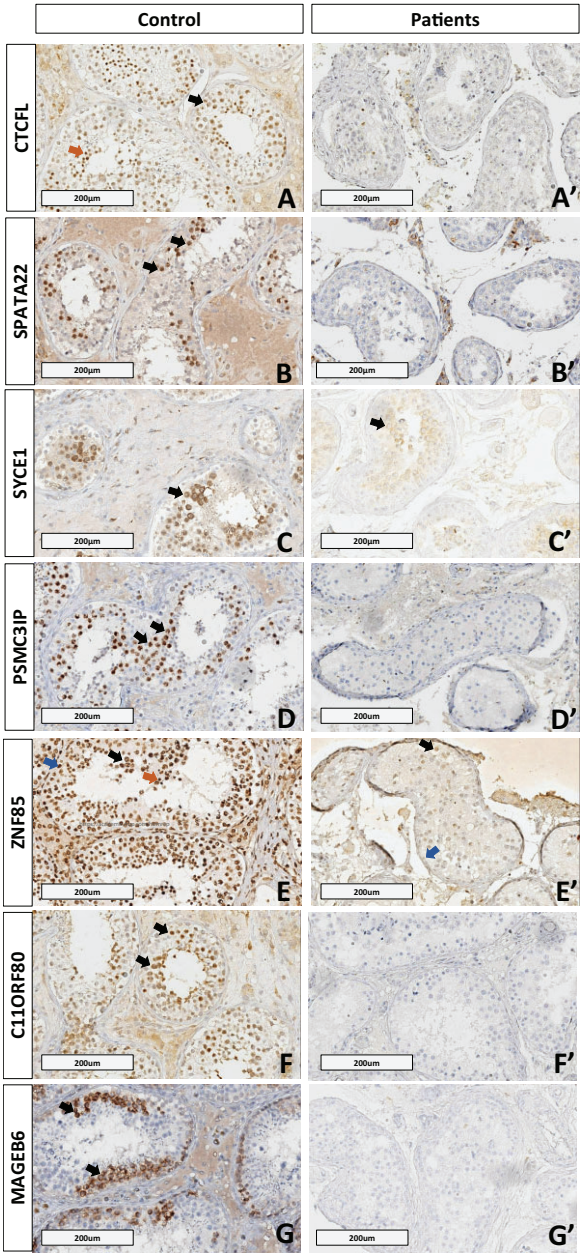

Figure 1. Immunohistochemical analysis results for candidate variants in testicular seminiferous tubules of controls (X) and patients (X') with maturation arrest. A and A': CTCFL, CCCTC-binding factor like; **B and B':** SPATA22, spermatogenesis associated 22; **C and C':** SYCE1, synaptonemal complex central element protein 1; **D and D':** PSMC3IP, PSMC3 interacting protein; **E and E':** ZNF85, zinc finger protein 85; **F and F':** C11ORF80, chromosome 11 open reading frame 80; **G and G':** MAGEB6, MAGE family member B6; **H and H':** EXO1, exonuclease 1; **I and I':** CCDC36, coiled-coil domain-containing protein 36; **J and J':** FANCM, FA complementation group M; **K and K':** MOV10L1, mov10 like RISC complex RNA helicase 1; **L and L':** TEX11, testis expressed 11; **M and M':** MCMDC2, minichromosome maintenance domain containing 2; **N and N':** ARL2, ADP ribosylation factor like GTPase 2; **O and O':** MEI1, meiotic double-stranded break formation protein 1. Blue arrows = spermatogonia, black arrows = spermatocytes and orange arrows = round spermatids. Scale bars = 200 µm. ARL2, ADP ribosylation factor like GTPase 2; C11ORF80, chromosome 11 open reading frame 80; CCDC36, coiled-coil domain-containing protein 36; CTCFL, CCCTC-binding factor like; EXO1, exonuclease 1; FANCM, FA complementation group M; MAGEB6, MAGE family member B6; MCMDC2, minichromosome maintenance domain containing 2; MEI1, meiotic double-stranded break formation protein 1; MOV10L1, mov10 like RISC complex RNA helicase 1; PSMC3IP, PSMC3 interacting protein; SPATA22, spermatogenesis associated 22; SYCE1, synaptonemal complex central element protein 1; TEX11, testis expressed 11; ZNF85, zinc finger protein 85.

(Continued)

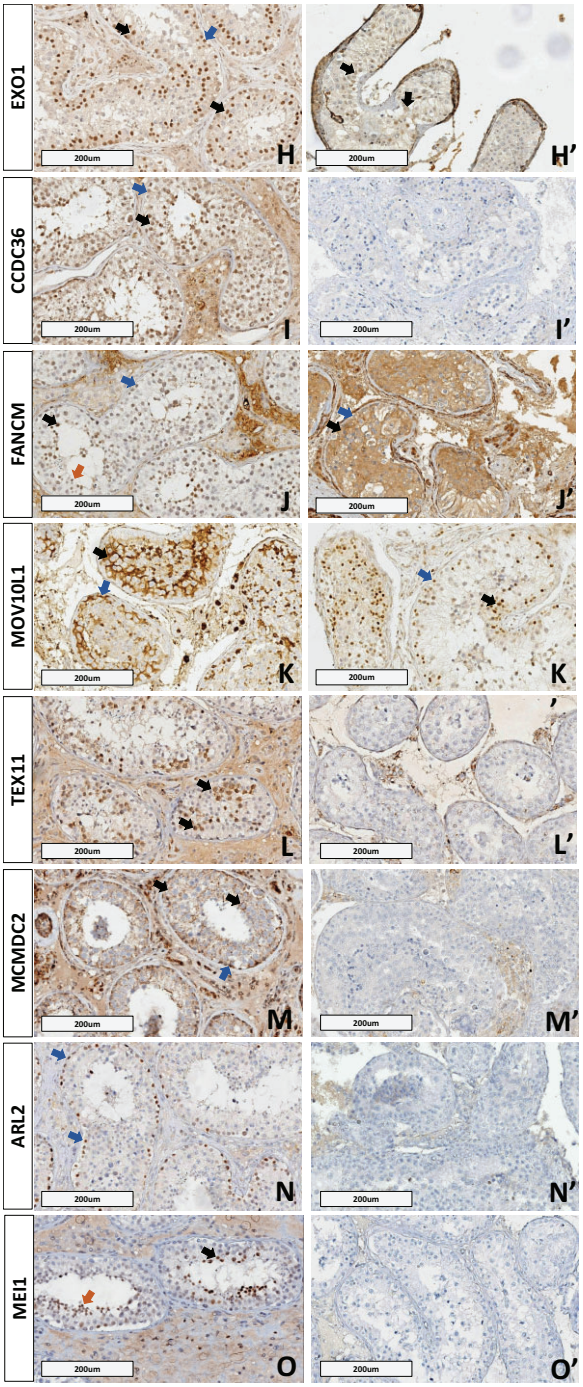

Figure 1. (Continued)

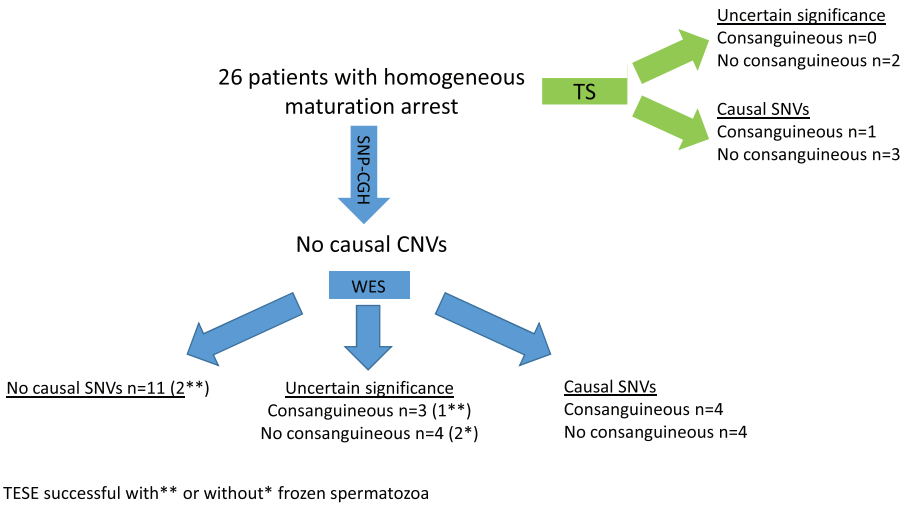

Figure 2. Summary of results obtained in our case series, according to two different genetic strategies: target sequencing or whole-exome sequencing. CNV, copy number variations; SNV, single-nucleotide variant; TESE, testicular sperm extraction; TS, target sequencing; WES, whole-exome sequencing.

MAGEB6 (MAGE family member B6) (P11), TEX11 (P20)), or during the spermatocyte and round spermatid stages (*CTCFL*, encoding CCCTC-binding factor like (P2)) and MEI1 (P24).

Overall interpretation

Considering our results as a whole (Fig. 2 and Table III), we identified clearly deleterious variants in eight patients (all in the TESE− group) and variants of uncertain significance (VUS) in seven patients (four in the TESE− group and three in the TESE+ group). P21 appeared to lack two proteins (MCMDC2 and ARL2). The genetic diagnosis rate was 57.7% (15 out of 26) overall: 100% for consanguineous patients, and 42% for non-consanguineous patients ($P = 0.010$ when comparing the latter two groups). The rates in the TESE+ group (3 out 5; 60%) and the TESE− group (12 out of 21; 57.1%) did not differ significantly.

Discussion

Since 2013, whole-genome analysis has identified more than 40 candidate genes for idiopathic NOA. However, some limitations should be noted. First, only a few of these genes (*SYCE1*, *TEX11* and *STAG3*) have been identified in more than one study (Houston *et al.*, 2021). Second, a number of different testicular histology phenotypes have been reported. Third, few of the studies focused on a particular, homogeneous phenotype. To date, the MA phenotype has been studied most frequently, and over 25 candidate genes have been linked to various stages of spermatogenesis. However, most of these variants were identified in consanguineous families and were not confirmed in independent cohorts. The present study is the first to have assessed a selected group of men with complete MA, TESE results and a testicular biopsy. The participants were included after a double-blind histological assessment (i.e. the examiners were not aware of the TESE outcome); hence, we were able to evaluate the utility of performing WES prior to TESE.

Use of an SNP-CGH array in cases of MA

First, we used an SNP-CGH array to identify new CNVs, screen for the recurrent *TEX11* deletion (Yatsenko *et al.*, 2015), restrict the WES to regions with LOH, and confirm the impact of consanguinity (likely to be high for a putative autosomal recessive disease). Only a few heterozygous deletions or duplications were observed for genes involved in spermatogenesis. No deleterious variants on the second allele were observed with WES for any of these genes. Furthermore, no *TEX11* gene deletions were found. Apart from the regions with LOH in patients with a consanguineous family background, the SNP-CGH array did not reveal any deleterious CNVs; the technique does not appear to be powerful enough in this context. It might be more useful to identify CNVs through next-generation sequencing (NGS). Several NGS tools have been developed in this respect, although each has strengths and weaknesses in terms of applicability; no single tool can detect the entire range of DNA variations (Pirooznia *et al.*, 2015). The development of new WES/whole-genome sequencing tools should soon enable the identification of smaller CNVs. Although many researchers have analysed CNVs in men with NOA, most of the studies revealed heterozygous CNVs, CNVs that only predisposed to infertility, or duplications of spermatogenesis genes that cannot alone account for the NOA phenotype (Sharma *et al.*, 2021).

Use of WES to identify SNPs

We successfully performed WES in a highly selected group of patients. The candidate genes were selected testis-specific genes and/or genes identified in a knock-out mouse model of male infertility with spermatogenesis MA. We identified 16 variants that might be responsible for MA in our patients. Only five of the affected genes have already been

Table III Detailed confirmation result for each variant and overall interpretation.

Patient	Gene	Variant		Confirmation				Conclusive	Previously reported gene	Actionable	
		Protein	Type	Sanger	Immunohistochemical assessment					Before TESE	After TESE
	Name				Human protein atlas	Expression in the control	Expression in the patient				
P2	*CTCFL*	C654R Y115C	Homozygote Homozygote	Yes Yes	Nuclear expression in spermatogonia, preleptotene spermatocytes (high) and pachytene spermatocytes (low)	Nuclear in germ cells, with the highest expression in spermatocytes and round spermatids	None	Yes	No	No	Yes
P3	*SPATA22*	G35R	Homozygote	Yes	No data	Nuclear expression in spermatocytes	None	Uncertain	No	No	No
P4	*MEI1*	intronic A317V	Heterozygote Heterozygote	Yes Yes	Cytoplasmic/membrane expression only in elongated or late spermatids (medium)	Nuclear in spermatocytes and round spermatids	No sample	Uncertain	Yes	Debatable	Yes
P6	*SYCE1*	R125G	Homozygote	Yes	Cytoplasmic/membrane expression in pachytene spermatocytes and round or early spermatids	Nuclear expression in spermatocytes	None	Yes	Yes	Debatable	Yes
P7	*PSMC3IP*	R111fs	Homozygote	Yes	Nuclear expression in spermatogonia, preleptotene and pachytene spermatocytes. and elongated spermatids/cytoplasmic expression in round spermatids	Nuclear expression in spermatocytes	None	Yes	Yes	Yes	Yes
P9	*ZNF85*	T245fs	Homozygote	Yes	Nuclear/cytoplasmic expression in germ cells and Leydig cells	Nuclear expression in all germ cells	None or very low	Uncertain	No	No	Probably
P10	*C11orf80*	intronic	Homozygote	Yes	Cytoplasmic/membrane expression in spermatogonia, preleptotene and pachytene spermatocytes, and in late and round spermatids	Nuclear expression in spermatocytes	None	Yes	No	No	Yes
P11	*MAGEB6*	T96fs	Heterozygote	Yes	Cytoplasmic/membrane expression in pachytene spermatocytes	Membrane expression in spermatocytes	None	Uncertain	No	No	No
P12	*EXO1*	E109K intronic	Heterozygote Heterozygote	Yes Yes	No data		None or very low	Uncertain	No	No	No

(continued)

Table III Continued

Patient	Gene	Variant		Confirmation				Conclusive	Previously reported gene	Actionable	
		Protein	Type	Sanger	Immunohistochemical assessment					Before TESE	After TESE
	Name				Human protein atlas	Expression in the control	Expression in the patient				
						Nuclear expression in spermatogonia (low) and spermatocytes (high)					
P15	*CCDC36*	N7S K351fs	Heterozygote Heterozygote	Yes Yes	Cytoplasmic/membrane expression in pachytene spermatocytes, and round and elongated spermatids	Nuclear expression in spermatogonia and spermatocytes	None	Yes	No	No	Yes
P18	*FANCM*	I208M R372W	Heterozygote Heterozygote	Yes Yes	No data	Nuclear expression in spermatogonia. spermatocytes and round spermatids	Cytoplasmic expression in germ cells	Uncertain	Yes	No	Debatable
P19	*MOV10L1*	A829T	Homozygote	Yes	Cytoplasmic/membrane expression in germ cells (weak) and leydig cells	Cytoplasmic expression in spermatogonia and spermatocytes	Nuclear expression in spermatogonia and spermatocytes	Uncertain	No	No	Debatable
P20	*TEX11*	A495T	Homozygote	Yes	Nuclear expression in spermatogonia, preleptotene and pachytene spermatocytes	Nuclear expression in spermatocytes	None	Yes	Yes	Yes	Yes
P21	*MCMDC2*	intronic	Homozygote	Yes	No data	Cytoplasmic expression in spermatogonia and spermatocytes	None	Yes	No	No	Yes
	ARL2	R96C	Homozygote	Yes	Nuclear expression in spermatogonia, preleptotene and pachytene spermatocytes. and elongated expression	Nuclear expression in spermatogonia	None	Yes	No	No	Yes
P24	*MEI1*	V39G D232V L329P	Heterozygote Heterozygote Heterozygote	No No No	Cytoplasmic/membrane expression only in elongated or late spermatids (medium)	Nuclear in spermatocytes and round spermatids	None	Yes	Yes	No	Yes

TESE, testicular sperm extraction.

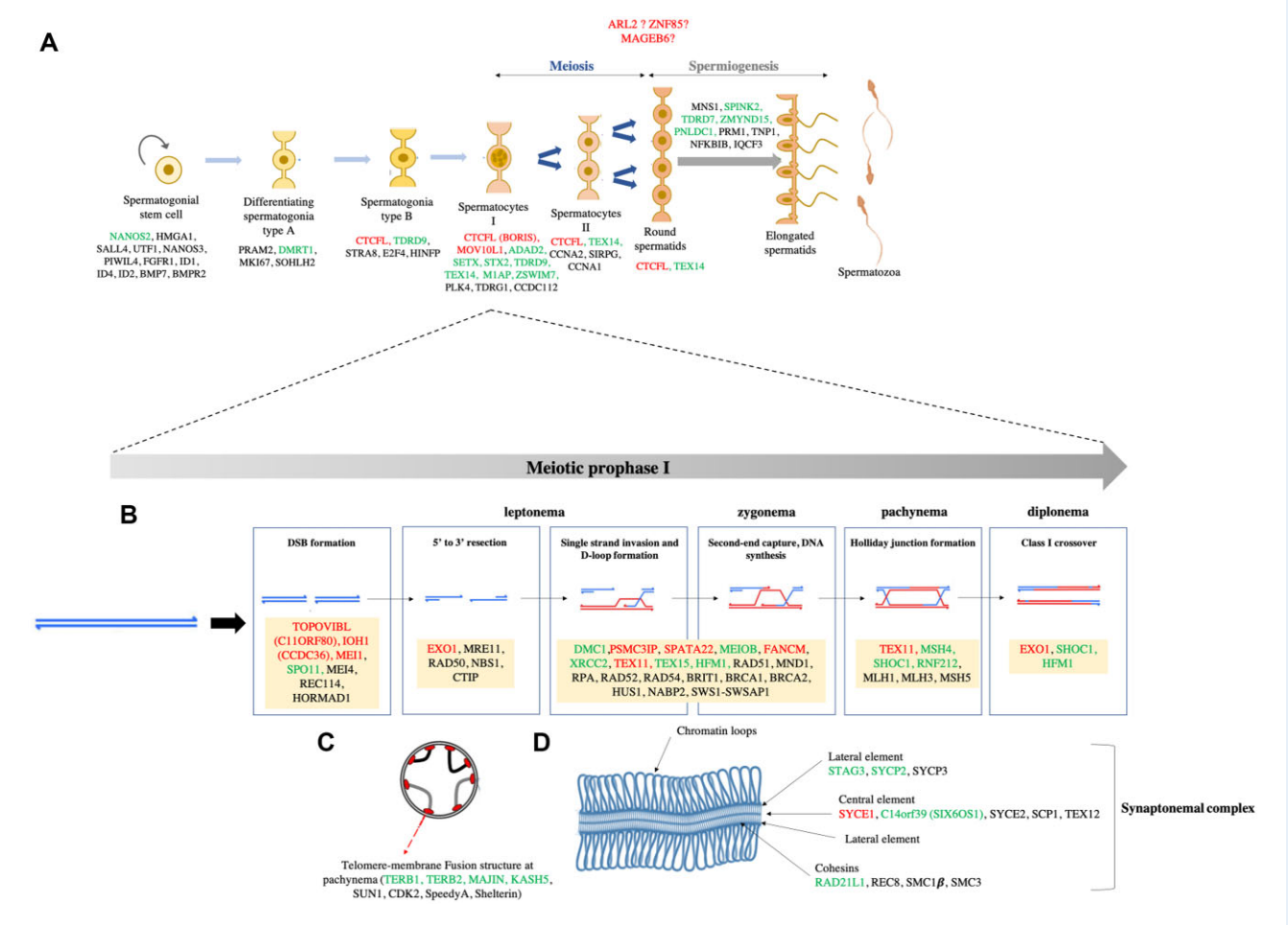

Figure 3. A schematic representation of the genes involved in the major events in spermatogenesis. Candidate genes involved in maturation arrest (MA) in the present study or in other studies are shown in red and green, respectively. Other spermatogenesis-associated genes not yet described in human MA are presented in black. (**A**) The spermatogenesis process. Human MA genes can be involved in the proliferation of spermatogonial stem cells, the differentiation of spermatogonia, epigenetic modifications during meiosis, the protection of spermatocytes from retrotransposons and spermiogenesis. (**B**) Homologous recombination during prophase I comprises four substages: leptonema, zygonema, pachynema and diplonema. A series of events occurs throughout this period, including double-strand break (DSB) formation, 5′-to-3′ resection, strand invasion and crossover (double Holliday junction) formation and resolution. MA-causative genes can affect DSB formation and repair, and pairing, synapsis and recombination between homologous chromosomes. (**C**) Structure of the fusion between chromosome telomeres and the nuclear membrane during pachynema meiosis. During meiosis, telomeres attach to the inner nuclear membrane (INM) and drive the chromosome movement required for homolog pairing and recombination. Human MA-associated genes form complexes to accomplish this meiotic task. (**D**) The tripartite structure of the synaptonemal complex (SC), consisting of two parallel lateral elements and a central element. The SC normally forms between homologs during meiotic prophase I. Human MA-associated genes are involved in the formation of the SC complex and synapsis between homologous chromosomes.

implicated in azoospermia or spermatogenesis MA. We notably highlighted a novel deleterious missense mutation (p.A495T) in the *TEX11* gene. More than 40 *TEX11* variants have been reported in the literature (Ghieh *et al.*, 2019)—confirming the high incidence of *TEX11* gene mutations in MA patients. We also identified novel variants in other genes reported previously as being associated with NOA: *MEI1*, *SYCE1*, *PSMC3IP*, *SPATA22* and *FANCM*. Various researchers have suggested the inclusion of these genes in a target panel for analysis prior to TESE (de Vries et al., 2014; Al-Agha et al., 2018; Ben Khelifa *et al.*, 2018; Kasak *et al.*, 2018; Wu *et al.*, 2021). Our WES analysis also highlighted deleterious variants in *CTCFL*, *C11orf80*, *EXO1*, *CCDC36*, *MOV10L1* and *MCMDC2*. All these genes have been described previously as acting during meiosis. The genes' roles in spermatogenesis are summarized in Supplementary Results. As explained in the description of our prioritization strategy for variant identification, knock-out mouse models have been developed and described for eight of these genes. The observed defects confirm the variants' potential effects on spermatogenesis. The roles of other candidate gene variants identified in this study have not been described unambiguously in the literature. There are no literature data on the functions of *ZNF85* and *ARL2*

genes and proteins during spermatogenesis in the mouse or in other animals. The role of *MAGEB6* during spermatogenesis has not yet been elucidated.

WES versus targeted sequencing: which is the best strategy?

When considering the immunohistochemical analyses, the TESE results and the variants reported in the present study, we were able to provide a definitive genetic diagnosis for eight of our patients. For the other seven patients, the immunohistochemical analyses (n = 4), the TESE outcome (n = 2) or both (n = 1) left room for uncertainty (Fig. 2). All the mutated genes described in the present study are probably involved in meiotic progression (Fig. 3) and are perhaps involved in MA.

Considering the complexity of the main steps in spermatogenesis and the large number of genes involved, one would expect the MA phenotype to show a high degree of genetic heterogeneity. Most of the genetic defects reported in the literature are involved in the early stages of meiosis, such as chromosome synapsis and double-strand break repair (Fig. 3). One could therefore expect whole-genome sequencing to identify a large number of meiotic gene variants that would impair pairing and/or recombination.

A targeted sequencing strategy has been suggested for the above-mentioned genes (Krausz *et al.*, 2020; Cannarella *et al.*, 2021). Had we applied a targeted strategy for genes known to be involved in MA or NOA, the genetic diagnosis rate would have been 27% (i.e. 7 of the 26 patients). Although this value is higher than for other diseases and disorders (such as cardiovascular disease (Richard *et al.*, 2019), epilepsy (Good *et al.*, 2021), intellectual deficiency (Bruel *et al.*, 2020) and premature ovarian insufficiency (Tucker *et al.*, 2016)), the application of WES more than doubled the genetic diagnosis rate (to 57%). When considering solely our seven consanguineous patients, a targeted strategy yielded a diagnosis for two of them. In contrast, the application of WES gave a genetic diagnosis rate of 100% (Fig. 2).

The present study focused on genes expressed in the testis because the patients' consent only covered the identification of genetic defects related to azoospermia; one can therefore consider that this focus corresponds to a broad-panel-based approach. In fact, WES can also generate incidental findings. Thus, the ACMG has issued guidelines on reporting incidental findings in clinical exome and genome sequencing (https://www.ncbi.nlm.nih.gov/clinvar/docs/acmg/). A list of more than 50 genes has been suggested for cardiomyopathy and cancer predisposition screening. The list only includes actionable genes, i.e. those associated with a clear medical strategy. Actionability should be the goal of WES analysis, and this is in line with our objective of avoiding unnecessary TESEs in men with NOA. Although the detection of a larger number of genetic variants increases the cost and complexity of analysis and interpretation, these factors have tended to fall over time. We believe that WES could become the gold standard genetic analysis in the near future—especially for NOA (Ghieh *et al.*, 2021a). However, a WES analysis must always be preceded by counselling. The implementation of WES is currently limited by the requirement for genetic counselling before and after the analysis, the complexity of the data, the patient's medical history and the implications for the patient's family. All these issues will need to be resolved in the near future, and genetic counselling should also address the issue of incidental findings to obtain the patient's consent for the identification of genetic defects not related to azoospermia. It should also be born in mind that only a few variants are reportedly of importance in human male meiosis arrest, azoospermia (Maor-Sagie *et al.*, 2015), and female premature ovarian insufficiency: for example *STAG3* (Jaillard *et al.*, 2020), *PSMC3IP* (Al-Agha et al., 2018) and *SYCE1* (de Vries et al., 2014).

Thus, the large number of genes involved in spermatogenesis (Matzuk and Lamb, 2008) and forthcoming improvements in genetic analysis software (which should report only class 4 and 5 variants, according to the ACMG's guidelines (Richards *et al.*, 2015)) will probably limit the value of targeted sequencing relative to WES. In fact, the pipeline that we used to identify candidate variants is, in a sense, a targeted approach because we only considered genes expressed in the testis. However, according to the human Protein Atlas (https://www.proteinatlas.org/humanproteome/tissue/testis), over 75% of all genes are expressed in the testis, and 10% are predominantly or exclusively expressed there.

Should WES be recommended before or after TESE?

For patients with azoospermia, the main objective of genetic testing (other than obtaining an aetiological diagnosis) is to better predict the outcome of TESE. At present, TESE is only contraindicated for individuals with a 46,XX karyotype or an AZFa and/or AZFb microdeletion (Patrat *et al.*, 2010). Similarly, an *AURKC* gene defect contraindicated IVF in patients with macrozoospermia (Dieterich *et al.*, 2007). In other spermatogenesis defects (e.g. globozoospermia (Celse *et al.*, 2021) and multiple morphological abnormalities of the flagellum (Touré *et al.*, 2021)), the diagnosis is solely aetiological. However, the contraindication of TESE after WES in patients with azoospermia requires strong arguments, such as the identification of a frequent pathologic variant. In contrast, when the variant is first reported, genetic counselling should include a discussion about the likelihood of a successful TESE result, relative to the literature data. Here, the TESE outcome was not a selection criterion. This gave us an opportunity to blindly evaluate the true impact of variant identification on spermatogenesis and, more specifically, to provide the best advice on the likelihood of sperm retrieval in TESE. A candidate variant was identified in 12 of the TESE− patients. When considering only genes repeatedly described in the literature, the WES data alone were conclusive and actionable for two patients (P7 with a PMSC3IP variant, and P20 with a TEX11 variant; Table III). When combined with the immunohistochemistry results, the WES data were conclusive for the 10 other patients (actionable after TESE; Table III). Among the TESE+ patients, VUS were identified for *SPATA22, MAGEB6*, and *EXO1*; only one gene has been described previously as having a role in human MA (*SPATA22* (Wu *et al.*, 2021)). Here, the genetic variants were part of the overall phenotype, and we cannot rule out effects exerted by combinations of genetic factors (Ghieh *et al.*, 2021b) and/or defects in the testicular environment.

Thus, the TESE-WES-immunohistochemistry sequence might be one of the best ways of avoiding repeated, futile TESEs after an initial negative outcome. As has been reported for TESE+ patients, the identification of a gene variant does not necessarily mean that TESE will be negative. When combined with an immunohistochemical assessment, WES is a powerful tool for classifying patients after TESE. Although an immunohistochemical assessment might improve the characterization

of meiotic arrest in metaphase cells, the functional validation of an identified variant is still the most powerful tool. However, cellular models of meiosis are not available, and the development of a specific mouse model poses ethical problems.

For men with a consanguineous family background, WES could be recommended; the identification of a frameshift mutation or a deleterious mutation already known to cause meiotic arrest in humans might help to avoid futile TESEs. More caution is probably required for non-consanguineous men.

Conclusion

We reported the results of an SNP-CGH array analysis and then WES in a series of infertile men with MA. The deleterious nature of some (but not all) of the highlighted mutations was confirmed in immunohistochemical analyses. We conclude that: most cases of MA (and especially those in consanguineous men) have a genetic aetiology; WES is more informative than small-scale, targeted sequencing for genes involved in spermatogenesis; WES might be a powerful tool after an initial negative TESE attempt in cases where a deleterious variant is associated with histologically homogeneous MA and would contraindicate recurrent TESE if the phenotype matches the genotype; and WES before TESE should be limited to consanguineous men. When a certain or probable pathologic variant is identified, consanguinity could be considered as a pejorative additive argument. The identification of gene defects will facilitate discussion with the patient about the risk/benefit ratio of TESE. Although contraindicating TESE following the identification of a pathologic variant will take more of the practitioner's time, we believe that this approach can markedly improve clinical practice.

Given first our postulate whereby MA is probably the best histological phenotype for indicating WES, and second the high probability of a single gene defect (corresponding to the great majority of gene defects described in MA), more data are probably needed before this approach can be extended to all men with NOA. However, our present results suggest that use of this approach can be broadened.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Data availability

All CGH and WES data could be available by request to the corresponding author.

Acknowledgements

We thank the staff at the UVSQ's genomics facility.

Authors' roles

F.G.: participated in data analysis, SNP-CGH array experiments, immunohistochemical analyses, Sanger sequencing and the writing of the first draft. S.F. and A.L.B.: revised the manuscript and validated the testicular histopathological phenotype. C.L., N.S.-B., M.A., D.M.-G., M.B., V.I., F.M. and J.P.: included participants and/or performed testicular biopsies. J.F., C.H., C.G. and M.D.: participated in and provided technical support for immunohistochemical analyses, CGH, WES and/or Sanger sequencing. S.J.: revised the manuscript. V.S.: supervised the genetic analyses. H.J.G.: bioinformatics analyses of WES data. A.L.: supervised the immunohistochemical analyses. B.M.-P.: project supervision and revision of the various versions of the manuscript. F.V.: supervised and coordinated the project and helped to draft the various versions of the manuscript.

Funding

The study was funded by the Fondation Maladies Rares (Paris, France), Merck (Kenilworth, NJ, USA), IRSF (Montigny le Bretonneux, France) and Agence de la Biomédecine (St Denis, France). F.G. received a doctoral fellowship from the Ministère Délégué à l'Enseignement Supérieur et à la Recherche, Paris-Saclay University (https://www.universite-paris-saclay.fr/en), Université Versailles St Quentin (UVSQ) and the PHASE department of Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE).

Conflict of interest

None.

References

Al-Agha AE, Ahmed IA, Nuebel E, Moriwaki M, Moore B, Peacock KA, Mosbruger T, Neklason DW, Jorde LB, Yandell M *et al.* Primary ovarian insufficiency and azoospermia in carriers of a homozygous PSMC3IP stop gain mutation. *J Clin Endocrinol Metab* 2018;**103**:555–563.

Alhathal N, Maddirevula S, Coskun S, Alali H, Assoum M, Morris T, Deek HA, Hamed SA, Alsuhaibani S, Mirdawi A *et al.* A genomics approach to male infertility. *Genet Med* 2020;**22**:1967–1975.

ASRM. The management of obstructive azoospermia: a committee opinion. *Fertil Steril* 2019;**111**:873–880.

Barbotin A-L, Dauvergne A, Dumont A, Ramdane N, Mitchell V, Rigot J-M, Boitrelle F, Robin G. Bilateral versus unilateral cryptorchidism in nonobstructive azoospermia: testicular sperm extraction outcomes. *Asian J Androl* 2019;**21**:445–451.

Barbotin A-L, Dossou Gbete F, Prasivoravong J, Marcelli F, Rigot J-M, Robin G, Mitchell V. New insights into the morphological and hormonal characteristics of spermatogenic arrest. *Andrologia* 2018;**50**:e13136.

Bellil H, Ghieh F, Hermel E, Mandon-Pepin B, Vialard F. Human testis-expressed (TEX) genes: a review focused on spermatogenesis and male fertility. *Basic Clin Androl* 2021;**31**:9.

Ben Khelifa M, Ghieh F, Boudjenah R, Hue C, Fauvert D, Dard R, Garchon HJ, Vialard F. A MEI1 homozygous missense mutation associated with meiotic arrest in a consanguineous family. *Hum Reprod* 2018;**33**:1034–1037.

Bruel A-L, Vitobello A, Tran Mau-Them F, Nambot S, Sorlin A, Denommé-Pichon A-S, Delanne J, Moutton S, Callier P, Duffourd Y, *et al.* Next-generation sequencing approaches and challenges in the

diagnosis of developmental anomalies and intellectual disability. *Clin Genet* 2020;**98**:433–444.

Cannarella R, Condorelli RA, Paolacci S, Barbagallo F, Guerri G, Bertelli M, La Vignera S, Calogero AE. Next-generation sequencing: toward an increase in the diagnostic yield in patients with apparently idiopathic spermatogenic failure. *Asian J Androl* 2021;**23**:24–29.

Cardoso-Moreira M, Halbert J, Valloton D, Velten B, Chen C, Shao Y, Liechti A, Ascenção K, Rummel C, Ovchinnikova S *et al.* Gene expression across mammalian organ development. *Nature* 2019; **571**:505–509.

Celse T, Cazin C, Mietton F, Martinez G, Martinez D, Thierry-Mieg N, Septier A, Guillemain C, Beurois J, Clergeau A, *et al.* Genetic analyses of a large cohort of infertile patients with globozoospermia, DPY19L2 still the main actor, GGN confirmed as a guest player. *Hum Genet* 2021;**140**:43–57.

Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;**34**:i884–i890.

Corona G, Minhas S, Giwercman A, Bettocchi C, Dinkelman-Smit M, Dohle G, Fusco F, Kadioglou A, Kliesch S, Kopa Z *et al.* Sperm recovery and ICSI outcomes in men with non-obstructive azoospermia: a systematic review and meta-analysis. *Hum Reprod Update* 2019;**25**:733–757.

Darde TA, Sallou O, Becker E, Evrard B, Monjeaud C, Le Bras Y, Jégou B, Collin O, Rolland AD, Chalmel F. The ReproGenomics Viewer: an integrative cross-species toolbox for the reproductive science community. *Nucleic Acids Res* 2015;**43**:W109–W116.

De Braekeleer M, Dao TN. Cytogenetic studies in male infertility: a review. *Hum Reprod* 1991;**6**:245–250.

de Vries L, Behar DM, Smirin-Yosef P, Lagovsky I, Tzur S, Basel-Vanagaite L. Exome sequencing reveals SYCE1 mutation associated with autosomal recessive primary ovarian insufficiency. *J Clin Endocrinol Metab* 2014;**99**:E2129–E2132.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, Angel G, del Rivas MA, Hanna M *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;**43**:491–498.

Dieterich K, Soto Rifo R, Faure AK, Hennebicq S, Ben Amar B, Zahi M, Perrin J, Martinez D, Sèle B, Jouk P-S *et al.* Homozygous mutation of AURKC yields large-headed polyploid spermatozoa and causes male infertility. *Nat Genet* 2007;**39**:661–665.

Dorosh A, Tepla O, Zatecka E, Ded L, Koci K, Peknicova J. Expression analysis of MND1/GAJ, SPATA22, GAPDHS and ACR genes in testicular biopsies from non-obstructive azoospermia (NOA) patients. *Reprod Biol Endocrinol* 2013;**11**:42.

Ferguson-Smith MA. X-Y chromosomal interchange in the aetiology of true hermaphroditism and of XX Klinefelter's syndrome. *Lancet* 1966;**288**:475–476.

Finsterbusch F, Ravindranathan R, Dereli I, Stanzione M, Tränkner D, Tóth A. Alignment of homologous chromosomes and effective repair of programmed DNA double-strand breaks during mouse meiosis require the minichromosome maintenance domain containing 2 (MCMDC2) protein. *PLoS Genet* 2016;**12**:e1006393.

Frost RJA, Hamra FK, Richardson JA, Qi X, Bassel-Duby R, Olson EN. MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. *Proc Natl Acad Sci USA* 2010;**107**:11847–11852.

Fu Q, Pandey RR, Leu NA, Pillai RS, Wang PJ. Mutations in the MOV10L1 ATP hydrolysis motif cause piRNA biogenesis failure and male sterility in mice. *Biol Reprod* 2016;**95**:103.

Ghieh F, Barbotin A-L, Leroy C, Marcelli F, Swierkowsky-Blanchard N, Serazin V, Mandon-Pepin B, Vialard F. Will whole-genome sequencing become the first-line genetic analysis for male infertility in the near future? *Basic Clin Androl* 2021a;**31**:21.

Ghieh F, Barbotin A-L, Prasivoravong J, Ferlicot S, Mandon-Pepin B, Fortemps J, Garchon H-J, Serazin V, Leroy C, Marcelli F *et al.* Azoospermia and reciprocal translocations: should whole-exome sequencing be recommended? *Basic Clin Androl* 2021b;**31**:27.

Ghieh F, Mitchell V, Mandon-Pepin B, Vialard F. Genetic defects in human azoospermia. *Basic Clin Androl* 2019;**29**:4.

Good J-M, Atallah I, Castro Jimenez M, Benninger D, Kuntzer T, Superti-Furga A, Tran C. NGS-based diagnosis of treatable neurogenetic disorders in adults: opportunities and challenges. *Genes* 2021;**12**:695.

Halder A, Kumar P, Jain M, Iyer VK. Copy number variations in testicular maturation arrest. *Andrology* 2017;**5**:460–472.

Hays E, Majchrzak N, Daniel V, Ferguson Z, Brown S, Hathorne K, La Salle S. Spermatogenesis associated 22 is required for DNA repair and synapsis of homologous chromosomes in mouse germ cells. *Andrology* 2017;**5**:299–312.

Houston BJ, Riera-Escamilla A, Wyrwoll MJ, Salas-Huetos A, Xavier MJ, Nagirnaja L, Friedrich C, Conrad DF, Aston KI, Krausz C *et al.* A systematic review of the validated monogenic causes of human male infertility: 2020 update and a discussion of emerging gene-disease relationships. *Hum Reprod Update* 2021;**28**:15–29.

Jaillard S, McElreavy K, Robevska G, Akloul L, Ghieh F, Sreenivasan R, Beaumont M, Bashamboo A, Bignon-Topalovic J, Neyroud A-S *et al.* STAG3 homozygous missense variant causes primary ovarian insufficiency and male non-obstructive azoospermia. *Mol Hum Reprod* 2020;**26**:665–677.

Jelinic P, Stehle J-C, Shaw P. The testis-specific factor CTCFL cooperates with the protein methyltransferase PRMT7 in H19 imprinting control region methylation. *PLoS Biol* 2006;**4**:e355.

Jungwirth A, Giwercman A, Tournaye H, Diemer T, Kopa Z, Dohle G, Krausz C; European Association of Urology Working Group on Male Infertility. European Association of Urology guidelines on Male Infertility: the 2012 update. *Eur Urol* 2012;**62**: 324–332.

Kasak L, Punab M, Nagirnaja L, Grigorova M, Minajeva A, Lopes AM, Punab AM, Aston KI, Carvalho F, Laasik E *et al.*; GEMINI Consortium. Bi-allelic recessive loss-of-function variants in FANCM cause non-obstructive azoospermia. *Am J Hum Genet* 2018;**103**: 200–212.

Kosaka-Suzuki N, Suzuki T, Pugacheva EM, Vostrov AA, Morse HC, Loukinov D, Lobanenkov V. Transcription factor BORIS (Brother of the Regulator of Imprinted Sites) directly induces expression of a cancer-testis antigen, TSP50, through regulated binding of BORIS to the promoter. *J Biol Chem* 2011;**286**:27378–27388.

Krausz C, Riera-Escamilla A, Moreno-Mendoza D, Holleman K, Cioppi F, Algaba F, Pybus M, Friedrich C, Wyrwoll MJ, Casamonti E *et al.* Genetic dissection of spermatogenic arrest through exome analysis: clinical implications for the management of azoospermic men. *Genet Med* 2020;**22**:1956–1966.

La Salle S, Palmer K, O'Brien M, Schimenti JC, Eppig J, Handel MA. Spata22, a novel vertebrate-specific gene, is required for meiotic progress in mouse germ cells. *Biol Reprod* 2012;**86**:45.

Lee AK, Potts PR. A comprehensive guide to the MAGE family of ubiquitin ligases. *J Mol Biol* 2017;**429**:1114–1142.

Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv13033997 Q-Bio* [Internet] 2013. Available from: http://arxiv.org/abs/1303.3997.

Maor-Sagie E, Cinnamon Y, Yaacov B, Shaag A, Goldsmidt H, Zenvirt S, Laufer N, Richler C, Frumkin A. Deleterious mutation in SYCE1 is associated with non-obstructive azoospermia. *J Assist Reprod Genet* 2015;**32**:887–891.

Matzuk MM, Lamb DJ. The biology of infertility: research advances and clinical challenges. *Nat Med* 2008;**14**:1197–1213.

McLachlan RI, Rajpert-De Meyts E, Hoei-Hansen CE, de Kretser DM, Skakkebaek NE. Histological evaluation of the human testis—approaches to optimizing the clinical value of the assessment: mini review. *Hum Reprod* 2007;**22**:2–16.

McNairn AJ, Rinaldi VD, Schimenti JC. Repair of meiotic DNA breaks and homolog pairing in mouse meiosis requires a minichromosome maintenance (MCM) paralog. *Genetics* 2017;**205**:529–537.

McQuillan R, Leutenegger A-L, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, Smolej-Narancic N, Janicijevic B, Polasek O, Tenesa A *et al*. Runs of homozygosity in European populations. *Am J Hum Genet* 2008;**83**:359–372.

Necsulea A, Soumillon M, Warnefors M, Liechti A, Daish T, Zeller U, Baker JC, Grützner F, Kaessmann H. The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* 2014;**505**:635–640.

Patrat C, Bienvenu T, Janny L, Faure A-K, Fauque P, Aknin-Seifer I, Davy C, Thiounn N, Jouannet P, Lévy R. Clinical data and parenthood of 63 infertile and Y-microdeleted men. *Fertil Steril* 2010;**93**:822–832.

Pirooznia M, Goes FS, Zandi PP. Whole-genome CNV analysis: advances in computational approaches. *Front Genet* 2015;**6**:138.

Poncelet DA, Bellefroid EJ, Bastiaens PV, Demoitié MA, Marine JC, Pendeville H, Alami Y, Devos N, Lecocq P, Ogawa T *et al*. Functional analysis of ZNF85 KRAB zinc finger protein, a member of the highly homologous ZNF91 family. *DNA Cell Biol* 1998;**17**:931–943.

Richard P, Ader F, Roux M, Donal E, Eicher J-C, Aoutil N, Huttin O, Selton-Suty C, Coisne D, Jondeau G *et al*. Targeted panel sequencing in adult patients with left ventricular non-compaction reveals a large genetic heterogeneity. *Clin Genet* 2019;**95**:356–367.

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E *et al.*; ACMG Laboratory Quality Assurance Committee. 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**:405–424.

Rivero-Hinojosa S, Pugacheva EM, Kang S, Méndez-Catalá CF, Kovalchuk AL, Strunnikov AV, Loukinov D, Lee JT, Lobanenkov VV. The combined action of CTCF and its testis-specific paralog BORIS is essential for spermatogenesis. *Nat Commun* 2021;**12**:3846.

Robert T, Nore A, Brun C, Maffre C, Crimi B, Bourbon H-M, de Massy B. The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. *Science* 2016;**351**:943–949.

Sanchez A, Adam C, Rauh F, Duroc Y, Ranjha L, Lombard B, Mu X, Wintrebert M, Loew D, Guarné A *et al*. Exo1 recruits Cdc5 polo kinase to MutLγ to ensure efficient meiotic crossover formation. *Proc Natl Acad Sci USA* 2020;**117**:30577–30588.

Sarkardeh H, Totonchi M, Asadpour O, Sadighi Gilani MA, Zamani Esteki M, Almadani N, Borjian Boroujeni P, Gourabi H. Association of MOV10L1 gene polymorphisms and male infertility in azoospermic men with complete maturation arrest. *J Assist Reprod Genet* 2014;**31**:865–871.

Schaetzlein S, Chahwan R, Avdievich E, Roa S, Wei K, Eoff RL, Sellers RS, Clark AB, Kunkel TA, Scharff MD *et al*. Mammalian Exo1 encodes both structural and catalytic functions that play distinct roles in essential biological processes. *Proc Natl Acad Sci USA* 2013;**110**:E2470–E2479.

Sharma A, Jain M, Halder A, Kaushal S. Identification of genomic imbalances (CNVs as well as LOH) in sertoli cell only syndrome cases through cytoscan microarray. *Gene* 2021;**801**:145851.

Sleutels F, Soochit W, Bartkuhn M, Heath H, Dienstbach S, Bergmaier P, Franke V, Rosa-Garrido M, van de Nobelen S, Caesar L *et al*. The male germ cell gene regulator CTCFL is functionally different from CTCF and binds CTCF-like consensus sites in a nucleosome composition-dependent manner. *Epigenetics Chromatin* 2012;**5**:8.

Stanzione M, Baumann M, Papanikos F, Dereli I, Lange J, Ramlal A, Tränkner D, Shibuya H, Massy B, de Watanabe Y *et al*. Meiotic DNA break formation requires the unsynapsed chromosome axis-binding protein IHO1 (CCDC36) in mice. *Nat Cell Biol* 2016;**18**:1208–1220.

Sujit KM, Sarkar S, Singh V, Pandey R, Agrawal NK, Trivedi S, Singh K, Gupta G, Rajender S. Genome-wide differential methylation analyses identifies methylation signatures of male infertility. *Hum Reprod* 2018;**33**:2256–2267.

Sund KL, Zimmerman SL, Thomas C, Mitchell AL, Prada CE, Grote L, Bao L, Martin LJ, Smolarek TA. Regions of homozygosity identified by SNP microarray analysis aid in the diagnosis of autosomal recessive disease and incidentally detect parental blood relationships. *Genet Med* 2013;**15**:70–78.

Suzuki T, Kosaka-Suzuki N, Pack S, Shin D-M, Yoon J, Abdullaev Z, Pugacheva E, Morse HC, Loukinov D, Lobanenkov V. Expression of a testis-specific form of Gal3st1 (CST), a gene essential for spermatogenesis, is regulated by the CTCF paralogous gene BORIS. *Mol Cell Biol* 2010;**30**:2473–2484.

Touré A, Martinez G, Kherraf Z-E, Cazin C, Beurois J, Arnoult C, Ray PF, Coutton C. The genetic architecture of morphological abnormalities of the sperm tail. *Hum Genet* 2021;**140**:21–42.

Tucker EJ, Grover SR, Bachelot A, Touraine P, Sinclair AH. Premature ovarian insufficiency: new perspectives on genetic cause and phenotypic spectrum. *Endocr Rev* 2016;**37**:609–635.

Vogt P, Chandley AC, Hargreave TB, Keil R, Ma K, Sharkey A. Microdeletions in interval 6 of the Y chromosome of males with idiopathic sterility point to disruption of AZF, a human spermatogenesis gene. *Hum Genet* 1992;**89**:491–496.

Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;**38**:e164.

Wei K, Clark AB, Wong E, Kane MF, Mazur DJ, Parris T, Kolas NK, Russell R, Hou H, Kneitz B *et al*. Inactivation of Exonuclease 1 in mice results in DNA mismatch repair defects, increased cancer susceptibility, and male and female sterility. *Genes Dev* 2003;**17**:603–614.

Wierenga KJ, Jiang Z, Yang AC, Mulvihill JJ, Tsinoremas NF. A clinical evaluation tool for SNP arrays, especially for autosomal recessive conditions in offspring of consanguineous parents. *Genet Med* 2013;**15**:354–360.

Wu Y, Li Y, Murtaza G, Zhou J, Jiao Y, Gong C, Hu C, Han Q, Zhang H, Zhang Y *et al*. Whole-exome sequencing of consanguineous families with infertile men and women identifies homologous mutations in SPATA22 and MEIOB. *Hum Reprod* 2021;**36**: 2793–2804.

Yatsenko AN, Georgiadis AP, Röpke A, Berman AJ, Jaffe T, Olszewska M, Westernströer B, Sanfilippo J, Kurpisz M, Rajkovic A, *et al*. X-linked TEX11 mutations, meiotic arrest, and azoospermia in infertile men. *N Engl J Med* 2015;**372**:2097–2107.

Zheng K, Xiol J, Reuter M, Eckardt S, Leu NA, McLaughlin KJ, Stark A, Sachidanandam R, Pillai RS, Wang PJ. Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway. *Proc Natl Acad Sci USA* 2010; **107**:11841–11846.

Zhou C-X, Wang Y, Shi L-Y, Wang Z-B, Ma Y, Li C-R, Zhang N-N, Zhang Y-X, Zhang F, Zhang D *et al*. GTPases Arf5 and Arl2 function partially distinctly during oocyte meiosis. *J Cell Biochem* 2021; **122**:198–208.