ANDROLOGY

ISSN: 2047-2919

ANDROLOGY

ORIGINAL ARTICLE

Correspondence:
Corinna Friedrich, Institute of Human Genetics, University of Münster, Vesaliusweg 12-14, 48149 Münster, Germany.
E-mail: Corinna.Friedrich@ukmuenster.de

*These authors contributed equally to the study and should be considered joint first authors.

Keywords:
candidate genes, male infertility, non-obstructive azoospermia, whole exome sequencing

Received: 9-Jul-2019
Revised: 20-Aug-2019
Accepted: 26-Aug-2019

doi: 10.1111/andr.12704

Sequence analysis of 37 candidate genes for male infertility: challenges in variant assessment and validating genes

[1,*]T. F. Araujo, [2,*]C. Friedrich, [1]C. H. P. Grangeiro, [1]L. R. Martelli, [1]J. D. Grzesiuk, [2]J. Emich, [2]M. J. Wyrwoll, [3]S. Kliesch, [1]A. L. Simões and [2]F. Tüttelmann

[1]Department of Genetics, Ribeirão Preto Medical School, University of São Paulo (USP), Ribeirão Preto, SP, Brazil, [2]Institute of Human Genetics, University of Münster, Münster, Germany, and [3]Centre of Reproductive Medicine and Andrology, Department of Clinical and Surgical Andrology, University Hospital Münster, Münster, Germany

ABSTRACT

Background: The routine genetic analysis for diagnosing male infertility has not changed over the last twenty years, and currently available tests can only determine the etiology of 4% of unselected infertile patients. Thus, to create new diagnostic assays, we must better understand the molecular and genetic mechanisms of male infertility. Although next-generation sequencing allows for simultaneous analysis of hundreds of genes and the discovery of novel candidates related to male infertility, so far only a few gene candidates have enough sound evidence to support the gene–disease relationship.

Objective: Since complementary studies are required to validate genes, we aimed to analyze the presence of potentially pathogenic rare variants in a set of candidate genes related to azoospermia in a hitherto understudied South American population.

Subjects and Methods: We performed whole exome sequencing in a group of 16 patients with non-obstructive azoospermia from Ribeirão Preto, Brazil. Based on a recent systematic review of monogenic causes of male infertility, we selected a set of 37 genes related to azoospermia, Sertoli-Cell-Only histology, and spermatogenic arrest to analyze. The identified variants were confirmed by Sanger sequencing, and their functional consequence was predicted by *in silico* programs.

Results: We identified potential pathogenic variants in seven genes in six patients. Two variants, c.671A>G (p.(Asn224Ser)) in *DMRT1* and c.91C>T (p.(Arg31Cys)) in *REC8*, have already been described in association with azoospermia. We also found new variants in genes that already have moderate evidence of being linked to spermatogenic failure (*TEX15, KLHL10*), in genes with limited evidence (*DNMT3B, TEX14*) and in one novel promising candidate gene that has no evidence so far (*SYCE1L*).

Discussion: Although this study included a small number of patients, the process of rationally selecting genes allowed us to detect rare potentially pathogenic variants, providing supporting evidence for validating candidate genes associated with azoospermia.

INTRODUCTION

Male infertility is a multifactorial condition associated with variable phenotypes, and it affects approximately 7% of men worldwide (Krausz & Riera-Escamilla, 2018). Diagnosing this condition is difficult not only because several environmental and genetic factors interfere with spermatogenesis (Cocuzza *et al.*, 2013), but also because of social taboos. In some countries of Africa, men unable to father children are so marginalized that some may even commit suicide (Zarrabi & Kruger, 2018). Moreover, a man is usually only analyzed after the woman has undergone an extensive workup (Cocuzza *et al.*, 2013).

The cornerstone of male infertility diagnostics is semen analysis, which evaluates sperm number, motility, and morphology following World Health Organization (WHO) guidelines (WHO, 2010). The most severe phenotype of male infertility is non-obstructive azoospermia (NOA), characterized by the absence of spermatozoa in the ejaculate even after centrifugation. In men with NOA, testicular biopsy reveals histological variations, including Sertoli-Cell-Only histology (SCO), maturation arrest at any stage of spermatogenesis, and mixed testicular atrophy. These variations in histopathology might be a consequence of the complex process that regulates spermatogenesis (Gershoni *et al.*, 2017).

Among several possible etiologies, those of genetic origin may be transmitted to one's offspring if a child can be conceived, for

[Copyright line changed after the first online publication on 3 September 2019]

example, by assisted reproduction. Therefore, genetic screening is relevant not only because of its diagnostic value but also for its potential prognostic value for clinical decision-making and appropriate genetic counseling (Kumari *et al*., 2015; Belva *et al*., 2016; Krausz & Riera-Escamilla, 2018). Since Klinefelter syndrome and other chromosomal aberrations are the major genetic causes of male infertility, the first genetic examination performed is karyotype analysis. This examination detects numerical chromosomal abnormalities, mosaicisms, and some structural chromosomal abnormalities. Patients with severe oligozoospermia ($<5 \times 10^6$ spermatozoa per mL semen) or azoospermia usually undergo molecular analysis to identify AZF deletions on chromosome Y (Krausz *et al*., 2013). Beyond these genetic examinations, gene sequencing in a clinical setting is currently only performed in the very rare condition of hypogonadotropic hypogonadism or in men with congenital bilateral absence of vas deferens (CBAVD) (Tüttelmann *et al*., 2018).

This routine diagnostic workup has not changed over the last twenty years, and the currently available genetic tests can only determine the etiology of 4% of unselected infertile men and 20% of azoospermic men (Tüttelmann *et al*., 2018). Thus, developing new diagnostic assays requires understanding the molecular and genetic mechanisms of male infertility.

Determining the genetic component of male infertility is difficult because >2,000 genes are involved in human spermatogenesis (Krausz *et al*., 2015). Since the late 1990s, an intense search has been underway to identify genetic risk factors of male infertility via associations with single nucleotide polymorphisms (SNP) and copy number variation (CNV) (Krausz *et al*., 2015). Despite the more than 300 SNPs distributed in 123 genes reported to be associated with male infertility (Krausz *et al*., 2015) and some CNVs (Tüttelmann *et al*., 2011; Krausz *et al*., 2012; Nakamura *et al*., 2017), most data remain controversial, as patients' ethnic and geographical origin seem to influence the phenotypic expression of these genomic variations (Krausz *et al*., 2015; Krausz & Riera-Escamilla, 2018).

A relatively novel approach to identify potential pathogenic genes is next-generation sequencing (NGS), which allows for rapid and cost-effective sequencing of whole exomes or genomes. As such, this approach can help identify novel genetic factors in different complex diseases, including male infertility (Krausz & Riera-Escamilla, 2018). In fact, whole exome sequencing (WES) has already been employed to identify novel genes related to the phenotype of non-obstructive azoospermia in humans (Gershoni *et al*., 2017; Fakhro *et al*., 2018), but the number of genes that can be confidently linked to this condition is still low (Oud *et al*., 2019). Without stronger evidence, some findings may lead to distorted conclusions, resulting in inappropriate subsequent research or false understanding of genetic pathways (Oud *et al*., 2019).

Although several variants in numerous genes have been described in association with male infertility, few studies have been independently validated or have provided functional evidence that the variants identified are pathogenic. Furthermore, only a fraction of the identified genes have supporting biological evidence (e.g., a knockout mouse model showing male infertility) (Tüttelmann *et al*., 2018). Regarding genes that do fulfill multiple levels of evidence for an association with male infertility, an unstructured assessment reportedly found three: *NR5A1*, *DMRT1*, and *TEX11* (Tüttelmann *et al*., 2018). Very recently, an extensive literature review and structured clinical validity assessment of a large number of genes showed that only a few have 'strong' evidence of being associated with male infertility, while many more have 'moderate' or 'limited' evidence (Oud *et al*., 2019).

Since studies in male infertility are still lacking (Barratt *et al*., 2017), our objective was to analyze the presence of rare, potentially pathogenic variants in a set of candidate genes associated with non-obstructive azoospermia and provide evidence for their clinical validity.

SUBJECTS AND METHODS

Ethical approval

This study was conducted in accordance with the ethical standards of the involved institutions. Written informed consent for genetic testing was obtained from all subjects participating, according to the approved Ethics Committee of the Hospital das Clínicas of Ribeirão Preto (reference number HCRP_n°8715/2013), and Ethics Committee of the State Medical Board and the Medical Faculty Münster (Kennzeichen 2010-578-f-S).

Patients

The study group consisted of 16 unrelated infertile patients, below 45 years of age at first evaluation, seeking advice for couple infertility at Human Reproduction Division of Hospital das Clínicas of Ribeirão Preto (HCRP, Brazil). These patients were recruited in a previous study (Grangeiro, 2018). All had azoospermia and reduced testicular volumes (<15 ml) measured by the same physician (CHPG) using an orchidometer. Twelve patients were confidently diagnosed as having NOA based on testicular biopsy. Six of them had SCO and six had germ cell maturation arrest. Two of the remaining patients had elevated FSH levels, which together with the reduced testis lead to the diagnosis of NOA. The diagnoses in the last two patients (M1807 and M1809) solely relied on reduced testis size. The normal routine diagnostic workups included GTG-banding with exclusion of chromosomal mosaicism (Hook, 1977) and AZF deletions according to the EAA and EMQN protocol (Krausz *et al*., 2013). Clinical information regarding these patients is summarized in Table 1.

We excluded patients showing characteristics known to be associated with male infertility phenotype, such as (i) endocrinopathy (hypogonadotropic hypogonadism, hyperprolactinemia, hyper/hypothyroidism, use of anabolic steroids); (ii) pathologies or procedures that affect the production of spermatozoa (chemo- or radiotherapy, traumatic or infectious orchitis, varicocoele (≥ grade II), bilateral cryptorchidism); (iii) post-testicular dysfunctions (erectile dysfunction/loss of libido/anorgasmia, retrograde ejaculation/anejaculation, congenital absence of the vas deferens, vasectomy, mechanical obstruction resulting from trauma or infection); or (iv) testicular volumes >15 ml.

For one variant filtering step, we utilized an already available in-house control group, which is part of our larger scale Male Reproductive Genomics (MERGE) study comprising >700 exomes. The control group consisted of 17 unrelated healthy men below 45 years attending the Centre of Reproductive Medicine and Andrology (CeRA) in Münster, Germany. They presented normal semen parameters according to WHO criteria (World Health Organization, 2010).

Table 1 Clinical and routine genetic characterization of NOA patients

Patient ID	Age (years)	Testosterone ng/mL (2.49–8.36)	FSH mIE/mL (<7.0)	LH mIE/mL (1.7–8.6)	Karyotype (GTG banding)	AZF deletions	Testicular histology
M1802	34	3.18	**16.7**	8.1	46, XY[100]	Negative	SCO
M1803	33	4.20	3.7	2.0	46, XY[50]	Negative	Mat.A
M1804	32	6.81	**7.7**	3.0	46, XY[50]	Negative	SCO
M1805	29	3.92	**23**	7.9	46, XY[50]	Negative	SCO
M1806	35	**1.83**	**13.1**	4.2	46, XY[50]	Negative	SCO
M1807	33	3.95	3.2	**1.0**	46, XY[50]	Negative	–
M1808	30	4.11	**21**	6.7	46, XY[50]	Negative	Mat.A
M1809	24	4.80	4.3	5.0	46, XY[50]	Negative	–
M1810	31	2.96	1.7	2.6	46, XY[50]	Negative	Mat.A
M1811	39	3.25	**19.7**	6.5	46, XY[100]	Negative	–
M1812	34	3.05	**11.5**	5.9	46, XY[50]	Negative	Mat.A
M1813	37	5.20	**7.7**	2.8	46, XY[50]	Negative	Mat.A
M1814	34	4.45	**7.7**	2.7	46, XY[50]	Negative	–
M1815	26	3.20	**9.4**	3.7	46, XY[100]	Negative	Mat.A
M1816	36	3.00	**17.5**	6.5	46, XY[50]	Negative	SCO
M1817	31	3.50	**10.1**	5.5	46, XY[50]	Negative	SCO

ID, Identity; FSH, Follicle-stimulating hormone; LH, Luteinizing hormone; AZF, azoospermia factor; SCO, Sertoli-Cell-Only histology; Mat.A, germ cell maturation arrest. Hormone values outside the normal range are marked in bold.

Whole exome sequencing and bioinformatic analysis

Genomic DNA was isolated from peripheral blood according to standard procedures and as described previously (Röpke *et al.*, 2013). For WES, samples were prepared and enriched according to the protocol of Agilent's SureSelect^{QXT} Target Enrichment for Illumina Multiplexed Sequencing Featuring Transposase-Based Library Prep Technology (Agilent Technologies, Inc., Santa Clara, USA). The libraries were index tagged using appropriate pairs of index primers for multiplexed sequencing. To capture libraries, SureSelect^{XT} Human All Exon V6 was used. Quantity and quality of the libraries were determined with Agilent's TapeStation 2200, and the final concentration was adjusted to 1.6 pM. Sequencing was performed on the Illumina NextSeq® 500 System using the NextSeq 500 V2 High-Output Kit (300 cycles).

After trimming, the remaining adapter sequences and primers were removed with Cutadapt v1.15 (Martin, 2011). Sequence reads were aligned against the reference genome GRCh37.p13 using BWA Mem v0.7.17 (Li & Durbin, 2010). Duplicate reads and reads that mapped to multiple locations in the genome were excluded from further analysis. Single nucleotide variations and small insertions/deletions (Indels) were identified and quality-filtered using GATK toolkit v3.8 with haplotype caller, according to the best practice recommendations (McKenna *et al.*, 2010). Briefly, haplotype caller performs local *de novo* assembly in active regions and calculates haplotype likelihoods at potential variant sites. Resulting variants were annotated with *Ensembl* Variant Effect Predictor (McLaren *et al.*, 2016).

Selection of genes for analysis

Initially, we analyzed the WES data for the three genes *NR5A1*, *DMRT1*, and *TEX11* from our unstructured assessment (Tüttelmann *et al.*, 2018). Subsequently, we selected a set of genes based on a systematic review that established the clinical validity of candidate genes related to male infertility (Oud *et al.*, 2019). We considered all genes with definitive, strong, moderate, and limited evidence of monogenetic association with male infertility and selected those leading to the phenotypes presented by our patients, such as NOA, SCO, and germ cell maturation arrest (see Table S1).

Variant filtering

We filtered for coding variants with potentially severe consequences for protein function—namely those affecting splice sites (donor/acceptor), start loss, stop gain, frameshift, and missense. In order to rule out common polymorphisms, we excluded all variants with a frequency higher than 0.05 in any subpopulation in the Exome Sequencing Project (ESP, https://evs.gs.washington.edu/EVS/), 1000 Genomes Browser (http://phase3browser.1000genomes.org/index.html), and the genome aggregation database gnomAD (https://gnomad.broadinstitute.org/; including 11,304 Latino individuals) (modified from Riera-Escamilla *et al.*, 2019). For genes associated with autosomal dominant inheritance or linked to X- and Y-chromosomes, we excluded variants with a frequency above 0.01. As an additional filtering step, variants identified in German controls were also excluded, because these will most likely neither cause NOA in any other population. For missense variants, we excluded those predicted to be tolerated by all of the *in silico* programs considered: SIFT, PolyPhen-2, MutationTaster, and CADD (score < 10). Additionally, we excluded single variants present in genes associated with autosomal recessive inheritance.

Confirmation of variants with Sanger sequencing

All variants remaining after filtering were confirmed by Sanger sequencing as described previously (Tewes *et al.*, 2014). The primers used for DNA amplification are shown in Table S2.

RESULTS

In our set of 37 candidate genes (Table S1), we identified 35 variants in the group of 16 Brazilian NOA patients (Table S3). Applying further selection criteria, we excluded 16 single recessive variants, three combinations of recessive variants without prediction of pathogenicity, one dominant variant with a frequency of 0.023 in the general population, and two dominant variants also found in normozoospermic men from the German controls. Along with six variants predicted as pathogenic, we considered three variants with uncertain prediction of pathogenicity (concomitant classification as benign and deleterious) and one predicted as benign in *TEX15* gene (c.7118G>A).

Hence, ten missense variants in seven candidate genes remained for subsequent in-depth assessment (Table 2).

We chose to report the heterozygous variant c.7118G>A in ***TEX15***, despite it being predicted as benign, because it was found in combination with another heterozygous variant (c.9223G>A ClinVar SCV000987204.1), predicted as deleterious (Table 2). Moreover, *TEX15* has moderate evidence of association with male infertility (Oud *et al*., 2019). Hence, we report a potentially compound heterozygous patient (M1804) (Fig. 1A) that, in contrast to previous reports, is affected by azoospermia due to SCO. This patient also presents mildly increased FSH levels (Table 1).

Two heterozygous variants in the gene ***TEX14***, c.727C>G and c.4297G>A (ClinVar SCV000987207.1), were identified in another potentially compound heterozygous patient (M1810) (Fig. 1B). This man presented germ cell maturation arrest in the testicular histology and normal hormonal parameters, comparable to patients with missense variants in *TEX14* previously reported (Fakhro *et al*., 2018).

A single variant was identified in the gene ***REC8***, c.91C>T (ClinVar SCV000987205.1), predicted to lead to an amino acid substitution p.(Arg31Cys), in an azoospermic patient with normal hormonal parameters (Patient M1809) (Fig. 1C). Despite the scarcity of studies in humans (Oud *et al*., 2019), the same variant was previously described in a patient with NOA, although then it was not considered relevant for male infertility (Griffin *et al*., 2008; Hann *et al*., 2011). In the same patient, we identified a potentially pathogenic homozygous variant (c.22C>G, ClinVar SCV000987206.1) in the gene ***SYCE1L*** (Fig. 1C) during the analysis of its paralogue *SYCE1*. The *SYCE1L* gene has no evidence of association with infertility so far (Oud *et al*., 2019), but variants in *SYCE1* have been described in azoospermic brothers presenting NOA due to meiotic arrest (Maor-Sagie *et al*., 2015).

Two single variants were found in the gene ***KHLH10***. The variant c.887T>C (ClinVar SCV000987210.1) was identified in patient M1816 (Fig. 1D), who presented SCO and increased FSH levels, while c.242A>T (ClinVar SCV000987208.1) was identified in patient M1811 (Fig. 1E), who presented NOA, increased FSH levels but who unfortunately did not have biopsy results. So far, missense and splice variants in *KHLH10* have only been described in association with oligozoospermia in humans (Yatsenko *et al*., 2006; Miyamoto *et al*., 2016) and not in the context of azoospermia or SCO as observed in this study (Patients M1816 and M1811).

Patient M1816 also carries a variant in the gene ***DMRT1*** (c.671A>G, ClinVar SCV000987211.1) (Fig. 1D). Although both variants (c.887T>C in *KLHL10*, c.671A>G in *DMRT1*) have a higher prevalence than our initial threshold of 0.01 in the general population, we included them because they were close to the limit (0.0102 and 0.0124, respectively) and in the ethnically matched Latino population even lower (not described and 0.0012, respectively). Furthermore, the variant c.671A>G was previously described in two patients with SCO (Tewes *et al*., 2014).

A single heterozygous variant, c.2452G>A (ClinVar SCV000987209.1), was identified in the gene ***DNMT3B*** in a patient with NOA and increased FSH levels (Patient M1814) (Fig. 1F). Despite limited evidence of an association with male infertility (Oud *et al*., 2019), rare missense variants within this gene have been described in patients with NOA (Li *et al*., 2015).

DISCUSSION

Regardless of the social factors hindering it, diagnosing male infertility is an enormous challenge because it is a complex, multifactorial condition resulting in heterogeneous phenotypes and involving thousands of genes (Tüttelmann *et al*., 2018). Moreover, once a deleterious mutation in a reproduction-related gene directly affects fitness, it tends to remain in the population at low frequencies, such that researchers must screen multiple genes in a large group of patients to find novel candidate genes (Oud *et al*., 2019).

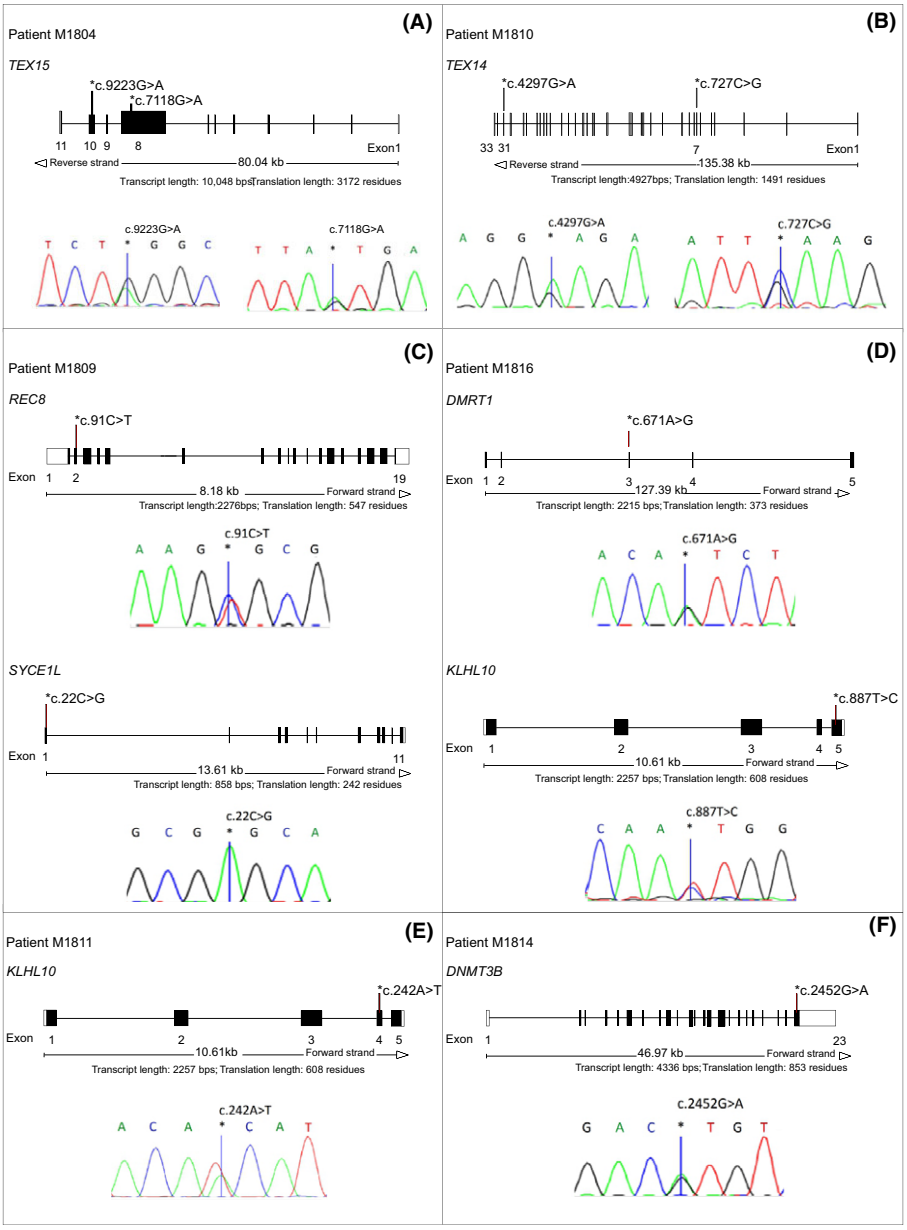
Testing multiple genes in multiple patients is now feasible as genomics has advanced and next-generation sequencing is available. Concerning idiopathic spermatogenic failure, WES has previously been successfully employed to identify novel candidate genes (Gershoni *et al*., 2017; Fakhro *et al*., 2018). However, the interpretation of genetic data is not simple, because most candidate genes have been found via mutations in single patients or single families, as is expected in the search for rare variants (Oud

Table 2 Variants identified by WES and confirmed by Sanger sequencing

Patient		Variant						*In silico* programs		Maximum allele frequency	Allele frequency
ID	Phenotype	Gene	Transcript	Inh.	cDNA	Protein	Genotype	CADD ≥ 10 damage	PolyPhen/SIFT/ MutationTaster	In any sub-population[a]	In Latino population[b]
M1804	SCO	*TEX15*	NM_001350162.1	AR	c.9223G>A	p.(Gly3075Arg)	1/0	24.9	P/D/-	0.005	0.0035
					c.7118G>A	p.(Ser2373Asn)	1/0	9.1	B/T/-	0.0257	0.0111
M1809	NOA	*REC8*	NM_001048205.1	AD	c.91C>T	p.(Arg31Cys)	1/0	35	P/-/D	0.0052	0.0010
		SYCE1L	NM_001129979.1	AR	c.22C>G	p.(Leu8Val)	1/1	6.4	P/-/-	0.0386	0.0016
M1810	Mat.A	*TEX14*	NM_001201457.1	AR	c.4297G>A	p.(Glu1433Lys)	1/0	26.9	B/D/D	0.0039	0
					c.727C>G	p.(Gln243Glu)	1/0	25.9	P/D/D	0.012	0.0005
M1811	NOA	*KLHL10*	NM_001329595.1	AD	c.242A>T	p.(Asn81Ile)	1/0	19.2	B/T/D	0.002	0.0013
M1814	NOA	*DNMT3B*	NM_006892.3	AD	c.2452G>A	p.(Val818Met)	1/0	34	P/D/-	0.0001	0
M1816	SCO	*KLHL10*	NM_001329595.1	AD	c.887T>C	p.(Ile396Thr)	1/0	23	B/T/D	0.0102[c]	–
		DMRT1	NM_021951.2	AD	c.671A>G	p.(Asn224Ser)	1/0	25.4	P/T/D	0.0124[c]	0.0012

ID, Patient Identity; Inh, inheritance pattern; SCO, Sertoli-Cell-Only histology; Mat.A, germ cell maturation arrest; AD, autosomal dominant inheritance; AR, autosomal recessive inheritance; 1/0, variant found in heterozygous status; 1/1, variant found in homozygous status; B, benign prediction; P, possibly or probably pathogenic; T, tolerated; D, deleterious or damaging; N, neutral. [a]ESP, 1000 Genomes, gnomAD. [b]gnomAD. [c]Dominant variants with frequency > 0.01 (see text for reasons why they were included).

Figure 1 Exon localization of variants identified in six NOA patients (adapted from http://www.ensembl.org).



et al., 2019). Moreover, the amount of new variants identified by NGS exceeds researchers' capacities to perform functional studies, which makes it difficult to distinguish between variants that cause disease from variants that are rare but benign (Walsh *et al.*, 2014). Thus, independent genetic studies and functional studies are necessary to provide robustness to the clinical validity of candidate genes linked to male infertility (Oud *et al.*, 2019).

The present study provides evidence for the clinical validity of some genes associated with azoospermia. Among 37 genes analyzed, we found variants in seven. We attribute our success of finding variants, even in a small number of patients, to the rational selection of candidate genes to analyze. Briefly, we considered only highly penetrant genes that exclusively affect fertility, and we selected those linked to the phenotypes presented by our patients by only considering genes with evidence of association to male infertility. To minimize the chance of finding benign variants, we considered only those with severe consequences, such as start loss, splicing, frameshift, and missense. We excluded common variants with allele frequencies higher than 0.05 in general population, and we selected those predicted as pathogenic and that were absent in the control group. We cannot confidently rule out obstructive azoospermia (OA) in the two patients with normal FSH levels, but also these presented with reduced testicular volume. We did, however, exclude mutations in the *CFTR* and *ADGRG2* genes, which are the most common causes for OA. With our approach, we identified potentially pathogenic variants in six of sixteen non-related Brazilian patients with NOA (~38%). This is the first comprehensive screening of South-America NOA men.

Two heterozygous variants, c.9223G>A(;)7118G>A, were identified in patient M1804, in ***TESTIS EXPRESSED GENE 15*** (***TEX15***, OMIM 605795). This gene, exclusively expressed in testis, is

essential for DNA double-strand break repair in germ cells, ensuring normal chromosome synapsis and meiotic recombination during spermatogenesis (Yang *et al*., 2008; Wang *et al*., 2018). In mice, mutations in this gene cause drastic reductions in testis size and meiotic arrest (Yang *et al*., 2008), whereas in humans, variants are associated with severe oligozoospermia, cryptozoospermia, and NOA due to maturation arrest (Okutman *et al*., 2015; Colombo *et al*., 2017; Wang *et al*., 2018). However, so far, no variants in *TEX15* have been reported in SCO patients (Okutman *et al*., 2015; Colombo *et al*., 2017; Wang *et al*., 2018). Although compound heterozygous patients have been reported (Colombo *et al*., 2017), most variants described in literature were nonsense and were found in homozygosis (Okutman *et al*., 2015; Wang *et al*., 2018). Of the *TEX15* variants we identified, one of the variants, c.7118G>A, was considered benign by four prediction programs and is relatively prevalent in the general population (0.0257). Moreover, as we did not have DNA samples from the parents to perform segregation analysis, we could not confirm if the variants were on different alleles. Thus, for now, the diagnostic value of both variants is uncertain. Considering the moderate evidence that *TEX15* is associated with male infertility (Oud *et al*., 2019), it would be worthwhile to perform *in vitro* functional studies to confirm these variants' pathogenicity.

Additionally, two variants, c.727C>G(;)4297G>A, were identified in another probably compound heterozygous patient (M1810) in ***TESTIS EXPRESSED GENE 14*** (***TEX14***, OMIM 605792). In this case, both variants were predicted as damaging. Studies in mice have shown that *Tex14*, exclusively expressed in the testis, is required for the formation of intercellular bridges and subsequent cell division of germ cells. In its absence, the spermatogonia do not complete the first cytokinesis (Greenbaum *et al*., 2006), which explains the germ cell maturation arrest phenotype presented by our patient. Missense variants in *TEX14* have already been described in azoospermic brothers presenting maturation arrest (Fakhro *et al*., 2018), while a homozygous splice site variant was identified in a patient with SCO (Fakhro *et al*., 2018), and frameshift variants were associated with both phenotypes (Gershoni *et al*., 2017; Fakhro *et al*., 2018). Therefore, we assume that the missense variants we found are probably causative for the infertility in patient M1810, reinforcing the evidence that mutations in *TEX14* may be responsible for many NOA cases.

In patient M1816, who presented SCO and increased FSH levels, we identified potentially damaging variants in two different genes (***DMRT1*** and ***KLHL10***), hampering our ability to define the causal factor of azoospermia. ***DOUBLESEX- AND MAB3-RELATED TRANSCRIPTION FACTOR 1*** (*DMRT1*; OMIM 602424) is a conserved sex-determination transcription factor that regulates genes in Sertoli cells and pre-meiotic germ cells during postnatal testis differentiation (Macdonald *et al*., 2018). ***KELCH HOMOLOG 10*** (*KLHL10*; OMIM 615081) is a highly evolutionarily conserved gene in mammals, exclusively expressed in the cytoplasm of spermatids. Its haplo-insufficiency causes a block in the elongation stage of spermiogenesis in mice, leading to asynchronous spermatid maturation (Yan *et al*., 2004). Deletions and missense mutations in *DMRT1* are associated with a wide spectrum of phenotypes, from XY gonadal dysgenesis to disorders of spermatogenesis such as cryptozoospermia, SCO and meiotic arrest (Tewes *et al*., 2014; Lima *et al*., 2015; Tüttelmann *et al*., 2018). The same variant we found, c.671A>G, has already been reported in two patients with SCO (Tewes *et al*., 2014), whereas missense and splicing variants in *KLHL10* have only been associated with oligozoospermia in humans (Yatsenko *et al*., 2006; Miyamoto *et al*., 2016). Hence, we assume that the variant c.671A>G in *DMRT1* is causative of SCO in patient M1816, reinforcing the association of this gene with male infertility. However, we cannot rule out the possibility that the variant c.887T>C in *KLHL10* also plays a role in the man's infertility. We report a second NOA patient carrying the potentially pathogenic missense variant c.242A>T in *KLHL10* (Patient M1811). This adds to the list of variants in *KLHL10* which may be associated with male infertility and suggests that they may also be responsible for NOA in humans, because the haplo-insufficiency of this gene disturbs the maturation of germ cells in mice.

A variant predicted as deleterious, c.2452G>A, was identified in ***DNMT3B (DNA METHYLTRANSFERASE 3 BETA)***, in patient M1814, who presented azoospermia and increased FSH levels. *Dnmt3b* encodes a *de novo* methyltransferase that establishes DNA methylation and, therefore, is classified as an epigenetic regulator of spermatogenesis in mice (Okano *et al*., 1999). *DNMT3B* is predominantly expressed in spermatogonia and spermatocytes. Recently, it was demonstrated that expression of DNMT proteins differs in spermatogenic cell types among NOA groups of patients, including those presenting hypospermatogenesis, round spermatid arrest, spermatocyte arrest, and SCO. Decreased expression of these proteins causes changes in global DNA methylation levels in spermatogenic cells, which may contribute to the development of male infertility in NOA patients (Uysal *et al*., 2019). Based on these functional studies and considering that missense variants have already been described in patients with NOA (Li *et al*., 2015), we presume that the variant we found plays a role in the azoospermia of patient M1814. Also, as ours is the second study reporting potentially pathogenic variants in *DNMT3B* in NOA patients (Oud *et al*., 2019), we offer more evidence supporting the association of this gene with male infertility.

Like *DNMT3B,* ***RECOMBINATION PROTEIN 8*** (***REC8***, GCID: GC14P024171) has limited evidence of association with male infertility (Oud *et al*., 2019). This gene encodes a specific component of the cohesin axis needed for assembly of the synaptonemal complex (SC) (Fukuda *et al*., 2014); the SC binds sister chromatids, preventing their local separation, thus assuring genetic exchange by crossing over and subsequent accurate segregation of homologous chromosome during meiosis (Agostinho *et al*., 2016; Ishiguro & Watanabe, 2016). While functional studies have reported REC8's important role during meiosis, they have failed to identify causal mutations (Hann *et al*., 2011). Coincidentally, the same variant we found in patient M1809, c.91C>T, has already been reported in a single patient with NOA (Griffin *et al*., 2008). When this variant was described, it was considered as polymorphism due to its high frequency of 0.5 in the European population (Griffin *et al*., 2008) but, according to ESP, 1000 genomes, and gnomAD, its maximum population frequency is 0.005. As it is indeed a rare variant, described for the second time in a patient with NOA, we propose this is a causal variant of male infertility.

We also identified in patient M1809 a homozygous variant, c.22C>G, in ***SYCE1L***. It is assumed that the ***SYNAPTONEMAL COMPLEX CENTRAL ELEMENT PROTEIN 1 LIKE*** **gene** (*SYCE1L*, GID: GC16P077233) has similar functions to its paralogue *SYCE1* (OMIM 611486) (https://www.uniprot.org). SYCE1 is one of the central components of the SC, which is formed by

lateral, transverse, and central elements in mammals (Gómez *et al.*, 2016). Disruption of *Syce1* also leads to infertility in mice due to failure in the SC formation and, consequently, meiosis arrest (Dunne & Davies, 2019). The variant c.22C>G found in *SYCE1L* is predicted as pathogenic, but it has a relatively high frequency in the general population (0.0386). Although mutations in this gene have not yet been reported, mutations in the paralogue, *SYCE1*, have been described in azoospermic brothers presenting meiotic arrest (Maor-Sagie *et al.*, 2015). Therefore, according to the 'paralogue annotation' principle (Walsh *et al.*, 2014), we hypothesize that mutations in *SYCE1L* can also cause azoospermia. From this perspective, we think *SYCE1L* should be considered in further investigations of male infertility. Taken together, for patient M1809, the plausible causative variant of NOA is c.91C>T, identified in the *REC8* gene, but the variant in *SYCE1L* might produce a synergistic effect, as both genes are closely related and are required for correct formation of the synaptonemal complex during meiosis.

Some limitations should be considered concerning this study. We did not examine the functional impact of identified variants but assessed our results based on published observations. Further *in vitro* and/or *in vivo* experiments would need to be carried out to evaluate the pathogenicity of the variants. Furthermore, the study cohort is relatively small. Nevertheless, we were able to identify variants in genes that were previously proposed to be associated with infertility in men.

In conclusion, we analyzed 37 candidate genes for male infertility in the first WES study of 16 unrelated NOA patients from Brazil and identified variants in seven genes. One candidate gene, *DMRT1*, is already known to play a critical role in spermatogenesis, and the variant we found, c.671A>G, has previously been described in NOA patients. Another variant, c.91AC>T in the *REC8* gene, has also been described in an NOA patient, but the literature regarding this gene is still rare. Moreover, we identified new variants in two genes, *KLHL10* and *TEX15,* with moderate evidence of their association with spermatogenic failure, and in two genes, *DNMT3B* and *TEX14,* with limited evidence. In addition, we described, for the first time, a homozygous variant in *SYCE1L*, which might be a novel candidate gene for male infertility. Even though our study included a small number of patients, the rational selection of genes allowed us to detect rare potentially pathogenic variants, providing supporting evidence to validate candidate genes associated with azoospermia.

ACKNOWLEDGMENTS

The authors thank all the patients who participated in this study and the physicians from Human Reproduction Division of Hospital das Clínicas of Ribeirão Preto (HCRP, Brazil) and the Centre of Reproductive Medicine and Andrology (Münster, Germany) who took care of them. We gratefully acknowledge Carolin Ruckert, Christina Burhöi and Nils Köckerling for the technical support, Dr. Albrecht Röpke for his scientific advice, Christian Ruckert for the great support in bioinformatic concerns and Vagner Ramon Rodrigues Silva for the assistance with the design of the figure.

FUNDING

This study was supported by CAPES (Coordenação de Aperfeiçoamento de Pessoal de nível Superior, Brazil. Process number: 88881.186971/2018-01) and was carried out within the frame of the DFG Clinical Research Unit 'Male Germ Cells: from Genes to Function' (CRU326 to FT).

CONFLICT OF INTEREST

The authors declared no conflict of interest.

AUTHORS' CONTRIBUTIONS

TFA designed and performed the research, analyzed data, and wrote the manuscript. CF designed and supported the research study and contributed to writing the manuscript. CHPG and LRM selected the azoospermic patients and performed the infertility protocol that includes physical examination, blood collection, cytogenetic analysis, AZF exclusion, and further clinical assessment. JDG performed DNA extractions for part of the patients and the multiplex PCR for AZF deletion. JE and MJW contributed with data analysis and lab work. SK selected the normozoospermic patients and collected their blood samples and clinical information. ALS assisted in the design of the study and correction of the manuscript. FT supervised the study, provided financial support, and essential reagents. All authors revised and approved the final version of the manuscript.

REFERENCES

Agostinho A, Manneberg O, von Schendel R, Hernández-Hernández A, Kousznetsova A, Blom H, Brismar H & Höög C. (2016) High density of REC8 constrains sister-chromatid axes and prevents illegitimate synaptonemal complex formation. *EMBO Rep* 17, 901–913.

Barratt CLR, Björndahl L, De Jonge CJ, Lamb DJ, Martini FO, McLachlan R, Oates RD, van der Poel S, St John B, Sigman M, Sokol R & Tournaye H. (2017) The diagnosis of male infertility: an analysis of the evidence to support the development of global WHO guidance – challenges and future research opportunities. *Hum Reprod Update* 23, 660–680.

Belva F, Bonduelle M, Roelants M, Michielsen D, Van Steirteghem A, Verheyen G & Tournaye H. (2016) Semen quality of young adult ICSI offspring: the first results. *Hum Reprod* 31, 2811–2820.

Cocuzza M, Alvarenga C & Pagani R. (2013) The epidemiology and etiology of azoospermia. *Clinics* 68(Suppl 1), 15–26.

Colombo R, Pontoglio A & Bini M. (2017) Two novel TEX15 mutations in a family with nonobstructive azoospermia. *Gynecol Obstet Investigation* 82(3), 283–286.

Dunne OM & Davies OR. (2019) Molecular structure of human synaptonemal complex protein SYCE1. *Chromosoma* 00, 1–14.

Fakhro KA, Elbardisi H, Arafa M, Robay A, Rodriguez-Flores JL, Al-Shakaki A, Syed N, Mezey JG, Abi Khalil C, Malek JA, Al-Ansari A, Al Said S & Crystal RG. (2018) Point-of-care whole-exome sequencing of idiopathic male infertility. *Genet Med* 20, 1365–1373.

Fukuda T, Fukuda N, Agostinho A, Hernández-Hernández A, Kouznetsova A & Höög C. (2014) STAG3-mediated stabilization of REC8 cohesin complexes promotes chromosome synapsis during meiosis. *EMBOjounal* 33, 1243–1256.

Gershoni M, Hauser R, Yogev L, Lehavi O, Azem F, Yavetz H, Pietrokovski S & Kleiman SE. (2017) Original research article A familial study of azoospermic men identifies three novel causative mutations in three new human azoospermia genes Original research article. *Genet Med* 19, 998–1006.

Gómez HL, Felipe-Medina N, Sánchez-Martín M, Davies O, Ramos I, García-Tunón I, de Rooij DG, Dereli I, Tóth A, Barbero JL, Benavente R, Llano E & Pendas AM. (2016) C14ORF39/SIX6OS1 is a constituent of the synaptonemal complex and is essential for mouse fertility. *Nat Commun* 7, 13298.

Grangeiro CHP. (2018) Genomic assessment of idiopathic male infertility by nonobstructive azoospermia, PhD Thesis. Department of Genetics, Medical Faculty of Ribeirão Preto, University of São Paulo.

Greenbaum MP, Yan W, Wu M, Lin Y, Agno JE, Sharma M, Braun RE, Rajkovic A & Matzuk MM. (2006) TEX14 is essential for intercellular bridges and fertility in male mice. *PNAS* 103, 4982–4987.

Griffin J, Emery BR, Christensen GL & Carrell DT. (2008) Analysis of the meiotic recombination gene REC8 for sequence variations in a population with severe male factor infertility. *Syst Biol Reprod Med* 54, 163–165.

Hann MC, Lau PE & Tempest HG. (2011) Meiotic recombination and male infertility: from basic science to clinical reality? *Asian J Androl* 13, 212–218.

Hook EB. (1977) Exclusion of chromosomal mosaicism: tables of 90%, 95%, and 99% confidence limits and comments on use. *Am J Num Genet* 29, 94–97.

Ishiguro K & Watanabe Y. (2016) The cohesin REC 8 prevents illegitimate inter-sister synaptonemal complex assembly. *EMBO Rep* 17, 783–784.

Krausz C & Riera-Escamilla A. (2018) Genetics of male infertility. *Nat Rev Urol* 15, 369–384.

Krausz C, Giachini C, Lo Giacco D, Daguin F, Chianese C, Ars E, Ruiz-Castane E, Forti G & Rossi E. (2012) High resolution X chromosome-specific array-CGH detects new CNVs in infertile males. *PLoS ONE* 7, e44887.

Krausz C, Hoefsloot L, Simoni M & Tüttelmann F. (2013) EAA/EMQN best practice guidelines for molecular diagnosis of Y-chromosomal microdeletions: state-of-the-art 2013. *Andrology* 2, 5–19.

Krausz C, Escamilla AR & Chianese C. (2015) Genetics of male infertility: from research to clinic. *Reproduction* 150, 159–174.

Kumari A, Yadav SK, Misro MM, Ahmad J & Ali S. (2015) Copy number variation and microdeletions of the Y chromosome linked genes and loci across different categories of Indian infertile males. *Sci Rep* 5, 17780.

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

Li Z, Yi H, Li H, Hu J, Liu X, Jiang T, *et al.* (2015) Excess of rare variants in genes that are key epigenetic regulators of spermatogenesis in the patients with non-obstructive azoospermia. *Sci Rep* 5, 8785.

Lima A, Carvalho F, Gonçalves J, Fernandes S, Marques P, Sousa M, Barros A, Seixas S, Amorim A, Conrad DF & Lopes AM. (2015) Rare double sex and mab-3-related transcription factor 1 regulatory variants in severe spermatogenic failure. *Andrology* 3, 825–833.

Macdonald J, Kilcoyne KR, Sharpe RM, Kavanagh Á, Anderson RA, Brown P, Smith LB, Jørgensen A & Mitchell RT. (2018) DMRT1 repression using a novel approach to genetic manipulation induces testicular dysgenesis in human fetal gonads. *Hum Reprod* 33, 2107–2121.

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

Martin M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 17, 10–12.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M & DePristo MA. (2010) The genome analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20, 1297–1303.

McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P & Cunningham F. (2016) The Ensembl variant effect predictor. *Genome Biol* 17, 122.

Miyamoto T, Minase G, Shin T, Ueda H, Okada H & Sengoku K. (2016) Human male infertility and its genetic causes. *Reprod Med Biol* 16, 81–88.

Nakamura S, Miyado M, Saito K, Katsumi M, Nakamura A, Kobori Y, Tanaka Y, Ishikawa H, Yoshida A, Okada H, Hata K, Nakabayashi K, Okamura K, Ogata H, Matsubara Y, Ogata T, Nakai H & Fukami M. (2017) Next-generation sequencing for patients with non-obstructive azoospermia: implications for significant roles of monogenic/oligogenic mutations. *Andrology* 5, 824–831.

Okano M, Bell DW, Haber DA & Li E. (1999) DNA Methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247–257.

Okutman O, Muller J, Baert Y, Serdarogullari M, Gultomruk M, Piton A, Rombaut C, Benkhalifa M, Teletin M, Skory V, Bakircioglu E, Goossens E, Bahceci M & Viville S. (2015) Exome sequencing reveals a nonsense mutation in TEX15 causing spermatogenic failure in a Turkish family. *Hum Mol Genet* 24, 5581–5588.

Oud MS, Volozonoka L, Smits RM, Vissers LELM, Ramos L & Veltman JA. (2019) A systematic review and standardized clinical validity assessment of male infertility genes. *Hum Reprod* 34, 932–941.

Riera-Escamilla A, Enguita-Marruedo A, Moreno-Mendoza D, Chianese C, Sleddens-Linkels E, Contini E, Benelli M, Natali A, Colpi GM, Ruiz-Castañé E, Maggi M, Baarends WM & Krausz C. (2019) Sequencing of a 'mouse azoospermia' gene panel in azoospermic men: identification of RNF212 and STAG3 mutations as novel genetic causes of meiotic arrest. *Hum Reprod* 34, 978–988.

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

Tewes AC, Ledig S, Tüttelmann F, Kliesch S & Wieacker P. (2014) DMRT1 mutations are rarely associated with male infertility. *Fertil Steril* 102, 816–820.

Tüttelmann F, Simoni M, Kliesch S, Ledig S, Dworniczak B, Wieacker P & Röpke A. (2011) Copy number variants in patients with severe oligozoospermia and sertoli-cell-only syndrome. *PLoS ONE* 6, e19426.

Tüttelmann F, Ruckert C & Röpke A. (2018) Disorders of spermatogenesis. *Med Genet* 30, 12–20.

Uysal F, Akkoyunlu G & Ozturk S. (2019) Decreased expression of DNA methyltransferases in the testes of patients with non-obstructive azoospermia leads to changes in global DNA methylation levels. *Reprod Fertil Dev* 31, 1386–1394.

Walsh R, Peters NS, Cook SA & Ware JS. (2014) Paralogue annotation identifies novel pathogenic variants in patients with Brugada syndrome and catecholaminergic polymorphic ventricular tachycardia. *J Med Genet* 51, 35–44.

Wang X, Jin HR, Cui YQ, Chen J, Sha YW & Gao ZL. (2018) Case study of a patient with cryptozoospermia associated with a recessive TEX15 nonsense mutation. *Asian J Androl* 20, 101–102.

World Health Organization. (2010) *WHO Laboratory Manual for Examination and Processing of Human Semen*, 5th edn. WHO Press, Geneva.

Yan W, Ma L, Burns KH & Matzuk MM. (2004) Haploinsufficiency of kelch-like protein homolog 10 causes infertility in male mice. *Proc Natl Acad Sci USA* 101, 7793–7798.

Yang F, Eckardt S, Leu NA, Mclaughlin KJ & Wang PJ. (2008) Mouse TEX15 is essential for DNA double-strand break repair and chromosomal synapsis during male meiosis. *J Cell Biol* 180, 673–679.

Yatsenko AN, Roy A, Chen R, Ma L, Murthy LJ, Yan W, Lamb DJ & Matzuk MM. (2006) Non-invasive genetic diagnosis of male infertility using spermatozoal RNA: KLHL10 mutations in oligozoospermic patients impair homodimerization. *Hum Mol Genet* 15, 3411–3419.

Zarrabi AD & Kruger TF. (2018) The challenges of supporting male infertility treatment in South Africa. *Nat Rev Urol* 15, 719–720.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Set of candidate genes analyzed in this study.
Table S2. List of primer sequences used in Sanger sequencing for validation of variants found by WES.
Table S3. List of 35 variants identified in 37 candidate genes.