Asian Journal of Andrology (2022) 24, 607–614
www.asiaandro.com; www.ajandrology.com

Open Access

ORIGINAL ARTICLE

Male Infertility

DEFB126 polymorphisms and association with idiopathic asthenozoospermia in China

Jiao-Yu He[1,2], Jian-Ying Peng[1,2], Qiu-Fu Li[1,2], Xiao-Li Lin[1,2], Yan-Ru Cui[1,2], Shi-Yu Ma[1,2], Shi-Yun Fan[1,2], Yi-Ran Liu[1,2], Zhi-Lin Song[1,2], Jun-Hang Deng[1,2], Xia Wei[1,2], Xian-Ping Ding[1,2]

Idiopathic asthenozoospermia, a common factor in male infertility, is characterized by altered sperm motility function in fresh ejaculate. Although the β-defensin 126 (DEFB126) protein is associated with asthenozoospermia, *DEFB126* gene polymorphisms have not been extensively studied. Therefore, the association between *DEFB126* gene polymorphisms and asthenozoospermia requires further investigation. Screening was performed by semen analysis, karyotype analysis, and Y microdeletion detection, and 102 fertile men and 106 men with asthenozoospermia in Chengdu, China, were selected for *DEFB126* gene sequence analyses. Seven nucleotide mutations and two nucleotide deletions in the *DEFB126* gene were detected. rs11467417 (317–318 *del/del*), rs11467497 (163–166 *wt/del*), c.152T>C, and c.227A>G were significantly different between the control and asthenozoospermia groups, likely representing high-risk genetic factors for asthenozoospermia among males. DEFB126 expression was not observed in sperm with rs11467497 homozygous deletion and was unstable in sperm with rs11467417 homozygous deletion. The rs11467497 four-nucleotide deletion leads to truncation of DEFB126 at the carboxy-terminus, and the rs11467417 binucleotide deletion produces a non-stop messenger RNA (mRNA). The above deletions may be responsible for male hypofertility and infertility by reducing DEFB126 affinity to sperm surfaces. Based on *in silico* analysis, the amino acids 51M and 76K are located in the highly conserved domain; c.152T>C (M51T) and c.227A>G (K76R) are predicted to be damaging and capable of changing alternative splice, structural and posttranslational modification sites of the RNA, as well as the secondary structure, structural stability, and hydrophobicity of the protein, suggesting that these mutations are associated with asthenozoospermia.

Asian Journal of Andrology (2022) **24**, 607–614; doi: 10.4103/aja2021115; published online: 05 April 2022

Keywords: β-defensin 126; idiopathic asthenozoospermia; *in silico* analysis; single-nucleotide polymorphisms

INTRODUCTION

Approximately 10%–15% of couples worldwide are affected by infertility, of which male factors account for nearly half.[1,2] Male infertility is a multifactorial medical syndrome influenced by intricate genetics, age, cryptorchidism, nutrition, weight, psychological stress, smoking, drugs, alcohol, and environmental exposures.[3,4] Chromosomal abnormalities, Y chromosome microdeletions, and gene polymorphisms are intricate genetic changes responsible for 15%–30% of male infertility.[5,6] Idiopathic asthenozoospermia is a common cause of male infertility and is characterized by reduced sperm motility (grade A + B sperm motility <50% or A <25%); it has been reported to be associated with genetic mutations or deletions.[7,8] A broad definition of asthenozoospermia is altered sperm function through motility.[9]

Some candidate genes associated with asthenozoospermia, such as *tektin-t*,[2] methylenetetrahydrofolate reductase (*MTHFR*),[10] solute carrier family 26 (*SLC26A8*),[11] N-acetyl-galactosaminyltransferase-like protein 5 (*GALNTL5*),[12] and cation channel, sperm-associated 2 (*CATSPER2*),[11] have been identified, though most causes and mechanisms remain unknown. Recently, more than 6000 different proteins associated with male infertility were identified by proteomics.[13] Human β-defensin 126 (DEFB126, formerly named epididymis-specific protein ESP13.2 in macaques) is a unique sperm surface-coating protein that belongs to the beta-defensin family and the innate immune system; until capacitation, the entire sperm surface is covered by the protein.[14,15] DEFB126 is synthesized by the principal cells of the epididymal epithelium and is maximally expressed in the distal corpus and proximal cauda.[14] Macaque DEFB126 has been reported to play a role in reproductive function by participating in sperm penetration through the cervical mucus and sperm-zona pellucida (sperm-ZP) recognition and binding.[16,17] DEFB126 is a multifunctional glycoprotein with a conserved β-defensin core motif, which contains six β-defensin cysteine residues with critical lipopolysaccharide (LPS) binding activity, anti-inflammatory effects, and an intracellular regulatory unique C-terminal glycosylated peptide tail with twenty O-glycosylation sites (serine and threonine residues)[18] that bind oligosaccharides to form a glycocalyx.[19,20] A long glycosylated peptide tail has been proven to be the key to effective sperm transport in the female reproductive tract; it interferes with immune recognition and promotes sperm penetration into the cervical mucus, conserving sperm resources and promoting fertilization. Coating the sperm surface with other protein components renders it essentially unrecognizable by the immune system.[14,17] Sperm lacking DEFB126 exhibit frequently interrupted movement

[1]Key Laboratory of Bio-Resources and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610065, China; [2]Chongqing Nanchuan Biotechnology Research Institute, Bio-resource Research and Utilization Joint Key Laboratory of Sichuan and Chongqing, Chongqing 400000, China.
Correspondence: Dr. XP Ding (brainding@scu.edu.cn)
Received: 26 March 2021; Accepted: 23 December 2021

and tortuous swimming paths when penetrating the cervical mucus, which is completely reversed when soluble DEFB126 is added to the sperm surface.[16] Moreover, the long glycosylated peptide tail mediates sperm attachment to oviduct epithelia, conserves resources until arrival at the oocyte,[16] and facilitates the delivery of a capacitated sperm to the site of fertilization.

Approximately 20% of Asian and European males of reproductive age carry a 2-nucleotide deletion in the open reading frame (ORF) of *DEFB126* (rs140685149), thus producing an uninterrupted mRNA;[21] rs11467497 is another 4-nucleotide frameshift mutant that leads to premature termination of translation.[22] For men with rs140685149 and rs11467497 mutations, a longer time is needed for impregnation. Therefore, *DEFB126* mutations are more likely to be associated with impaired sperm function and subfertility.[23] Considering the important biological function of DEFB126 for male infertility, *DEFB126* polymorphisms and their association with idiopathic asthenozoospermia in China urgently need to be studied. In future clinical applications, *DEFB126* polymorphisms may even serve as a target for detecting male infertility and may facilitate the development of new clinical treatment protocols.

PATIENTS AND METHODS

Patients with idiopathic asthenozoospermia and controls

In this study, 119 patients with idiopathic asthenozoospermia and 102 control samples were collected from the Affiliate Reproductive Hospital Genitalia Hygiene Research Center (Chengdu, China) between April 2017 and September 2018; semen and peripheral samples were collected for later experiments. Patients with abnormal karyotype, Y chromosome microdeletion, cryptorchidism, leukocytospermia, orchitis, varicocele, or endocrine hypogonadism were excluded (**Supplementary Figure 1**). Patients who reported drug abuse, alcoholic abuse, substance abuse, or tobacco use were also excluded. The controls were all men with normal sperm function who had at least one child without any assisted reproduction technology. Informed consent was obtained from all participants according to the relevant regulations, and this study has been ethically approved by the Education and Research Committee and Ethics Committee of Sichuan University (Chengdu, China; approval No. CYK20170311001).

Semen and karyotype analysis

Semen samples were obtained by masturbation into sterile containers after a period of 3–5 days of abstinence. The sperm concentration, total motility, viability, and round cell concentration were evaluated according to the WHO fourth edition laboratory manual. Samples with the following semen parameters were included: sperm concentration more than 20×10^6 ml^{-1}; rapid forward progressive motility sperm grade A <25%; and total progressive motile sperm grade A + B <50% in fresh ejaculate.

Standard G-banding was used for karyotyping analysis using peripheral blood samples from all subjects. At least 50 metaphases were analyzed for each patient to confirm any abnormalities; such samples were excluded from this study.

Genomic DNA extraction

A Whole Genome Extraction Kit (Transgen, Beijing, China) was used to extract total DNA of human peripheral blood. Briefly, peripheral blood was resuspended and mixed with lysis solution (containing 10 μl RNaseA, 500 μl binding buffer, and 20 μl proteinase K), which was incubated at room temperature for 10 min. Then, lysates were added to a centrifugal column for DNA binding, and the bound DNA was washed with 500 μl clean buffer and then twice with wash buffer. Subsequently, 100 μl of preheated elution buffer (EB; Transgen) was added, the solution was eluted using a centrifugal column, and the extracted genomic DNA was quantitatively analyzed by spectrophotometry (Bio-Rad, Shanghai, China) and stored at −20°C until examination.

Y microdeletions detection

Multiplex polymerase chain reaction was used to detect microdeletions in the azoospermia factor (AZF) region in the patients with asthenozoospermia. Seven sequence-tagged site markers in the AZF region were selected to detect microdeletions, including *sY86* and ubiquitin specific peptidase 9, Y-linked (*USP9Y*) in AZFa, *sY127* and *sY134* in AZFb, *sY254*, and *sY255* in AZFc.[24] *sY14* (sex-determining region of Y-chromosome [*SRY*]) was detected as a polymerase chain reaction (PCR) internal control. The primers used are shown in **Supplementary Table 1**. The PCR products were verified by agarose gel electrophoresis.

PCR amplification and genotyping

Primers were designed according to the *DEFB126* reference sequence. Two exons and flanking intronic parts of the *DEFB126* gene were amplified by PCR using an A200 Gradient Thermal Cycler (Long Gene, Hangzhou, China); the primers used are shown in **Supplementary Table 2**. The total PCR volume was 25 μl, including 0.25 mmol l^{-1} deoxy-ribonucleoside triphosphate (dNTPs), 1.5 mmol l^{-1} Mg^{2+}, 2 U DNA polymerase, 0.1 μmol l^{-1} of each primer, and 200 ng genomic DNA. The conditions were as follows: 95°C (5 min, initial denaturation), 35 cycles of 94°C (1 min, denaturation), 59°C/61°C (1 min, annealing), 72°C (1 min, extension), and 72°C (5 min, final extension). The PCR products were analyzed by electrophoresis at 120 V for 40 min at room temperature. Subsequently, the PCR products were sequenced using an ABI3730XL DNA Analyzer (Applied Biosystems, Foster City, CA, USA).

Expression analysis of the wild-type and deletion mutant DEFB126 proteins

Liquefied semen was centrifuged to obtain sperm cells, which were treated with 10× lysis buffer (2% sodium dodecyl sulfate [SDS]), 100 mmol l^{-1} Tris/HCl, pH 7.6 (Beyotime, Shanghai, China) at 95°C for 5 min and centrifuged at 16 000*g* for 5 min (Hitachi, Beijing, China). The supernatant was used as the total cell protein extract. The protein concentration was determined by the bicinchoninic acid (BCA) method, and Western blot was performed to analyze the DEFB126 protein in patients and controls (antibody incubation with rabbit anti-DEFB126 IgG; Abcam, Shanghai, China) and sleep antirabbit horseradish peroxidase (HRP) IgG (Abcam).

Bioinformatic analysis

Analysis of DEFB126 protein conservation in eight different species (*Homo sapiens*, *Pan troglodytes*, *Macaca fascicularis*, *Equus caballus*, *Theropithecus gelada*, *Pteropus alecto*, *Piliocolobus tephrosceles*, and *Rhinolophus sinicus*) was performed, as displayed by sequence logo.[25] The sequences were obtained from the National Center for Biotechnology Information (NCBI).

The damaging effect of nonsynonymous mutations on protein structure and function was predicted by Polymorphism Phenotyping version 2 (PolyPhen-2),[26] Sorting Tolerant From Intolerant (SIFT),[27] Mutation Taster (https://www.mutationtaster.org/),[28] FATHMM-XF (http://fathmm.biocompute.org.uk/fathmm-xf/),[29] SNPs&GO (http://snps.biofold.org/snps-and-go),[30] screening for nonacceptable

polymorphisms (SNAP[2]; http//www.rostlab.org/services/SNAP),[31] Protein Variation Effect Analyzer (PROVEAN; http://provean.jcvi.org),[32] MutPred2 (http://mutpred.mutdb.org/),[33] and Protein ANalysis THrough Evolutionary Relationships (PANTHER; http://panther.celera.com).[34]

The effect of nonsynonymous mutations on RNA structural stability and RNA splicing pattern was evaluated by RNAsnp software (http://rth.dk/resources/rnasnp/)[35] and Alternative Splice Site Predictor (ASSP) Online software (http://wangcomputing.com/assp/), respectively.[36]

The effect of nonsynonymous SNPs (nsSNPs) on posttranslational modification of the human DEFB126 protein was analyzed by the Mod Pred tool (http://www.modpred.org).[37] The secondary structure and stability of wild-type and mutant DEFB126 were predicted by the self-optimized prediction method server (SOPMA; https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsasopma.html)[38] and I-Mutant 3.0 (http://folding.biofold.org/i-mutant/i-mutant2.0.html),[39] respectively. The three-dimensional (3D) structure of the DEFB126 protein was examined using Protein Model Portal (http://www.proteinmodelportal.org).[40] The effect of non-synonymous mutations on DEFB126 protein hydrophobicity and ligand binding sites was evaluated by the ExPASy-ProtScale (https://web.expasy.org/protscale/)[41] and FTsite tools (http://ftsite.bu.edu),[42] respectively.

Statistical analyses

The Chi-squared test was used to identify Hardy–Weinberg equilibrium and estimate genotype frequencies between the patient and control groups. Unconditional logistic regression analysis was used to calculate odds ratios (ORs), and 95% confidence intervals (CIs) were employed to measure the risk associated with variant genotypes. All P values were two-tailed, and $P < 0.05$ was considered statistically significant. These analyses were carried out with Statistical Package for the Social Sciences software version 20.0 (SPSS Inc., Chicago, IL, USA).

RESULTS

After comprehensive screening by semen analysis, karyotype analysis, and detection of Y microdeletions (**Supplementary Table 3**), 106 patients with idiopathic asthenozoospermia and 102 controls were selected. The descriptive comparison of our study population is shown in **Supplementary Table 4**.

Polymorphism analysis of human DEFB126

The two exons and flanking intronic sequences of *DEFB126* in asthenozoospermia and control samples were sequenced, and seven nucleotide mutations, c.152T>C (M51T, rs75442118), c.164A>G (Q55R, rs200901241), c.193C>T (R65C, rs375427290), c.216T>C (rs77783981), c.227A>G (K76R, rs74717966), c.288T>C (rs77106211), and c.308C>T (A103V, rs149330491) and two nucleotide deletions 317–318 *del/del* (rs11467417) and 163–166 *wt/del* (rs11467497), were detected (**Supplementary Table 5**). However, c.164A>G (Q55R), c.193C>T (R65C), and c.308C>T (A103V) only occurred in the patients with asthenozoospermia. After screening the 102 controls and analyzing an online database, we confirmed that the frequencies of genotypes c.152T>C (M51T), c.227A>G (K76R), 163–166 *wt/del* (rs11467497), and 317–318 *del/del* (rs11467417) showed significant differences between the asthenozoospermia and fertile control groups (**Table 1**, **Supplementary Figure 2**, and **Supplementary Table 5**).

Expression of the wild-type and deletion mutant DEFB126 proteins

In general, homozygous deletion mutation of *DEFB126* alters protein expression in normal sperm. For example, DEFB126 protein expression was significantly different between sperm with the wild-type sequence and sperm with the rs11467497 heterozygote deletion ($P < 0.01$). DEFB126 protein expression was not observed in sperm with the rs11467497 homozygous deletion. No significant difference in DEFB126 protein expression was observed between sperm with the wild-type and rs11467417 heterozygote deletion. Unstable DEFB126 protein expression was observed in sperm with rs11467417 homozygous deletion (**Supplementary Figure 3** and **4**).

Conservation and pathogenicity analysis of the DEFB126 protein sequence

Analysis of DEFB126 protein conservation in eight species (*Homo sapiens, Pan troglodytes, Macaca fascicularis, Equus caballus, Theropithecus gelada, Pteropus alecto, Piliocolobus tephrosceles,* and *Rhinolophus sinicus*) was performed, and the results indicated that amino acids 51M and 76K are located in a highly conserved domain of the DEFB126 protein (**Supplementary Figure 5**).

Computational analysis using online variant pathogenicity prediction tools (PolyPhen-2, SIFT, Mutation Taster, FATHMM-XF, SNPs&GO, SNAP[2], PROVEAN, MutPred2 score, PANTHER) suggested that c.T152C and c.A227G are most likely disease-causing mutations (**Table 2**).

Effects of c.152T>C and c.227A>G on alternative splice sites

c.152T>C (M51T) and c.227A>G (K76R) are located in exon 2. Neural Network Splice Site Prediction Tool was used to evaluate the strength of altered splice sites, and the differences were observed in IntronGC, Alt./Cryptic, Constitutive, and Confidence functions. Additionally, c.227A>G is located in the receptor sequence (mark with asterisk), with a significant effect on selective splicing. These data indicate enhancement of the exon 2 receptor site caused by c.227A>G (K76R) (**Supplementary Table 6**–**8**).

Effect of c.152T>C and c.227A>G on RNA structure and posttranslational modification sites

The data from the RNAsnp web server showed that c.152T>C (M51T) and c.227A>G (K76R) have a significant effect on the RNA structure (c.152T>C [M51T]: distance = 0.0234, $P = 0.6586$; c.227A>G [K76R]: distance = 0.0383, $P = 0.5521$, respectively). c.152T>C (M51T) changes the minimum free energy of the RNA from −69.50 kcal mol^{-1} to −69.80 kcal mol^{-1}; the T/C transition results in an RNA secondary structure change between nucleotide 12 and 159 region. c.A227G (K76R) changes the minimum free energy of the RNA from −60.90 kcal mol^{-1} to −64.80 kcal mol^{-1}, though the structure is similar (**Figure 1**).

The ModPred tool was used to analyze the effect of nsSNPs on the posttranslational modification process of the human DEFB126 protein, showing that K76R is involved in proteolytic cleavage. However, c.T152C has no influence on posttranslational modification of DEFB126 (**Table 3**).

Effects of c.152T>C (M51T) and c.227A>G (K76R) on protein secondary structure, structural stability, and hydrophilicity

Based on secondary structure prediction of wild-type and mutant DEFB126, M51T likely produces a shortened random coil and an increase in the adjacent alpha helix, beta turn, which may modify the 3D structure (**Supplementary Figure 6** and **7**). K76R has no influence on the secondary structure of the protein (**Supplementary Figure 6**).

The influence of M51T and K76R on protein stability, in terms of free energy, was predicted by I-Mutant 3.0, and the results indicated that although M51T and K76R cause instability, M51T is most damaging due to its lowest free energy of −0.58 kcal mol^{-1} (**Table 3**).

Table 1: Genotype and allele frequencies for DEFB126 c.152T>C, c.227A>G, rs11467417 and rs11467497 among idiopathic asthenozoospermia and control group

Genotype/allele	*Control (total=102), n (%)*	*Asthenozoospemia (total=106), n (%)*	*OR (95% CI)*	*P*
152TT	95 (93.2)	85 (80.2)	NS	
152TC	7 (6.9)	21 (19.8)	3.353 (1.358–8.281)	0.006**
152CC	0 (0)	0 (0)	NS	
152T	197 (96.6)	191 (90.1)	NS	
152C	7 (3.4)	21 (9.9)	3.094 (1.286–7.447)	0.008**
152 (TC+CC)	7 (6.9)	21 (19.9)	3.353 (1.358–8.281)	0.006**
227AA	94 (92.2)	86 (81.1)	NS	
227AG	8 (7.8)	20 (18.9)	2.733 (1.114–6.526)	0.020*
227GG	0 (0)	0 (0)	NS	
227A	196 (96.1)	192 (90.6)	NS	
227G	8 (3.9)	20 (9.4)	2.552 (1.098–5.934)	0.025*
227 (AG+GG)	8 (7.8)	20 (18.9)	2.733 (1.114 6.526)	0.020*
163–166 *wt*/*wt* (CAAA/CAAA)	93 (91.2)	86 (81.1)	NS	
163–166 *wt*/*del*	8 (7.8)	19 (17.9)	2.568 (1.069–6.171)	0.031*
163–166 *del*/*del*	1 (1.0)	1 (0.9)	1.081 (0.067–17.558)	0.956
163–166 *wt*	194 (95.1)	191 (90.1)	NS	
163–166 *del*	10 (4.9)	21 (9.9)	2.133 (0.979–4.649)	0.052
163–166 (*wt*/*del*+*del*/*del*)	9 (8.8)	20 (18.9)	2.403 (1.038–5.564)	0.037*
317–318 *wt*/*wt* (CC/CC)	46 (45.1)	35 (33.0)	NS	
317–318 *wt*/*del*	41 (40.2)	45 (42.5)	1.443 (0.784–2.655)	0.239
317–318 *del*/*del*	15 (14.7)	26 (24.5)	2.278 (1.052–4.934)	0.035*
317–318 *wt*	133 (65.2)	115 (54.3)	NS	
317–318 *del*	71 (34.8)	97 (45.8)	1.580 (1.064–2.345)	0.023*
317–318 (*wt*/*del*+*del*/*del*)	56 (54.9)	71 (67.0)	1.666 (0.950–2.923)	0.074

*P<0.05 and **P<0.01. DEFB126: β-defensin 126; CI: confidence interval; NS: no significance; OR: odds ratio

Table 2: Bioinformatic analysis of the DEFB126 variant

Variation	*Amino acid variation*	*dbSNP ID*	*PolyPhen-2*[a]	*SIFT*[b]	*Mutation Taster*[c]	*FATHMM-XF*[d]	*SNPs&GO*[e]	*SNAP2*[f]	*PROVEAN*[g]	*MutPred2 score*[h]	*PANTHER*[i]
c.T152C	p.M51T	rs75442118	Benign	Tolerated	Polymorphism	Tolerated	Neutral	Effect	Deleterious	0.077	Probably benign
c.A227G	p.K76R	rs74717966	Possibly-damaging	Tolerated	Polymorphism	Tolerated	Neutral	Effect	Neutral	0.044	Probably benign

[a]PolyPhen-2: Polymorphism Phenotyping version 2, prediction scores range from 0 to 1 with high scores indicating probably or possibly damaging. [b]SIFT: scores vary between 0 and 1, variants with scores close or equal to 0 are predicted to be damaging. [c]Mutation taster: the probability value is the probability of the prediction, a value close to 1 indicates a high "security" of the prediction. [d]FATHMM-XF: values above 0.5 are predicted to be deleterious, while those below 0.5 are predicted to be neutral or benign. [e]SNPs&GO: probability, disease probability (if >0.5, mutation is predicted disease). [f]SNAP2: predict a score (ranges from −100 strong neutral prediction to +100 strong effect prediction) that reflects the likelihood of this specific mutation to alter the native protein function. [g]PROVEAN: screening for nonacceptable polymorphisms, variants with scores lower than −2.5 (cut off) are predicted to be deleterious. [h]MutPred2 score: >0.5 could be considered as "harmful", and a score >0.75 should be treated as a high confidence "harmful" prediction. [i]PANTHER: Protein ANalysis THrough Evolutionary Relationships. SNPs: single nucleotide polymorphism; dbSNP: dbSNP: the Single Nucleotide Polymorphism Database; SIFT: sorting Tolerant From Intolerant; SNAP2: screening for nonacceptable polymorphisms 2; PROVEAN: Protein Variation Effect Analyzer; GO: Gene Ontology

Table 3: Effect of M51T and K76R on posttranslational modification and protein stability

dbSNP ID	*Mutation site*	*ModPred analysis*		*I-Mutant 2.0*		
		Modified type	*Score*	*Stability*	*RI*	*ΔΔG (kcal mol^{-1})*
rs74717966	p.K76R	Proteolytic cleavage	0.52	Decrease	4	-0.20
rs75442118	p.M51T	None	-	Decrease	7	-0.58

dbSNP: the Single Nucleotide Polymorphism Database; RI: reliability index; -: no reliable value

The effect of M51T and K76R on the protein hydropathy plot was predicted by ExPASy-ProtScale, demonstrating an imbalance in hydrophobicity caused by M51T and K76R, which may result in protein structural modification (**Figure 2**).

Ligand binding site prediction by FTsite

The FTsite algorithm identifies binding sites using apo structures from two established test sets. The FTsite tool identified three different ligand binding sites in DEFB126. Ligand binding site 1 consists of 11 residues, site 2 of 8 residues, and site 3 of 9 residues. However, no mutations in those ligand binding sites were detected in this study (**Supplementary Table 9**).

DISCUSSION

Defensin was originally thought to help the reproductive system defend against invading pathogens, but it has been shown to be associated with specific functions of sperm priming, motility, and capacitation. β-defensin is referred to as an innate effector and antimicrobial peptide (AMP), and it is involved in immune regulation, cancer, wound healing, cell migration, angiogenesis, and male reproductive function. Various β-defensins with regional-specific expression have been found in the human and mouse epididymis, suggesting that these secretory defense peptides may contribute to sperm maturation in different epididymis microenvironment segments. Many β-defensins are associated with sperm motility and fertility, such as β-defensin 22,[18] β-defensin 15,[43] Bin1b,[44] Defb15,[43] DEFB1,[45] β-defensin 1,[45] β-defensin 114,[46] and DEFB126.[18] The β-defensin DEFB126 is a multifunctional glycoprotein

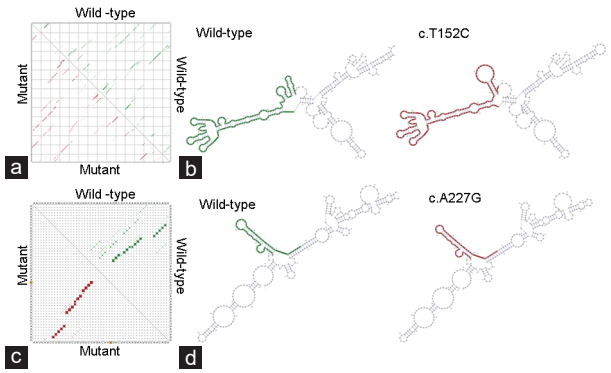


Figure 1: The effect of c.T152C, c.A227G on RNA structure. Enclosed area with marked differences for 152TT and 152CC, 227AA and 227GG genotypes. (**a**) Probabilities of wild-type and mutated sequences are presented in the upper and lower triangle of the plot, respectively, and polymorphic nucleotide is shown by yellow color. (**b**) Optimal secondary structure of RNA sequence (highlighted from 12 nt to 159 nt) in the 152TT genotype with minimum free energy of –69.50 kcal mol^{-1} and 152CC genotype with minimum free energy of –69.80 kcal mol^{-1}. (**c**) Enclosed area with marked differences for 227AA and 227GG genotypes. Probabilities of wild-type and mutated sequences are presented in the upper and lower triangle of the plot, respectively, and polymorphic nucleotide is shown by yellow color. (**d**) Optimal secondary structure of RNA sequence (highlighted from 73 nt to 222 nt) in the 227AA genotype with minimum free energy of –60.90 kcal mol^{-1} and 227GG genotype with minimum free energy of –64.80 kcal mol^{-1}. RNA: ribonucleic acid; nt: nucleotide.

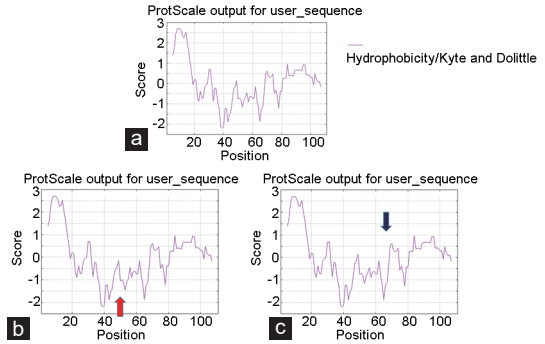


Figure 2: Hydropathy plot for the DEFB126 protein prepared in the Expasy ProtScale Website according to the Kyte and Doolittle algorithm. The hydrophobicity of (**a**) the wild-type DEFB126 protein, (**b**) the DEFB126 protein with M51T mutation, and (**c**) the DEFB126 protein with K76R mutation. The results showed that the M51T and K76R of the DEFB126 protein had lower hydrophobicity scores than the wild-type. DEFB126: β-defensin 126.

that has been mapped to a cluster on chromosome 20p13;[18] and it has two exons that encode 111 amino acids. The protein consists of a conserved β-defensin core and a C-terminal glycosylated peptide tail. DEFB126 appears to interact with the lipid membrane as a covalently linked dimer via sequences of hydrophobic amino acids associated with the defensin-like core.[14,16] The C-terminal glycosylated peptide tail is important for normal sperm function. The tolerance of sperm can be enhanced by covering the DEFB126 protein, masking its unique testicular and epididymal surface antigens; sialic acid residues on the DEFB126 protein are hydrolyzed by sialidase NMase to eliminate immune protection on the sperm surface.[16,17,47]

Evidence suggests that DEFB126 promotes sperm penetration of the cervical mucus and mediates sperm attachment to oviductal epithelia.[16] Expression of *DEFB126* mRNA in the epididymis head is downregulated in patients with nonobstructive azoospermia, and the proportion of DEFB126-positive sperm in men with normal fertility is significantly different from that in men with infertility and varicocele. Coculture of immature human testicular sperm with DEFB126-transfected H9C2 cells increases sperm motility by approximately 15%. The DEFB126, one of the potential targets for the development of posttesticular male contraception,[48] has an immunoprotective function, and various mutations affect total sperm motility and fertilization ability and are thus closely related to asthenospermia.

Chromosomal aberration is one of the most common causes of male infertility, and the literature has demonstrated that structural chromosomal abnormalities in males can lead to abnormal sperm concentrations and motility.[49,50] Karyotype analysis is a powerful diagnostic tool that provides valuable information for genetic counselling and thus should be used in infertility screening.[51] Spermatogenic failure is another form of male infertility, and microdeletions in the AZF region are detected in 5%–15% of males with spermatogenic failure.[52] Therefore, men with karyotype abnormalities and/or Y microdeletions were excluded from this study. In our study, 119 asthenozoospermia patients were screened by karyotype analysis and Y microdeletion testing, and 6 of them were found to have chromosomal aberrations, including two Klinefelter's syndrome instances, two translocations and two small Y microdeletions, and seven Y microdeletions (**Supplementary Table 3**). Patients with asthenozoospermia due to one of the above causes were excluded from this study.

Protein structure affects function, and mutations in DEFB126 may be associated with sperm dysfunction and decreased motility. Two nucleotide deletions, rs11467417 (317–318 *del/del*) and rs11467497 (163–166 *wt/del*), and seven nucleotide mutations, c.152T>C (M51T, rs75442118), c.164A>G (Q55R, rs200901241), c.193C>T (R65C, rs375427290), c.216T>C (rs77783981), c.227A>G (K76R, rs74717966), c.288T>C (rs77106211), and c.308C>T (A103V, s149330491), were detected in this study. Among them, rs11467417 (317–318 *del/del*, $P = 0.035$), rs11467497 (163–166 *wt/del*, $P = 0.031$), c.152T>C ($P = 0.006$), and c.227A>G ($P = 0.020$) showed statistically significant differences between the control and asthenospermia groups.

Mutations in the DEFB126 carboxy-terminal ORF alter the composition of the DEFB126 peptide (**Supplementary Figure 4**), which may lead to the synthesis and secretion of nonfunctional proteins associated with reduced sperm binding ability.[53] For rs11467417, a binucleotide deletion produces a nonstop mRNA that may result in protein reduction due to translational inhibition and proteasomal destruction.[54] *DEFB126* mRNA expression in the epididymis tissues of patients with the rs11467417 homozygous deletion genotype was lower than that of patients with the wild-type homozygous genotype.[21] Immunohistochemical staining showed that in males with the rs11467417 homozygous deletion, only a small number of DEFB126 protein capsules were observed on the sperm surface.[55] DEFB126 protein expression was unstable in those with the homozygous deletion genotype in this study, thus demonstrating that DEFB126 homozygous deletion is a cause of asthenospermia.[23]

The rs11467497 four-nucleotide deletion (151 base pair [bp] upstream of rs11467417), which has been found in more than one HapMap population (Japanese, Chinese, Americans [Utah], and Yoruba),[56] leads to premature termination of translation and production of peptides truncated at the carboxyl-terminus. As mentioned above, the carboxyl-terminus plays an important role, and this deletion may prevent sperm from being transported in the female reproductive tract, ultimately leading to a decline in fertility. In this study, a significant dosage effect for rs11467497 heterozygotes was observed in patients with

asthenozoospermia (OR = 2.568, 95% CI: 1.069–6.171, $P = 0.031$), which is consistent with previous observations. In addition, rs11467497 shows a significant correlation with weak sperm disease, and a lack of DEFB126 protein expression occurs in the sperm of patients with homozygous deletion, also consistent with previous research.[37] Round cells in semen are composed of spermatogenic cells and nonspermatogenic cells. Neutrophils are present in nonspermatogenic round cells, which are often indicative of male reproductive tract infection and/or subsequent inflammatory response,[57] that may reduce sperm motility, fertilization capacity, and reduce male fertility, and rs11467497-*del* was significantly associated with a higher number of round cells.[58–60] These results indicate that rs11467497 is indeed associated with asthenospermia and *DEFB126* deletion caused by rs11467497-*del* increases round cells number and leads to infertility. Moreover, these findings suggest that two common *DEFB126* deletion variants (rs11467497 and rs11467417) may be responsible for male hypofertility and infertility by reducing the affinity of DEFB126 for sperm surfaces.

c.152T>C (M51T) and c.227A>G (K76R) are found on exon 2, and both are located in conserved sequences, indicating that these two sites are essential for protein function. The DEFB126 protein sequence alignment results for eight species showed that M51 and K76 are highly conserved (**Supplementary Figure 5**), indicating that these two sites may be closely related to the role of the β-defensin core. Amino acid substitution experiments showed that the first amino acid residues, six cysteine amino acid residues, and adjacent residues of the β-defensin peptide are key of the *Bubalus bubalis* β-defensin functional element.[61] M51T alters protein secondary structure (random curl shortened, adjacent α-helix, and β-fold increase) and may affect the formation of three disulfide bonds in the defensin core. K76R may affect glycation of the DEFB126 carboxylic terminus (due to the presence of threonine on both sides), thereby affecting the role of DEFB16 in cervical mucus interaction, immune protection, sperm-fallopian tube interactions, and sperm release from the fallopian isthmus sperm bank. Clusters of carboxy-terminal threonine and serine are reported to form a large hydrating shell to protect sperm, and K76R may weaken this protective effect.[62]

Selective splicing enriches proteome diversity and regulates developmental and tissue-specific processes by generating multiple transcripts from a single gene.[63,64] As pre-mRNA splicing disorders play a role in human diseases, accurate prediction of selective splicing events is of great significance in the study of gene function and disease treatment. ASSP results showed that c.T152C changes the parameters of donor sites 132, 185, and 192 and that c.227A>G changes the parameters of receptor site 223 in the constituent receptor sequence, which may influence gene expression and be involved in posttranslational modification, leading to DEFB16 level changes in the testis.

The function of many noncoding RNA genes and mRNA *cis*-regulatory elements is largely determined by their structure, which is dependent on their sequence. Single-nucleotide polymorphisms (SNPs) may disrupt RNA structure, interfere with molecular function, and lead to phenotypic effects. Variants and wild-type sequences were evaluated using RNA folding algorithms, and c.T152C alters the free energy of the DEFB126 optimal secondary structure but has no effect on the general structure. c.A227G changes the minimum free energy of the RNA structure from −51.50 kcal mol^{-1} to −55.40 kcal mol^{-1}. The A/G transformation changes the secondary structure of RNA in this region, resulting in an additional ring structure between nucleotides 71 and 123, and such secondary structure change may affect RNA binding sites.

Posttranslational protein modification (PTMS) is an important biochemical event that regulates various cellular functions.[65] PTMS-based regulation can be achieved through modification of a single amino acid or through a combination of the same or different modification sites. PTMS can also occur at the interaction interface and affect protein–protein binding.[66] ModPre prediction indicated that K76R is the cleavage site that might affect the stability of the DEFB126 protein.

Protein structure classification describes important properties of the protein folding process and can provide important information for protein structure analysis, functional analysis, drug design, and many other biomedical applications.[67,68] As detected in this study, M51T results in the shortening of random coils in the DEFB126 secondary structure but increases the α-helix and β-fold. M51 is located in the corner region of the DEFB126 three dimensional (3D) structure, and M51T may lead to a protein 3D structure change.

Nonsynonymous SNPs in the coding region can alter amino acid composition and affect protein function, expression, conformation, or stability. Approximately 58% of exon SNPs estimated in the human genome can lead to amino acid sequence changes, known as nonsynonymous SNPs (nsSNPs).[69] Experiments have shown that nearly 1/3 of nsSNP mutations are harmful to human health; thus, identifying harmful nsSNP mutations has important implications for protein function and disease diagnosis.[70] An important contribution to protein stability is the solvent separation capacity of protein hydrophobic residues. The effect of mutations on protein stability can be evaluated by the thermal stability of the protein system, and a ΔΔG value below zero indicates instability. K76R and M51T are hydrophobic amino acid substitutions that change the hydrophobicity and decrease the free energy and protein stability. M51T is most destructive due to its lowest free energy of −0.58 kcal mol^{-1}. Therefore, the destructive effects of the K76R and M51T substitutions in asthenozoospermia may be caused by changes in *DEFB126* gene expression.

The polarity and charged amino acid residues on the protein surface determine the binding specificity and affinity of a ligand, and DEFB126 is uniformly distributed across the sperm surface, which may be associated with its binding to the lipid bilayer. Meanwhile, c.152T>C is highly hydrophilic and may interact with the environment or other proteins, thus affecting interaction of the adjacent highly hydrophobic region (amino acids 58–72) with the lipid bilayer. β-defensins contain large amounts of lysine, which seems to promote their interaction with the negatively charged plasma membrane of microorganisms through electrostatic interactions. K76R in DEFB126 may affect its antibacterial, antiviral, and antifungal activities.

Many proteins must interact with ligand molecules to perform their biological functions, and localization of ligand binding sites has important application value in protein structure prediction, protein functional relationship elucidation, protein engineering, and drug design. Evolutionarily conserved binding site residues are important not only for binding to ligands but also for preventing binding to other molecules, thus affecting their selectivity and specificity.[71,72] In addition to electrostatic interactions, hydrophobic interactions play a crucial role in ligand binding.[71] The NOP2/Sun domain family, member 7 (*Nsun7*) 7th exon mutation, enhances protein hydrophobicity and changes the ligand binding site and protein–ligand interaction without affecting protein folding, resulting in impaired Nsun7 protein function and weakened sperm motility.[73] The FTsite predicts three different ligand binding sites of DEFB126, though the sites of mutation found in this study did not include ligand binding sites, indicating that mutation may not cause low sperm motility by affecting ligand binding sites.

The samples in our study were screened using the WHO 4th criteria. While the criteria in WHO 5th/WHO 6th for sperm morphologic anomalies for teratozoospermia are stricter,[9,74] some sperm flagellum

defects cannot be identified by the WHO 4th criteria, and flagellum disorder due to gene defects not only causes tail defects but also underlies asthenozoospermia. Thus, gene defects result in a variety of functional (and morphological) flagellum anomalies that also underlie asthenozoospermia,[75,76] and the idiopathic asthenozoospermia caused by *DEFB126* polymorphisms may also occur by affecting sperm flagellar function and structure.

Sperm with *DEFB126* homozygous deletion display dysfunctional behavior in the female reproductive tract based on conventional semen assessments. In our study, rs11467417 (317–318 *del/del*) and rs11467497 (163–166 *wt/del*) nucleotide deletions and c.152T>C and c.227A>G nucleotide mutations were significantly different between the control and asthenospermia groups, are predicted to be damaging, and might be associated with asthenozoospermia in Chengdu, China. Therefore, detection of the *DEFB126* genotype can be added to early clinical infertility assessment of patients with asthenospermia to identify the main cause of asthenospermia for efficient diagnosis and treatment. In sperm with homozygous rs11467417 deletion, expression of the DEFB126 protein was unstable, and no DEFB126 protein was detected in sperm with homozygous rs11467497 deletion. Sperm lacking DEFB126 showed severely weakened activity in cervical mucus, while the addition of DEFB126 to the sperm of patients with homozygous deletion significantly improved the ability of human sperm to penetrate hyaluronic acid gel and completely restored their ability to penetrate this viscoelastic reproductive tract medium. Therefore, adding DEFB126 to sperm with *DEFB126* homozygous deletion genotypes for less invasive vaginal and cervical insemination or the delivery of DEFB126 to sperm via vaginal gel for sperm therapy may have a role in the treatment of clinical infertility caused by sperm with *DEFB126* homozygous deletion. The nucleotide mutations c.152T>C and c.227A>G are predicted to affect the structure and function of DEFB126. For patients with defining mutation types, timely use of assisted reproductive technology, such as intrauterine insemination (IUI) and *in vitro* fertilization (IVF), can reduce the time and monetary costs of infertile couples.

CONCLUSION

The present study is about the association between polymorphisms of *DEFB126* and asthenozoospermia in a Chinese population. The data in our study indicate that *DEFB126* mutations (rs11467417 317–318 *del/del*; rs11467497 163–166 *wt/del*; c.T152>C M51T; and c.227A>G K76R) are likely associated with asthenozoospermia. Genotyping of *DEFB126* polymorphisms in men with asthenospermia can help determine the most effective fertility intervention. The results of our study point to the underlying causes of lower male fertility, thus opening the possibility of new clinical diagnosis and treatment. Sperm surface proteomics provides researchers with data for identifying proteins on the surface of sperm and may lead to insight into the environment in which sperm travel from the testis to the fallopian tube. DEFB126 has important effects on the sperm surface, and other sperm membrane surface proteins (such as heat shock protein A4L) may also be important for sperm migration to the fallopian tube. Functional studies of the diversity of this genetic material may provide research directions for male infertility in the future.

AUTHOR CONTRIBUTIONS

JYH, JYP, XLL, and XPD conceived and designed the study; JYH, XPD, JHD, QFL, XLL, ZLS, XW, and YRL collected the samples; JYH, JYP, XLL, SYM, YFS, and YRC performed the experiments; JYH and XPD wrote the manuscript. All authors read and approved the final manuscript.

COMPETING INTERESTS

All authors declare no competing interests.

ACKNOWLEDGMENTS

All authors thank the Affiliate Reproductive Hospital Genitalia Hygiene Research Center (Chengdu, China) and all patients and control subjects for their participation in this study.

Supplementary Information is linked to the online version of the paper on the *Asian Journal of Andrology* website.

REFERENCES

1 Abid S, Maitra A, Meherji P, Patel Z, Kadam S, *et al*. Clinical and laboratory evaluation of idiopathic male infertility in a secondary referral center in India. *J Clin Lab Anal* 2008; 22: 29–38.
2 Zhang SH, Zhang JH, Ding XP, Zhang S, Chen HH, *et al*. Association of polymorphisms in *tektin-t* gene with idiopathic asthenozoospermia in Sichuan, China. *J Assist Reprod Genet* 2016; 33: 181–7.
3 Moawad AR, Fernandez MC, Scarlata E, Dodia C, Feinstein SI, *et al*. Deficiency of peroxiredoxin 6 or inhibition of its phospholipase A2 activity impair the *in vitro* sperm fertilizing competence in mice. *Sci Rep* 2017; 7: 1–13.
4 Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, *et al*. Rate of *de novo* mutations and the importance of father's age to disease risk. *Nature* 2012; 488: 471–5.
5 Smith R, Kaune H, Parodi D, Madariaga M, Rios R, *et al*. Increased sperm DNA damage in patients with varicocele: relationship with seminal oxidative stress. *Hum Reprod* 2006; 21: 986–93.
6 Carra E, Sangiorgi D, Gattuccio F, Rinaldi AM. Male infertility and mitochondrial DNA. *Biochem Biophys Res Commun* 2004; 322: 333–9.
7 Zuccarello D, Ferlin A, Cazzadore C, Pepe A, Garolla A, *et al*. Mutations in dynein genes in patients affected by isolated non-syndromic asthenozoospermia. *Hum Reprod* 2008; 23: 1957–62.
8 Zuccarello D, Ferlin A, Garolla A, Pati MA, Moretti A, *et al*. A possible association of a human tektin-t gene mutation (A229V) with isolated non-syndromic asthenozoospermia: case report. *Hum Reprod* 2008; 23: 996–1001.
9 World Health Organization. WHO Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction. 4th edition. Cambridge: Cambridge University Press; 1999.
10 Silva NN, Sabino AD, Tafuri A, Lima AA. Lack of association between methylenetetrahydrofolate reductase C677T polymorphism, HPV infection and cervical intraepithelial neoplasia in Brazilian women. *BMC Med Genet* 2019; 20: 1–8.
11 Zhang Y, Malekpour M, Al-Madani N, Kahrizi K, Zanganeh M, *et al*. Sensorineural deafness and male infertility: a contiguous gene deletion syndrome. *J Med Genet* 2007; 44: 233–40.
12 Takasaki N, Tachibana K, Ogasawara S, Matsuzaki H, Hagiuda J, *et al*. A heterozygous mutation of *GALNTL5* affects male infertility with impairment of sperm motility. *Proc Natl Acad Sci U S A* 2014; 111: 1120–5.
13 Li JY, Liu FJ, Wang HY, Liu X, Liu J, *et al*. Systematic mapping and functional analysis of a family of human epididymal secretory sperm-located proteins. *Mol Cell Proteomics* 2010; 9: 2517–28.
14 Yudin AI, Tollner TL, Li MW, Treece CA, Overstreet JW, *et al*. ESP13.2, a member of the β-defensin family, is a macaque sperm surface-coating protein involved in the capacitation process. *Biol Reprod* 2003; 69: 1118–28.
15 Rodríguez-Jiménez FJ, Krause A, Schulz S, Forssmann WG, Conejo-Garcia JR, *et al*. Distribution of new human β-defensin genes clustered on chromosome 20 in functionally different segments of epididymis. *Genomics* 2003; 81: 175–83.
16 Tollner TL, Yudin AI, Treece CA, Overstreet JW, Cherr GN. Macaque sperm coating protein DEFB126 facilitates sperm penetration of cervical mucus. *Hum Reprod* 2008; 23: 2523–34.
17 Tollner TL, Yudin AI, Treece CA, Overstreet JW, Cherr GN. Macaque sperm release ESP13.2 and PSP94 during capacitation: the absence of ESP13.2 is linked to sperm-zona recognition and binding. *Mol Reprod Dev* 2004; 69: 325–37.
18 Liu H, Yu H, Gu Y, Xin A, Zhang Y, *et al*. Human beta-defensin DEFB126 is capable of inhibiting LPS-mediated inflammation. *Appl Microbiol Biotechnol* 2013; 97: 3395–408.
19 Julenius K, Mølgaard A, Gupta R, Brunak S. Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology* 2005; 15: 153–64.
20 Yudin AI, Generao SE, Tollner TL, Treece CA, Overstreet JW, *et al*. Beta-defensin 126 on the cell surface protects sperm from immunorecognition and binding of anti-sperm antibodies. *Biol Reprod* 2005; 73: 1243–52.
21 Barbagli G, Lazzeri M. Re: tissue-engineered autologous urethras for patients who need reconstruction: an observational study. *Eur Urol* 2011; 60: 1303–4.

22 Duan S, Shi C, Chen G, Zheng JF, Wu B, *et al*. Another functional frame-shift polymorphism of *DEFB126* (rs11467497) associated with male infertility. *J Cell Mol Med* 2015; 19: 1077–84.
23 Xin A, Cheng L, Diao H, Wu Y, Zhou S, *et al*. Lectin binding of human sperm associates with *DEFB126* mutation and serves as a potential biomarker for subfertility. *Sci Rep* 2016; 6: 1–12.
24 Navarro-Costa P, Plancha CE, Gonaçlves J. Genetic dissection of the AZF regions of the human Y chromosome: thriller or filler for male (in)fertility? *J Biomed Biotechnol* 2010; 2010: 936569.
25 Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res* 2004; 14: 1188–90.
26 Zamenhof S. Mutations. *Am J Med* 1963; 34: 609–26.
27 Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003; 31: 3812–4.
28 Schwarz J, Cooper D, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* 2014; 11: 361–2.
29 Rogers MF, Shihab HA, Mort M, Cooper DN, Gaunt TR, *et al*. FATHMM-XF: accurate prediction of pathogenic point mutations via extended features. *Bioinformatics* 2017; 34: 511–3.
30 Capriotti E, Calabrese R, Fariselli P, Martelli PL, Altman RB, *et al*. WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation. *BMC Genomics* 2013; 14 Suppl 3: S6.
31 Hecht M, Bromberg Y, Rost B. Better prediction of functional effects for sequence variants from VarI-SIG 2014: identification and annotation of genetic variants in the context of structure, function and disease. *BMC Genomics* 2016; 16: 1–12.
32 Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* 2015; 31: 2745–7.
33 Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, *et al*. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. *Nat Commun* 2020; 11: 1–28.
34 Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, *et al*. PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res* 2003; 31: 334–41.
35 Sabarinathan R, Tafer H, Seemann SE, Hofacker IL, Stadler PF, *et al*. RNAsnp: efficient detection of local RNA secondary structure changes induced by SNPs. *Hum Mutat* 2013; 34: 546–56.
36 Wang M, Marín A. Characterization and prediction of alternative splice sites. *Gene* 2006; 366: 219–27.
37 Pejaver V, Hsu WL, Xin F, Dunker AK, Uversky VN, *et al*. The structural and functional signatures of proteins that undergo multiple events of post-translational modification. *Protein Sci* 2014; 23: 1077–93.
38 Geourjon C, Deléage G. Sopma: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics* 1995; 11: 681–4.
39 Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res* 2005; 33: 306–10.
40 Arnold K, Kiefer F, Kopp J, Battey JN, Podvinec M, *et al*. The protein model portal. *J Struct Funct Genomics* 2009; 10: 1–8.
41 Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, *et al*. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* 1999; 112: 531–52.
42 Ngan CH, Hall DR, Zerbe B, Grove LE, Kozakov D, *et al*. FtSite: high accuracy detection of ligand binding sites on unbound protein structures. *Bioinformatics* 2012; 28: 286–7.
43 Zhao Y, Diao H, Ni Z, Hu S, Yu H, *et al*. The epididymis-specific antimicrobial peptide β-defensin 15 is required for sperm motility and male fertility in the rat (*Rattus norvegicus*). *Cell Mol Life Sci* 2011; 68: 697–708.
44 Zhou CX, Zhang YL, Xiao L, Zheng M, Leung KM, *et al*. An epididymis-specific β-defensin is important for the initiation of sperm maturation. *Nat Cell Biol* 2004; 6: 458–64.
45 Diao R, Fok KL, Chen H, Yu MK, Duan Y, *et al*. Deficient human β-defensin 1 underlies male infertility associated with poor sperm motility and genital tract infection. *Sci Transl Med* 2014; 6: 249ra108.
46 Yu H, Dong J, Gu Y, Liu H, Xin A, *et al*. The novel human β-defensin 114 regulates lipopolysaccharide (LPS)-mediated inflammation and protects sperm from motility loss. *J Biol Chem* 2013; 288: 12270–82.
47 Balding DJ. A tutorial on statistical methods for population association studies. *Nat Rev Genet* 2006; 7: 781–91.
48 Sipilä P, Jalkanen J, Huhtaniemi IT, Poutanen M. Novel epididymal proteins as targets for the development of post-testicular male contraception. *Reproduction* 2009; 137: 379–89.
49 Stouffs K, Seneca S, Lissens W. Genetic causes of male infertility. *Ann Endocrinol (Paris)* 2014; 75: 109–11.
50 Suganya J, Kujur SB, Selvaraj K, Suruli MS, Haripriya G, *et al*. Chromosomal abnormalities in infertile men from Southern India. *J Clin Diagnostic Res* 2015; 9: GC05–10.
51 Poli MN, Miranda LA, Gil ED, Zanier GJ, Iriarte PF, *et al*. Male cytogenetic evaluation prior to assisted reproduction procedures performed in Mar del Plata, Argentina. *J Bras Reprod Assist* 2016; 20: 62–5.
52 Vogt PH, Edelmann A, Kirsch S, Henegariu O, Hirschmann P, *et al*. Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. *Hum Mol Genet* 1996; 5: 933–43.
53 Aram R, Chan PT, Cyr DG. Beta-defensin126 is correlated with sperm motility in fertile and infertile men. *Biol Reprod* 2020; 102: 92–101.
54 Ito-Harashima S, Kuroha K, Tatematsu T, Inada T. Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. *Genes Dev* 2007; 21: 519–24.
55 Boroujeni PB, Ebrahimian S, Abedini M, Chayjan MR, Hassani M, *et al*. The role of DEFB126 variation in male infertility and medically assisted reproduction technique outcome. *Reprod Biomed Online* 2019; 39: 649–57.
56 Altshuler DL, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, *et al*. A map of human genome variation from population-scale sequencing. *Nature* 2010; 467: 1061–73.
57 Johanisson E, Campana A, Luthi R, De Agostini A. Evaluation of "round cells" in semen analysis: a comparative study. *Hum Reprod Update* 2000; 6: 404–12.
58 Purvis K, Christiansen E. Infection in the male reproductive tract. Impact, diagnosis and treatment in relation to male infertility. *Int J Androl* 1993; 16: 1–13.
59 Rusz A, Pilatz A, Wagenlehner F, Linn T, Diemer T, *et al*. Influence of urogenital infections and inflammation on semen quality and male fertility. *World J Urol* 2012; 30: 23–30.
60 Shi TY, Chen G, Huang X, Yuan Y, Wu X, *et al*. Effects of reactive oxygen species from activated leucocytes on human sperm motility, viability and morphology. *Andrologia* 2012; 44: 696–703.
61 Batra V, Maheshwarappa A, Dagar K, Kumar S, Soni A, *et al*. Unusual interplay of contrasting selective pressures on β-defensin genes implicated in male fertility of the Buffalo (*Bubalus bubalis*). *BMC Evol Biol* 2019; 19: 1–19.
62 Jentoft N. Why are proteins O-mannosylated? *Trends Biochem Sci* 1990; 15: 291–4.
63 Smith CW, Valcárcel J. Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem Sci* 2000; 25: 381–8.
64 Graveley BR. Alternative splicing: increasing diversity in the proteomic world. *Trends Genet* 2001; 17: 100–7.
65 Nussinov R, Tsai CJ, Xin F, Radivojac P. Allosteric post-translational modification codes. *Trends Biochem Sci* 2012; 37: 447–55.
66 Nishi H, Hashimoto K, Panchenko AR. Phosphorylation in protein-protein binding: effect on stability and function. *Structure* 2011; 19: 1807–15.
67 Zhou GP, Assa-Munt N. Some insights into protein structural class prediction. *Proteins Struct Funct Genet* 2001; 44: 57–9.
68 Chou KC. Structural bioinformatics and its impact to biomedical science and drug discovery. *Front Med Chem* 2012; 3: 455–502.
69 Tennessen JA, Bigham AW, O'connor TD, Fu W, Kenny EE, *et al*. Evolution and functional impact of rare coding variation from deep sequencing of human exomes *Science* 2012; 337: 64–9.
70 Tokuriki N, Tawfik DS. Stability effects of mutations and protein evolvability. *Curr Opin Struct Biol* 2009; 19: 596–604.
71 Liou JW, Chang FT, Chung Y, Chen WY, Fischer WB, *et al*. *In silico* analysis reveals sequential interactions and protein conformational changes during the binding of chemokine CXCL-8 to its receptor CXCR1. *PLoS One* 2014; 9: 17–23.
72 Najmanovich RJ. Evolutionary studies of ligand binding sites in proteins. *Curr Opin Struct Biol* 2017; 45: 85–90.
73 Khosronezhad N, Colagar AH, Jorsarayi SG. T26248G-transversion mutation in exon7 of the putative methyltransferase *Nsun7* gene causes a change in protein folding associated with reduced sperm motility in asthenospermic men. *Reprod Fertil Dev* 2015; 27: 471–80.
74 Cao XW, Lin K, Li CY, Yuan CW. [A review of WHO Laboratory Manual for the examination and processing of human semen (5th edition)]. *Zhonghua Nan Ke Xue* 2011; 17: 1059–63. [Article in Chinese].
75 Coutton C, Escoffier J, Martinez G, Arnoult C, Ray PF. Teratozoospermia: spotlight on the main genetic actors in the human. *Hum Reprod Update* 2014; 21: 455–85.
76 He BM, Chen R, Sun TQ, Yang Y, Zhang CL, *et al*. Prostate cancer risk prediction models in Eastern Asian populations: current status, racial difference, and future directions. *Asian J Androl* 2020; 22: 158–61.

Supplementary Table 1: Primers used for detection of Y microdeletions

Gene	*Primer*	*Primer sequence (5' to 3')*	*Size (bp)*	*PCR product size (bp)*
AZFa	sY86-F	AGACTATGCTTCAGCAGGTC	20	156
	sY86-R	CCAGTCTTTGGGATTTCTTT	20	
	USP9Y-F	AGCCTGTTTCAGTGTTTGGG	20	169
	USP9Y-R	ATACCCTCAAGGAACCTCAAG	21	
AZFb	sY127-F	AGCACCCACTGGAATCTACC	20	195
	sY127-R	CATGGCTACACAGACAGGGA	20	
	sY134-F	GTCTGCCTCACCATAAAACG	20	301
	sY134-R	ACCACTGCCAAAACTTTCAA	20	
AZFc	sY254-F	GGGTGTTACCAGAAGGCAAA	20	380
	sY254-R	GAACCGTATCTACCAAAGCAGC	22	
	sY255-F	GTTACAGGATTCGGCGTGAT	20	123
	sY255-R	CTCGTCATGTGCAGCCAC	18	
sY14 (SRY)	SRY-F	GAATATTCCCGCTCTCCGGA	20	472
	SRY-R	GCTGGTGCTCCATTCTTGAG	20	

PCR: polymerase chain reaction; SRY: sex determining region of Y-chromosome

Supplementary Table 2: Primers used for the amplification of the 2 exons of the DEFB126 gene

Number	*Exon number*	*PCR product size*	*Sequence (5′ → 3′)*
1	Exon-1	206 bp	F - TCATACTGAATAGAGACTTCTGGAC
			R - GAAGCCAGACCCCATTTCCA
2	Exon-2	578 bp	F - TGGCCCCTGGTAAGTGTTT
			R - CTTTGCTTTAATGAGTCGGGGA

F means forward primer; R means reverse primer. PCR: polymerase chain reaction

Supplementary Table 3: Karyotype analysis and detection of Y microdeletions in men with asthenozoospermia

Screening methods	*Mutation*	*Quantity*	*Percentage*
Karyotyping	Klinefelter's syndrome	2	1.68
	Translocation	2	1.68
	Small Y	2	1.68
Y microdeletions	All sites	1	0.84
	sY254 and sY255	5	4.20
	sY254, sY255, sY127, sY134	1	0.84

Supplementary Table 4: Comparison of age and semen parameters between idiopathic asthenozoospermic group and controls

Clinical parameters	*Asthenozoospermic group (n=106)*[a]	*Control group (n=102)*[a]	*P*[b]
Age (year)	32.3±5.1	32.0±5.5	0.528
Sperm concentration ($\times 10^6$ ml^{-1})	35.2±23.6	89.2±12.8	<0.001
Rapid progressive motility (%)	8.6±6.1	37.9±5.5	<0.001
Total progressive motility (%)	18.5±10.1	70.5±7.4	<0.001

[a]Data are presented as mean±s.d.; [b]The comparison between groups was done with the Student's *t*-test; $P<0.05$ was considered statistically significant. s.d.: standard deviation

Supplementary Table 5: Genotypes and allele frequencies for DEFB126 among idiopathic asthenozoospermia and controls

dbSNP ID	*Genotype/allele*	*Controls (102)*	*Asthenozoospermia (106)*	*OR (95% CI)*	*P*
rs75442118	152TT	95	85	3.353 (1.358–8.281)	0.006
	152TC	7	21		
	152CC	0	0		
rs200901241	164AA	102	105	1.010 (0.991–1.028)	0.325
	164AG	0	1		
	164GG	0	0		
rs375427290	193CC	102	105	1.010 (0.991–1.028)	0.325
	193CT	0	1		
	193TT	0	0		
rs77783981	216TT	92	94	1.174 (0.484–2.852)	0.722
	216TC	10	12		
	216CC	0	0		
rs74717966	227AA	94	86	2.733 (1.144–6.526)	0.02
	227AG	8	20		
	227GG	0	0		
rs77106211	288TT	96	92	2.435 (0.897–6.607)	0.073
	288TC	6	14		
	288CC	0	0		
rs149330491	308CC	102	105	1.010 (0.991–1.028)	0.325
	308CT	0	1		
	308TT	0	0		
rs11467417	317_318CC/CC	35	46		
	317_318CC/--	41	45	1.443	0.239
	317_318--/--	26	15	2.278	0.035
rs11467497	163_166CAAA/CAAA	93	86		
	163_166CAAA/----	8	19	2.568 (1.069–6.171)	0.031
	163_166----/----	1	1	1.081 (0.067–17.558)	0.956

OR: odds ratio, CI: confidence interval, dbSNP: the Single Nucleotide Polymorphism database

Supplementary Table 6: The selective splicing sites of DEFB126 wild-type

Position (bp)	*Putative splice site*	*Sequence*	*Score*	*Intron GC*	*Activations*		*Confidence*
					Alt./cryptic	*Constitutive*	
58	Alt.isoform/cryptic donor	TTGGTCTCAGgtaattggta	10.974	0.371	0.553	0.357	0.355
65	Alt.isoform/cryptic donor	CAGGTAATTGgtatgtgaaa	9.684	0.386	0.938	0.042	0.955
132	Alt.isoform/cryptic donor	AGAGATGCATgtaaagaatg	7.064	0.486	0.902	0.069	0.924
185	Altisoform/cryptic acceptor	tgtgttccagCTGACAGACG	5.461	0.471	0.732	0.260	0.645
192	Alt.isoform/cryptic acceptor	cagctgacagACGTGCTAAT	3.811	0.486	0.909	0.087	0.904
223	Constitutive acceptor	ctgtgtccagACAAAGACTA	7.890	0.500	0.413	0.559	0.262

Intron GC: intron GC values correspond to 70 nt of the neighboring intron

Supplementary Table 7: The selective splicing sites of DEFB126 with c.T152C

Position (bp)	*Putative splice site*	*Sequence*	*Score*	*Intron GC*	*Activations*		*Confidence*
					Alt./cryptic	*Constitutive*	
58	Alt.isoform/cryptic donor	TTGGTCTCAGgtaattggta	10.974	0.371	0.553	0.357	0.355
65	Alt.isoform/cryptic donor	CAGGTAATTGgtatgtgaaa	9.684	0.386	0.938	0.042	0.955
132	Alt.isoform/cryptic donor	AGAGATGCATgtaaagaatg	7.064	0.500	0.918	0.058	0.937
185	Alt.isoform/cryptic acceptor	tgtgttccagCTGACAGACG	5.461	0.486	0.732	0.260	0.644
192	Alt.isoform/cryptic acceptor	cagctgacagACGTGCTAAT	3.811	0.500	0.909	0.087	0.904
223	Constitutive acceptor	ctgtgtccagACAAAGACTA	7.890	0.500	0.413	0.559	0.262

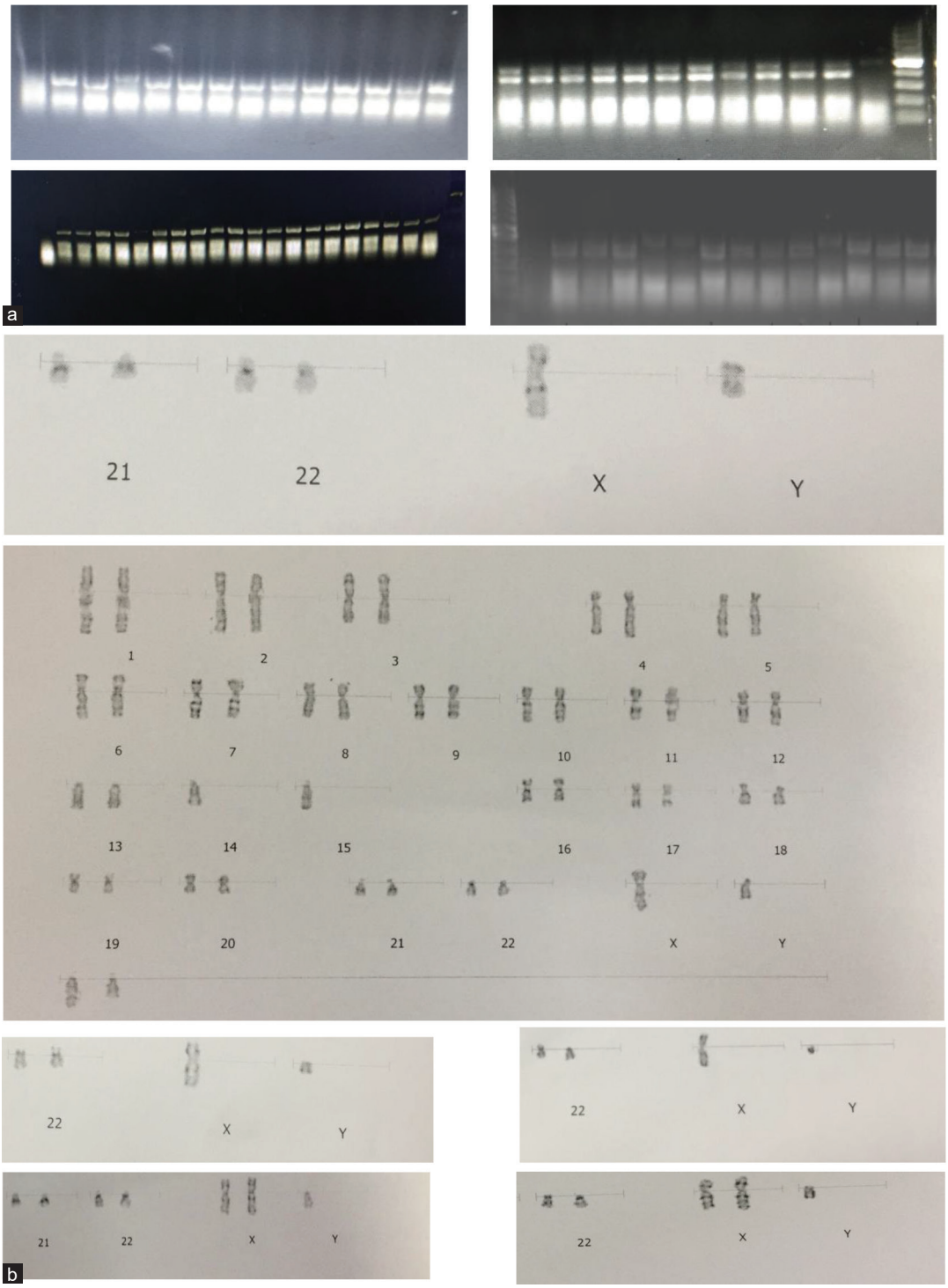
Intron GC: intron GC values correspond to 70 nt of the neighboring intron

Supplementary Table 8: The selective splicing sites of DEFB126 with c.A227G

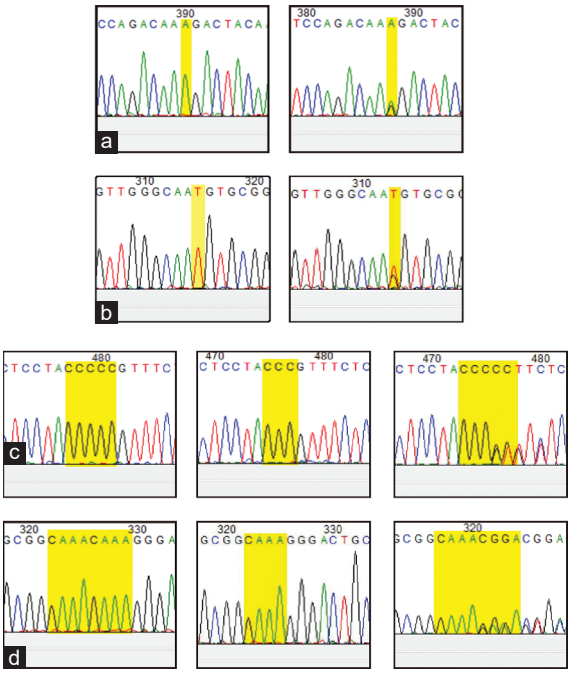
Position (bp)	*Putative splice site*	*Sequence*	*Score*	*Intron GC*	*Activations*		*Confidence*
					Alt./Cryptic	*Constitutive*	
58	Alt.isoform/cryptic donor	TTGGTCTCAGgtaattggta	10.974	0.371	0.553	0.357	0.355
65	Alt.isoform/cryptic donor	CAGGTAATTGgtatgtgaaa	9.684	0.386	0.938	0.042	0.955
132	Alt.isoform/cryptic donor	AGAGATGCATgtaaagaatg	7.064	0.486	0.902	0.069	0.924
185	Alt.isoform/cryptic acceptor	tgtgttccagCTGACAGACG	5.461	0.471	0.732	0.260	0.645
192	Alt.isoform/cryptic acceptor	cagctgacagACGTGCTAAT	3.811	0.486	0.909	0.087	0.904
223	Constitutive acceptor	ctgtgtccagACAAGGACTA	7.890	0.500	0.373	0.599	0.376

Supplementary Table 9: The amino acid residues of DEFB126 ligand binding sites

Location 1	*Location 2*	*Location 3*
SER X 3	SER X 3	SER X 3
CYS X 34	PHE X 6	THR X 7
VAL X 60	THR X 7	LEU X 8
PRO X 61	VAL X 10	PHE X 11
ALA X 62	VAL X 60	CYS X 27
ASP X 63	PHE X 71	ASP X 30
TYR X 68	CYS X 72	VAL X 31
PRO X 69	VAL X 73	GLY X 32
VAL X 70		VAL X 60
PHE X 71		
VAL X 73		

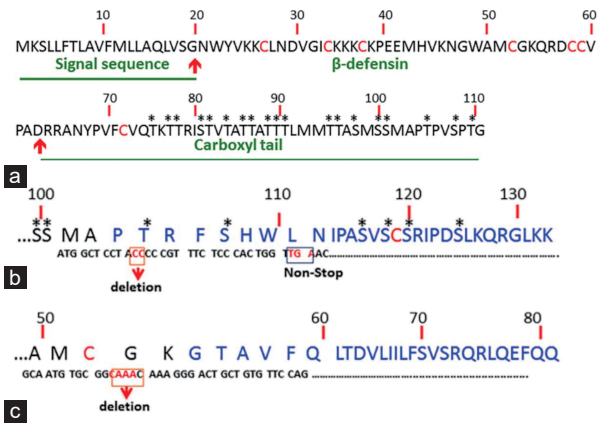


Supplementary Figure 1: A comprehensive mutation screen in men with asthenozoospermia. (**a**) The result of Y microdeletions. (**b**) The result of aberrant karyotype: translocation, small Y, Klinefelter's syndrome.

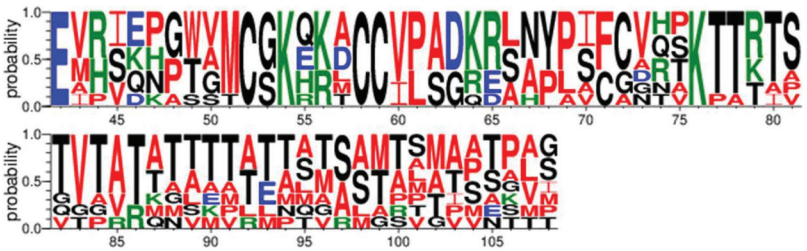


Supplementary Figure 2: Sequencing results of c.A227G, c.T152C, rs11467417 and rs11467497. (**a**) c.A227G mutation site analysis, from left to right, wild type (AA), mutant heterozygote (AG); (**b**) c.T152C mutation site analysis, from left to right, wild type (TT), mutant heterozygote (TC); (**c**) rs11467417 deletion site analysis, from left to right: wild type (CC/CC), deleted homozygote (--/--), deleted heterozygote (CC/--); (**d**) rs11467497 deletion site analysis, from left to right: wild type (CAAA/CAAA), deleted homozygote (----/----), deleted heterozygote (CAAA/----). Mutations and deletions are highlighted in yellow.

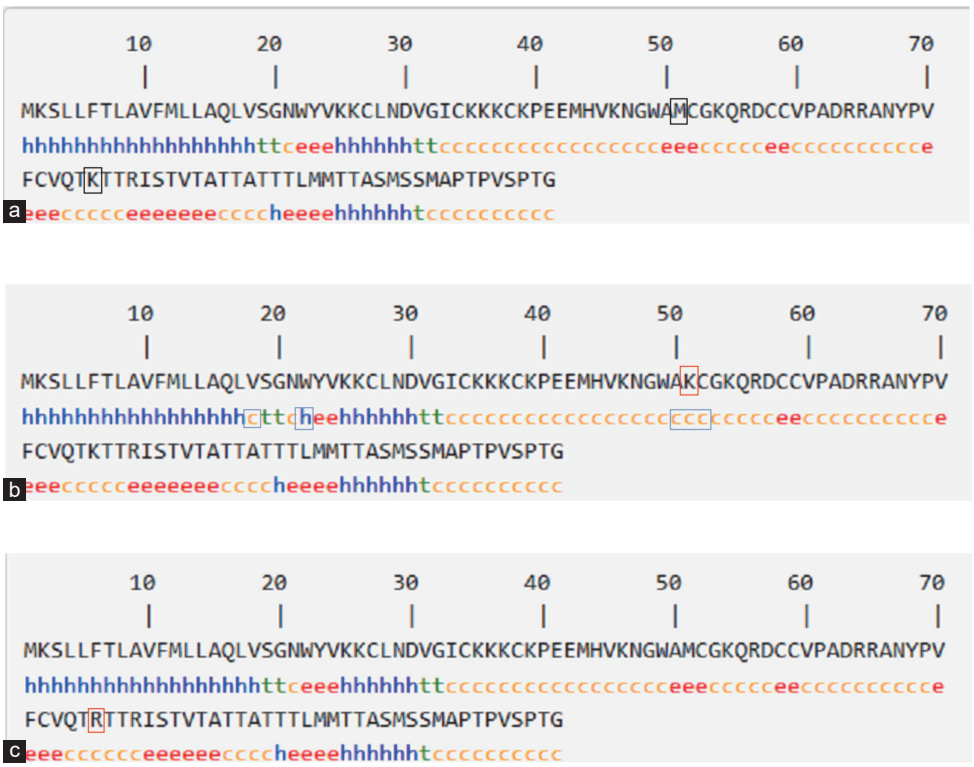
Supplementary Figure 3: DEFB126 expression in sperm of different genotypes. (**a**) The expression of DEFB126 in sperm of different genotypes of rs11467497, from left to right: wild type (CAAA/CAAA), deleted homozygote (----/----), deleted heterozygote (CAAA/----). (**b**) The expression of DEFB126 in sperm of different genotypes of rs11467417, from left to right: wild type (CC/CC), deleted homozygote (--/--), deleted heterozygote (CC/--). DEFB126: β-defensin 126.



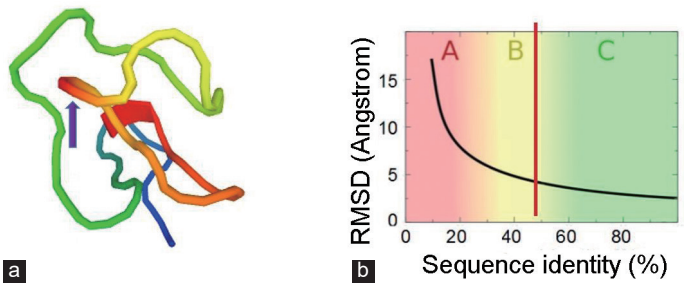
Supplementary Figure 4: Schematic diagram of human DEFB126 peptide sequence. DEFB126 peptide contains the signal sequence, beta-defensin core, and carboxyl tail, and the carboxyl terminal includes an unpaired cysteine and several potential O-glycosylation sites. *Serines and threonines amino acid. (**a**) The DEFB126 wild type, (**b**) the DEFB126 2-del sequence, and (**c**) the DEFB126 4-del sequence. DEFB126: β-defensin 126.



Supplementary Figure 5: Conservation analysis of DEFB126 protein Sequence in eight different species. DEFB126: β-defensin 126.



Supplementary Figure 6: Secondary structure prediction of the DEFB126 protein. (**a**) Secondary structure prediction result of the wild-type DEFB126 protein. (**b**) Secondary structure prediction result of the DEFB126 protein with M51T (red box), the mutant DEFB126 protein reveals that threonine51 (red box) is located in a shortened random coil and the primary random coil turn into alpha helix and beta turn (gray box). (**c**) Secondary structure prediction result of the DEFB126 protein with K76R (red box). DEFB126: β-defensin 126.



Supplementary Figure 7: Protein structure model of DEFB126 was given from the Protein Model Portal. Protein structure model of DEFB126 was given from the Protein Model Portal. (**a**) the location of mutation (M51T) in DEFB126 protein 3D structure; (**b**) this model is based on target-template sequence alignment of 48% sequence identity. A, B, C is used to indicate the difficulty of modeling tasks, A (red area of the graph) means a difficult modeling task, requiring close inspection of the results, B (yellow area of the graph) means a moderate difficult modeling task), C (green area of the graph) means a rather straightforward modeling task. DEFB126: β-defensin 126.