Journal of Assisted Reproduction and Genetics (2022) 39:1195–1203
https://doi.org/10.1007/s10815-022-02433-z

GENETICS

Identification of a new splice-acceptor mutation in HFM1 and functional analysis through molecular docking in nonobstructive azoospermia

Neda Saebnia[1] · Reza Ebrahimzadeh-Vesal[2] · Aliakbar Haddad-Mashhadrizeh[3] · Nazanin Gholampour-Faroji[4] · Albert Schinzel[5] · Zeinab Neshati[1,6] · Mohsen Azimi-Nezhad[2,7]

Received: 7 December 2021 / Accepted: 7 February 2022 / Published online: 29 April 2022

Abstract

Purpose To investigate the genetic cause of nonobstructive azoospermia (NOA).

Methods We performed whole exome sequencing (WES) on the proband who had three relatives suffering from NOA. We used a list of candidate genes which have high expression level in testis and their mutations have been reported in NOA. Sanger sequencing verified the identified variant and its structural and functional consequence was evaluated by protein three-dimensional (3D) structure prediction and protein-ligand docking.

Results WES revealed a novel splice-acceptor mutation (c.1832-2A>T) in helicase for meiosis 1 (HFM1) gene, which co-segregated with the NOA in this family. 3D structural models were generated and verified. Molecular docking indicated that the c.1832-2A>T mutation affects not only the ADP binding residues but also the hydrogen bond interactions. The ADP binding site will be lost in the mutant protein, potentially causing defective crossover and synapsis.

Conclusion We report that the c.1832-2A>T mutation is the likely cause of NOA in the family studied. Regarding that many reported NOA genes are involved in the formation of crossovers and synapsis and have critical roles in the production of germ cells, we suggest that such genes should be considered for screening of infertility among large cohorts of infertile individuals.

Keywords Male infertility · Nonobstructive azoospermia · Whole exome sequencing · Protein modeling · Molecular docking · HFM1

Introduction

Nonobstructive azoospermia (NOA), defined by the absence of sperm in the semen, occurs in ~1% of all men and ~10% of infertile men [1, 2]. NOA is a result of spermatogenic failure and includes 60% of azoospermia cases [3].

Currently, analysis of karyotype and microdeletions of the Y chromosome are two main genetic examinations for NOA. These two genetic tests can only determine the cause of ~20% of NOA patients. Therefore, it is necessary to detect further genetic mechanisms in order to have a more accurate diagnosis in male infertility [4]. In the past decade, the

✉ Zeinab Neshati
neshati@um.ac.ir

✉ Mohsen Azimi-Nezhad
aziminm@gmail.com

[1] Department of Biology, Faculty of Science, Ferdowsi University of Mashhad, Mashhad, Iran

[2] Non-Communicable Diseases Research Center, Neyshabur University of Medical Sciences, Neyshabur, Iran

[3] Industrial Biotechnology Research Group, Institute of Biotechnology, Ferdowsi University of Mashhad, Mashhad, Iran

[4] Biotechnology Department, Iranian Research Organization for Science and Technology (IROST), Tehran, Iran

[5] Institute of Medical Genetics, University of Zurich, Zurich, Switzerland

[6] Novel Diagnostics and Therapeutics Research Group, Institute of Biotechnology, Ferdowsi University of Mashhad, Mashhad, Iran

[7] UMR INSERM U 1122, IGE-PCV, Interactions Gène-Environment En Physiopathologie Cardiovascular Université De Lorraine, Nancy, France

development of high-throughput sequencing has permitted the exploration of causes in genetically heterogeneous pathologies such as NOA [5].

It has been previously reported that mutations in ZMM proteins (an acronym for yeast proteins Zip1/Zip2/Zip3/Zip4, Msh4/Msh5, and Mer3 and also known as the synapsis initiation complex (SIC)), such as testis expressed 11 (TEX11), helicase for meiosis 1 (HFM1), and shortage in chiasmata 1 (SHOC1), affect spermatogenesis. ZMM proteins play a role as a link between recombination and formation of synaptonemal complex (SC) [6]. HFM1, the human homolog of yeast Mer3, encodes an ATP-dependent DNA helicase which is necessary for homologous recombination and has specific expression in germ cells [7].

HFM1-null mice of both genders are infertile due to incomplete synapsis between homologous chromosomes in meiosis I and subsequent apoptosis at diakinesis [8]. A recent study identified compound heterozygous variants of this gene in two sisters and another unrelated woman with primary ovarian insufficiency (POI) [9]. HFM1 variants have also been reported in azoospermia or severe oligozoospermia [10].

In this study, we analyzed the genetic pathology of NOA in an Iranian family comprising four infertile males (Fig. 1). Whole exome sequencing (WES) revealed a novel splice-acceptor mutation within intron 15/38 of HFM1 gene (c.1832-2A>T), which co-segregated with the phenotype in this family. Additionally, we investigated the effect of this splice-acceptor mutation on the HFM1 protein structure and function in silico. The three-dimensional (3D) structures of wild-type HFM1 and mutants were predicted by molecular modeling with the I-TASSER method [11]. In order to assess the structural differences in interacting behaviors of wild-type HFM1 and mutants, we performed a protein-ligand docking strategy [12].

Materials and methods

Subjects

A 28-year-old Iranian patient affected with NOA (Fig. 1, IV-4) joined our study for genetic analysis. He had three other affected relatives (Fig. 1, III-6, IV-7, and IV-12) and his family stated that many other infertile males live in their village. Since the patients were born to healthy parents in a village with a high degree of consanguineous marriage, we suspected that the mode of inheritance of azoospermia in this family was autosomal recessive. Written informed consent was obtained from all participants.

The proband IV-4 had no history of varicocele, cryptorchidism, genital trauma or infection, and chemo- or radiotherapy. Levels of follicle-stimulating hormone (FSH), luteinizing hormone (LH), and testosterone were normal. Semen analysis according to WHO guidelines indicated azoospermia and testicular sperm extraction (TESE) confirmed NOA in this individual. Karyotyping and Y chromosome microdeletion (YCMD) tests were undertaken for patient IV-4.

Whole exome sequencing and data analysis

Genomic DNA was isolated from blood by a salting out method. DNA samples were sent to the Novogene company for obtaining genetic data. Data processing of WES was performed according to the Celse et al. study [13]. Variant Effect Predictor v92 (VEP) was used to annotate the variants and predict their impact [14]. We excluded low-impact variants and prioritized high-impact ones (e.g., missense, frameshift, stop-gain, stop-lost, start-lost, and splice variants). The expression profile was retrieved

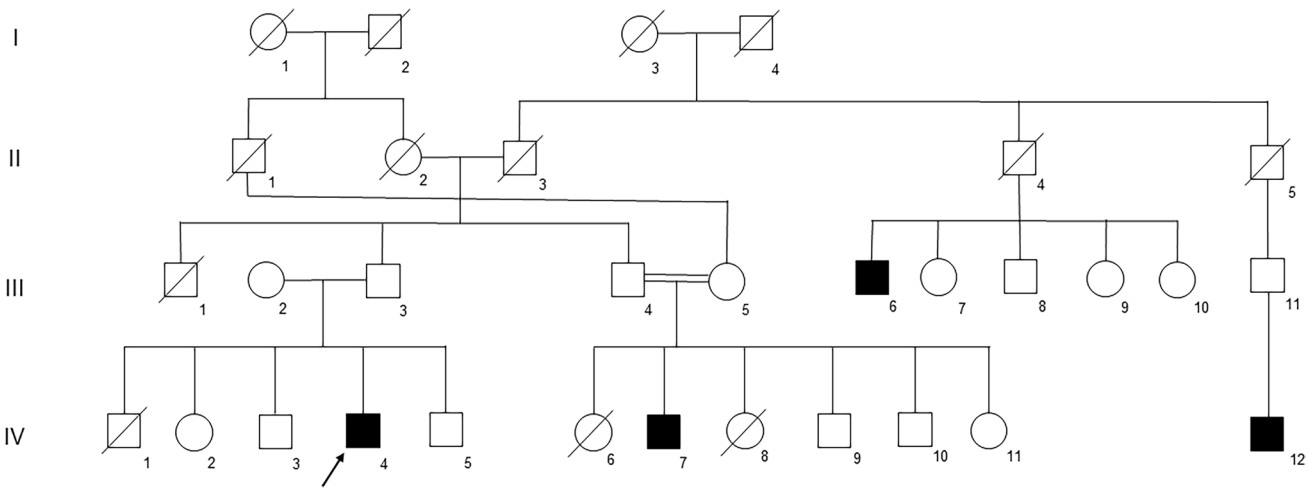

Fig. 1 Pedigree of the family with four cases of azoospermia. Black squares indicate infertile men. Arrow shows proband

from the Genotype-Tissue Expression project (GTEx v7). We excluded variants with a minor allele frequency (MAF) >0.01 in gnomAD v2.0 and the Iranome database, >0.03 in 1000 Genomes Project, and >0.05 in NHLBI Exome Sequencing Project, or observed in an in-house exome dataset (250 healthy individuals or patients with a different phenotype). Finally, we selected the variants within any of the genes in a list of 142 NOA candidate genes established by Ray et al. (submitted to *European Urology* in 2021) (Fig. S1; Table S1). These 142 genes present an elevated expression in testis and their variants have been reported either in human, knock-out mice, or more distant animal models exhibiting NOA. We used NNSPLICE 0.9 version (https://www.fruitfly.org/seq_tools/splice.html) for the assessment of deleterious splicing variants.

Sanger verification of candidate variant

Sanger verification of the candidate variant c.1832-2A>T in the HFM1 gene in proband and all available family members was carried out using an Applied Biosystems 3130XL Genetic Analyzer. Analyses were performed using Chromas 2.6.6 software. Sequences of the primers were as follows: forward 5′-AAAAGGAAAAAGCGCAAGTG-3′ and reverse 5′-CGACCAGCTCTACCAATCATC-3′.

In silico analysis and tertiary structure prediction

NNSPLICE was used for the prediction of probable cryptic splice-acceptor sites (3′CSS) within intron 15 and exon 16 of the HFM1 gene. Modeling of the 3D structures of wild-type and two probable mutant variants (exon 16 skipping and intron 15 inclusion) was predicted by I-TASSER web server.

I-TASSER employs numerous threading approaches to look PDB for recognizable proof of structural template and builds a full-length atomic model through iterative template fragment assembly simulations. The server is ranked the best for protein structure prediction method by Critical Assessment of Protein Structure Prediction [11, 15, 16]. I-TASSER normalized *Z*-score >1 illustrates a template with a good alignment quality. Confidence-score (*C*-score) (values of −5 to 2) reflects the confidence of each model. A higher *C*-score value means a higher confidence. The template modeling score (TM-score) >0.5 indicates correct topology of the model, while score <0.17 corresponds to a model of random topology [15]. These scores are used to select the best ranked model. Therefore, models with the highest normalized *Z*-score, *C*-score, and TM-score were chosen for further evaluations.

Verification of the predicted 3D structures

The modeled structures were verified by structural analysis and verification server (SAVES v6.0) (https://saves.mbi.ucla.edu/). Among different programs of SAVES metaserver, two programs: ERRAT [17–19] and Ramachandran plot analysis [20, 21] were used. The precision of the predicted models was assessed by the PSi/Phi Ramachandran plot analysis.

Binding site prediction and molecular docking

COACH web server predicted the functions of both wild-type and mutant variants based on the I-TASSER-predicted structures [22]. The interaction between the HFM1 and mutant variants with adenosine diphosphate (ADP) was assessed by molecular docking. AutoDock Vina [23] was used by blind docking approach. The ADP crystal structures were acquired via complex structure (PDB: 4kit). AutoDock Vina on PyRx platform was used to find the binding energy of the HFM1 and its mutant variants with ligand [12]. Ten best poses were generated and scored using AutoDock Vina scoring functions. PyMOL Molecular Graphics System v.1.3 was used to investigate the interactions of the complexes and the interaction features were illustrated (Schrodinger, LLC).

Results

Routine genetic tests (analysis of karyotype and YCMD) for the proband (IV-4), showed a normal 46, XY male.

Whole exome sequencing

Variant filtering resulted in 80 homozygous and 1019 heterozygous variants in patient IV-4. Two of these variants were predicted as “HIGH impact” variants (not expected to produce functional protein) by VEP. Just one, the variant in HFM1, was among 142 NOA candidate genes (Table S1). HFM1 maps to chromosome 1 and patient IV-4 carried a homozygous splice-acceptor mutation (c.1832-2A>T) in this gene, which was predicted deleterious by both ada- and rf-scores (from dbscSNV). NNSPLICE predicted that this variant would disrupt the splice-acceptor site (3′SS) of intron 15 in the HFM1 gene (Table 1).

Table 1 Splice site prediction for wild-type and mutant HFM1 by NNSPLICE 0.9 version

Donor site predictions for wild-type HFM1					Donor site predictions for mutant HFM1				
Start	End	Score	Exon	Intron	Start	End	Score	Exon	Intron
146	160	0.24	tgctaca**gt**aaattc		146	160	0.24	tgctaca**gt**aaattc	
237	251	0.94	atttaag**gt**gtgtta		237	251	0.94	atttaag**gt**gtgtta	
Acceptor site predictions for wild-type HFM1					Acceptor site predictions for mutant HFM1				
Start	End	Score	Intron	Exon	Start	End	Score	Intron	Exon
148	188	0.12	ctacagtaaattctttgg**tag**tgtttagttctttatcatat		148	188	0.12	ctacagtaaattctttggt**ag**tgtttagttctttatcatat	
223	263	0.16	taataggtttttggatttaa**gg**tgtgttactattttagatt		223	263	0.16	taataggtttttggatttaa**gg**tgtgttactattttagatt	
240	280	0.26	taaggtgtgttactattttt**ag**attgtaatgctctttttct		240	280	0.26	taaggtgtgttactattttt**ag**attgtaatgctctttttct	
379	419	0.16	ctatgtcttaatctgatgc**ag**ttactaccagtactttagct		-	-	-	-	
397	437	0.11	cagttactaccagtactttt**ag**ctatgggagtaaatttgcct		397	437	0.23	ctgttactaccagtactttt**ag**ctatgggagtaaatttgcct	
430	470	0.13	atttgcctgctcacctagt**ag**ttataaaatctacaatgcat		430	470	0.13	atttgcctgctcacctagt**ag**ttataaaatctacaatgcat	

Exon/intron boundary shown in bold

https://www.fruitfly.org/seq_tools/splice.html

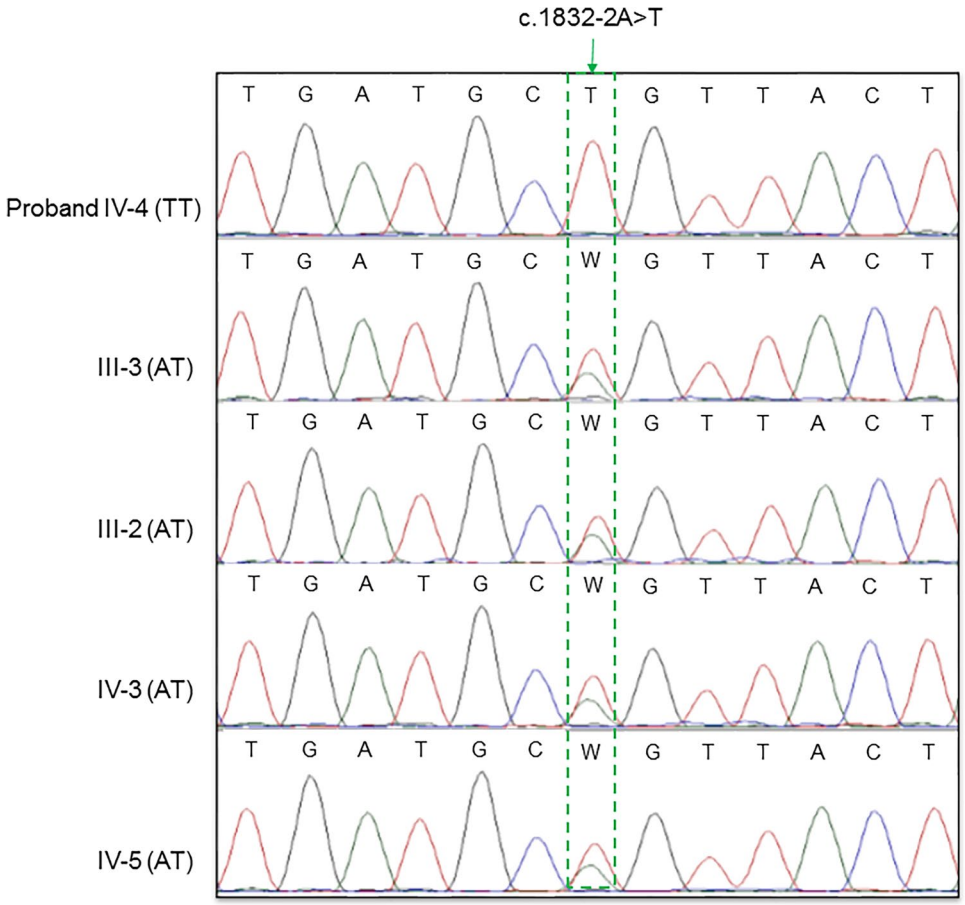

Fig. 2 Sanger sequencing electropherograms for the c.1832-2A>T mutation in the HFM1 gene. TT, homozygous mutation; AT, heterozygous mutation

Validation of the variant

Sanger sequencing verified the mutation c.1832-2A>T in the HFM1 gene in the family members. The proband was homozygous, while his parents and the two healthy brothers were heterozygous, consistent with autosomal recessive inheritance (Fig. 2).

Prediction of cryptic splice sites

NNSPLICE predicted five 3′SSs within intron 15 and exon 16 other than the authentic 3′SS, but none of them had the necessary conditions to be used as 3′CSS (Table 1).

Prediction and validation of 3D structures

I-TASSER identified human Brr2 Helicase Region S1087L (PDB: 4f92B) and crystal structure of yeast full-length Brr2 in complex with Prp8 Jab1 domain in *Saccharomyces cerevisiae* S288C (PDB: 5dcaA) as the most similar structural templates for modeling of wild-type and mutant variants of HFM1, respectively. Top ranking models with the highest *C*-scores: −1.53, −0.95, and −1.07, and TM-scores: 0.53±0.15, 0.59±0.14, and 0.48±0.15 for wild-type, exon skipping mutant and intron inclusion mutant are shown in Fig. 3 a–c, respectively. These three models were further verified through overall quality factor of the ERRAT and Ramachandran plot (Table 2).

Identification of binding site residues

The COACH server was utilized to predict the ligand binding sites in the wild-type HFM1 protein. This server showed very close match with 4kitB (*C*-score = 0.41) and predicted the following putative ADP binding site residues: TYR280, GLN285, GLY308, LYS309, THR310, GLU412, ASN621. The COACH server detected no ADP binding site for exon 16 skipping and intron 15 inclusion mutants.

Docking analysis of wild-type HFM1 and mutants with ADP

The interactions and ligand binding sites in the wild-type HFM1 and mutants were firstly identified using the COACH server. For further verification and analysis of the interactions of wild-type HFM1 and mutant structures with ADP, ligand docking interactions were assessed using AutoDock Vina scoring functions. The best poses from AutoDock are shown in Fig. 4. The active cavity of the wild-type HFM1 and mutants is being characterized with various amino acid residues within 4Å around ADP ligand which the numbers were 20 for wild-type HFM1 and 17 for both mutants. Table 3 compares the amino acid residues within 4Å around ADP as the active cavity of wild-type HFM1 and exon 16 skipping and intron 15 inclusion mutants. The ADP ligand forms hydrogen bonds with residues GLY306, ARG346, ASP349, and GLN650 in the binding site of wild-type HFM1 and residues TYR83, SER85, THR87, ASP102, GLY108, and ASP145 in the binding site of exon 16 skipping mutant and residues THR87, GLN101, and LEU104 in the binding site of intron 15 inclusion mutant. As shown in Table 3 and Fig. 4, the interactions between ligand and mutant

Table 2 ERRAT and Ramachandran analysis for wild-type HFM1 and mutant forms

Protein	ERRAT score* (%)	RAMPAGE		
		Favored	Allowed	Outliers
Wild-type	84.69	65.3%	28.4%	2.4%
Exon 16 skipping mutant	76.13	65.2%	28.1%	2.7%
Intron 15 inclusion mutant	86.73	70.9%	24.1%	3.4%

*Overall quality factor generated by ERRAT server

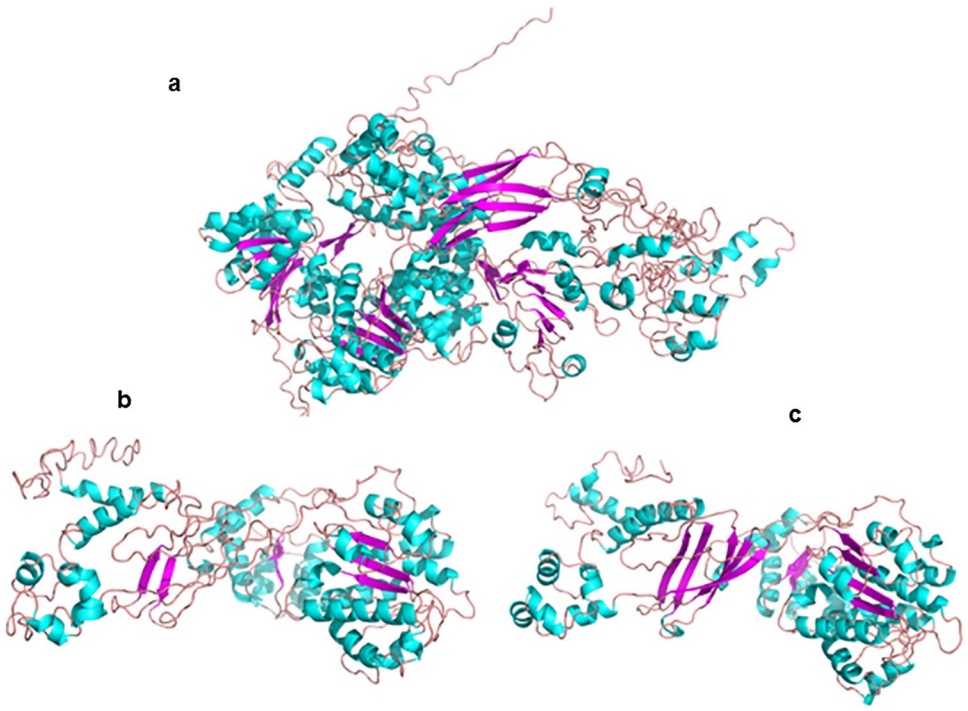

Fig. 3 I-TASSER structural models of wild-type HFM1 and mutants. 3D structures of **a** wild-type HFM1; **b** exon 16 skipping mutant; and **c** intron 15 inclusion mutant

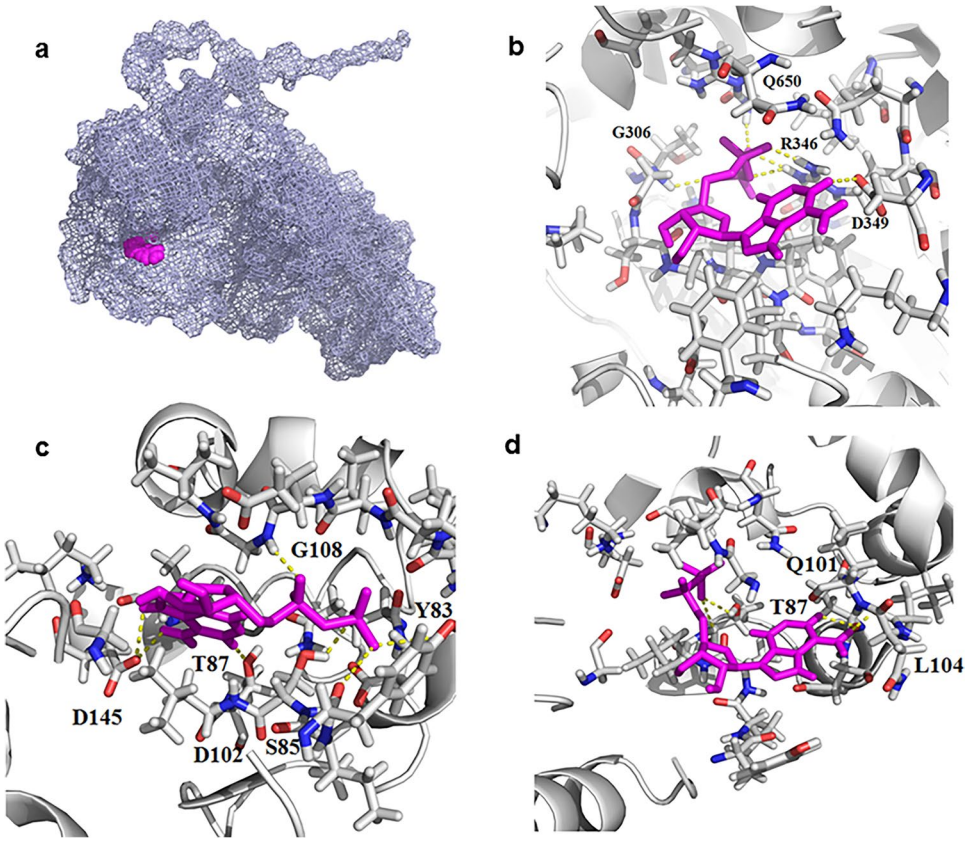

Fig. 4 Structures of the wild-type HFM1 and exon skipping and intron inclusion mutants in complex with ADP. **a** Mesh model of the 3D structure of wild-type HFM1 in complex with ADP ligand. **b** Structure of the wild-type HFM1 in complex with ADP. **c** Structure of the exon skipping mutant in complex with ADP. **d** Structure of the intron inclusion mutant in complex with ADP. Parts **b**, **c**, and **d** show the structure of active cavity of wild-type HFM1 and exon skipping and intron inclusion mutants in complex with ADP. Contact residues within 4Å of the ligand are shown in stick form. In each complex, the residues which formed hydrogen bonds with ADP are labeled. The ADP is in magenta. The hydrogen bonds are shown in the form of a yellow dashed line

Table 3 Contact residues within 4Å of the ADP ligand. The amino acid residues participating in the interaction at the binding pocket of wild-type HFM1 and mutants. The residues in bold underlined font were used as binding site residues that formed hydrogen bond for docking

Protein	Contact residues within 4Å of the ADP ligand
Wild-type	PHE278, GLN285, THR305, **GLY306**, SER307, GLY308, LYS309, THR310, VAL311, PHE315, LEU342, GLN345, **ARG346**, **ASP349**, LYS353, ASP411, PRO488, ASN621, **GLN650**, ARG654
Exon 16 skipping	**TYR83**, ILE84, **SER85**, LEU86, **THR87**, GLN101, **ASP102**, LEU104, ASN105, LEU106, GLU107, **GLY108**, V109, V142, **ASP145**, LYS147, GLY232
Intron 15 inclusion	TYR83, ILE84, SER85, LEU86, **THR87**, **GLN101**, ASP102, **LEU104**, ASN105, LEU106, GLY108, V109, ASN111, V142, ASP145, LYS147, LY232

Table 4 The positional binding energy interactions of wild-type HFM1 and mutant forms with the ADP as ligand

Pose mode	Binding affinity (kcal/mol)		
	Wild-type	Exon 16 skipping	Intron 15 inclusion
1	−7.9	−8.1	−7.7
2	−7.5	−7.9	−7.6
3	−7.4	−7.9	−7.5
4	−7.4	−7.7	−7.4
5	−7.4	−7.5	−7.3
6	−7.4	−7.1	−7.3
7	−7.3	−7.0	−7.3
8	−7.1	−7.0	−7.2
9	−7.0	−7.0	−7.2

protein residues were different from those in wild-type HFM1. The number of interacting residues and hydrogen bonds with ADP were different between wild-type and the two probable mutants. The binding affinities for the best binding conformation with ADP for wild-type HFM1 and exon 16 skipping and intron 15 inclusion mutants were −7.9, −8.1, and −7.7 kcal/mol, respectively (Table 4). According to Table 4 and due to the best positional binding energy, the built model of exon 16 skipping mutant had a higher binding affinity and the mutant of intron 15 inclusion had a lower binding affinity to ADP, in comparison with the interaction of wild-type HFM1.

Discussion

Known genetic factors are implicated in 21–29% of NOA. However, 12–41% of this genetically heterogeneous disorder is idiopathic [24]. WES is a powerful method for the identification of novel genes in genetic diseases. Many studies have reported causal genes such as TEX11, SHOC1, HFM1, SYCE1, SYCP2, and SYCP3 in azoospermia which are involved in the formation of crossover and synapsis between homologous chromosomes during meiosis [25–30].

In this study, we identified a novel mutation (c.1832-2A>T) in the HFM1 gene which co-segregated with the NOA in a family with four affected individuals. The proband was homozygous, while his parents and two brothers were heterozygous.

HFM1 encodes an ATP-dependent DNA helicase with specific expression in the testis and ovary which is required to form crossovers and proper synapsis between homologous chromosomes [7]. The transcript containing variant c.1832-2A>T has 39 exons and its protein comprises 1435 amino acids (aa). Hitherto, pathogenic variants of HFM1 are observed in POI, azoospermia, and oligozoospermia [9, 10]. HFM1-null mice of both genders are sterile, because of maturation arrest at meiosis I and apoptosis at diakinesis [8].

Most splice site mutations lead to three probable consequences: (1) exon skipping, (2) intron inclusion (for small introns), or (3) using a CSS [31]. About the transition of A>T in the 3′SS of intron 15/38 of the HFM1 gene, we evaluated the probability of using 3′CSS. A CSS must be strong enough and be mapped within the 100 nucleotide region from the authentic splice site to be activated following the mutation in the authentic splice site [32, 33]. Regarding c.1832-2A>T mutation, none of the predicted 3′SSs by NNSPLICE had the property to be used instead of authentic 3′SS (Table 1). Therefore, the variant c.1832-2A>T in the HFM1 gene would cause just two probable consequences: (1) exon 16 skipping, or (2) intron 15 inclusion (a small intron with 399 nucleotides) (Fig. 5).

Skipping of exon 16 or inclusion of intron 15 of the HFM1 gene would lead to in-frame stop codons, TGA or TAA, respectively. These premature termination codons (PTCs) are resided ≥ 50–55 nucleotides upstream of the exon-exon junction. These events are predicted to result in two probable consequences: (1) nonsense-mediated mRNA decay (NMD), or (2) defective proteins (611aa or 612aa in exon skipping or intron inclusion, respectively). As the helicase C-terminal domain of the HFM1 protein is located between amino acids 519 and 720 and the asparagine 621

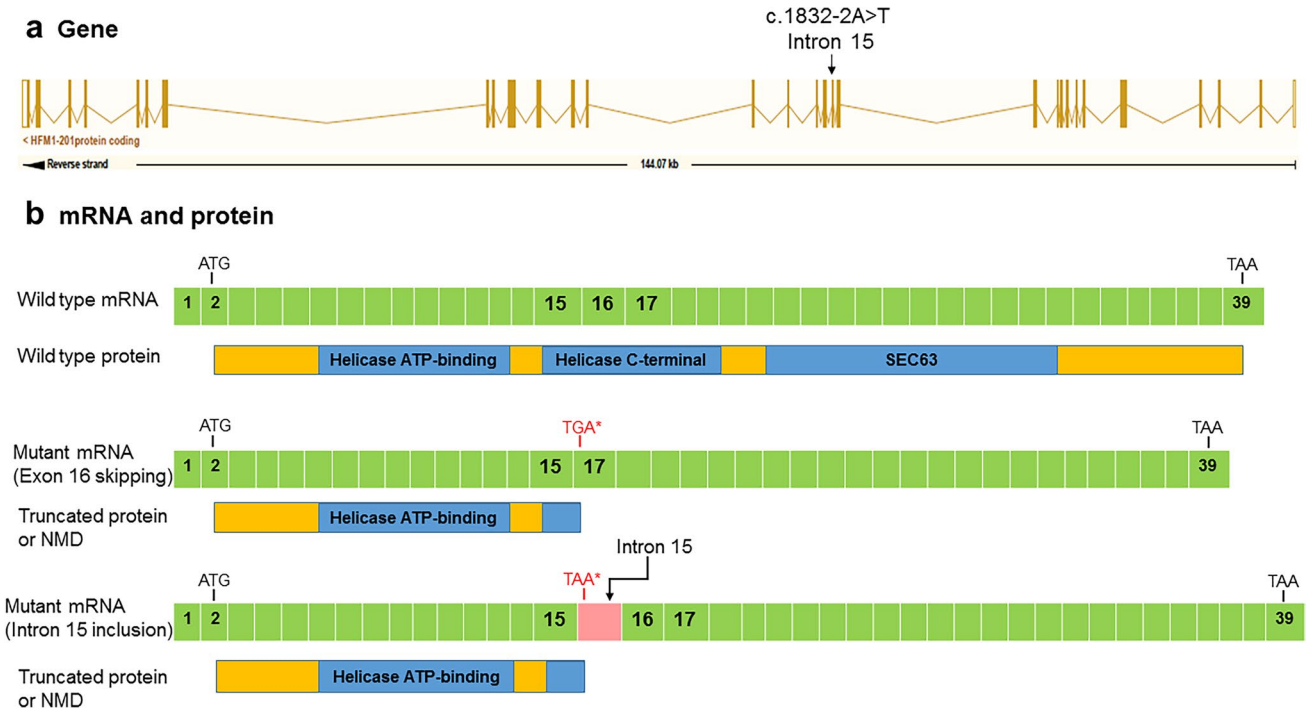

Fig. 5 Schematic view of the HFM1 gene, mRNA, and protein. **a** Location of variant c.1832-2A>T within the HFM1 gene. **b** Schematic view of wild-type and mutant HFM1 mRNAs and proteins represented with domains. The boxes show exons and the numbers indicate the exons numbers. The stars show premature termination codons (TGA and TAA) in two probable mutant mRNAs (exon skipping and intron inclusion)

(N621) is predicted as a residue in the ADP binding site of the wild-type HFM1 protein by both COACH and molecular docking approaches, the hydrolysis of ATP in the reaction [ATP + H_2O <=> ADP + phosphate], which is necessary for helicase activity of HFM1, cannot take place (https://www.uniprot.org/uniprot/A2PYH4/protvista). Therefore, the mutant protein will not be able to bind to ADP and form crossover and complete synapsis between homologous chromosomes during meiosis. Ultimately, defective sperm production and NOA may ensue.

We showed that a heterogeneous pathology such as NOA can be well diagnosed by WES. Here we provided another evidence for the importance of prescription of WES to NOA patients. Regarding that the genes involved in the formation of crossovers and synapsis have critical roles in the production of germ cells, we suggest that such genes be further studied in cases of infertility.

We performed a computational analysis by modeling and molecular docking approaches to investigate the importance of structural and functional features of the HFM1 gene. For the structural analysis, we modeled the wild-type HFM1 and mutant structures by I-TASSER (Fig. 3). Before performing the docking analysis, ligand binding sites of the wild-type HFM1 and two probable mutants were identified by the COACH server [22]. Then, docking analysis was done between the wild-type HFM1 and mutant structures with ADP. Binding energies between the mutant proteins and ligand indicated a lower binding affinity of the intron 15 inclusion mutant and higher binding affinity of the exon 16 skipping mutant to the ADP ligand compared to the wild-type HFM1. It is worth noting that ADP interacting residues differed in both mutant proteins compared to the wild-type HFM1 (Table 3; Fig. 4). These results demonstrate that the c.1832-2A>T mutation leads to structural changes in the wild-type HFM1. Functional property of the protein was revealed by the docking of ADP with wild-type HFM1 and mutants, indicating the effect of mutation on protein-ligand complex formation. The aim of molecular docking was to evaluate the binding affinity, hydrogen bonds, and interactions between the ADP ligand with the 3D structures of the wild-type HFM1 and mutants. The results indicated that c.1832-2A>T mutation in wild-type HFM1 reduced the number of hydrogen bonds with ADP from 6 to 4 in the intron 15 inclusion mutant and increased the number of hydrogen bonds with ADP from 6 to 7 in the exon 16 skipping mutant. It was highlighted that c.1832-2A>T mutation affects not only the ADP binding residues but also the hydrogen bond interactions.

The computational approaches such as protein modeling and protein-ligand docking provide the opportunity for the prediction of the 3D structure of biological compounds, their interactions, and the involved biological pathways and even simulating the physiological conditions of these biological pathways [34].

Conclusion

We identified a novel splice-acceptor mutation (c.1832-2A>T) in the HFM1 gene. The c.1832-2A>T mutation affects ADP binding residues, potentially causing defective crossover and synapsis. Regarding that many reported NOA genes are involved in the formation of crossovers and synapsis and have critical roles in the production of germ cells, we suggest that such genes can be considered for studying of infertility among large cohorts of infertile cases.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10815-022-02433-z.

Acknowledgements The authors thank the patients and their family members who participated in this study. We appreciate the contribution of Ferdowsi University of Mashhad and Neyshabur University of Medical Sciences to provide facilities to accomplish this study. We also appreciate Novin Fertility Center for introducing patients to our study. We gratefully acknowledge Professor Pierre F. Ray and Dr. Reza Maroofian for the financial support, Nicolas Thierry-Mieg for his support in bioinformatics concerns, and Abolfazl Rad for his help in extraction and sending DNA samples for WES. The authors thank the management of advanced computational center, Khayyam Innovation Ecosystem, Mashhad, Iran for providing the facilities and encouragement to carry out this research work.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Rowe PJ, et al. WHO manual for the standardized investigation, diagnosis and management of the infertile male / Patrick J. Rowe ... [et al.]. Geneva: World Health Organization; 2000.
2. Tournaye H, Krausz C, Oates RD. Novel concepts in the aetiology of male reproductive impairment. Lancet Diabetes Endocrinol. 2017;5(7):544–53.
3. Saebnia N, Neshati Z, Bahrami A. Role of microRNAs in etiology of azoospermia and their application as non-invasive biomarkers in diagnosis of azoospermic patients. J Gynecol Obstet Hum Reprod. 2021:102207.
4. Araujo TF, et al. Sequence analysis of 37 candidate genes for male infertility: challenges in variant assessment and validating genes. Andrology. 2020;8(2):434–41.
5. Kherraf Z-E, et al. SPINK2 deficiency causes infertility by inducing sperm defects in heterozygotes and azoospermia in homozygotes. EMBO Mol Med. 2017;9:1132–49.
6. Lynn A, Soucek R, Börner GV. ZMM proteins during meiosis: crossover artists at work. Chromosom Res. 2007;15(5):591–605.
7. Tanaka K, et al. HFM1, the human homologue of yeast Mer3, encodes a putative DNA helicase expressed specifically in germline cells. DNA Seq. 2006;17(3):242–6.

8. Guiraldelli MF, et al. Mouse HFM1/Mer3 is required for crossover formation and complete synapsis of homologous chromosomes during meiosis. PLoS Genet. 2013;9(3):e1003383.
9. Wang J, et al. Mutations in HFM1 in recessive primary ovarian insufficiency. N Engl J Med. 2014;370(10):972–4.
10. Zhang W, et al. Association analysis between HFM1 variations and idiopathic azoospermia or severe oligozoospermia in Chinese men. Sci China Life Sci. 2017;60(3):315–8.
11. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008;9(1):1–8.
12. Gholampour-Faroji N, et al. Modeling, stability and the activity assessment of glutathione reductase from Streptococcus thermophilus; insights from the in-silico simulation study. Comput Biol Chem. 2019:107121.
13. Celse T, et al. Genetic analyses of a large cohort of infertile patients with globozoospermia, DPY19L2 still the main actor, GGN confirmed as a guest player. Hum Genet. 2021;140(1):43–57.
14. McLaren W, et al. The Ensembl variant effect predictor. Genome Biol. 2016;17(1):122.
15. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc. 2010;5(4):725–38.
16. Yang J, et al. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 2015;12(1):7–8.
17. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 1993;2(9):1511–9.
18. Bowie JU, Lüthy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure. Science. 1991;253(5016):164–70.
19. Lüthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. Nature. 1992;356(6364):83–5.
20. Carugo O, Djinovic-Carugo K. Half a century of Ramachandran plots. Acta Crystallogr Sect D. 2013;69(8):1333–41.
21. Pontius J, Richelle J, Wodak SJ. Deviations from standard atomic volumes as a quality measure for protein crystal structures. J Mol Biol. 1996;264(1):121–36.
22. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Res. 2015;43:W174–81.
23. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31(2):455–61.
24. Arafat M, et al. Mutation in TDRD9 causes non-obstructive azoospermia in infertile men. J Med Genet. 2017;54(9):633–9.
25. Maor-Sagie E, et al. Deleterious mutation in SYCE1 is associated with non-obstructive azoospermia. J Assist Reprod Genet. 2015;32(6):887–91.
26. Schilit SLP, et al. SYCP2 Translocation-mediated dysregulation and frameshift variants cause human male infertility. Am J Hum Genet. 2020;106(1):41–57.
27. Miyamoto T, et al. Azoospermia in patients heterozygous for a mutation in SYCP3. Lancet. 2003;362(9397):1714–9.
28. Yang F, et al. Meiotic failure in male mice lacking an X-linked factor. Genes Dev. 2008;22(5):682–91.
29. Yatsenko AN, et al. X-linked TEX11 mutations, meiotic arrest, and azoospermia in infertile men. N Engl J Med. 2015;372(22):2097–107.
30. Guiraldelli MF, et al. SHOC1 is a ERCC4-(HhH)2-like protein, integral to the formation of crossover recombination intermediates during mammalian meiosis. PLoS Genet. 2018;14(5):e1007381.
31. Anna A, Monika G. Splicing mutations in human genetic disorders: examples, detection, and confirmation. J Appl Genet. 2018;59(3):253–68.
32. Kananura C, et al. A splice-site mutation in GABRG2 associated with childhood absence epilepsy and febrile convulsions. Arch Neurol. 2002;59(7):1137–41.
33. de Boer M, et al. Activation of cryptic splice sites in three patients with chronic granulomatous disease. Mol Genet Genomic Med. 2019;7(9):e854–4.
34. Gholampour-Faroji N, et al. Modeling, mutagenesis and in-silico structural stability assay of the model of superoxide dismutase of Lactococcus lactis subsp. cremoris MG1363. Iran J Biotechnol. 2020;18(1):32–41.