ARTICLE

# Bi-allelic Recessive Loss-of-Function Variants in *FANCM* Cause Non-obstructive Azoospermia

Laura Kasak,[1] Margus Punab,[2] Liina Nagirnaja,[3] Marina Grigorova,[1] Ave Minajeva,[1] Alexandra M. Lopes,[4,5] Anna Maria Punab,[1] Kenneth I. Aston,[6] Filipa Carvalho,[5,7] Eve Laasik,[1] Lee B. Smith,[8,9] GEMINI Consortium, Donald F. Conrad,[3,10] and Maris Laan[1,*]

Infertility affects around 7% of men worldwide. Idiopathic non-obstructive azoospermia (NOA) is defined as the absence of spermatozoa in the ejaculate due to failed spermatogenesis. There is a high probability that NOA is caused by rare genetic defects. In this study, whole-exome sequencing (WES) was applied to two Estonian brothers diagnosed with NOA and Sertoli cell-only syndrome (SCOS). Compound heterozygous loss-of-function (LoF) variants in *FANCM* (Fanconi anemia complementation group M) were detected as the most likely cause for their condition. A rare maternally inherited frameshift variant p.Gln498Thrfs*7 (rs761250416) and a previously undescribed splicing variant (c.4387−10A>G) derived from the father introduce a premature STOP codon leading to a truncated protein. FANCM exhibits enhanced testicular expression. In control subjects, immunohistochemical staining localized FANCM to the Sertoli and spermatogenic cells of seminiferous tubules with increasing intensity through germ cell development. This is consistent with its role in maintaining genomic stability in meiosis and mitosis. In the individual with SCOS carrying bi-allelic *FANCM* LoF variants, none or only faint expression was detected in the Sertoli cells. As further evidence, we detected two additional NOA-affected case subjects with independent *FANCM* homozygous nonsense variants, one from Estonia (p.Gln1701*; rs147021911) and another from Portugal (p.Arg1931*; rs144567652). The study convincingly demonstrates that bi-allelic recessive LoF variants in *FANCM* cause azoospermia. *FANCM* pathogenic variants have also been linked with doubled risk of familial breast and ovarian cancer, providing an example mechanism for the association between infertility and cancer risk, supported by published data on *Fancm* mutant mouse models.

## Introduction

Male infertility is a complex pathology affecting approximately 7% of men worldwide, with the highest prevalence (up to 12%) among Europeans.[1] Azoospermia (MIM: 415000), defined as the total absence of spermatozoa in the ejaculate, represents the most extreme form of male infertility. Despite its severe clinical phenotype, the condition is relatively common among cases diagnosed with male factor infertility. In a recent prospective clinical-epidemiological study among Estonian men attending the Andrology Centre over a 9-year period, azoospermia-affected case subjects represented 22% of 1,737 men detected with reduced total sperm counts.[2] The two major currently assigned causes of azoospermia are seminal tract obstruction due to physical blockage in the genital tract (25% of case subjects) and known genetic factors such as chromosomal aberrations, Y chromosome microdeletions, and *CFTR* (MIM: 602421) mutations (25%). Today's andrological workup is able to assign a clinical cause for ~80% of azoospermic men.[2] For the unexplained non-obstructive azoospermia (NOA)-affected case subjects (20%), rare genetic defects affecting the complex process of spermatogenesis are highly likely. A mutation in any of the more than 1,000 genes estimated to be male germ cell specific[3,4] could cause defective spermatogenesis in otherwise healthy men. An additional challenge to identify genetic mutations causing NOA stems from their extreme clinical effect, i.e., inability of affected persons to contribute to the next generation. Thus, family-based analysis that has been successfully utilized for mapping mutations causing other rare monogenetic diseases is not widely applicable for azoospermia-affected case subjects.

Whole-exome sequencing (WES) is a state-of-the-art method that has shifted genetic mutation screens from single candidate genes to the entire coding portion of the genome. It is an especially relevant approach for NOA with an expected wide palette of mutated genes. WES data analysis resulting in the full mutational spectrum across an individual exome has the potential to uncover novel genetic contributors to azoospermia. So far, nine WES analyses of azoospermia-affected case subjects have been performed in 18 consanguineous pedigrees revealing recessive homozygous family-specific variants in a diverse spectrum of genes, such as *TAF4B* (MIM: 601689), *ZMYND15* (MIM: 614312), *TEX15* (MIM: 605795), *NPAS2* (MIM: 603347), *MEIOB* (MIM: 617670), *TEX14* (MIM: 605792; two independent studies), *DNAH6* (MIM: 603336), *DMC1* (MIM: 602721), *TDRD9* (MIM: 617963), *NANOS2* (MIM: 608228), *SPO11* (MIM: 605114), *CCDC155*, and *SPINK2* (MIM: 605753).[5–13] Of these, *TEX15* is the only gene in which recurrent

[1]Institute of Biomedicine and Translational Medicine, University of Tartu, Tartu 50411, Estonia; [2]Andrology Centre, Tartu University Hospital, Tartu 50406, Estonia; [3]Department of Genetics, Washington University School of Medicine, St. Louis, MO 63110, USA; [4]IPATIMUP - Institute of Molecular Pathology and Immunology of the University of Porto, Porto 4200-135, Portugal; [5]i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto 4200-135, Portugal; [6]Division of Urology, Department of Surgery, University of Utah School of Medicine, Salt Lake City, UT 84132, USA; [7]Serviço de Genética, Departamento de Patologia, Faculdade de Medicina da Universidade do Porto, Porto 4200-319, Portugal; [8]MRC Centre for Reproductive Health, University of Edinburgh, Edinburgh EH16 4TJ, UK; [9]School of Environmental and Life Sciences, University of Newcastle, Callaghan, NSW 2308, Australia; [10]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA
*Correspondence: maris.laan@ut.ee
https://doi.org/10.1016/j.ajhg.2018.07.005.

mutations have been identified also in a non-consanguineous family, where compound heterozygous variants were detected in two siblings diagnosed with NOA.[14] In addition, WES approach has uncovered X-linked mutations in *MAGEB4* (MIM: 300153), *TEX11* (MIM: 300311), and *WNK3* (MIM: 300358) causative to NOA in single families.[7,13,15] Although some of the WES studies have included functional work and/or screening for additional case subjects,[6,8,11,12] the majority represent reports of a single affected individual or family without confirming the functional effect of the claimed mutation. Only for *TEX11*, there is further evidence from independent studies.[16,17]

In the current study, we applied WES to identify the genetic cause of infertility in two Estonian brothers with NOA and Sertoli cell-only syndrome (SCOS [MIM: 305700, 400042]). We detected compound heterozygous loss-of-function (LoF) variants in *FANCM* (Fanconi anemia complementation group M [MIM: 609644]) as the most likely cause for azoospermia in both brothers. One of the variants represents an extremely rare nonsense variant (p.Gln498Thrfs*7; rs761250416; gnomAD database: minor allele frequency [MAF] = $7.22 \times 10^{-5}$) and the other a previously undescribed variant with predicted pathogenic effect on splicing (c.4387−10A>G [p.Arg1436_Ser1437insLeuLeu*]). Immunohistochemical staining of testicular tissue sections revealed none or minimal FANCM protein expression in the seminiferous tubules of the index case subject compared to abundant levels detected in the Sertoli and spermatogenic cells of control subjects. As further evidence, we detected two additional NOA-affected case subjects with independent *FANCM* homozygous nonsense variants, one from Estonia (p.Gln1701* [c.5101C>T]; rs147021911; one homozygote among 138,488 gnomAD exomes/genomes; MAF = $1.34 \times 10^{-3}$) and another from Portugal (p.Arg1931* [c.5791C>T]; rs144567652; no homozygotes/138,536 individuals; MAF = $1.03 \times 10^{-3}$). These data further strengthen the causative link between bi-allelic recessive *FANCM* variants and NOA.

## Subjects and Methods

### Ethics Statement

The study was approved by the Ethics Review Committee of Human Research of the University of Tartu, Estonia (approval date 254/M-17, 21.12.2015) and PTDC/SAU-GMG/101229/2008, approved by the INSA Ethics Committee (Lisbon, Portugal on 6 November 2007). Written informed consent for evaluation and use of their clinical data for scientific purposes was obtained from each person (and their parents, if available) prior to recruitment. The study was carried out in compliance with the Helsinki Declaration.

### Clinical Profiling of Two Brothers with Idiopathic Male Infertility

The two index case subjects are brothers (B.1, B.2) diagnosed by M.P. with idiopathic non-obstructive azoospermia (NOA [MIM: 415000]) at the Andrology Centre, Tartu University Hospital (AC-TUH), Estonia. The applied routine andrological pipeline at the AC-TUH to document the epidemiological, laboratory, and clinical examination data of men attending the center has been described recently in detail.[2] Andrological workup of the two index case subjects excluded the following known causal factors for male infertility underlying their condition: seminal tract obstruction, chromosomal abnormalities, *AZF* microdeletions, hypogonadotropic hypogonadism, testicular diseases, androgen abuse, severe trauma or operation in the genital area, or chemo- and radiotherapy. Further details of the clinical assessment procedures are provided in Supplemental Subjects and Methods.

Family history and essential andrological data of the father of the index case subjects were collected by the managing clinical team at the AC-TUH. The two parents were native Estonians but originated from different regions of the country. Familial anamnesis excluded consanguinity. Paternal and maternal genomic DNA was extracted from the EDTA-blood in order to analyze the parent-of-origin status of the variants detected in the index brothers. In addition to standard karyotyping, the genomes of B.1, B.2, and their parents were also profiled for copy number variants (CNVs) using genome-wide SNP microarrays (Supplemental Subjects and Methods).

### Histological Profiling of Testicular Tissue in Estonian Index Case Subjects

Analysis of testicular histology of the two index case subjects was performed independently at different hospitals in two countries by the clinical pathology service centers. Testes biopsies of brother 1 (case subject B.1) were analyzed at the Pathology Centre, Diagnostic Clinic of the East Tallinn Central Hospital (Tallinn, Estonia). Brother 2 (case subject B.2) was subjected to a testicular sperm microextraction (mTESE) operation and testicular histology profiling at the Centre for Reproductive Medicine, Universitair Ziekenhuis Brussel (Brussels, Belgium). Testicular biopsy was performed as an open surgical biopsy. In both cases multiple biopsies were taken bilaterally. Tissue samples were fixed in Stieve fixative and tissue sections were stained with hematoxylin and eosin. Histologic evaluation was performed by trained histopathology experts and results classified as recommended previously.[18]

### Whole-Exome Sequencing

Whole-exome sequencing (WES) of the index case subjects B.1 and B.2 was performed at the Illumina CSPro lab, The Estonian Genome Center Core Facility, University of Tartu. Full details are provided in Supplemental Subjects and Methods. In brief, genomic DNA sequencing libraries were prepared with the Nextera Rapid Capture Exome sample preparation kit, targeting >214,000 coding exons and following the protocol provided by the manufacturer (Illumina). Nextera Rapid Capture Exome libraries were sequenced on HiSeq2500 with 100-bp paired-end reads. For >90% of sequenced bases, the achieved quality score was higher than Q30. FastQ files were subsequently trimmed and aligned to the human reference genome (hg19, GRCh37) with the Burrows-Wheeler Aligner (BWA, v.0.6.1).[19] SAMtools (v.0.1.18)[20] was used to filter out reads marked as PCR duplicates and reads not in a proper read pair. For both samples, the mapping efficiency of generated WES reads was 99% and the coverage of exonic regions with ≥10 reads was 93%. Mean coverage was 85× (B.1) and 80× (B.2). Single-nucleotide substitutions and small indel variants were called with Genome Analysis Toolkit 1.6-7-g2be5704 (GATK).[21] Variant sites >50 bp away from the

nearest exome sequence capture target or with low quality score (<Q20) were filtered out.

### Data Analysis and Variant Prioritization using PSAP Pipeline

Population sampling probability (PSAP) pipeline[22] was applied in order to prioritize potential causative variants from the WES data. It is a model-based framework to evaluate the significance of genotypes ascertained from a nuclear family or a single case subject (referred to as the "n-of-one" problem). This single-sided testing framework determines the probability of sampling a genotype or set of genotypes based on the pathogenicity scores and frequencies of variants observed in the unaffected population. Input file format for the PSAP pipeline is Variant Call Format (VCF). Currently, the data used to create the PSAP null distributions do not contain variants outside of the GencodeV19 CDS.[22]

In the current study, a PSAP option "individual analysis pipeline" was applied. After filtering of the VCF files (details in Supplemental Subjects and Methods), the pipeline resulted in ~13,000 variants in both index case subjects. The variants were further prioritized to satisfy the following criteria: (1) low minor allele frequency (MAF $\leq$ 0.01) or a previously undescribed variant; (2) the PSAP statistical significance value, popScore $\leq$ 0.005; (3) the CADD score $\geq$ 20, i.e., including variants among the top 1% of deleterious variants in the human genome; and (4) detected in both index case subjects (B.1, B.2). All variants prioritized by PSAP pipeline that had a missing allele frequency (NA) in the ExAC, ESP6500, and 1000GP were double-checked in dbSNP build 151 and variants with MAF > 1% were excluded. In addition, visual inspection of the quality of sequencing reads was performed using The Integrative Genomics Viewer (IGV) software.[23,24] Low-confidence variants due to unreliable reads (typically within read ends), a position within a long mononucleotide track and/or location in genes reported susceptible for sequencing errors[25] were discarded. In order to identify variants with potentially severe effect on early spermatogenesis, all the finally retained variants were manually inspected in parallel by two researchers (L.K., M.L.) using relevant information available in the literature and databases: PubMed, Ensembl, The Human Protein Atlas,[26] and gnomAD.[27] Based on the family history of no other reported infertility cases, large maternal pedigree, and fertile brother of the father, we considered the following disease models: autosomal-recessive effect (homozygous or compound heterozygous), autosomal-dominant *de novo* mutation derived from either of the parents (possible gonadal mosaicism in a parent), or maternally inherited X-linked variant.

Potential pathogenic effect of novel intronic variants in the vicinity of exon-intron boundaries on the RNA splicing was assessed using software Human Splicing Finder (v.3.0) that combines 12 different algorithms to identify and predict mutations' effect on splicing motifs including the acceptor and donor splice sites, the branch point and auxiliary sequences known to either enhance or repress splicing: Exonic Splicing Enhancers (ESE) and Exonic Splicing Silencers (ESS).[28] The transcript database utilized by this splicing prediction tool is based on Ensembl Release 70 (January 2013).

### Experimental Validation by Sanger Sequencing

Variants in *FANCM* (MIM: 609644), *CEP19* (MIM: 615586), and *CEP57* (MIM: 607951), prioritized in the PSAP analysis of the two Estonian azoospermic brothers (B.1, B.2), were experimentally verified using Sanger sequencing. The parent-of-origin of the variants was determined by sequencing maternal and paternal genomic DNA extracted from blood samples. To validate the *FANCM* mutations in additional azoospermic individuals, case subject 3 (Estonian) and case subject 4 (Portuguese), only the affected individuals were Sanger sequenced as parental DNA samples were not available. Full details are provided in Supplemental Subjects and Methods.

### Immunohistochemical Localization of FANCM Protein

To identify and localize the FANCM protein in human testicular tissue, 3 μm paraffin-embedded human testis samples from two control subjects (Supplemental Subjects and Methods) and the index case subject B.1 were analyzed. In addition, FANCM localization was investigated in two normal prostate samples (Supplemental Subjects and Methods). All biopsies had been collected and handled at the Pathology Centre, Diagnostic Clinic of the East Tallinn Central Hospital (Tallinn, Estonia).

For immunostaining, the sections were deparaffinized and subsequently incubated for antibody retrieval in Dako Target Retrieval Solution for 15 min at 96°C, blocked with Dako REAL Peroxidase-Blocking solution for 5 min. Subsequently, the tissue sections were incubated with the specific anti-human FANCM antibody (aa 1507–1679; 1:500; CV5.1; Novus biologicals). Primary antibody reactions were performed for 1 hr at room temperature under humid conditions. After several washings, the antigen-antibody complex was visualized by using DAKO REAL EnVision Detection System, Peroxidase/DAB+, Rabbit/Mouse. All solutions and buffers were provided by DAKO (Agilent Technologies). Slides were counterstained with hematoxylin, dehydrated, and coverslipped with Permount (Fisher Scientific) for light microscopy. Every experiment included negative controls without exposure to the primary antibody. Imaging was performed with Olympus BX60 microscope using Olympus DP71 digital camera and CellA imaging software (Olympus Optical). Microscope magnifications ×40 were used. All measurements were acquired at the same light intensity and processed the same way for the light background correction.

### Identification of Additional Azoospermia-Affected Case Subjects with *FANCM* Bi-allelic LoF Variants

Exome sequencing of an additional Estonian individual (case subject 3) was performed by the NGS Service at the Institute for Molecular Medicine Finland (FIMM) Helsinki, Finland. WES for this case subject was opted due to his clinical data referring to extreme testicular damage with a potential genetic cause. Wet-lab processing, base calling of the raw sequencing data, primary sequence analysis, and variant calling followed the standard local pipeline. The respective details are provided in Supplemental Subjects and Methods. Variant prioritization of the WES data was performed using the PSAP pipeline as described above.

The Portuguese NOA sample set (n = 296) was subjected to WES in the framework of the GEnetics of Male INfertility Initative (GEMINI) consortium aiming to map the genetic architecture of unexplained NOA phenotype (L.N., unpublished data). The current study utilized this resource to perform a targeted lookup of LoF variants in *FANCM*. A description of the consortium and its contributors, as well as the ongoing generation of a large WES dataset of worldwide NOA-affected case subjects is provided in Supplemental Subjects and Methods. The variants identified among the Portuguese NOA study group through WES were subjected to similar prioritization methods as applied in this study, including implementation of the PSAP software (L.N., unpublished data).

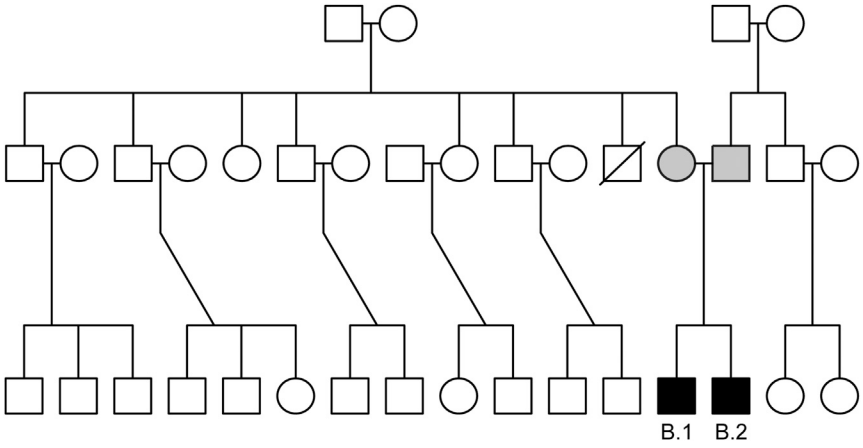


**Figure 1. Pedigree of the Index Family**
Black boxes indicate affected siblings who underwent whole-exome sequencing. Genome-wide CNV profiling and Sanger sequencing on all the selected variants from the WES analysis was applied in the two siblings as well as their parents (gray boxes). The kinship2 R package[31] was used to plot the pedigree.

Additionally, a high-confidence list of genes was compiled for the analyzed case subjects using STRING protein-protein interaction data, highlighting functionally connected genes that are most likely to affect the disease under study. The lookup of mutations within the *FANCM* locus was undertaken based on this generated high-confidence gene list. In the Portuguese samples, the calling of CNVs was carried out using the generated WES dataset and the XHMM[29] pipeline.

### Clinical Phenotyping of Azoospermic Case Subjects 3 and 4

Case subject 3 (Estonian) visited the national andrological service at the age of 52 years upon his own initiative to carry out a prophylactic andrological check-up. The routine andrological pipeline applied for the clinical phenotyping and assessment of the known causes of azoospermia was identical to the index case subjects B.1 and B.2. He was diagnosed with idiopathic NOA and hypergonadotropic hypogonadism. His general health records included a minor myocardial infarction at the age of 49 years.

The Portuguese NOA study group (n = 296) had been phenotyped and recruited at the Genetics Department from INSA-IP (Lisbon) and at the Genetics Department of FMUP (Porto). Individuals with known causes of infertility, including karyotype abnormalities and Y chromosome microdeletions in the *AZF* regions, were excluded. Many of these samples (n = 110) had been previously analyzed by SNP microarrays.[30] Case subject 4 identified in the current study with the *FANCM* homozygous nonsense variant presented at the Urology consult at CHSJ at age 36 with NOA. He was additionally tested for known *CFTR* mutations.

## Results

### Characterization of the Index Case Subjects with NOA and Sertoli Cell-only Syndrome

The index case subjects were two brothers (B.1 and B.2) diagnosed with idiopathic non-obstructive azoospermia (NOA) at the Andrology Centre, Tartu University Hospital, Estonia (by M.P.). The index family represented native Estonians and the reported family history data excluded consanguinity. No fertility problems have been described previously in the family history (Figure 1).

The brothers, aged 26 and 29 years at phenotyping, were both in good physical shape with general health and metabolic parameters within the normal range (Table 1). No chronic diseases were recorded in their medical history. Reflecting their infertility status and severe testicular damage, they presented nearly 50% reduced total testicular volume (22–30 mL versus median 47 mL in Estonian fertile men[2]). Both brothers had elevated serum FSH levels (31.0 and 16.0 versus reference 1.5–12.4 IU/L) and brother 1 (B.1) had also supranormal LH level (9.6 versus reference 1.7–8.6 IU/L). In addition, B.1 had cryptorchidism (MIM: 219050) reported at birth with spontaneous descent of testicles by the end of the first year of life. Both brothers and their father had grade 1–2 left-sided varicocele. The father of the index case subjects presented normal andrological parameters (Table 1).

For B.1, histopathological analysis of the testes detected Sertoli cell hyperplasia, nearly missing germinal epithelium, and complete lack of sperm, leading to the diagnosis of Sertoli cell-only syndrome (SCOS) (Figure S1A). Similarly, for brother 2 (B.2), who was subjected to a testis biopsy and microsurgical testicular sperm extraction (mTESE) operation at the Centre for Reproductive Medicine in Brussels (Belgium), no spermatozoa were recovered. His testicular histopathology likewise showed Sertoli cell-only syndrome (SCOS) with exclusive sclerosis and tubular atrophy (Figure S1B). His extended clinical examination showed that epididymis and vas deferens were bilateral with no obvious defects.

In addition to standard karyotyping, the genomes of the index brothers (B.1, B.2) and their parents were also profiled for submicroscopic chromosomal copy number variants (CNVs) using genome-wide SNP microarrays. No inherited or *de novo* deletions or duplications with clearly pathogenic effect on spermatogenesis were detected in the genomes of the analyzed azoospermic brothers either on autosomes or the X chromosome (Table S1).

### Prioritization of Potentially Pathogenic Variants from the WES Dataset

Genomic DNA extracted from blood samples of the two index brothers were subjected to whole-exome sequencing (WES) to detect the underlying causative mutation(s) for their extreme infertility. The VCF files resulting from the WES data were analyzed with PSAP pipeline,

**Table 1. Clinical Characteristics of the Estonian Study Subjects at the Clinical Assessment**

| | Brother 1 | Brother 2 | Father | Case 3 | Reference Values[a] |
|---|---|---|---|---|---|
| **General Parameters** | | | | | |
| Age (years) | 26 | 29 | 53 | 52 | – |
| Height (cm) | 190 | 187 | ND | 180 | – |
| Weight (kg) | 85 | 84 | ND | 85 | – |
| BMI (kg/m2) | 23.5 | 24.0 | ND | 26.1 | 18.5–24.9 |
| SBP/DBP (mmHg) | 138/76 | 127/86 | 168/96 | 145/91 | <140/90 |
| **Fertility Parameters** | | | | | |
| Total testes volume (mL)[b] | 30 | 20 | 42 | 19 | 50.0 (33.0–70.0)[c] |
| Testicular volume: left | 14 | 10 | 18 | 9 | – |
| Testicular volume: right | 16 | 10 | 24 | 10 | – |
| Varicocele testis (left side) | grade 2 | grade 1 | grade 2 | no | – |
| FSH (IU/L) | 31.0 | 16.0 | 6.3 | 42.3 | 1.5–12.4 |
| LH (IU/L) | 9.6 | 4.1 | 5.2 | 18.6 | 1.7–8.6 |
| Testosterone (nmol/L) | 14.1 | 14.8 | 22.6 | 7.6 | 11.4–27.9 |
| Estradiol (pmol/L) | 41.9 | 44.9 | 61.8 | 58.2 | <159 |
| SHBG (nmol/L) | 40.5 | 25.5 | 49.0 | 26.2 | 18.3–54.1 |
| PSA (μg/L) | 0.37 | 0.47 | ND | 0.30 | <1.4 |
| Semen volume (mL) | 4.0 | 4.6 | ND | 1.95 | >1.5 |
| Total sperm count ($10^6$/ejaculate) | 0 | 0 | ND | 0 | >39 |
| Karyotype | 46,XY | 46,XY | 46,XY | 46,XY | – |
| *AZFa*, *AZFb*, and *AZFc* microdeletions | no | no | no | no | – |
| **Metabolic Parameters** | | | | | |
| FAT % | 15.2 | 20.8 | ND | 28.2 | <24.5 |
| Glucose (mmol/L) | 4.8 | 4.9 | ND | ND | 3.3–5.5 |
| Cholesterol (mmol/L) | 4.3 | 4.8 | ND | 6.4 | <5.0 |
| HDL (mmol/L) | 1.38 | 1.16 | ND | 1.70 | >1.0 |
| LDL (mmol/L) | 2.81 | 3.07 | ND | 4.88 | <3.0 |
| Triglyceride (mmol/L) | 0.59 | 1.55 | ND | 0.83 | <1.7 |

Abbreviations: *AZFa-c*, azoospermia factor a-c; BMI, body mass index; FSH, follicle-stimulating hormone; HDL, high-density lipoprotein; LDL, low-density lipoprotein; LH, luteinizing hormone; ND, no data; PSA, prostate-specific antigen; SBP/DBP, systolic/diastolic blood pressure; SHBG, sex hormone binding globulin.
[a]Reference values from the United Laboratories of Tartu University Hospital and WHO 2010.
[b]Physical examination for the assessment of testicular size was performed with orchidometer (made of birch wood, Pharmacia & Upjohn, Denmark).
[c]Baltic young male cohort, data presented as median (5th–95th%).[58]

a computational approach for the identification of pathogenic genetic variants.[22] The input VCF files contained >60,000 variants per exome and the PSAP output resulted in ~13,000 variants per person (Figure S2). The shared variants by B.1 and B.2, exhibiting low MAF (≤0.01) or being previously undescribed, with the PSAP popScore ≤ 0.005 and the CADD score ≥ 20, were prioritized. After these filtering steps, 25 candidate variants (all heterozygous) were initially retained and were further subjected to manual inspection for the quality of variant prediction, information on testicular expression, and relevance to male infertility (Table S2). Manual filtering discarded 15 variants due to unreliable reads, location in loci reported to be susceptible for sequencing errors,[25] or location in genes with specific/enhanced expression in tissues other than testis. For seven of the remaining ten genes, the available literature evidence did not support a possible role in male germ cell development and gonadal function (*PRKAG2* [MIM: 602743], *CNIH4* [MIM: 617483], *COBLL1* [MIM: 610318], *NR1D2* [MIM: 602304], *TMEM2* [MIM: 605835], *PIGQ* [MIM: 605754], *FRG1BP*; Table S2).

Heterozygous variants in the genes *FANCM* (encoding Fanconi anemia complementation group M), *CEP19* (Centrosomal Protein 19 [MIM: 615586]), and *CEP57*

**Table 2. Summary of Potentially Pathogenic Variants Detected in the Analyzed Case Subjects with Idiopathic Non-obstructive Azoospermia**

| | | | | | gnomAD Dataset[a] | | Genotype | | |
|---|---|---|---|---|---|---|---|---|---|
| **Position (hg38)** | **ID** | **Gene** | **cDNA Change** | **Variant Type** | **MAF** | **Carriers** | **Case(s)** | **Mother** | **Father** |
| **Variants Prioritized from the WES Dataset of the Two Estonian Brothers with NOA and SCOS[b]** | | | | | | | | | |
| Chr3:196707729 | rs748622284 | *CEP19* | c.326T>C | p.Leu109Pro | $1.80 \times 10^{-5}$ | Het: 5; Hom: 0 | T/C | T/T | T/C |
| Chr11:95818897 | NA | *CEP57* | c.692C>T | p.Ala231Val | NA | NA | C/T | C/C | C/T |
| Chr14:45159189 | rs761250416 | *FANCM* | c.1491dupA | p.Gln498Thrfs*7 | $7.22 \times 10^{-5}$ | Het: 20; Hom: 0 | A/AA | A/AA | A/A |
| Chr14:45183764 | NA | *FANCM* | c.4387−10A>G | p.Arg1436_ Ser1437insLeuLeu* | NA | NA | A/G | A/A | A/G |
| **Variant Identified in the Estonian NOA-Affected Case Subject 3 using WES** | | | | | | | | | |
| Chr14:45189123 | rs147021911 | *FANCM* | c.5101C>T | p.Gln1701* | $1.34 \times 10^{-3}$ | Het: 370; Hom: 1 | T/T | NA | NA |
| **Variant Identified in the Portuguese Case Subject 4 using Targeted Candidate Gene Analysis among 296 NOA-Affected Case Subjects** | | | | | | | | | |
| Chr14:45198718 | rs144567652 | *FANCM* | c.5791C>T | p.Arg1931* | $1.03 \times 10^{-3}$ | Het: 285; Hom: 0 | T/T | NA | NA |

Abbreviations: MAF, minor allele frequency; NA, not available; NOA, non-obstructive azoospermia; SCOS, Sertoli cell-only syndrome; WES, whole-exome sequencing.
[a]138,610; 138,516; 138,488; and 138,536 exome datasets for *CEP19* p.Leu109Pro, *FANCM* p.Gln498Thrfs*7, *FANCM* p.Gln1701*, and *FANCM* p.Arg1931*, respectively.
[b]Both index brothers (B.1, B.2) have identical heterozygous genotype for these four variants.

(Centrosomal Protein 57 [MIM: 607951]) passed all manual filtering steps. An A-nucleotide insertion in exon 9 of *FANCM* creates a frameshift at position p.Gln498Thr and consequently, a new reading frame that ends shortly with a STOP codon, causing loss of function (LoF) (c.1491dupA [p.Gln498Thrfs*7]; Table 2; Figure 2; Table S2). *FANCM* is predominantly expressed in the testis and this variant located in the conserved helicase domain of FANCM protein is very rare (gnomAD database: MAF = $7.22 \times 10^{-5}$; 20 hetero- and no homozygotes). Pathogenic heterozygous missense variants were detected in *CEP19* (rs748622284, p.Leu109Pro [c.326T>C]; gnomAD MAF = $1.80 \times 10^{-5}$; SIFT score = 0, PolyPhen score > 0.93) and *CEP57* (previously undescribed variant; p.Ala231Val [c.692C>T]). Both genes are critical for the mitotic spindle formation, *CEP19* is mainly transcribed in the testis, and *CEP57* displays high testicular expression.

As the three identified pathogenic and potentially causative variants were heterozygous, we manually screened the generated VCF files for possible additional variants in *CEP19*, *CEP57*, and *FANCM* that may have been missed by the conservative PSAP filtering criteria. No additional variants shared among the index case subjects B.1 and B.2 were detected in *CEP19* (Table S3). For *CEP57*, the detected intronic homozygous variant rs5793747 is unlikely to cause NOA as its allele frequency in worldwide populations is 2%–9% and there is no clearly predicted functional effect. In *FANCM*, in total 11 exonic variants shared by B.1 and B.2 were identified, including 9 common (MAF > 0.1) and 2 rare variants. In addition to the rare frameshift variant c.1491dupA (rs761250416), a novel splice region variant (c.4387−10A>G) was detected in intron 16, immediately (10 bp) upstream of exon 17. The variant is predicted to have a severe pathogenic effect on mRNA splicing by potentially activating an intronic cryptic acceptor site and thus extending exon 17 (Figures 2 and S3). The 9 bp stretch of intronic DNA (TTA-TTA-TAG) included into the *FANCM* mRNA encodes "leucine-leucine-STOP" and leads to a premature truncation of the protein (p.Arg1436_Ser1437insLeuLeu*) and potentially nonsense-mediated mRNA decay resulting in loss of function. Also, new silencer motifs and broken enhancer motifs were predicted.

### Compound Heterozygous LoF Variants in *FANCM* as the Most Likely Cause of NOA in the Index Case Subjects

Sanger sequencing was used to experimentally validate the carrier status of the prioritized variants in *FANCM*, *CEP19*, and *CEP57* for the index case subjects and to determine their parental inheritance (Figure 2; Table 2). Both brothers were detected as compound heterozygous for the two *FANCM* LoFs. The frameshift causing variant (p.Gln498Thrfs*7) had been inherited from the mother and the novel splicing affecting variant (c.4387−10A>G) from the father (Figures 2 and S4). Heterozygous missense variants in *CEP19* and *CEP57* were both inherited from the father and thus their causative nature on severe male infertility was excluded.

The two *FANCM* LoFs were identified as the most likely cause of NOA and SCOS in the two Estonian brothers B.1 and B.2. FANCM protein has a vital role in regulating meiotic crossovers and DNA repair in replication,[32] and

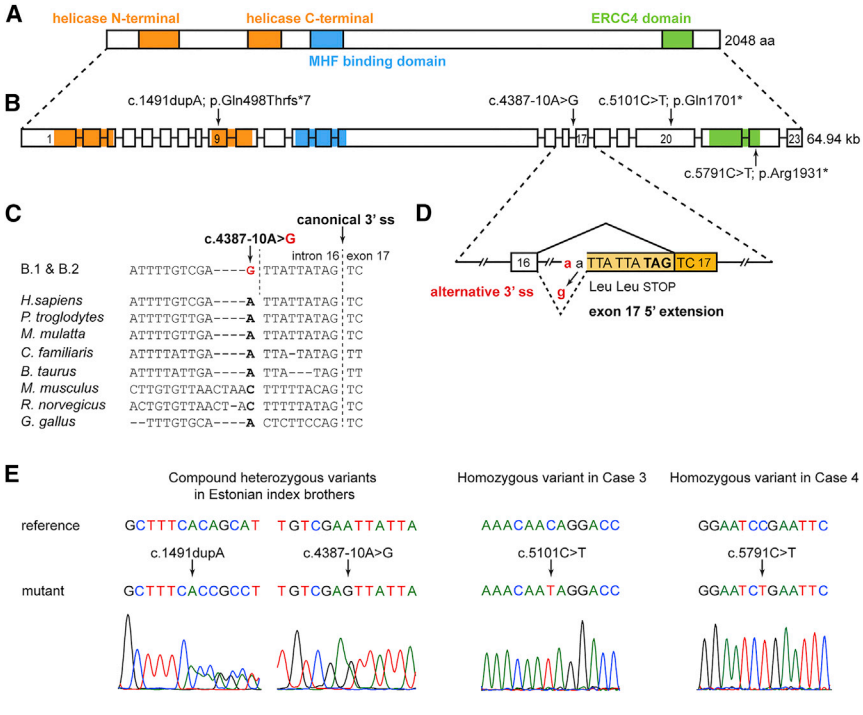


**Figure 2. Structure of the FANCM Protein and the Genetic Context of the *FANCM* Loss-of-Function Variants Detected in Four Men Diagnosed with NOA**

(A) FANCM protein has three critical functional domains. Helicase and ERCC4 DNA repair endonuclease domain are important for DNA binding. The helicase domain is the primary component that allows FANCM to transfer the core complex along DNA during repair and it is highly conserved among all eukaryotes, including yeast.[32,33] MHF stimulates the double-strand DNA-binding activity of FANCM and stabilizes its interaction with chromatin, whereas the MHF-FANCM complex supports genomic stability during cell division.[34] ERCC4 domain binds to FAAP24 protein, creating a heterodimer that is critical in protecting cells from interstrand crosslinks.[35]

(B) Relative positions of the variants detected in the Estonian index brothers B.1 and B.2 (c.1491duplA [p.Gln498Thrfs*7] and c.4387−10A>G), a third Estonian case subject 3 (c.5101C>T [p.Gln1701*]), and Portuguese case subject 4 (c.5791C>T [p.Arg1931*]).

(C) DNA sequence alignment of human *FANCM* with its homologs in seven other vertebrate species showing the conservation of the c.4387−10A>G residue mutated in the index brothers.

(D) Predicted alternative 3′ ss (splice site) likely created by the A>G variant upstream of the normal exon 17 3′ ss, which adds nine nucleotides in frame between exons 16 and 17, leading to a 3 amino acid insertion and premature termination.

(E) Experimental validation of *FANCM* variants identified from the WES datasets using Sanger sequencing. Additional experimental confirmation by cloning and Sanger sequencing of heterozygous mutations of Estonian index cases B.1 and B.2 is provided in Figure S4.

the available literature for human and mouse supports its critical function in male and female fertility and in early spermatogenesis (Table 3). *Fancm* mutant homozygous male mice present hypogonadism, abnormal spermatogonia, reduced testis size, and azoospermia in a fraction of seminiferous tubules or progressive loss of germ cells due to elevated level of meiotic DNA damage.

### Localization of FANCM Protein in Male Reproductive Tissues

FANCM protein expression was investigated in the testicular tissue sections from the index brother B.1 compared to two control subjects, using immunohistochemistry (IHC) (Figure 3). In normal spermatogenesis, FANCM protein localized to the cytoplasm of Sertoli cells and spermatogenic cells in the seminiferous tubules. The staining intensity of spermatogenic cells depended on the stage of maturation. The spermatogonia showed faint FANCM positivity, while primary spermatocytes to spermatids exhibited increased expression (Figure 3, I, II, IV, V). In tubules with mature spermatozoa, staining intensity dropped (Figure 3, III). FANCM expression was also present in the interstitial Leydig cells (Figure 3, III, IV). These results are consistent with FANCM protein localization in mouse testis[36] (Table 3). Seminiferous tubules of the index case subject B.1 containing only Sertoli cells showed missing or only faint staining intensity compared to control subjects (Figure 3, VII–IX).

IHC staining was also studied on two normal prostate tissue sections revealing rather strong staining intensity of the FANCM protein in the fibromuscular stroma (Figure S5).

### Additional Azoospermia-Affected Case Subjects with Independent *FANCM* Homozygous Nonsense Variants

During WES targeted to Estonian men with clinical signs of extreme testicular damage, we identified an additional NOA-affected individual (case subject 3) with an independent homozygous nonsense variant in *FANCM* (p.Gln1701*; Table 2). After prioritization and filtering of the PSAP results as described above for the brothers, this was the only homozygous LoF mutation of the 28 variants retained (Table S4). The identified variant creates a STOP codon in exon 20 shortly before the conserved ERCC4 endonuclease domain of the FANCM protein. At the first visit to the clinic, the individual was 52 years old and in addition to azoospermia he presented substantially reduced total testicular volume (19 mL versus median 47 mL in Estonian fertile middle-aged men[2]) and hypergonadotropic hypogonadism. He displayed extremely elevated serum FSH (42.3 IU/L) and LH (18.6 IU/L) levels and low testosterone levels (7.6 nmol/L) (Table 1). The medical history of case subject 3 excluded any chronic diseases to explain the andrological profile. Notably, a recent study reported two consanguineous Finnish sisters with non-syndromic primary ovarian insufficiency to carry homozygous nonsense

**Table 3. Supportive Information from Literature and Databases for the Relevance of *FANCM* Variants in Infertility**

| |
|---|
| ***FANCM Function***[32] |
| interstrand cross-link removal, anti-crossover function in mitotic cells |
| replication fork repair by catalyzing fork regression |
| regulation of meiotic crossovers in *D. melanogaster*, *A. thaliana*, and *S. cerevisiae* |
| ***FANCM Expression in Human*** |
| GTEx, HPA: Testis enhanced (RNA tissue category) |
| protein present in the nuclei of oogonia, stronger in pachytene stage oocytes in human fetal ovaries[37] |
| ***FANCM Expression Data in Mouse***[36] |
| Fancm mRNA expression levels in testicular germ cells increase through all the post-natal development and maintain high during adulthood |
| Fancm protein localizes in the cytoplasm of Sertoli cells, the cytoplasm, and Golgi of spermatogonia through to early pachytene spermatocytes, immunostaining faint or absent beyond pachytene spermatocytes |
| ***FANCM-Deficient Male Mice***[a] **(MGI)**[45] |
| Abnormal spermatogonia and seminiferous tubule morphology |
| Azoospermia; some seminiferous tubules lack spermatogonia and contain no sperm |
| Small testes, hypogonadism, Leydig cell hyperplasia |
| Elevated mitotic sister chromatid exchange |
| Reduced overall and tumor-free survival |
| ***FANCM-Mutant Male Mice***[b,42] |
| Near Sertoli cell-only-like phenotype in many tubules |
| Reduced germ cells in neonates and progressive loss of germ cells in adult males |
| Elevated level of meiotic DNA damage and spontaneous sister chromatid exchange |
| Reduced testis size |
| Premature senescence and increased tumor susceptibility |
| ***FANCM-Mutant Female Mice*** **(MGI)**[45,46] |
| Abnormal ovary morphology |
| Depletion of primordial follicles and decreased mature ovarian follicle number |
| Transmission ratio distortion: fewer than expected females |
| Reduced overall and tumor-free survival |
| ***FANCM Loss-of-Function Mutation in Females***[37] |
| irregular cycles, hot flashes, spaniomenorrhea |
| high FSH and low AMH levels |
| primary ovarian insufficiency |

Abbreviations: GTEx, The Genotype-Tissue Expression; HPA, The Human Protein Atlas; MGI, Mouse Genome Informatics; FSH, follicle-stimulating hormone; AMH, anti-müllerian hormone.
[a]Homozygous deletion of exon 2, leading to frameshift and premature stop in exon 3.
[b]Homozygous p.Cys142Arg mutation in conserved DEXDc domain.

mutation of the same p.Gln1701* variant.[37] Despite being the most prevalent *FANCM* nonsense variant in the gnomAD database (MAF = $1.34 \times 10^{-3}$) with the majority of the heterozygous carriers (341/368) identified among Europeans, only one individual (also Finnish) out of 138,488 sequenced exomes/genomes was detected homozygous for this nonsense variant.

Additionally, a mutation screening targeted to *FANCM* was carried out in the Portuguese NOA study group (n = 296) using a WES dataset generated in the framework of the GEMINI Project (Supplemental Data). One Portuguese azoospermic individual (case subject 4) was detected with a homozygous nonsense variant in exon 22, causing a premature stop in the conserved ERCC4 endonuclease domain of the FANCM protein (p.Arg1931*; Tables 2 and S5). No homozygous individuals have been reported previously for this variant (MAF = $1.03 \times 10^{-3}$) across the 138,536 sequenced exomes/genomes (gnomAD database).

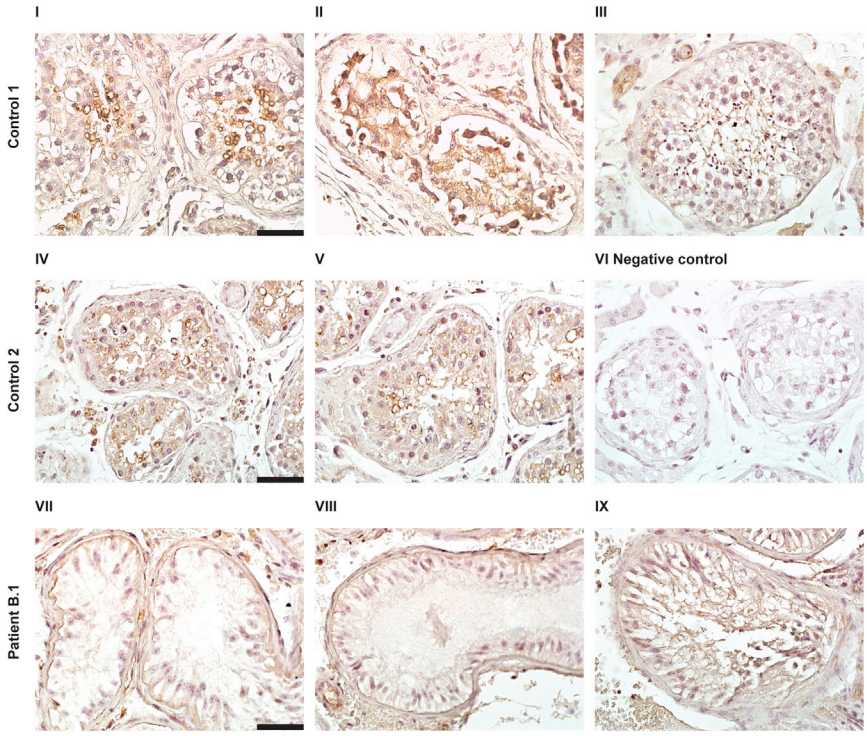


**Figure 3. Immunohistochemical Detection of FANCM in Testicular Tissue Sections Representing Normal Spermatogenesis and Sertoli Cell-only Syndrome**

The localization of FANCM protein in normal testicular tissue sections was analyzed in control samples obtained from men with either seminoma (31 years old) or cardiac insufficiency (68 years old). Both control subjects exhibited active spermatogenesis (I–V). FANCM protein localized in the cytoplasm of Sertoli cells and spermatogenic cells. Staining intensity was gradually enhanced along germ cell maturation stages: in spermatogonia only faint FANCM positivity was detected, whereas spermatocytes and spermatids showed increased expression (I–IV). In tubules with mature spermatozoa, staining intensity dropped (III). In seminiferous tubules of the index brother B.1 (26 years old), no spermatogenic cells were detected and only Sertoli cells were present (VII–IX). FANCM expression in the cytoplasm of Sertoli cells was missing or showed only faint intensity compared to controls. Additionally, focal unspecific staining was observed in tubular basement membrane, myoepithelial cells, and surrounding stroma, only in the tissue sections of B.1. In the negative control performed without the FANCM antibody no staining was detected (VI). Brown color indicates chromogen-labeled antibody and violet color hematoxylin nuclear staining. Scale bar, 50 μm.

*In silico* analysis of WES data detected no deletions spanning *FANCM* in case subject 4, excluding hemizygosity.

Homozygous carrier status of the two nonsense variants in case subjects 3 and 4 were confirmed by Sanger sequencing (Figure 2). As testicular histology was not available for these men, the ultimate causative link between *FANCM* LoF variants and SCOS is still to be verified in future clinical case subjects.

## Discussion

The data of the current study convincingly demonstrate that bi-allelic loss-of-function variants in *FANCM* cause non-obstructive azoospermia (NOA) and male infertility. We present several lines of evidence to support our claims. First, whole-exome sequencing (WES) was applied to analyze a familial case of two Estonian brothers with NOA (case subjects B.1, B.2) and Sertoli cell-only syndrome (SCOS). The outcome data provided strong evidence that compound heterozygous loss-of-function (LoF) variants in *FANCM* are likely causative for their condition (Table 2). Both a rare maternally inherited frameshift variant in exon 9 (p.Gln498Thrfs*7) and a previously undescribed splicing variant in intron 16 (c.4387−10A>G) derived from the father introduce a premature STOP codon leading to nonsense-mediated decay and/or to a truncated protein missing critical domains (Figure 2). Second, we demonstrated a substantial presence of the FANCM protein in the spermatogenic and Sertoli cells of seminiferous tubules in healthy testicular tissue, whereas in the index case subject B.1 it was nearly absent (Figure 3). Third, we identified two additional NOA-affected case subjects with *FANCM* homozygous nonsense mutations (Table 2). The p.Gln1701* variant (exon 20) identified in the Estonian case subject 3 leads to a truncated protein, resulting in higher occurrence of chromosome breakages and rearrangements.[37] The p.Arg1931* (exon 22) variant detected in the NOA-affected case subject from Portugal produces a truncated FANCM protein that lacks endonuclease (ERCC4) and FAAP24-interacting domains, and thereof DNA repair activity.[38] Thus, the total number of identified European NOA-affected case subjects with *FANCM* bi-allelic LoF mutations was increased to four men, including two homozygotes with STOP mutations. Notably, in the gnomAD database (for most variants, >120,000 subjects analyzed) only three homozygotes have been reported across all identified high-confidence LoFs in *FANCM* and two of them originate from South Asia. Finally, a WES study of two consanguineous Finnish sisters with non-syndromic primary ovarian insufficiency (POI) detected the same homozygous nonsense variant in *FANCM* that was identified in the Estonian NOA-affected case subject 3 (p.Gln1701*).[37] POI is the corresponding female phenotype of oligo-/azoospermia. This recent collective information adds *FANCM* to the list of the few genes involved in both male and female infertility.[39]

In the current study, the FANCM protein localized in the cytoplasm of Sertoli cells and spermatogenic cells and the staining intensity in normal testis tissue sections was

dependent on the stage of sperm maturation (Figure 3). The analysis of the testicular biopsy from the brother B.1 showed either no or minimal presence of the FANCM protein. Weak staining detected in some seminiferous tubules indicates that the splice site mutation most likely does not entirely ablate antibody recognition; however, the function and stability of the protein are likely disrupted. Our human data are in agreement with the mouse *Fancm* gene expression pattern (Table 3). Whereas *Fancm* transcripts were enriched in haploid germ cells of the testis, the protein has been mainly localized to the cytoplasm and Golgi of the spermatogonia and early pachytene spermatocytes in mouse[36] (Table 3). It has been suggested that the synthesized FANCM protein is stored in the Golgi in order to be distributed to sites of DNA damage at appropriate time points of germ cell development. Normal prostate tissue sections displayed high expression of FANCM in the fibromuscular stroma (Figure S5) and its role in prostate function is yet to be elucidated.

FANCM is the most conserved protein in the Fanconi anemia (FA) family[40] and one of only two proteins in the human genome that comprises both a helicase/ATPase and ERCC4 endonuclease domain in a single protein.[41] It has multitasking functions with the most diverse and complex biological roles evolved in human lineage, including involvement in DNA replication, repair, and in anti-crossover function to maintain genomic stability.[32,42] If the FANCM protein is non-functional, the FA core complex cannot be recruited to chromatin and resolve DNA damage.[43] Defects in DNA repair pathways are known to cause arrest of spermatogenesis and abnormal recombination, resulting in male infertility.[44] Consistently, the majority of the conducted WES studies for azoospermia have identified variants in genes with enriched testicular expression and involvement in basic cellular function, including DNA damage response, control of transcription, and the stages of mitosis and meiosis of spermatogenesis (Table S6).

Reproductive phenotypes of the two mouse models with a deletion or missense mutation in exon 2 (N-terminal part of the helicase domain) are consistent with our human data (Table 3). *Fancm*-deficient male mice display reduced testis size and abnormal morphology of spermatogonia and seminiferous tubules, including completely missing germ cells in a subset of tubules.[45] Several male $Fancm^{-/-}$ mice were azoospermic. Also, *Fancm* mutant mice exhibit hypogonadism and progressive spermatogenesis defects that are initiated already during embryogenesis, including SCOS phenotype in many tubules.[46] These features are consistent with the clinical characteristics of the two index case subjects B.1-B.2 (SCOS) and case subject 3 (hypogonadism). Most probably, hypergonadotropic hypogonadism at the age of 52 years reflects a downstream consequence of impaired testicular function across his lifetime.

Consistent with the mouse models exhibiting increased tumor susceptibility, numerous recent studies in humans have linked homo- and heterozygous LoF variants in the *FANCM* with doubled risk to familial breast and ovarian cancer.[38,47–51] These reports include the maternally inherited frameshift variant (p.Gln498Thrfs*7) detected in the Estonian index brothers,[52] the p.Arg1931* identified in the Portuguese case subject 4, and the p.Gln1701* nonsense mutation present in the Estonian case subject 3 and the two Finnish sisters with POI.[48,51] Unfortunately, the available family data on the case subjects in the current study did not enable assessment of the risk of different tumors. The mother of B.1 and B.2 (heterozygous carrier of p.Gln498Thrfs*7) was healthy in her early fifties; no genotype data or detailed medical records, including age, were available for her two sisters and five brothers (Figure 1). Notably, apart from breast and ovarian cancer, various other cancer types had been diagnosed in some families with *FANCM* mutations, including prostate cancer in the father of a female with breast cancer.[52] As our data displayed high expression of FANCM protein in the fibromuscular stroma of prostate, the relevance of *FANCM* mutations in the context of prostate cancer is alerted.

Although the *FANCM* gene mutations were initially linked with the diagnosis of Fanconi anemia (FA), this direct causative link has later been rejected[53] as the reported index case subject actually carried a mutation in *FANCA*, a well-established FA gene.[54] Neither of the *Fancm* mutant mouse models showed any impairment of the hematopoietic system phenotype.[45,46] Consistent with this, a population-based study of rare variants specific to Finns investigated retrospective clinical records of two *FANCM* homozygous LoF variant carriers, but no evidence for blood diseases had been recorded in their medical histories.[55]

What could be the contribution of *FANCM* loss-of-function variants among men diagnosed with NOA? The gnomAD database reports 94 high-confidence *FANCM* LoF variants and in total 880 alleles (including only 3 homozygotes) that have been detected across mean 113,442 exomes/genomes (as of April 2018). Thus, the estimated number of *FANCM* LoF heterozygous variant carriers in the general population, including men and women, is roughly 1/130 subjects. A bi-allelic LoF carrier is expected with a prevalence of 1/67,600. It has been estimated that azoospermia affects about 1% of men, and fewer than 5%–10% of these men have SCOS,[56] i.e., 1/2,000–1/1,000. A large fraction of these cases are already explained by currently well-known causes, including major genetic factors.[2] For the 20% unexplained azoospermia cases, a targeted screen of bi-allelic *FANCM* LoF variants exhibits an attractive clinical perspective, especially for the men diagnosed with SCOS.

As a final observation, most intriguingly, our index family displays a significantly lower number of girls born on the maternal side of the family (Figure 1). The mother has two sisters (one of them has no children) and five brothers (one died at an early age), and in the next generation there are 12 sons and only 2 daughters ($p = 3 \times 10^{-3}$, one-sided binomial test that there is an equal number of sons and daughters). Unfortunately, DNA samples of the

extended family were not available for the current study. The observed skew toward male offspring replicates the described mouse phenotype. An underrepresentation of live-born female *Fancm* knockout mice has been reported, suggesting a reduced viability in early development for females with bi-allelic loss-of-function mutations.[45] We speculate that elevated spontaneous sister chromatid exchange in XX conceptuses with a *FANCM* LoF variant, including heterozygous carriers, may interfere with X chromosome inactivation and/or affect normal chromosomal segregation in the early human embryo. The described female homozygous knock-out mice[46] and human case subjects presented abnormal ovarian morphology and decreased mature ovarian follicle number.[37]

In summary, we report the phenotypes of the human male knock-outs of *FANCM* and show that bi-allelic recessive LoF variants in this gene represent a previously undescribed genetic cause for NOA, possibly linked to SCOS phenotype. These data are consistent with the reported female *FANCM* knock-outs exhibiting POI. Importantly, the accumulated data on *FANCM* functional diversity provides an example mechanism for the known epidemiological association between infertility and cancer risk.[57] Potential implications of the carrier status of *FANCM* pathogenic variants for somatic health warrant further targeted studies.

## Accession Numbers

Novel SNVs have been added to dbSNP (ss2137543924, ss2137543926; handle MLAAN, dbSNP Build B152).

## Supplemental Data

Supplemental Data include five figures, Supplemental Subjects and Methods, Consortium Membership, and seven tables and can be found with this article online at https://doi.org/10.1016/j.ajhg.2018.07.005.

## Acknowledgments

We are thankful to the patients and their family for their trust and motivation to seek the ultimate cause of their clinical condition. Herman Tournaye from the Centre for Reproductive Medicine, Universitair Ziekenhuis Brussel is thanked for detailed comments on the testicular histopathology. Eero Semjonov and Georgi Dzaparidze from the Pathology Department of East Tallinn Central Hospital are thanked for sampling and comments on pathoanatomical analysis. Mart Adler and Laura Kibena are thanked for technical assistance in genetic tests.

The study was supported by the European Union through the European Regional Development Fund (project HAPPY PREGNANCY, 3.2.0701.12-0047; for M.L. and M.P.), the Estonian Research Council (grants IUT34-12 for M.L. and PUT181 to M.P.), and the National Institutes of Health of the United States of America (R01HD078641 to D.F.C. and K.I.A. and R01MH101810 to D.F.C.). Some work on this project was conducted while D.F.C. was visiting the University of Tartu under the support of a grant from the United States Department of State (Fulbright Project #P000086). A.M.L. is funded by the Portuguese Government through FCT (IF/01262/2014). IPATIMUP integrates the i3S Research Unit, which is partially supported by FCT in the framework of the project "Institute for Research and Innovation in Health Sciences" (POCI-01-0145-FEDER-007274). L.B.S. is funded by a MRC (UK) Programme Grant number MR/N002970/1.

## Declaration of Interests

The authors declare no competing interests.

Received: May 10, 2018
Accepted: July 2, 2018
Published: August 2, 2018

## Web Resources

1000 Genomes, http://www.internationalgenome.org/
BioEdit Software, http://www.mbio.ncsu.edu/BioEdit/bioedit.html
Centre for Reproductive Medicine, Universitair Ziekenhuis Brussel, http://www.uzbrussel.be
dbSNP, https://www.ncbi.nlm.nih.gov/projects/SNP/
East Tallinn Central Hospital, https://www.itk.ee
Ensembl Genome Browser, http://www.ensembl.org/index.html
Estonian Genome Center, https://www.geenivaramu.ee/en/core-facility
gnomAD Browser, http://gnomad.broadinstitute.org/
GTEx Portal, https://www.gtexportal.org/home/
Human Splicing Finder, http://www.umd.be/HSF3/
Mouse Genome Informatics, http://www.informatics.jax.org/
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
OMIM, http://www.omim.org/
Primer-BLAST, https://www.ncbi.nlm.nih.gov/tools/primer-blast/
PSAP, https://github.com/awilfert/PSAP-pipeline
PubMed, http://www.ncbi.nlm.nih.gov/PubMed/
The Human Protein Atlas, http://www.proteinatlas.org/
UCSC In-Silico PCR, https://genome.ucsc.edu/cgi-bin/hgPcr
United Laboratories, Tartu University Hospital, https://www.kliinikum.ee/en/united-laboratories

## References

1. Agarwal, A., Mulgund, A., Hamada, A., and Chyatte, M.R. (2015). A unique view on male infertility around the globe. Reprod. Biol. Endocrinol. *13*, 37.
2. Punab, M., Poolamets, O., Paju, P., Vihljajev, V., Pomm, K., Ladva, R., Korrovits, P., and Laan, M. (2017). Causes of male infertility: a 9-year prospective monocentre study on 1737 patients with reduced total sperm counts. Hum. Reprod. *32*, 18–31.
3. Schultz, N., Hamra, F.K., and Garbers, D.L. (2003). A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. Proc. Natl. Acad. Sci. USA *100*, 12201–12206.
4. Djureinovic, D., Fagerberg, L., Hallström, B., Danielsson, A., Lindskog, C., Uhlén, M., and Pontén, F. (2014). The human testis-specific proteome defined by transcriptomics and antibody-based profiling. Mol. Hum. Reprod. *20*, 476–488.
5. He, W.B., Tu, C.F., Liu, Q., Meng, L.L., Yuan, S.M., Luo, A.X., He, F.S., Shen, J., Li, W., Du, J., et al. (2018). *DMC1* mutation that causes human non-obstructive azoospermia and

premature ovarian insufficiency identified by whole-exome sequencing. J. Med. Genet. *55*, 198–204.
6. Arafat, M., Har-Vardi, I., Harlev, A., Levitas, E., Zeadna, A., Abofoul-Azab, M., Dyomin, V., Sheffield, V.C., Lunenfeld, E., Huleihel, M., and Parvari, R. (2017). Mutation in TDRD9 causes non-obstructive azoospermia in infertile men. J. Med. Genet. *54*, 633–639.
7. Okutman, O., Muller, J., Skory, V., Garnier, J.M., Gaucherot, A., Baert, Y., Lamour, V., Serdarogullari, M., Gultomruk, M., Röpke, A., et al. (2017). A no-stop mutation in MAGEB4 is a possible cause of rare X-linked azoospermia and oligozoospermia in a consanguineous Turkish family. J. Assist. Reprod. Genet. *34*, 683–694.
8. Okutman, O., Muller, J., Baert, Y., Serdarogullari, M., Gultomruk, M., Piton, A., Rombaut, C., Benkhalifa, M., Teletin, M., Skory, V., et al. (2015). Exome sequencing reveals a nonsense mutation in TEX15 causing spermatogenic failure in a Turkish family. Hum. Mol. Genet. *24*, 5581–5588.
9. Gershoni, M., Hauser, R., Yogev, L., Lehavi, O., Azem, F., Yavetz, H., Pietrokovski, S., and Kleiman, S.E. (2017). A familial study of azoospermic men identifies three novel causative mutations in three new human azoospermia genes. Genet. Med. *19*, 998–1006.
10. Ramasamy, R., Bakırcıoğlu, M.E., Cengiz, C., Karaca, E., Scovell, J., Jhangiani, S.N., Akdemir, Z.C., Bainbridge, M., Yu, Y., Huff, C., et al. (2015). Whole-exome sequencing identifies novel homozygous mutation in NPAS2 in family with nonobstructive azoospermia. Fertil. Steril. *104*, 286–291.
11. Ayhan, Ö., Balkan, M., Guven, A., Hazan, R., Atar, M., Tok, A., and Tolun, A. (2014). Truncating mutations in TAF4B and ZMYND15 causing recessive azoospermia. J. Med. Genet. *51*, 239–244.
12. Kherraf, Z.E., Christou-Kent, M., Karaouzene, T., Amiri-Yekta, A., Martinez, G., Vargas, A.S., Lambert, E., Borel, C., Dorphin, B., Aknin-Seifer, I., et al. (2017). SPINK2 deficiency causes infertility by inducing sperm defects in heterozygotes and azoospermia in homozygotes. EMBO Mol. Med. *9*, 1132–1149.
13. Fakhro, K.A., Elbardisi, H., Arafa, M., Robay, A., Rodriguez-Flores, J.L., Al-Shakaki, A., Syed, N., Mezey, J.G., Abi Khalil, C., Malek, J.A., et al. (2018). Point-of-care whole-exome sequencing of idiopathic male infertility. Genet. Med. https://doi.org/10.1038/gim.2018.10.
14. Colombo, R., Pontoglio, A., and Bini, M. (2017). Two novel TEX15 mutations in a family with nonobstructive azoospermia. Gynecol. Obstet. Invest. *82*, 283–286.
15. Sha, Y., Zheng, L., Ji, Z., Mei, L., Ding, L., Lin, S., Wang, X., Yang, X., and Li, P. (2018). A novel TEX11 mutation induces azoospermia: a case report of infertile brothers and literature review. BMC Med. Genet. *19*, 63.
16. Yatsenko, A.N., Georgiadis, A.P., Röpke, A., Berman, A.J., Jaffe, T., Olszewska, M., Westernströer, B., Sanfilippo, J., Kurpisz, M., Rajkovic, A., et al. (2015). X-linked TEX11 mutations, meiotic arrest, and azoospermia in infertile men. N. Engl. J. Med. *372*, 2097–2107.
17. Yang, F., Silber, S., Leu, N.A., Oates, R.D., Marszalek, J.D., Skaletsky, H., Brown, L.G., Rozen, S., Page, D.C., and Wang, P.J. (2015). TEX11 is mutated in infertile men with azoospermia and regulates genome-wide recombination rates in mouse. EMBO Mol. Med. *7*, 1198–1210.
18. McLachlan, R.I., Rajpert-De Meyts, E., Hoei-Hansen, C.E., de Kretser, D.M., and Skakkebaek, N.E. (2007). Histological evaluation of the human testis–approaches to optimizing the clinical value of the assessment: mini review. Hum. Reprod. *22*, 2–16.
19. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.
20. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.; and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.
21. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.
22. Wilfert, A.B., Chao, K.R., Kaushal, M., Jain, S., Zöllner, S., Adams, D.R., and Conrad, D.F. (2016). Genome-wide significance testing of variation from single case exomes. Nat. Genet. *48*, 1455–1461.
23. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.
24. Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192.
25. Kwak, S.H., Chae, J., Choi, S., Kim, M.J., Choi, M., Chae, J.H., Cho, E.H., Hwang, T.J., Jang, S.S., Kim, J.I., et al. (2017). Findings of a 1303 Korean whole-exome sequencing study. Exp. Mol. Med. *49*, e356.
26. Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. Science *347*, 1260419.
27. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.
28. Desmet, F.O., Hamroun, D., Lalande, M., Collod-Béroud, G., Claustres, M., and Béroud, C. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. *37*, e67.
29. Fromer, M., and Purcell, S.M. (2014). Using XHMM software to detect copy number variation in whole-exome sequencing data. Curr. Protoc. Hum. Genet. *81*, 1–21.
30. Lopes, A.M., Aston, K.I., Thompson, E., Carvalho, F., Gonçalves, J., Huang, N., Matthiesen, R., Noordam, M.J., Quintela, I., Ramu, A., et al. (2013). Human spermatogenic failure purges deleterious mutation load from the autosomes and both sex chromosomes, including the gene DMRT1. PLoS Genet. *9*, e1003349.
31. Sinnwell, J.P., Therneau, T.M., and Schaid, D.J. (2014). The kinship2 R package for pedigree data. Hum. Hered. *78*, 91–93.
32. Xue, X., Sung, P., and Zhao, X. (2015). Functions and regulation of the multitasking FANCM family of DNA motor proteins. Genes Dev. *29*, 1777–1788.
33. Xue, Y., Li, Y., Guo, R., Ling, C., and Wang, W. (2008). FANCM of the Fanconi anemia core complex is required for both monoubiquitination and DNA repair. Hum. Mol. Genet. *17*, 1641–1652.

34. Tao, Y., Jin, C., Li, X., Qi, S., Chu, L., Niu, L., Yao, X., and Teng, M. (2012). The structure of the FANCM-MHF complex reveals physical features for functional assembly. Nat. Commun. *3*, 782.
35. Wang, Y., Han, X., Wu, F., Leung, J.W., Lowery, M.G., Do, H., Chen, J., Shi, C., Tian, C., Li, L., and Gong, W. (2013). Structure analysis of FAAP24 reveals single-stranded DNA-binding activity and domain functions in DNA damage response. Cell Res. *23*, 1215–1228.
36. Jamsai, D., O'Connor, A.E., O'Donnell, L., Lo, J.C., and O'Bryan, M.K. (2014). Uncoupling of transcription and translation of Fanconi anemia (FANC) complex proteins during spermatogenesis. Spermatogenesis *5*, e979061.
37. Fouquet, B., Pawlikowska, P., Caburet, S., Guigon, C., Mäkinen, M., Tanner, L., Hietala, M., Urbanska, K., Bellutti, L., Legois, B., et al. (2017). A homozygous *FANCM* mutation underlies a familial case of non-syndromic primary ovarian insufficiency. eLife *6*, e30490.
38. Peterlongo, P., Catucci, I., Colombo, M., Caleca, L., Mucaki, E., Bogliolo, M., Marin, M., Damiola, F., Bernard, L., Pensotti, V., et al.; GENESIS; kConFab; and SWE-BRCA (2015). FANCM c.5791C>T nonsense mutation (rs144567652) induces exon skipping, affects DNA repair activity and is a familial breast cancer risk factor. Hum. Mol. Genet. *24*, 5345–5355.
39. Harper, J.C., Aittomäki, K., Borry, P., Cornel, M.C., de Wert, G., Dondorp, W., Geraedts, J., Gianaroli, L., Ketterson, K., Liebaers, I., et al.; on behalf of the European Society of Human Reproduction and Embryology and European Society of Human Genetics (2018). Recent developments in genetics and medically assisted reproduction: from research to clinical applications. Eur. J. Hum. Genet. *26*, 12–33.
40. Rosado, I.V., Niedzwiedz, W., Alpi, A.F., and Patel, K.J. (2009). The Walker B motif in avian FANCM is required to limit sister chromatid exchanges but is dispensable for DNA crosslink repair. Nucleic Acids Res. *37*, 4360–4370.
41. Meetei, A.R., Medhurst, A.L., Ling, C., Xue, Y., Singh, T.R., Bier, P., Steltenpool, J., Stone, S., Dokal, I., Mathew, C.G., et al. (2005). A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. Nat. Genet. *37*, 958–963.
42. Yan, Z., Delannoy, M., Ling, C., Daee, D., Osman, F., Muniandy, P.A., Shen, X., Oostra, A.B., Du, H., Steltenpool, J., et al. (2010). A histone-fold complex and FANCM form a conserved DNA-remodeling complex to maintain genome stability. Mol. Cell *37*, 865–878.
43. Che, R., Zhang, J., Nepal, M., Han, B., and Fei, P. (2018). Multifaceted Fanconi anemia signaling. Trends Genet. *34*, 171–183.
44. Gunes, S., Al-Sadaan, M., and Agarwal, A. (2015). Spermatogenesis, DNA damage and DNA repair mechanisms in male infertility. Reprod. Biomed. Online *31*, 309–319.
45. Bakker, S.T., van de Vrugt, H.J., Rooimans, M.A., Oostra, A.B., Steltenpool, J., Delzenne-Goette, E., van der Wal, A., van der Valk, M., Joenje, H., te Riele, H., and de Winter, J.P. (2009). Fancm-deficient mice reveal unique features of Fanconi anemia complementation group M. Hum. Mol. Genet. *18*, 3484–3495.
46. Luo, Y., Hartford, S.A., Zeng, R., Southard, T.L., Shima, N., and Schimenti, J.C. (2014). Hypersensitivity of primordial germ cells to compromised replication-associated DNA repair involves ATM-p53-p21 signaling. PLoS Genet. *10*, e1004471.
47. Bogliolo, M., and Surrallés, J. (2015). Fanconi anemia: a model disease for studies on human genetics and advanced therapeutics. Curr. Opin. Genet. Dev. *33*, 32–40.
48. Neidhardt, G., Hauke, J., Ramser, J., Groß, E., Gehrig, A., Müller, C.R., Kahlert, A.K., Hackmann, K., Honisch, E., Niederacher, D., et al. (2017). Association between loss-of-function mutations within the FANCM gene and early-onset familial breast cancer. JAMA Oncol. *3*, 1245–1248.
49. Kiiski, J.I., Pelttari, L.M., Khan, S., Freysteinsdottir, E.S., Reynisdottir, I., Hart, S.N., Shimelis, H., Vilske, S., Kallioniemi, A., Schleutker, J., et al. (2014). Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer. Proc. Natl. Acad. Sci. USA *111*, 15172–15177.
50. Catucci, I., Osorio, A., Arver, B., Neidhardt, G., Bogliolo, M., Zanardi, F., Riboni, M., Minardi, S., Pujol, R., Azzollini, J., et al. (2018). Individuals with FANCM biallelic mutations do not develop Fanconi anemia, but show risk for breast cancer, chemotherapy toxicity and may display chromosome fragility. Genet. Med. *20*, 452–457.
51. Kiiski, J.I., Tervasmäki, A., Pelttari, L.M., Khan, S., Mantere, T., Pylkäs, K., Mannermaa, A., Tengström, M., Kvist, A., Borg, Å., et al. (2017). FANCM mutation c.5791C>T is a risk factor for triple-negative breast cancer in the Finnish population. Breast Cancer Res. Treat. *166*, 217–226.
52. Nguyen-Dumont, T., Myszka, A., Karpinski, P., Sasiadek, M.M., Akopyan, H., Hammet, F., Tsimiklis, H., Park, D.J., Pope, B.J., Slezak, R., et al. (2018). FANCM and RECQL genetic variants and breast cancer susceptibility: relevance to South Poland and West Ukraine. BMC Med. Genet. *19*, 12.
53. Bogliolo, M., Bluteau, D., Lespinasse, J., Pujol, R., Vasquez, N., d'Enghien, C.D., Stoppa-Lyonnet, D., Leblanc, T., Soulier, J., and Surrallés, J. (2018). Biallelic truncating FANCM mutations cause early-onset cancer but not Fanconi anemia. Genet. Med. *20*, 458–463.
54. Singh, T.R., Bakker, S.T., Agarwal, S., Jansen, M., Grassman, E., Godthelp, B.C., Ali, A.M., Du, C.H., Rooimans, M.A., Fan, Q., et al. (2009). Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M. Blood *114*, 174–180.
55. Lim, E.T., Würtz, P., Havulinna, A.S., Palta, P., Tukiainen, T., Rehnström, K., Esko, T., Mägi, R., Inouye, M., Lappalainen, T., et al.; Sequencing Initiative Suomi (SISu) Project (2014). Distribution and medical impact of loss-of-function variants in the Finnish founder population. PLoS Genet. *10*, e1004494.
56. Amin, M.B., and Tickoo, S.K. (2016). Diagnostic Pathology: Genitourinary, Second Edition (Elsevier).
57. Nagirnaja, L., Aston, K.I., and Conrad, D.F. (2018). Genetic intersection of male infertility and cancer. Fertil. Steril. *109*, 20–26.
58. Grigorova, M., Punab, M., Poolamets, O., Sõber, S., Vihljajev, V., Žilaitienė, B., Erenpreiss, J., Matulevičius, V., Tsarev, I., and Laan, M. (2013). Study in 1790 Baltic men: FSHR Asn680Ser polymorphism affects total testes volume. Andrology *1*, 293–300.