ORIGINAL ARTICLE

ANDROLOGY WILEY

The human sperm proteome—Toward a panel for male fertility testing

Thomas Greither[1] ⓘ | Mario Dejung[2] | Hermann M. Behre[1] | Falk Butter[3] |
Holger Herlyn[4] ⓘ

[1]Center for Reproductive Medicine and Andrology, Martin Luther University Halle-Wittenberg, Halle, Germany

[2]Proteomics Core Facility, Institute of Molecular Biology, Mainz, Germany

[3]Department of Quantitative Proteomics, Institute of Molecular Biology, Mainz, Germany

[4]Anthropology, Institute of Organismic and Molecular Evolution, Johannes Gutenberg University Mainz, Mainz, Germany

Correspondence
Thomas Greither, Center for Reproductive Medicine and Andrology, Martin Luther University Halle-Wittenberg, Halle, Germany.
Email: thomas.greither@medizin.uni-halle.de

Holger Herlyn, Anthropology, Institute of Organismic and Molecular Evolution, Johannes Gutenberg University Mainz, Mainz, Germany.
Email: hherlyn@uni-mainz.de

Abstract

Background: Although male factor accounts for 40%–50% of unintended childlessness, we are far from fully understanding the detailed causes. Usually, affected men cannot even be provided with a molecular diagnosis.

Objectives: We aimed at a higher resolution of the human sperm proteome for better understanding of the molecular causes of male infertility. We were particularly interested in why reduced sperm count decreases fertility despite many normal-looking spermatozoa and which proteins might be involved.

Material and methods: Applying mass spectrometry analysis, we qualitatively and quantitatively examined the proteomic profiles of spermatozoa from 76 men differing in fertility. Infertile men had abnormal semen parameters and were involuntarily childless. Fertile subjects exhibited normozoospermia and had fathered children without medical assistance.

Results: We discovered proteins from about 7000 coding genes in the human sperm proteome. These were mainly known for involvements in cellular motility, response to stimuli, adhesion, and reproduction. Numbers of sperm proteins showing at least threefold deviating abundances increased from oligozoospermia ($N = 153$) and oligoasthenozoospermia ($N = 154$) to oligoasthenoteratozoospermia ($N = 368$). Deregulated sperm proteins primarily engaged in flagellar assembly and sperm motility, fertilization, and male gametogenesis. Most of these participated in a larger network of male infertility genes and proteins.

Discussion: We expose 31 sperm proteins displaying deviant abundances under infertility, which already were known before to have fertility relevance, including ACTL9, CCIN, CFAP47, CFAP65, CFAP251 (WDR66), DNAH1, and SPEM1. We propose 18 additional sperm proteins with at least eightfold differential abundance for further testing of their diagnostic potential, such as C2orf16, CYLC1, SPATA31E1, SPATA31D1, SPATA48, EFHB (CFAP21), and FAM161A.

Conclusion: Our results shed light on the molecular background of the dysfunctionality of the fewer spermatozoa produced in oligozoospermia and syndromes including it. The male infertility network presented may prove useful in further elucidating the molecular mechanism of male infertility.

KEYWORDS
biomarker, disorder, label-free quantification, male infertility, sperm motility, sperm proteome

1 | INTRODUCTION

An estimated 8%–15% of couples are involuntarily childless worldwide, with male factor infertility accounting for 40%–50% of the cases.[1–3] Although the causes of the disorder are often suspected in reduced sperm quantity and/or quality, the proteome of the spermatozoon has not yet been conclusively elucidated. Instead of giving the full picture, the available data appear to mainly reflect the rising resolution of the analytical pipelines employed. For example, while the liquid chromatography–mass spectrometry (MS) study by Baker et al.[4] detected 1056 gene products in human spermatozoa, the pipeline employed by Wang et al.[5] enabled the determination of 4675 sperm proteins. Larger numbers of 6198 and 6238 human sperm proteins were gathered in subsequent reviews based on proteomics studies.[6,7] The recent compilation by Castillo et al.[8] even contained 6871 human sperm proteins. However, such a scale has not been confirmed in single MS analysis so far.

As with the sperm proteome, the molecular causes of male fertility impairment are only partially understood. Previous studies in this area were based on transcript abundances in testes and spermatozoa[9,10,11] or consulted a methodologically broad range of published data.[12,13] In addition, mutations, copy-number variations, and polymorphisms in the genome have been checked for associations with infertility.[14,15] Also, epigenetic deviations may be associated with fertility impairment[16]; however, this is still debated.[17] Others derived fertility relevance from parameters such as sequence conservation and phenotypes of murine knockouts,[18,19] and quantitative proteomics of spermatozoa is increasingly being incorporated.[20–22] Yet, the diversity of approaches has not made it easier to define a standard panel of male fertility markers for routine diagnostics. Rather, the corner stone of male fertility diagnostics remains to be the recording of spermiogram parameters.[7,23]

For obvious reasons, the absence of any spermatozoa in ejaculates or azoospermia causes infertility, whereas the presence of a certain amount of fully functional spermatozoa (normozoospermia) suggests fertility. But the question is as to why a reduction in sperm count can be critical to fertilization success, if there are still many motile and morphologically normal spermatozoa.[24] In fact, the criteria of oligozoospermia (O) (i.e., less than 39 million spermatozoa per ejaculate and less than 16 million spermatozoa per milliliter) still allow for the thousands of seemingly unimpaired spermatozoa per ejaculate. The same may be asked for asthenozoospermia, that is, a decrease in the proportion of progressively moving spermatozoa to less than 30%, and

teratozoospermia, that is, a decline in the percentage of spermatozoa with normal morphology to less than 4%.[3,25,26] This is contrasted by only 100–1000 functional sperm cells, which assumedly reach the egg cell in natural fertilization.[27] A possible explanation is that spermatozoa that appear to be fully functional are actually dysfunctional. Consistent with this, poorer semen parameters are accompanied by reduced success rates of assisted reproduction techniques (ARTs) despite the use of spermatozoa considered promising.[24] The cause of such a dysfunction would then have to be sought in the molecular field, which would blur the boundaries to idiopathic infertility or infertility without phenotypic evidence.[1,12,23,28] In fact, about 40% of the infertile men are idiopathic infertile,[14] which underlines the high need for progress in this field of diagnostics.

The present study aims at an in-depth reconstruction of the human sperm proteome and an elucidation of the molecular causes of male fertility losses. We were especially interested as to why fecundity is reduced in O and syndromes including it, and which proteins might be useful biomarkers for assessing the fertility status of respective men. To address these points, we carried out MS analysis of spermatozoa in an unprecedented sample of 76 men. Donors were either fertile and presented with normozoospermia or infertile and diagnosed with O, oligoasthenozoospermia (OA), or oligoasthenoteratozoospermia (OAT). In contrast to previous proteomic studies based on pools of sperm samples,[21,28–30] spermatozoa of single subjects were analyzed separately. From the differentially abundant proteins, we distilled out those, which should have elevated diagnostic potential. Finally, we demonstrate that non-overlapping sets of (candidate) markers of male fertility together establish the male infertility network.

2 | MATERIALS AND METHODS

2.1 | Study cohorts

Permission for sample collection was granted by the Ethics Committee of the Medical Faculty at the Martin Luther University Halle/Saale (approval number 218/14.04.10/2; date of decision: April 19, 2010; date of approved amendment: March 14, 2013). All participants were recruited from the fertility outpatient care of the University Hospital Halle (Saale) and gave written informed consent. Ejaculates were provided by 76 men (Table S1). Spermiogram parameters were re-assessed according to the updated version of the World Health Organization (WHO) laboratory manual for the examination and processing of

human semen (6th edition, 2021),[26] with no changes in the diagnoses of the study participants. For clarity reasons, we use descriptive terms for abnormal spermiograms (O, asthenozoospermia, teratozoospermia, and mixed forms) as proposed in previous WHO laboratory manuals (5th edition and earlier). Nonetheless, the authors acknowledge the fact that the cut-off values used for defining these terms do not necessarily determine the fertility status of an individual male. Thirty-eight infertile men (mean and median age = 32 years, each) presented with O (21), OA (7), or OAT (10), and were unintentionally childless, that is, their partners did not conceive despite regular unprotected intercourse within at least 12 months. In the 38 infertile patients, median sperm count per ejaculate was 36.5 mio (mean = 41.4 mio). The control cohort included 38 men (mean and median age = 30 years, each) presenting with normozoospermia, who already had fathered offspring without medical assistance. Median sperm count of the corresponding ejaculates was 155.0 mio (mean = 205.7 mio) (Table S1).

2.2 | Sample preparation for MS

Upon swim-up, spermatozoa were separated from seminal plasma by centrifugation for 5 min at 1000 g. After discarding the plasma, cells were resuspended twice in phosphate-buffered saline and re-pelleted by centrifugation at 2000 g for 5 min, each time succeeded by discard of the supernatant. Washed spermatozoa were shock-frosted in liquid nitrogen and stored at –80°C. Following thawing on ice, samples were boiled in 1× lithium dodecyl sulfate buffer and separated by polyacrylamide gel electrophoresis on a 4%–12% gradient NOVEX gel (Thermo Scientific) for 10 min at 180 V. The gel was fixated and stained with Coomassie blue for 30 min and subsequently washed with water overnight. In-gel digestion was performed as previously described.[31] In short, each sample was cut from the gel, minced, and transferred to a reaction tube. The gel pieces were destained with 50% EtOH/25 mM ammonium bicarbonate (ABC) buffer (pH 8.0). After dehydration of the gel pieces with pure acetonitrile (ACN), samples were dried for 5 min in a concentrator (Eppendorf) and afterward incubated with reduction buffer (10 mM dithiothreitol in 50 mM ABC) for 30 min at 56°C. The reduction buffer was removed, substituted with alkylation buffer (50 mM iodoacetamide in 50 mM ABC), and then subjected to 30 min incubation at room temperature in the dark. Gel pieces were completely dehydrated with pure ACN and covered in trypsin solution (1 µg trypsin in 50 mM ABC per sample). Proteins were digested overnight at 37°C. Tryptic peptides were extracted twice by incubation with extraction buffer (3% trifluoroacetic acid and 30% ACN) for 15 min and afterward with pure ACN. After concentration of the elution fraction to about 10%–20% in a concentrator (Eppendorf), the peptides were passed through a StageTip.[32] StageTips were prepared using two layers of C18 material (Empore), which was activated with methanol, washed with solution B (80% ACN, 0.1% formic acid), and equilibrated once with solution A (50 mM ABC, 0.1% formic acid). Extracted peptides were loaded on the StageTips and washed with solution A. Peptides were eluted with 30 µL solution B, ACN was removed by use

of a concentrator (Eppendorf), and samples were diluted with 6 µL solution A.

2.3 | MS measurement

The samples were injected via an autosampler into an uHPLC (EASY-nLC 1000, Thermo Scientific). Peptides were loaded on a 25 cm capillary (75 µm inner diameter; New Objective) packed in-house with Reprosil C18-AQ 1.9 µm resin (Dr. Maisch) for reverse-phase chromatography. The EASY-nLC 1000 HPLC system was directly mounted to a Q Exactive Plus mass spectrometer (Thermo Scientific). Peptides were eluted from the column with a 208 min optimized gradient from 2% to 40% ACN with 0.1% formic acid at a flow rate of 225 nL/min with a column oven set-up operating at 40°C (Sonation). The heated capillary temperature was set to 250°C. Spray voltage ranged from 2.2 to 2.4 kV. The mass spectrometer was operated in data-dependent acquisition mode with one MS full scan and up to 10 triggered MS/MS scans using higher energy collisional dissociation. MS full scans were obtained in the orbitrap at 70,000 resolution with a maximal injection time of 20 ms, while MS/MS scan resolution was set to 17,500 and maximal injection for 120 ms. Unassigned and charge state 1 were excluded from MS/MS selection and peptide match was preferred. Detailed information can be obtained from the measurement files available on the public ProteomeXchange repository (PXD037531).

2.4 | MS data analysis

Raw data analysis was performed with MaxQuant v1.5.2.8[33] with standard settings except label-free quantification (LFQ) and match between runs were activated. Time windows for matching and aligning were 0.7 and 20 min, respectively. MaxQuant analysis including detailed settings has been deposited at ProteomeXchange under the above accession number. After MaxQuant, potential contaminants, reverse database hits and proteins only identified by site were removed. Further, we filtered for a minimum of two identified peptides (minimum 1 unique). For quantitative proteome analyses, lacking values were imputed assuming a beta distribution within 0.2 and 2.5 percentiles of measured LFQ intensities per sample. Downstream comparisons of protein abundances focused on spermatozoa of infertile men diagnosed with reduced sperm count and fertile men with normal semen parameters. Differentially abundant sperm proteins were grouped according to infertility diagnoses, thus generating O list (O vs. normozoospermia), OA list (OA vs. normozoospermia), and OAT list (OAT vs. normozoospermia). For being contained in a list, a protein needed to be detected in at least two samples of a diagnosis, and abundances had to vary by a factor of three at minimum (false discovery rate [FDR] corrected p-value ≤ 0.05, Welch two sample t-test). This corresponds to a minimum difference of log2-transformed values of 1.585. We additionally highlight sperm proteins showing at least eightfold differential abundances, corresponding to a minimum difference of log2-transformed values of 3.000.

2.5 | Further analyses

We ran ShinyGO v0.76[34] on ENSEMBL stable IDs of the genes encoding MS-determined sperm proteins for deriving the lengths of coding sequences and transcripts. We also adopted from the ShinyGO output the grouping of genes according to higher level biological process gene ontologies (GOs). Moreover, we matched ENSEMBL stable gene IDs from present MS analysis with the genes to the sperm proteins collected by Castillo et al.[8] For the latter compilation, UniProt/SwissProt IDs had been converted into ENSEMBL IDs using BioMart (Ensembl Genes 106, Human genes, GRCh38.p13).

Overlaps between above O-associated lists of genes/proteins were rendered in a Venn diagram (https://bioinformatics.psb.ugent.be/webtools/Venn/). We tested for the overrepresentation of biological process GOs in O, OA, and OAT lists using ClueGO v2.5.9[35] within Cytoscape v3.9.1.[36] In addition, we matched the list for the most comprehensive infertility diagnosis, that is, the OAT list, with entries in NCBI's Online Mendelian Inheritance in Men (OMIM) database (state February 1, 2022) referencing male fertility impairment. Corresponding OMIM entries were related to O and non-obstructive azoospermia using all possible forms of spermatogenetic failure (SPGF) as search items: SPGF1, SPGF2, …, SPGF65, SPGFX1, SPGFX2, SPGFX3, SPGFY1, and SPGFY2. We further collected OMIM entries on asthenozoospermia, teratozoospermia, and male infertility. We screened the WWW for publications using the combination of symbol and infertil* as initial search items (e.g., APAF1 infertil*). We additionally queried The Human Protein Atlas v22 (http://proteinatlas.org/; accessed on January 10, 2023) for the expressional landscape of the genes behind differentially abundant sperm proteins.

Subsequently, we assessed if the candidate markers presented herein and additional ones emerging from external sources might be connected at a higher level. For this purpose, we reconstructed three protein–protein interaction (PPI) networks of growing complexity. Using standard settings in STRING v11.5,[37] the first network was confined to the biomarkers as inferred from present empirical data. The second network additionally considered candidate markers as emerging from a previous computational approach. Corresponding genes/proteins were calculated to have elevated fertility probability (≥0.35) given enhanced sequence conservation, heightened testicular transcript abundance, and increased connectivity in a body-wide PPI network. Furthermore, knockouts of their murine orthologs associated with male fertility impairment.[19] We extracted the aforementioned proteins from Dataset M at the PreFer Genes website (https://prefer-genes.uni-mainz.de/). A third network was reconstructed after further expansion of the sample by all proteins with previous mention in the context of male infertility in OMIM database (see above). The networks were compared by their overall coherence and average node degree, that is, the mean number of PPIs per protein or node. Each network was tested for PPI enrichment.

Raw p-values from chi-square tests, GO enrichment tests, and PPI enrichment tests were transformed into FDRs.[38] Further consideration of test results required an FDR of 0.001 or lower.

3 | RESULTS

3.1 | Fine-scale reconstruction of the human sperm proteome

Present MS analysis identified 5778 protein groups ($q \leq 0.05$, each) across sperm samples provided by 76 men and purified by swim-up (Table S1). These contained 33,191 gene products to which 6981 ENSEMBL stable gene IDs were available. Of the latter, 4545 had counterparts in the compilation of sperm proteins by Castillo et al.[8] Thus, 65% of the sperm proteins for which we found evidence in human spermatozoa can be regarded as reproduced (Table S2). The 6939 genes mapped by ShinyGO significantly differed from the rest in the human genome by having shorter coding sequences (FDR = 6E-15) and transcripts (FDR = 1.7E-20; t-test, each) (Figure 1). A closer look for the 50 biological process GOs occurring with highest frequencies corroborated the importance of sperm motility for fertilization. Processes such as cell proliferation and reproduction, adhesion, and responsiveness to stimuli were also represented, as were involvements in developmental processes and the immune system. The latter included genes coding for spermatogenesis-associated protein 20 (SPATA20), zona pellucida-binding protein 1 (ZPBP), and sperm-associated antigen 4 (SPAG4) (Table S3). Additional higher ranking terms were of more general nature such as "catabolic process" (Figure 2). The GOs of the remaining 50 smaller protein groups partially confirmed roles in male gametogenesis (meiotic cell cycle process, spermatid nucleus differentiation) and fertilization (acrosome reaction), while others referred to additional entanglements (Table S3).

3.2 | Exactness of MS analysis

For being considered, a protein had to be determined in at least one unique and one razor peptide across the 76 samples analyzed. However, much higher values were achieved in the validated and novel candidate markers detailed below (Tables 1–4 and S2). In case of the validated marker candidates, we had 4–246 (mean = 51, median = 32) determinations of unique and razor peptides. These were mostly unique peptides, as illustrated by corresponding mean and median numbers of 47 and 29, respectively. In the novel candidate markers, the span of unique and razor peptides ranged from 7 to 177 (mean = 33, median = 21). Again, these were mostly specific to a particular protein group, as indicated by mean (32) and median (21) values of unique peptides. We consider these counts as a confirmation of the reliability of the detections.

3.3 | Quantitative MS analysis

The pipeline applied unveiled numerous protein groups with at least threefold differential, mostly lowered, abundances (FDR ≤ 0.05, each; t-test) in spermatozoa of infertile men diagnosed with O, OA, and OAT,

TABLE 1 Validated candidate markers: sperm proteins exhibiting differential abundances that have previous mention in the context of male fertility impairment.

Symbol	ID	OAT	O	OA	Functional annotation (UniProt)	THPA
ACTL9*	Q8TC94	0.173	↓	↓	Acrosome biogenesis and perinuclear theca formation	I
CCIN*	Q13939	0.256	↓	↓	Cytoskeletal element in spermiogenesis	I
CFAP47*	Q6ZTR5	0.128	↓	↓	Flagellar formation and sperm motility	III
CFAP65*	Q6ZU64	0.236	↓	↓	Flagellar formation and sperm motility	III
WDR66*	Q8TBY9	0.254	↓	↓	Spermatozoa motility	III
DNAH1*	Q9P2D7	0.193	↓	↓	ATPase activity; sperm flagellum motility; formation of the inner dynein arms and biogenesis of the axoneme	III
SPEM1*	Q5F289	0.261	↓	↓	Cytoplasm removal during spermatogenesis	I
ACTL7A*	Q9Y615	0.238	↓		Acrosome biogenesis	I
MNS1*	Q9C512	0.133	↓		Control of meiotic division and germ cell differentiation (…) during meiosis; sperm flagellum assembly	II
PLCZ1*	Q86YW0	0.182	↓		Initiates embryonic development	I
PMFBP1*	Q8TBY8	0.270	↓		Spermatogenesis (…); maintenance of sperm head and tail integrity	II
QRICH2*	Q9H0J4	0.204	↓		Scaffold protein (…); maintenance of sperm head and tail integrity	II
SPAG17*	Q6Q759	0.184	↓		Component of the central pair apparatus of axonemes; function and structure of motile cilia; spermatogenesis; formation of sperm head and flagellum	III
TCTE1*	Q5JU00	0.199	↓		Nexin–dynein regulatory complex component; key regulator of ciliary/flagellar motility; microtubule sliding in motile axonemes; sperm motility	II
CFAP43*	Q8NDM7	0.208		↓	Sperm flagellum axoneme organization and function	III
AK7*	Q96M32	0.258			Ciliary structure and function	III
APAF1*	O14727	3.578			Apoptosis trigger	IV
CATIP*	Q7Z7H3	0.293			Ciliogenesis	IV
CATSPER2*	Q96P56	0.186			Voltage-gated channel essential for sperm hyperactivation, acrosome reaction, and chemotaxis toward the oocyte	III
CFAP44*	Q96MT7	0.224			Sperm flagellum axoneme organization and function	III
CFAP52*	Q8N1V2	0.283			Ciliary and flagellar beating	III
CFAP58*	Q5T655	0.291			Assembly and organization of the sperm flagellar axoneme	III
CFAP69*	A5D8W1	0.215			Sperm flagellum assembly and stability	IV
CFAP70*	Q5T0N1	0.232			Axoneme-binding protein; regulation of ciliary motility and cilium length	III
DNAH10*	Q8IVF4	0.268			ATPase activity; sperm motility; sperm flagellar assembly	III
DNAH17*	Q9UFH2	0.300			ATPase activity; outer dynein arms (…) in the sperm flagellum; sperm motility; sperm flagellar assembly and beating	III
FSIP2*	Q5CZC0	0.241			Spermatogenesis	II
SMCP*	P49901	0.167			Sperm motility	I
SPEF2*	Q9C093	0.215			Development of sperm axoneme and manchette; sperm head morphology; adapter for dynein-mediated protein transport in spermatogenesis	III
SUN5*	Q8TC36	0.258			Anchoring sperm head to the tail; attachment of the coupling apparatus to the sperm nuclear envelope	I
TTC29*	Q8NA56	0.252			Axonemal and/or peri-axonemal assembly and structure; sperm flagellum assembly and beating	II

Note: Listed are sperm proteins with previous mention in the context of male fertility impairment (NCBI Online Mendelian Inheritance in Men [OMIM]) that showed at least threefold deviating abundances in infertile men diagnosed with oligoasthenoteratozoospermia (OAT) (false discovery rate [FDR] ≤ 0.05, *t*-test). Arrows indicate the direction of abundance changes in infertile men diagnosed with oligozoospermia (O) or oligoasthenozoospermia (OA). Changes refer to abundances in normozoospermic spermatozoa of fertile men. WDR66 corresponds to CFAP251. Asterisks highlight proteins included in the largest connected component shown in Figure 6. Categories on the left refer to transcript abundances reported in The Human Protein Atlas v22 (THPA): male germline-specific (I), male germline-biased (II), elevated in male germline and somatic cell types, which unlikely (III) or possibly were contained in the samples (IV).

TABLE 2 Validated candidate markers: external evidence supporting male fertility relevance.

Symbol	Evidence	Panel
ACTL9	Variants: fertilization failure, male infertility[39]	Yes
CCIN	Mutations: teratozoospermia, male infertility[40] Sperm head-shaping factor essential for male fertility[41]	
CFAP47	Variants: asthenoteratozoospermia, male infertility[42]	Yes
CFAP65	Mutation: MMAF, male infertility[43] Mutations: MMAF, male infertility[44]	
WDR66 CFAP251	Deletion: MMAF, male infertility[45] Loss: immotile spermatozoa, lacking mitochondria, male infertility[46]	Yes
DNAH1	Mutations: MMAF[47]	
SPEM1	Predictive value for sperm retrieval in azoospermia[48] Lack in mice: aberrant cytoplasm rem., sperm deformation, male infertility[49]	
ACTL7A	Disruption: acrosomal defects, early embryonic arrest[50] Variants: total fertilization failure, male infertility[51]	
MNS1	Mutations: laterality defects, likely male infertility[52] Variant: laterality defects, male infertility[53]	
PLCZ1	Mutation: infertility, oocyte activation deficiency[54]	Yes
PMFBP1	Mutations: ASS[55] Mutations: ASS[56,57]	Yes
QRICH2	Mutations: MMAF[58] Mutations: MMAF, male infertility[59]	Yes
SPAG17	Mutation: asthenozoospermia[60]	
TCTE1	Variants: asthenozoospermia, male infertility[61]	
CFAP43	Mutations: MMAF, male infertility[62] Mutations: MMAF[63]	Yes
AK7	Mutation: MMAF, male infertility[64]	Yes
APAF1	Raised testicular expression: male subfertility[65]	
CATIP	Mutation: oligoasthenoteratozoospermia[66]	Yes
CATSPER2	Non-syndromic male infertility[67] Copy-number variant, reduced expression: idiopathic male infertility[68]	
CFAP44	Mutations: MMAF, male infertility[62] Reduced expression: potential effect on sperm motility and morphology[69]	
CFAP52	Exon 2 deletion: male infertility[70,71]	
CFAP58	Variants: asthenozoospermia, flagellar axoneme + mitoch. sheath defects[72]	Yes
CFAP69	Absence: MMAF, male infertility[73]	Yes
CFAP70	Mutations: asthenoteratozoospermia, male infertility[74]	Yes
DNAH10	Mutations: asthenoteratozoospermia, male infertility[75] Variants: asthenoteratozoospermia, male infertility[76]	Yes
DNAH17	Mutations: asthenozoospermia, male infertility[77] Variant: asthenozoospermia, flagella destabilization[78] Intronic deletion: perturbed splicing, defect sperm flagella, male infertility[79]	Yes
FSIP2	Mutations: MMAF[80,81]	Yes
SMCP	Diff. expression in non-obstructive azoospermia versus oligoasthenozoospermia[82] Deletion in mice: asthenozoospermia[83]	
SPEF2	Mutations: defects between flagella and cilia bridge, link of MMAF and PCD[84] In mice: germ cell differentiation[85]	Yes
SUN5	Mutations: ASS[86,87]	Yes
TTC29	Mutations: asthenozoospermia, male infertility[88]	Yes

Note: Evidence refers to men if not stated otherwise. "Yes" indicates inclusion in a male fertility panel of 81 genes (ID192.04).
Abbreviations: ASS, acephalic spermatozoa syndrome; Diff., differential; mitoch., mitochondrial; MMAF, multiple morphological abnormalities of the flagella; PCD, primary ciliary dyskinesia; rem., removal.

TABLE 3 Novel candidate markers: sperm proteins with the greatest decline in abundance.

Symbol	UniProt ID	OAT	O	OA	Functional annotation	THPA
C2orf16	Q68DN1	0.032	↓	↓	High-abundance sperm protein in normozoospermic samples that achieved pregnancy[20]	II
CYLC1*	P35663	0.056	↓	↓	Architectural role during spermatogenesis; spermatid differentiation (UniProt)	I
SPATA31E1*	Q6ZUB1	0.061	↓	↓	Spermatogenesis (UniProt)	II
SPATA31D1	Q6ZQQ2	0.090	↓	↓	Spermatogenesis (UniProt)	II
SPATA48*	A4D263	0.099	↓	↓	Essential for normal spermatogenesis (UniProt)	II
EFHB	Q8N7U6	0.118	↓	↓	Cilia- and flagella-associated protein 21 (HGNC)	II
FAM161A*	Q3B820	0.124	↓	↓	Ciliogenesis (UniProt)	III
CYLC2*	Q14093	0.086	↓		Architectural role during spermatogenesis; spermatid differentiation (UniProt)	I
CCDC173	Q0VFZ6	0.087	↓		Cilia- and flagella-associated protein 210 (UniProt)	II
SPEM3	A0A1B0GUW6	0.092	↓		Abundance varies with fertility differences in boars[89]	I
CFAP126*	Q5VTH2	0.098	↓		Cilium basal body docking and positioning (UniProt)	II
FAM71E2	Q8N5Q1	0.105	↓		Epigenetic transcriptional repression (UniProt)	II
C1orf194*	Q5T5A4	0.105	↓		Downregulated abundance in spermatozoa of patients with testicular cancer seminoma[90]	II
C1orf158*	Q8N1D5	0.110	↓		Upregulated testicular expression in non-obstructive azoospermia[91]	II
DCDC2C	A8MYV0	0.120	↓		Microtubule organization; sperm flagellum (QuickGO)	II
EQTN*	Q9NQ60	0.123	↓		Acrosomal membrane-anchored protein involved in (…) fertilization and in acrosome biogenesis (UniProt)	I
SYPL1	Q16563	0.107			Downregulated testicular expression in non-obstructive azoospermia[92]	IV
FAM205C	A6NFA0	0.118			Abundance varies with fertility differences in boars[89]	II

Note: Listed are sperm proteins without previous reference to male fertility impairment (NCBI Online Mendelian Inheritance in Men [OMIM]) that showed at least eightfold decreased abundances in oligoasthenoteratozoospermia (OAT) (false discovery rate [FDR] ≤ 0.05, t-test). Arrows indicate the direction abundance changes in spermatozoa of infertile men diagnosed with oligozoospermia (O) and oligoasthenozoospermia (OA). Changes refer to protein abundances in normozoospermic sperm proteins of fertile men. Asterisks highlight proteins included in the largest connected component shown in Figure 6. Categories on the left refer to transcript abundances reported in The Human Protein Atlas v22 (THPA): male germline-specific (I), male germline-biased (II), elevated in male germline and somatic cell types, which unlikely (III) or possibly were contained in the samples (IV).

relative to normozoospermic spermatozoa of normal-fertile men (Table S4). Thus, 133 of 153 (87%) differentially abundant protein groups exhibited downregulation in O spermatozoa (O list). With 154, the total number of protein groups varying in abundance was almost the same in the OA versus normozoospermia comparison (OA list), whereby most of these showed declines in abundance under infertility once more (N = 110; 71%). The OAT versus normozoospermia comparison (OAT list) revealed the highest number of differentially abundant protein groups (N = 368), with lowered levels predominating in the infertility cohort again (N = 315; 86%) (Table S4). Fifty-five protein groups were shared among the three O-associated lists. In addition, there was an increasing number of unique protein groups with growing complexity of the diagnosis, and thus from the O via the OA to the OAT list (Figure 3). Also, the number of overrepresented GOs (FDR ≤ 0.001, each) was much higher in the OAT list (N = 72) than in O (N = 26) and OA (N = 20) lists. Nevertheless, cilium-dependent cell motility yielded the largest fraction of GOs in all three infertility diagnoses. The three lists also had in common that 75% of the

overrepresentations pertained to the structure and function of cilia or the sperm flagellum. Additional overrepresentations referred to male gamete generation and fertilization (O, OA, and OAT lists), cAMP signaling (OA and OAT lists), and sodium–potassium exchange (OAT list) (Figure 4).

Twelve percent of the altogether 198 genes with OMIM entries relating to male fertility impairment had counterparts in the most comprehensive of our three lists, the OAT list (Table S5). Broken down to individual aberrations from normozoospermia, we observed the following pattern: the gene list compiled from OMIM entries on teratozoospermia matched to 10% (3 of 29) with genes or proteins in our OAT list. In OMIM lists for asthenozoospermia and teratozoospermia, 17% (12 of 70) and 35% of the genes (23 of 66) had matches in our OAT list. Thus, the proportion of matching genes in all three OMIM lists considerably exceeded the random expectation of <1%, as given by the total number of genes in single OMIM lists (N = 23–70) divided by 20,471 protein-coding human genes according to ENSEMBL Genes 106: GRCh38.p13.

TABLE 4 Novel candidate markers: external evidence supporting male fertility relevance.

| Symbol	Samples	Evidence	Direction
C2orf16	Spermatozoa: 13 teratozoospermia versus 8 normozoospermia	Microarray	↓[93]
	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
	178 x testis: meiotic arrest versus three other diagnoses	Microarray	↓[94]
CYLC1	Spermatozoa: 13 teratozoospermia versus 8 normozoospermia	Microarray	↓[93]
	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
	178 x testis: meiotic arrest versus three other diagnoses	Microarray	↓[94]
SPATA31E1	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
SPATA31D1	Spermatozoa: 5 less versus 5 more fertilization-competent[§]	PCR	↓[95]
SPATA48	Testis: 6 azoospermia versus 6 normal fertile	PCR	↓[96]
EFHB	Spermatozoa: 13 teratozoospermia versus 8 normozoospermia	Microarray	↓[93]
	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
	178 x testis: pre-meiotic arrest versus three other diagnoses	Microarray	↓[94]
FAM161A	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
CYLC2	Spermatozoa: 13 teratozoospermia versus 8 normozoospermia	Microarray	↓[93]
	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
	178 x testis: meiotic arrest versus three other diagnoses	Microarray	↓[94]
CCDC173	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
SPEM3	Spermatozoa: 4 less versus 4 more fertilization-competent[+]	Mass spectrometry	↓[89]
CFAP126	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
	178 x testis: meiotic arrest versus three other diagnoses	Microarray	↓[94]
FAM71E2	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
	178 x testis: meiotic arrest versus three other diagnoses	Microarray	↓[94]
C1orf194	Spermatozoa: 15 testicular cancer seminoma versus 15 control	Mass spectrometry	↓[90]
	Spermatozoa: 20 asthenozoospermia versus 20 normozoospermia	Microarray	↓[9]
C1orf158	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
	178 x testis: meiotic arrest versus three other diagnoses	Microarray	↓[94]
	Spermatozoa: 13 teratozoospermia versus 8 normozoospermia	Microarray	↓[93]
DCDC2C	Spermatozoa: 1 low-fertile male versus 1 high-fertile male[+]	Microarray	↓[97]
EQTN	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
	178 x testis: meiotic arrest versus three other diagnoses	Microarray	↓[94]
SYPL1	Spermatozoa: 13 teratozoospermia versus 8 normozoospermia	Microarray	↓[93]
	178 x testis: pre-meiotic arrest versus three other diagnoses	Microarray	↓[94]
FAM205C	178 x testis: post-meiotic arrest versus three other diagnoses	Microarray	↓[94]
	178 x testis: meiotic arrest versus three other diagnoses	Microarray	↓[94]

Note: Comparisons mostly refer to men and exceptionally to domestic pig (+) and horse (§).
Abbreviation: PCR, polymerase chain reaction.

3.4 | Candidate markers of male fertility

Thirty-one sperm proteins in our OAT list had entries in the OMIM database suggestive of an involvement in male fertility impairment (Table 1). Their GOs related to sperm flagellum assembly, structure and beating, as well as male gametogenesis. We refer to these proteins as to validated candidate markers in the following. Out of these, apoptosis trigger APAF1 was the only one showing higher abundance in OAT than normozoospermic spermatozoa (Table 1). The remaining proteins with previous mention relating to male fertility impairment exhibited
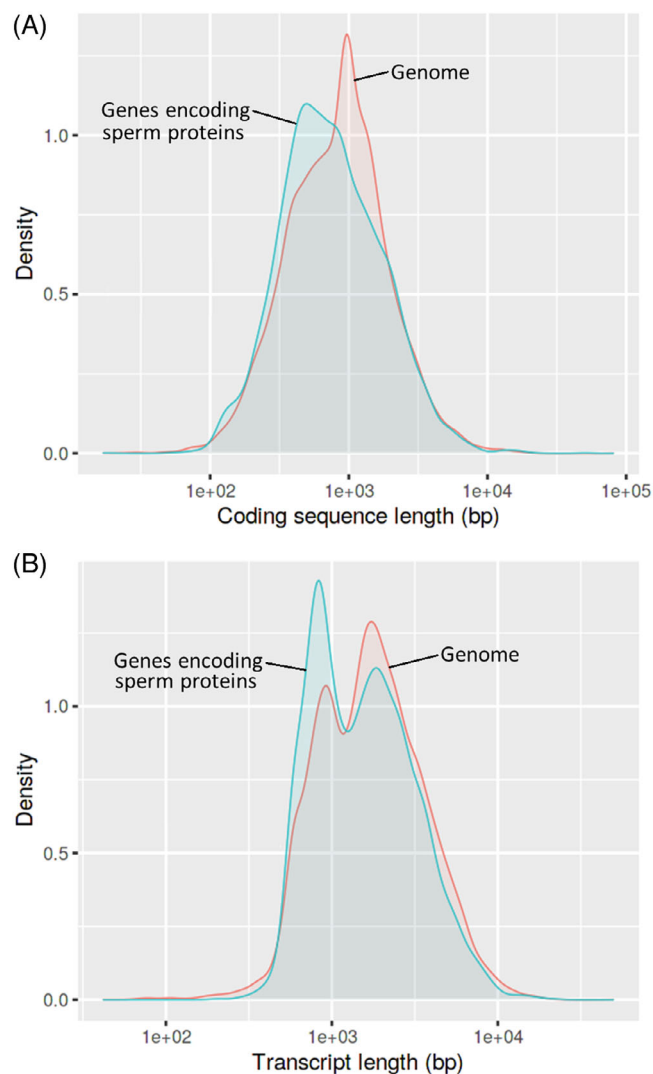
FIGURE 1 Distributions of coding sequence length (A) and transcript length (B) in human genes encoding sperm proteins, relative to the genome. Shifts to the left indicate overall lower values in 6939 sperm protein-coding genes mapped by ShinyGO v0.76 (Table S2). Chi-square tests contrast patterns in sperm protein-coding genes and all other genes were highly significant (false discovery rate [FDR] <0.001, each).

reduced levels under OAT. Seven of these displayed decreased abundances across all three O-associated diagnoses: actin-like protein 9 (ACTL9), calicin (CCIN), CFAP47, CFAP65, CFAP251 (WDR66), dynein axonemal heavy chain 1 (DNAH1), and sperm maturation protein 1 (SPEM1) (Table 1). In-depth search for previous evidence (Table 2) revealed that most validated candidate markers were previously associated with malformed and dysfunctional sperm flagella and impaired sperm motility. Corresponding terms included multiple morphological abnormalities of the flagella (MMAF) and asthenozoospermia ($N = 20$). This was followed by associations with morphological defects such as acephalic spermatozoa syndrome (ASS), acrosomal defects, and lacking mitochondria ($N = 10$). The least connections were found to reduced sperm count ($N = 3$). Notably, 18 of the validated candidate markers of

male fertility were covered by an 81 gene panel and hence are already in use in male fertility diagnostics (Table 2).

Lack of OMIM entries relating to male fertility impairment let us refer to a group of 18 sperm proteins with extremely deviating abundances under infertility as to novel candidate markers. These exhibited at least eightfold reduced abundances in the OAT–normozoospermia comparison (Table 3). According to the data available, the novel candidate markers engage in spermiogenesis, including ciliogenesis and acrosome biogenesis. For all novel candidate markers, we found previous indications of downregulated expression under reduced male fertility. For example, no *SPATA48* transcript was determined in testicular samples of six azoospermic men, while corresponding transcripts were highly abundant in healthy subjects (Table 4). Furthermore, seven proteins had significantly reduced abundances in O and OA spermatozoa, and, hence, might elucidate the deeper causes of male fertility impairment whenever reduced sperm count is suspected to play a role.

There were three protein groups, each containing two proteins, with decreased abundances in spermatozoa under male infertility (Table 5). Half of these proteins had OMIM entries relating to male fertility. Until clarification whether one or both group members occur at lowered abundances in spermatozoa of infertile men, the corresponding proteins lend themselves less readily as candidate markers.

Downstream, we checked for potential somatic expression of the genes to the proteins in Tables 1, 3, and 5. According to version 22 of The Human Protein Atlas, transcripts of most of the genes with validated male fertility relevance (Table 1) were specific to (category I: seven genes) or much more frequent in male germline (category II: six genes), or transcript abundance was elevated in male germline and somatic cell types, which will not have contaminated our samples (category III: 15 genes). The latter was exemplified in respiratory ciliated and endometrial ciliated cells (Figures 5 and S1). Genes of only few validated candidate markers showed the expected expression in male germline and additionally displayed elevated transcript abundances in somatic cell types that could have been contained in the samples analyzed (category IV: three genes). Prevalence of categories I–III was reproduced in genes encoding novel candidate markers (Table 3 and Figure S2). Most of the coding genes displayed exclusive (category I: four genes) or skewed expression in male germline (category II: 12 genes), and one gene additionally exhibited raised transcript amounts in somatic cells, which unlikely were contained in the samples (category III: one gene). The transcriptional profile of only one of the novel candidate markers allowed for the possibility that somatic cell types could have been included in the sample in addition to the targeted spermatozoa (category IV). The genes underlying protein groups even spread without exception across categories I–III (Table 5 and Figure S3). Thus, for only few of the proteins contamination with somatic cells might have contributed to the finding of differential abundance. But also, these few ones had OMIM entries indicating male fertility relevance (Table 1: APAF1, CATIP, CFAP69) or were already found to be downregulated in spermatozoa under male infertility (Table 3: SYPL1). We therefore assume that contaminations of our sperm samples with somatic cells should not have compromised present results.
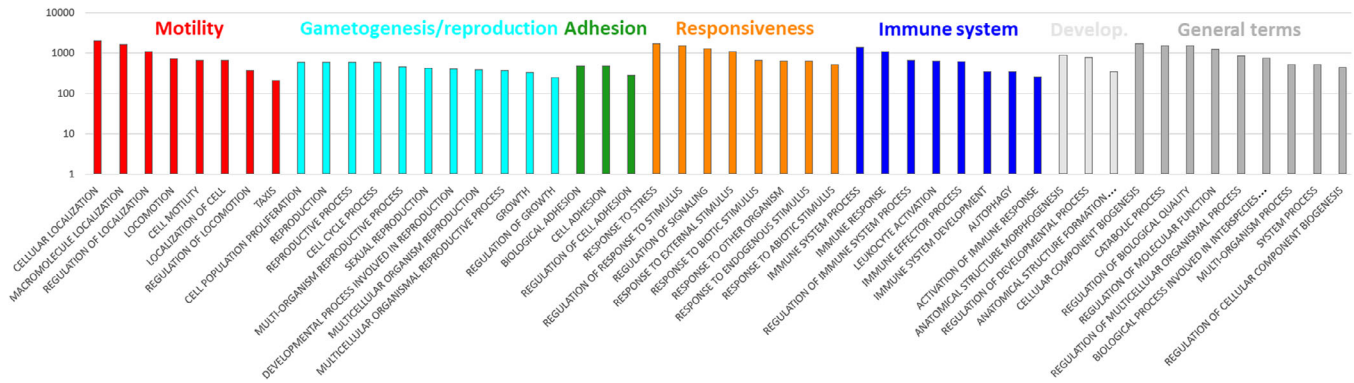
FIGURE 2 Representation of higher level biological process gene ontologies (GOs) in the human sperm proteome. Shown are 50 GOs with highest frequencies of occurrence. Included are 6939 sperm protein-coding genes mapped by ShinyGO v0.76. See Table S2 for the full terms of two GOs abbreviated in the above chart. The same table includes the remaining 50 GOs, which occur with lower frequency, and the genes subsumed under each GO.

TABLE 5 Protein groups exhibiting differential abundances.

Symbol	ID	OAT	O	OA	Functional annotation (UniProt)	THPA
SPATA31D3, SPATA31D4*	P0C874, Q6ZUB0	0.091	↓	↓	Spermatogenesis	I I
CFAP45*, SNHG28	Q9UL16, P0DPA3	0.101			Dynein ATPase-dependent ciliary and flagellar beating	III No entry
DNAH8*, DNAH14	Q96JB1, A0A804HLD3	0.318			ATPase activity; outer dynein arms (…) sperm flagellum; sperm flagellar assembly; sperm motility	II III

Note: The list contains protein groups showing at least three- or eightfold differential abundances in spermatozoa of infertile men diagnosed with oligoasthenoteratozoospermia (OAT) (false discovery rate [FDR] ≤ 0.05, t-test). Arrows indicate the direction of abundance changes in spermatozoa of infertile men diagnosed with oligozoospermia (O) and oligoasthenozoospermia (OA). Changes refer to abundances in normozoospermic spermatozoa of fertile men. Commas separate members of single protein groups. Asterisks highlight proteins included in the largest connected component shown in Figure 6. Categories on the left refer to transcript abundances reported in The Human Protein Atlas v22 (THPA): male germline-specific (I), male germline-biased (II), and elevated in male germline and somatic cell types, which unlikely were contained in the samples (III).

TABLE 6 Network statistics.

Dataset source	N_{Nodes}	N_{Edges} (N_{exp})	$N_{Nodes\ in\ LCC}$	$\%_{Nodes\ in\ LCC}$	Average nd	EET (FDR)
Present study	50	75 (2)	33	66	3.00	<0.001
Present study + FRP	91	208 (13)	75	82	4.57	<0.001
Present study + FRP + OMIM	225	891 (132)	192	85	7.92	<0.001

Note: Networks were reconstructed from the proteins carved out in the present study (Tables 1, 3, 5), previous inference of fertility relevance probability (FRP)[19], and database screening (NCBI Online Mendelian Inheritance in Men [OMIM]). Numbers of nodes refer to the proteins mapped by STRING v11.5. Abbreviations: EET, edge enrichment test; FDR, false discovery rate; LCC, largest connected component; N_{exp}, expected number of edges; nd, node degree.

3.5 | The male infertility network

For all protein samples considered, STRING gathered more PPIs than to be expected for the proteins mapped (FDR < 0.001, each). In addition, all network parameters rose with growing number of proteins included (Tables 6 and S6). Thus, STRING built a network of 50 nodes, thereof 66% contained in the largest connected component (LCC), and 75 edges when run on the proteins in Tables 1, 3, and 5. Higher coherence was reached upon addition of 41 candidate markers for which

we had previously calculated elevated fertility relevance probability.[19] Especially, the number of disconnected nodes decreased, despite an almost doubled number of proteins mapped (N = 91). With 208, this network had a disproportionally larger number of edges. Correspondingly, the number of PPIs per node was higher in the expanded network (node degree = 4.57) than in the one restricted to the proteins highlighted in the present study (node degree = 3.00). Addition of all other proteins with OMIM entries suggestive of male fertility relevance more than doubled the number of nodes to 225. The total number of PPIs
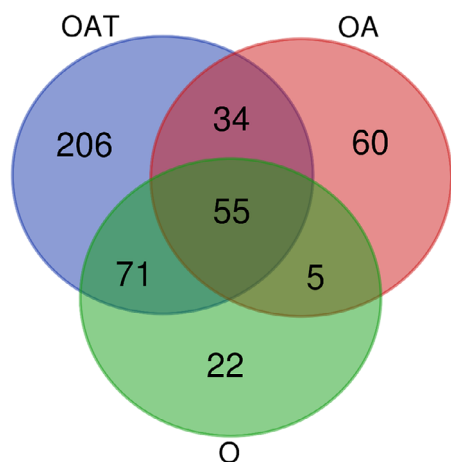
FIGURE 3 Numbers of shared and unique protein groups showing differential abundances in mass spectrometry (MS) analyses of oligozoospermic (O), oligoasthenozoospermic (OA), and oligoasthenoteratozoospermic spermatozoa (OAT). For consideration, a protein group had to show at least threefold differential abundance under infertility compared to levels in normozoospermic spermatozoa (false discovery rate [FDR] ≤ 0.05, *t*-test). The number of unique protein groups increased with growing complexity of the infertility diagnosis, and thus from O via OA to OAT. The Venn diagram was drawn using https://bioinformatics.psb.ugent.be/webtools/Venn/.

increased to 891, and the average node degree to 7.92 (Table 6). The corresponding LCC was comprised of 192 nodes (Figure 6), thereunder half of the novel candidate markers and all of the validated candidate markers. Also included in the LCC were 37 out of the abovementioned 41 proteins with increased fertility relevance probability,[19] and most of the proteins with OMIM entries pointing to male fertility relevance. We refer to this LCC as to the male infertility network.

4 | DISCUSSION

### 4.1 | The human sperm proteome and its coding genes

According to present MS analysis, about 6981 genes participate in the human sperm proteome (Tables S1 and S2). This represents a considerable increase compared with previous estimates from MS studies, which had determined 3444[21] and 4718[22] sperm proteins. The current number also exceeds the 6871 sperm proteins as gathered in a recent survey.[8] Merging the compilation from the latter study with present MS results, about 9300 genes might participate in the human sperm proteome. This would be ca. 45% of the approximately 20,500 coding human genes according to genome annotation ENSEMBL Gene 106: GRCh38.p13. The complexity of the sperm proteome will be higher, considering diversification by alternative splicing and post-translational modification.[29,98,99]

Overall shorter cDNAs, as observed in this study, could reflect adaptation to better swimming performance. In line with such possibility, there is previous evidence that selection for increased motil-
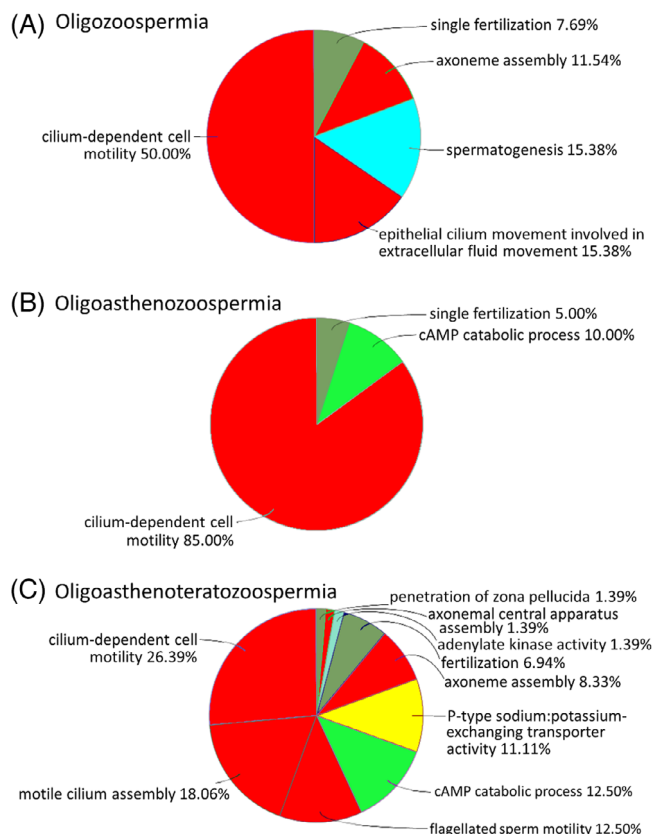
FIGURE 4 Representation of biological process gene ontologies (GOs) in sperm proteins exhibiting deviant abundances in oligozoospermia (A), oligoasthenozoospermia (B), and oligoasthenoteratozoospermia (C). Only highly significant results from overrepresentation test were considered (false discovery rate [FDR] ≤ 0.001). Highest diversity of overrepresented GOs was found in the most complex diagnosis, oligoasthenoteratozoospermia. GOs relating to sperm flagellum and motility made up more than two-thirds of the annotations in all diagnoses (red fill). Additional GOs overrepresented were relatable to male gametogenesis (turquoise), cAMP metabolism (light green), fertilization (olive green), and ion exchange (yellow). The lists tested were compiled from protein groups showing at least threefold differential abundances in spermatozoa of infertile men with respective diagnoses, relative to normozoospermic spermatozoa of fertile men (FDR ≤ 0.05, *t*-test). Analyses used ClueGO v2.5.9 within Cytoscape v3.9.1.

ity might drive changes in sperm morphology.[100,101] This might involve remodeling of protamines which are well-known for evolving under sexual selection.[102] In the present reconstruction of the human sperm proteome, importance of sperm motility was reflected in implications of numerous coding genes in cilium assembly and function (Figure 2). GOs referring to cellular responsiveness and adhesion were also conclusive considering the challenges spermatozoa have to cope with.[103,104] This might also be true for annotations referencing the immune system. In fact, male fertility and fertilization require evasion from the own immune response and the one of the partner.[99,105] Still, corresponding GO groups contained genes encoding proteins with clear fertility relevance, such as
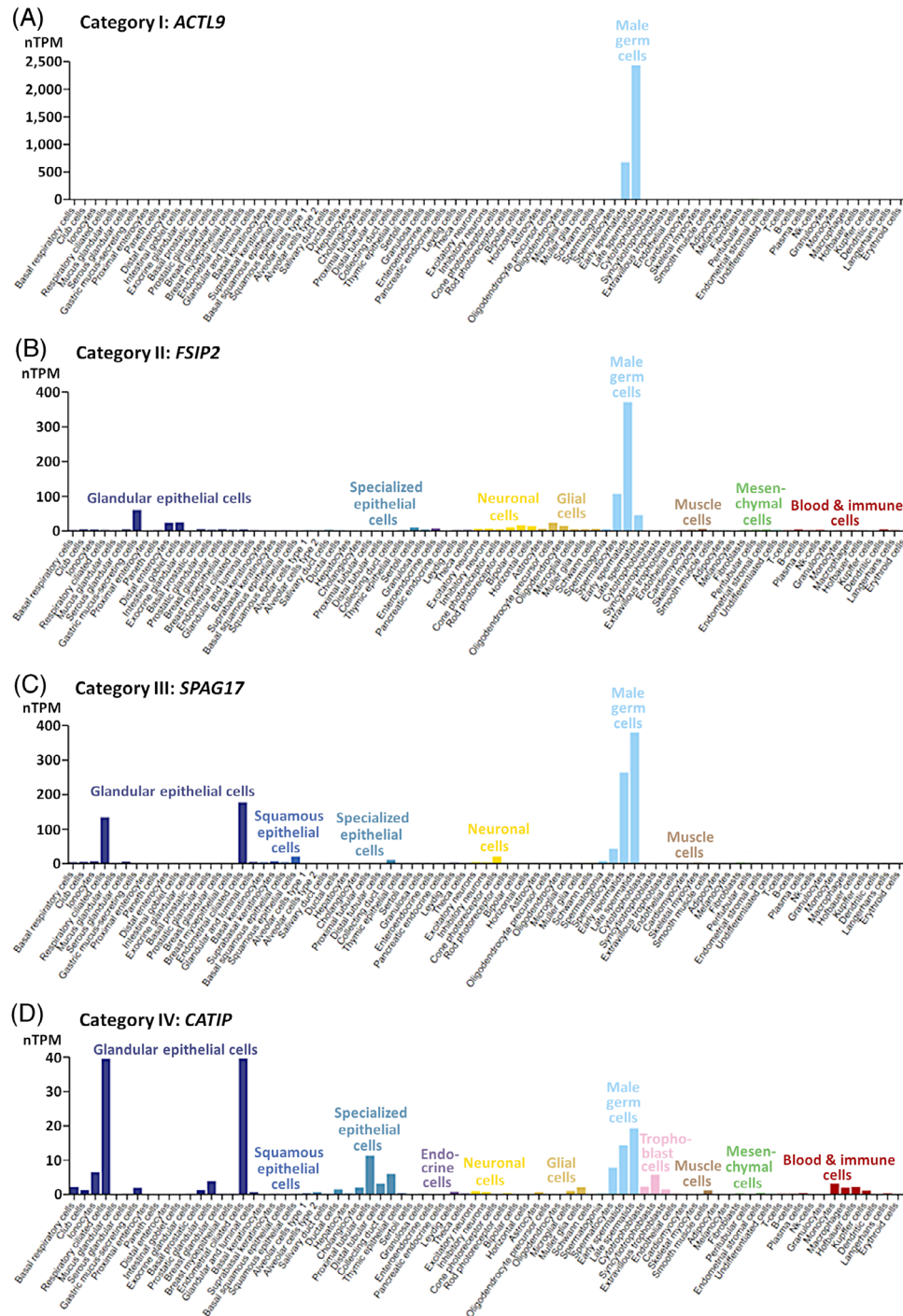
FIGURE 5 Examples for categorization of genes according to their transcriptional topology. Transcription was considered either specific to (category I), or strongly skewed toward male germ cells (II) or elevated in male germ cells in addition to somatic cell types which improbably (III) or possibly were contained in the samples analyzed (IV). The plots shown (A–D) are modified snapshots from The Human Protein Atlas v22. For expressional topologies of all genes with proteins in Tables 1, 3, and 5, see Figures S1–S3.

SPATA20,[11,106] ZPBP,[107] and SPAG4[108] (Table S3). Thus, reference to the immune system should not be overestimated in our opinion. Not least, present GO analyses (Figure 2) confirmed that spermatozoa contain proteins that are known primarily for their roles in male gametogenesis.[109]

4.2 | Molecular causes of male fertility impairment

Results of present overrepresentation tests revealed that the fewer spermatozoa produced under O, OA, and OAT carry signatures of disturbed gametogenesis. Functional annotations overrepresented in
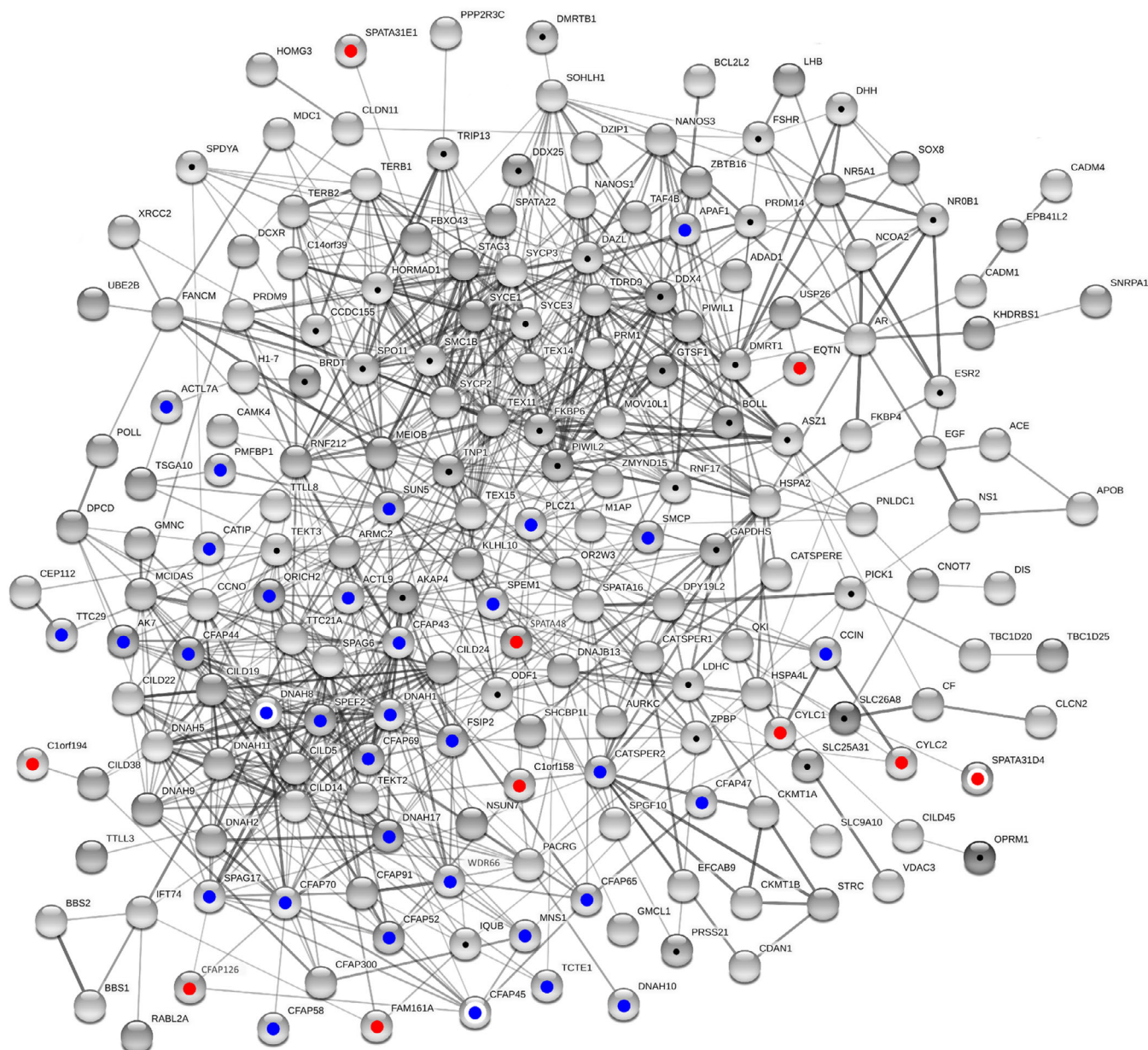
FIGURE 6 Male infertility network. Shown is the largest connected component from network reconstruction using standard settings in STRING v11.5. Balls give proteins and lines interactions. Thickness of lines corresponds to confidence levels. The reconstruction is based on a list merged from validated (blue dots) and novel candidate markers (red dots) according to the present study, proteins with elevated fertility relevance probability (black dots) according to previous calculations,[19] and further proteins with prior mention in the male fertility context (Online Mendelian Inheritance in Men [OMIM]). White circles surrounding colored filling highlight members of protein groups as specified in Table 5. In its current state, the male infertility network contains 192 proteins connected by 891 edges. For more network statistics, see Table 6.

differentially abundant proteins were further consistent with impairment of sperm motility and fertilization competence in infertile men (Figure 4). This includes proteins involved in cAMP signaling and sodium–potassium exchange. The first is crucial for all steps of sperm function in mammals,[110] and importance for motility and fertilization has been assigned to the latter.[111] Significance of proper sperm motility for reproduction[112] was additionally reflected in frequent references to the structure and function of the cilium or flagellum (Figure 4). The pattern was even more strongly reflected in the data on the individual proteins and coding genes. Thus, for the validated and novel candidate markers, there was evidence of involvements in malformations

such as ASS and MMAF and, related to the latter, impaired motility or asthenozoospermia (Tables 1–4). This suggests that reduced sperm count could be less decisive in O and syndromes including it. The ultimate reason might rather be that the fewer spermatozoa produced are dysfunctional to an increased proportion. This seems particularly noteworthy given that the samples analyzed had undergone a swim-up, meaning that the spermatozoa represented a selection of spermatozoa with increased fertilization competence. Even these would be functionally compromised according to our results, to an increased proportion at least. This could also explain why lowered spermiogram parameters can reduce success rates in ARTs.[113,114]

4.3 | Toward a marker panel of male fertility assessment

Quantitative MS analysis of the entire sperm proteome might complement diagnostics of the male fertility status in the future. If it comes to this point, special attention may be paid to the abundances of particular sperm proteins. Estimating the fertility status from differential abundances of sperm proteins on scales as found here ($N = 153–368$) should then be more feasible than based on numbers in the range of 1000 or higher as previously reported.[21,22] This was achieved by applying a more conservative threshold of differential abundance in the present study (at least 3×) than was the case in both referenced investigations (at least 1.5×).[21,22] Still, the definition of a further reduced number of male fertility markers remains to be a worthwhile goal for standardized fertility testing.[19,28,115] The proteins summarized in Tables 1 and 3, and to a lesser extent in Table 5, may contribute to this. It will be interesting to validate in future independent studies the corresponding proteins using alternative methods such as western blot.

Diagnostic potential should especially have the 31 sperm proteins with previous mention in the OMIM database hinting to roles in the etiology of male infertility (Table 1). Increased importance of these validated candidate markers for reproduction is underscored by transcript data. Indeed, we found expression of most of the coding genes to be tailored to male germline (Figure S1). Numerous associations with male fertility impairment lend additional support for their marker potential, as did their already-mentioned involvements in sperm flagellum structure and functioning, sperm head formation, and spermiogenesis (Tables 1 and 2). Against this background, it also seems plausible for most validated candidate markers that a decrease in abundance should be detrimental for male fertility maintenance. There were only two exceptions from the pattern. Exceptional in terms of function was 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase zeta-1 (PLCZ1). The protein proposedly belongs to the sperm factors transferred to the zygote and as such is likely to initiate the development of the zygote into the blastocyst.[116–118] Consistently, a *PLCZ1* variant in an infertile man was found to associate with oocyte activation deficiency.[54] Thus, infertility at PLCZ1 shortage could hint to disturbed development following fertilization of the ovum. Still, PLCZ1 seems to be also important for spermiogenesis considering spermatogenic defect in male *Plcz1* knockout mice.[119] The second exception from the above pattern refers to the direction of abundance change in reduced-relative to normal-fertile men. In fact, apoptotic protease-activating factor 1 (APAF1) was the only validated candidate marker showing increased abundance under infertility. Although exceptional in the present study, the present finding fits previous evidence of raised testicular abundance of *APAF1* transcript in men with impaired versus normal spermiogenesis.[65] It appears to be plausible for an apoptosis trigger[120] that oversupply negatively affects sperm number and viability.[20]

Seven of the proteins, to which we here refer to as validated candidate markers, displayed deviant abundances under O, OA, and OAT (Table 1): ACTL9, CCIN, CFAP47, CFAP65, CFAP251 (WDR66),

DNAH1, and SPEM1. Functional data would be in line with a predictive value of corresponding protein abundances whenever a reduction in sperm count is observed. Thus, SPEM1 should have importance for proper sperm formation including cytoplasm removal as suggested by deformed spermatozoa and infertility in male knockout mice.[49] Furthermore, ACTL9 plays a significant role in the fusion of proacrosomal vesicles and the anchoring of the acrosome to the perinuclear theca during spermiogenesis. Correspondingly, there are variants that associate with fertilization disorders and male infertility.[39] CCIN also localizes to the acrosomal area, thereby showing affinity to F-actin.[121] As a basic cytoskeleton protein of the mammalian sperm head, absence or dysregulation of CCIN is known to cause sperm malformation.[122] Relevance for proper flagellum formation and functioning is further inherent to three cilia and flagella-associated proteins (CFAP47, CFAP65, CFAP251/WDR66) and DNAH1.[42,43,45,46,91,123,124] Nevertheless, all sperm proteins in Table 1 could be implicated in the development of male fertility disorders, and hence might be further evaluated for their potential as male fertility markers. In support of such utilization, the coding genes of 18 out of the 31 validated candidate markers are already included in an 81 gene panel (ID192.04) for male fertility testing (e.g., www.zhma.de).

Diagnostic potential might also have 18 proteins displaying at least eightfold and thus extreme deregulation of abundance (Table 3). We refer to these proteins as novel candidate markers because OMIM search did not uncover respective previous entries. Nevertheless, downregulation of testicular and spermatozoic expression under reduced male fertility was documented for the coding genes to all novel candidate markers (Table 4). Predictive power of corresponding protein levels would also be consistent with the functional and transcriptional data available, as will be detailed below in the ones showing reduced levels in all three diagnoses including reduced sperm count. In case of SPATA31E1, SPATA31D1, and SPATA48, relevance for male fertility is reflected by transcript enrichment in the male germline (Figure S2). Transcriptional data additionally suggest that EF-hand domain family member B (EFHB, alias CFAP21), chromosome 2 open reading frame 16 (C2orf16, alias CB016), and family with sequence similarity 161 member A (FAM161A) participate in male fertility maintenance (Figure S2). Indeed, EFHB is a regulator of store-operated Ca^{2+} entry (SOCE) impairment of which causes male infertility in the murine model.[125,126] Furthermore, FAM161A is a component of the cilia–basal body complex[127] modulating mammalian sperm movement.[127,128] Another protein with strongly deregulated abundance in association with O, cylicin 1 (CYLC1), is involved in the formation of the sperm head.[129] As in several other novel candidate markers, *CYLC1* transcription is highly enriched in male germ line, suggestive of special importance for male fertility (Figure S2).

5 | CONCLUSIONS

Present high-resolution mass spectrometry analyses of sperm samples from 38 reduced- and 38 normal-fertile men uncovered proteins of

about 7000 genes (Table S2). The functional data available for differentially abundant proteins suggests that spermatozoa of infertile donors diagnosed with oligoasthenoteratozoospermia, oligoasthenozoospermia, and oligozoospermia carry proteomic signatures of disturbed gametogenesis. Our results further imply that the fewer spermatozoa in these diagnoses are dysfunctional to an increased proportion, with motility and fertilization ability being compromised. Thus, a reduction in sperm count might have been less decisive for the development of infertility in the corresponding subjects than the fact that the spermatozoa produced were dysfunctional to an increased proportion. While this was to be expected for both syndromes it was not for cases in which pure oligozoospermia was diagnosed. This sheds new light on the challenges in selecting proper spermatozoa for in vitro fertilization and intracytoplasmic sperm injection based on microscopic examination.[113,114]

Quantitative analysis led to the identification of sperm proteins the abundances of which might have predictive value in fertility diagnostics. As detailed above, the respective proteins had been determined in considerably high numbers of peptides, suggestive of highly confident identifications. Marker potential should especially have the ones for which the coding genes already had entries in the Online Mendelian Inheritance in Men (OMIM) database relating to roles in male fertility impairment (Tables 1 and 2 and Figure S1). Such applied prospect is matched by the fact that the coding genes of about two-thirds of the proteins are included in a genotyping panel used in fertility testing (ID192.04). Yet proteins for which we observed extremely reduced levels under infertility (Table 3) might also be considered for their potential in fertility diagnostics. In support of such usability, expression of the coding genes was previously found to be reduced in testes and spermatozoa of men with impaired fertility (Table 4; see also Figure S2). Members of deregulated protein groups might be evaluated for utilization in fertility testing too (Table 5 and Figure S3).

Our results further illustrate that most of the proteins highlighted above are interconnected with other players in the development of male infertility. Indeed, we observed a growing connectivity with increasing number of prospective male fertility markers included in network reconstructions (Figure 6 and Table 6). We expect the male infertility network to grow with the addition of further (candidate) markers.[9,21,130] For example, future network analyses may include male fertility markers in the seminal plasma proteome.[9,21,131,132] With the completion of the network, we should increasingly become able to comprehend the molecular mechanisms behind the development of male infertility. This may open new perspectives in the diagnosis and treatment of male infertility.[133]

AUTHOR CONTRIBUTIONS

Thomas Greither and Holger Herlyn conceived the study. Thomas Greither and Hermann M. Behre recruited and characterized the participants. Falk Butter, Mario Dejung, and Holger Herlyn performed the experiments. Holger Herlyn analyzed the data and drafted the manuscript. Thomas Greither, Hermann M. Behre, Falk Butter, and Holger Herlyn critically revised the manuscript. All authors approved the final version of the manuscript.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to disclose.

ORCID

Thomas Greither https://orcid.org/0000-0002-1888-7813
Holger Herlyn https://orcid.org/0000-0003-3157-7243

REFERENCES

1. Vogt PH. Molecular genetics of human male infertility: from genes to new therapeutic perspectives. *Curr Pharm Des.* 2004;10(5):471-500. doi:10.2174/1381612043453261
2. Ji G, Long Y, Zhou Y, Huang C, Gu A, Wang X. Common variants in mismatch repair genes associated with increased risk of sperm DNA damage and male infertility. *BMC Med.* 2012;10:49. doi:10.1186/1741-7015-10-49
3. Kumar N, Singh AK. Trends of male factor infertility, an important cause of infertility: a review of literature. *J Hum Reprod Sci.* 2015;8(4):191-196. doi:10.4103/0974-1208.170370
4. Baker MA, Reeves G, Hetherington L, Müller J, Baur I, Aitken RJ. Identification of gene products present in Triton X-100 soluble and insoluble fractions of human spermatozoa lysates using LC-MS/MS analysis. *Proteomics Clin Appl.* 2007;1(5):524-532. doi:10.1002/prca.200601013
5. Wang G, Guo Y, Zhou T, et al. In-depth proteomic analysis of the human sperm reveals complex protein compositions. *J Proteomics.* 2013;79:114-122. doi:10.1016/j.jprot.2012.12.008
6. Amaral A, Castillo J, Ramalho-Santos J, Oliva R. The combined human sperm proteome: cellular pathways and implications for basic and clinical science. *Hum Reprod Update.* 2014;20(1):40-62. doi:10.1093/humupd/dmt046
7. Jodar M, Soler-Ventura A, Oliva R. Semen proteomics and male infertility. *J Proteomics.* 2017;162:125-134. doi:10.1016/j.jprot.2016.08.018
8. Castillo J, Jodar M, Oliva R. The contribution of human sperm proteins to the development and epigenome of the preimplantation embryo. *Hum Reprod Update.* 2018;24(5):535-555. doi:10.1093/humupd/dmy017
9. Bansal SK, Gupta N, Sankhwar SN, Rajender S. Differential genes expression between fertile and infertile spermatozoa revealed by transcriptome analysis. *PLoS One.* 2015;10(5):e0127007. doi:10.1371/journal.pone.0127007
10. Han B, Wang L, Yu S, et al. One potential biomarker for teratozoospermia identified by in-depth integrative analysis of multiple microarray data. *Aging.* 2021;13(7):10208-10224. doi:10.18632/aging.202781
11. Omolaoye TS, Hachim MY, Du Plessis SS. Using publicly available transcriptomic data to identify mechanistic and diagnostic biomarkers in azoospermia and overall male infertility. *Sci Rep.* 2022;12(1):2584. doi:10.1038/s41598-022-06476-1
12. O'Flynn O'Brien KL, Varghese AC, Agarwal A. The genetic causes of male factor infertility: a review. *Fertil Steril.* 2010;93(1):1-12. doi:10.1016/j.fertnstert.2009.10.045

13. Panner Selvam MK, Agarwal A. Update on the proteomics of male infertility: a systematic review. *Arab J Urol.* 2018;16(1):103-112. doi:10.1016/j.aju.2017.11.016

14. Krausz C, Riera-Escamilla A. Genetics of male infertility. *Nat Rev Urol.* 2018;15(6):369-384. doi:10.1038/s41585-018-0003-3

15. Greither T, Behre HM, Herlyn H. Genome-wide association screening determines peripheral players in male fertility maintenance. *Int J Mol Sci.* 2022;24(1):524. doi:10.3390/ijms24010524

16. Tahmasbpour E, Balasubramanian D, Agarwal A. A multi-faceted approach to understanding male infertility: gene mutations, molecular defects and assisted reproductive techniques (ART). *J Assist Reprod Genet.* 2014;31(9):1115-1137. doi:10.1007/s10815-014-0280-6

17. Leitão E, Di Persio S, Laurentino S, et al. The sperm epigenome does not display recurrent epimutations in patients with severely impaired spermatogenesis. *Clin Epigenetics.* 2020;12(1):61. doi:10.1186/s13148-020-00854-0

18. Schumacher J, Zischler H, Herlyn H. Effects of different kinds of essentiality on sequence evolution of human testis proteins. *Sci Rep.* 2017;7:43534. doi:10.1038/srep43534

19. Greither T, Schumacher J, Dejung M, et al. Fertility relevance probability analysis shortlists genetic markers for male fertility impairment. *Cytogenet Genome Res.* 2020;160(9):506-522. doi:10.1159/000511117

20. Liu X, Liu G, Liu J, et al. iTRAQ-based analysis of sperm proteome from normozoospermic men achieving the rescue-ICSI pregnancy after the IVF failure. *Clin Proteomics.* 2018;15:27. doi:10.1186/s12014-018-9203-3

21. Liang J, Zheng Y, Zeng W, et al. Proteomic profile of sperm in infertile males reveals changes in metabolic pathways. *Protein J.* 2021;40(6):929-939. doi:10.1007/s10930-021-10013-w

22. Yang J, Liu Q, Yu B, Han B, Yang B. 4D-quantitative proteomics signature of asthenozoospermia and identification of extracellular matrix protein 1 as a novel biomarker for sperm motility. *Mol Omics.* 2022;18(1):83-91. doi:10.1039/d1mo00257k

23. Bracke A, Peeters K, Punjabi U, Hoogewijs D, Dewilde S. A search for molecular mechanisms underlying male idiopathic infertility. *Reprod Biomed Online.* 2018;36(3):327-339. doi:10.1016/j.rbmo.2017.12.005

24. Deutsches IVF Register (D.I.R.). D.I.R. Jahrbuch 2009: Supplement 1. *J Reproduktionsmed Endokrinol.* 2010;7(6):470-497.

25. Cooper TG, Noonan E, von Eckardstein S, et al. World Health Organization reference values for human semen characteristics. *Hum Reprod Update.* 2010;16(3):231-245. doi:10.1093/humupd/dmp048

26. World Health Organization (WHO). WHO Laboratory Manual for the Examination and Processing of Human Semen. *6th ed.* WHO; 2021.

27. Henkel R. Sperm preparation: state-of-the-art—physiological aspects and application of advanced sperm preparation methods. *Asian J Androl.* 2012;14(2):260-269. doi:10.1038/aja.2011.133

28. Panner Selvam MK, Agarwal A, Pushparaj PN, Baskaran S, Bendou H. Sperm proteome analysis and identification of fertility-associated biomarkers in unexplained male infertility. *Genes.* 2019;10(7):522. doi:10.3390/genes10070522

29. Parte PP, Rao P, Redij S, et al. Sperm phosphoproteome profiling by ultra performance liquid chromatography followed by data independent analysis (LC-MS(E)) reveals altered proteomic signatures in asthenozoospermia. *J Proteomics.* 2012;75(18):5861-5871. doi:10.1016/j.jprot.2012.07.003

30. Légaré C, Droit A, Fournier F, et al. Investigation of male infertility using quantitative comparative proteomics. *J Proteome Res.* 2014;13(12):5403-5414. doi:10.1021/pr501031x

31. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc.* 2006;1(6):2856-2860. doi:10.1038/nprot.2006.468

32. Rappsilber J, Mann M, Ishihama Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc.* 2007;2(8):1896-1906. doi:10.1038/nprot.2007.261

33. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol.* 2008;26(12):1367-1372. doi:10.1038/nbt.1511

34. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics.* 2020;36(8):2628-2629. doi:10.1093/bioinformatics/btz931

35. Bindea G, Mlecnik B, Hackl H, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics.* 2009;25(8):1091-1093. doi:10.1093/bioinformatics/btp101

36. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13(11):2498-2504. doi:10.1101/gr.1239303

37. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019;47(D1):D607-D613. doi:10.1093/nar/gky1131

38. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc.* 1995;57(1):289-300.

39. Dai J, Zhang T, Guo J, et al. Homozygous pathogenic variants in *ACTL9* cause fertilization failure and male infertility in humans and mice. *Am J Hum Genet.* 2021;108(3):469-481. doi:10.1016/j.ajhg.2021.02.004

40. Fan Y, Huang C, Chen J, et al. Mutations in *CCIN* cause teratozoospermia and male infertility. *Sci Bull.* 2022;67(20):2112-2123. doi:10.1016/j.scib.2022.09.026

41. Yan W. Calicin is a key sperm head-shaping factor essential for male fertility. *Sci Bull.* 2022;67(23):2395-2397. doi:10.1016/j.scib.2022.11.022

42. Liu C, Tu C, Wang L, et al. Deleterious variants in X-linked *CFAP47* induce asthenoteratozoospermia and primary male infertility. *Am J Hum Genet.* 2021;108(2):309-323. doi:10.1016/j.ajhg.2021.01.002

43. Zhang X, Shen Y, Wang X, Yuan G, Zhang C, Yang Y. A novel homozygous *CFAP65* mutation in humans causes male infertility with multiple morphological abnormalities of the sperm flagella. *Clin Genet.* 2019;96(6):541-548. doi:10.1111/cge.13644

44. Li W, Wu H, Li F, et al. Biallelic mutations in *CFAP65* cause male infertility with multiple morphological abnormalities of the sperm flagella in humans and mice. *J Med Genet.* 2020;57(2):89-95. doi:10.1136/jmedgenet-2019-106344

45. Kherraf Z-E, Amiri-Yekta A, Dacheux D, et al. A homozygous ancestral SVA-insertion-mediated deletion in *WDR66* induces multiple morphological abnormalities of the sperm flagellum and male infertility. *Am J Hum Genet.* 2018;103(3):400-412. doi:10.1016/j.ajhg.2018.07.014

46. Auguste Y, Delague V, Desvignes J-P, et al. Loss of calmodulin- and radial-spoke-associated complex protein CFAP251 leads to immotile spermatozoa lacking mitochondria and infertility in men. *Am J Hum Genet.* 2018;103(3):413-420. doi:10.1016/j.ajhg.2018.07.013

47. Zhuang B-J, Xu S-Y, Dong L, et al. Novel *DNAH1* mutation loci lead to multiple morphological abnormalities of the sperm flagella and literature review. *World J Mens Health.* 2022;40(4):551-560. doi:10.5534/wjmh.210119

48. Hashemi M-S, Mozdarani H, Ghaedi K, Nasr-Esfahani MH. Among seven testis-specific molecular markers, *SPEM1* appears to have a significant clinical value for prediction of sperm retrieval in azoospermic men. *Andrology.* 2018;6(6):890-895. doi:10.1111/andr.12528

49. Zheng H, Stratton CJ, Morozumi K, Jin J, Yanagimachi R, Yan W. Lack of *Spem1* causes aberrant cytoplasm removal, sperm deformation,

and male infertility. *Proc Natl Acad Sci U S A*. 2007;104(16):6852-6857. doi:10.1073/pnas.0701669104

50. Xin A, Qu R, Chen G, et al. Disruption in *ACTL7A* causes acrosomal ultrastructural defects in human and mouse sperm as a novel male factor inducing early embryonic arrest. *Sci Adv*. 2020;6(35):eaaz4796. doi:10.1126/sciadv.aaz4796

51. Wang J, Zhang J, Sun X, et al. Novel bi-allelic variants in ACTL7A are associated with male infertility and total fertilization failure. *Hum Reprod*. 2021;36(12):3161-3169. doi:10.1093/humrep/deab228

52. Ta-Shma A, Hjeij R, Perles Z, et al. Homozygous loss-of-function mutations in *MNS1* cause laterality defects and likely male infertility. *PLoS Genet*. 2018;14(8):e1007602. doi:10.1371/journal.pgen.1007602

53. Leslie JS, Rawlins LE, Chioza BA, et al. *MNS1* variant associated with situs inversus and male infertility. *Eur J Hum Genet*. 2020;28(1):50-55. doi:10.1038/s41431-019-0489-z

54. Wang F, Zhang J, Kong S, et al. A homozygous nonsense mutation of *PLCZ1* cause male infertility with oocyte activation deficiency. *J Assist Reprod Genet*. 2020;37(4):821-828. doi:10.1007/s10815-020-01719-4

55. Zhu F, Liu C, Wang F, et al. Mutations in *PMFBP1* cause acephalic spermatozoa syndrome. *Am J Hum Genet*. 2018;103(2):188-199. doi:10.1016/j.ajhg.2018.06.010

56. Nie H, Tang Y, Zhang X, Tan Y, Qin W. Novel mutations of *PMFBP1* in a man with acephalic spermatozoa defects. *Mol Genet Genomic Med*. 2022;10(9):e2020. doi:10.1002/mgg3.2020

57. Deng T-Q, Xie Y-L, Pu J-B, Xuan J, Li X-M. Compound heterozygous mutations in PMFBP1 cause acephalic spermatozoa syndrome: a case report. *World J Clin Cases*. 2022;10(34):12761-12767. doi:10.12998/wjcc.v10.i34.12761

58. Kherraf Z-E, Cazin C, Coutton C, et al. Whole exome sequencing of men with multiple morphological abnormalities of the sperm flagella reveals novel homozygous *QRICH2* mutations. *Clin Genet*. 2019;96(5):394-401. doi:10.1111/cge.13604

59. Shen Y, Zhang F, Li F, et al. Loss-of-function mutations in *QRICH2* cause male infertility with multiple morphological abnormalities of the sperm flagella. *Nat Commun*. 2019;10(1):433. doi:10.1038/s41467-018-08182-x

60. Xu X, Sha Y-W, Mei L-B, et al. A familial study of twins with severe asthenozoospermia identified a homozygous *SPAG17* mutation by whole-exome sequencing. *Clin Genet*. 2018;93(2):345-349. doi:10.1111/cge.13059

61. Zhou S, Wu H, Zhang J, et al. Bi-allelic variants in human *TCTE1/DRC5* cause asthenospermia and male infertility. *Eur J Hum Genet*. 2022;30(6):721-729. doi:10.1038/s41431-022-01095-w

62. Tang S, Wang X, Li W, et al. Biallelic mutations in *CFAP43* and *CFAP44* cause male infertility with multiple morphological abnormalities of the sperm flagella. *Am J Hum Genet*. 2017;100(6):854-864. doi:10.1016/j.ajhg.2017.04.012

63. Khan I, Shah B, Dil S, et al. Novel biallelic loss-of-function mutations in *CFAP43* cause multiple morphological abnormalities of the sperm flagellum in Pakistani families. *Asian J Androl*. 2021;23(6):627-632. doi:10.4103/aja.aja_26_21

64. Lorès P, Coutton C, El Khouri E, et al. Homozygous missense mutation L673P in adenylate kinase 7 (AK7) leads to primary male infertility and multiple morphological anomalies of the flagella but not to primary ciliary dyskinesia. *Hum Mol Genet*. 2018;27(7):1196-1211. doi:10.1093/hmg/ddy034

65. Cavalcanti MCO, Steilmann C, Failing K, et al. Apoptotic gene expression in potentially fertile and subfertile men. *Mol Hum Reprod*. 2011;17(7):415-420. doi:10.1093/molehr/gar011

66. Arafat M, Harlev A, Har-Vardi I, et al. Mutation in *CATIP* (C2orf62) causes oligoteratoasthenozoospermia by affecting actin dynamics. *J Med Genet*. 2020. doi:10.1136/jmedgenet-2019-106825

67. Avidan N, Tamary H, Dgany O, et al. CATSPER2, a human autosomal nonsyndromic male infertility gene. *Eur J Hum Genet*. 2003;11(7):497-502. doi:10.1038/sj.ejhg.5200991

68. Luo T, Chen H-Y, Zou Q-X, et al. A novel copy number variation in *CATSPER2* causes idiopathic male infertility with normal semen parameters. *Hum Reprod*. 2019;34(3):414-423. doi:10.1093/humrep/dey377

69. Kamel A, Saberiyan M, Mirfakhraie R, Teimori H. Reduced expression of CFAP44 and CFAP44-AS1 may affect sperm motility and morphology. *Andrologia*. 2022;54(7):e14447. doi:10.1111/and.14447

70. Ta-Shma A, Perles Z, Yaacov B, et al. A human laterality disorder associated with a homozygous WDR16 deletion. *Eur J Hum Genet*. 2015;23(9):1262-1265. doi:10.1038/ejhg.2014.265

71. Dougherty GW, Mizuno K, Nöthe-Menchen T, et al. CFAP45 deficiency causes situs abnormalities and asthenospermia by disrupting an axonemal adenine nucleotide homeostasis module. *Nat Commun*. 2020;11(1):5520. doi:10.1038/s41467-020-19113-0

72. He X, Liu C, Yang X, et al. Bi-allelic loss-of-function variants in *CFAP58* cause flagellar axoneme and mitochondrial sheath defects and asthenoteratozoospermia in humans and mice. *Am J Hum Genet*. 2020;107(3):514-526. doi:10.1016/j.ajhg.2020.07.010

73. Dong FN, Amiri-Yekta A, Martinez G, et al. Absence of *CFAP69* causes male infertility due to multiple morphological abnormalities of the flagella in human and mouse. *Am J Hum Genet*. 2018;102(4):636-648. doi:10.1016/j.ajhg.2018.03.007

74. Beurois J, Martinez G, Cazin C, et al. *CFAP70* mutations lead to male infertility due to severe astheno-teratozoospermia. A case report. *Hum Reprod*. 2019;34(10):2071-2079. doi:10.1093/humrep/dez166

75. Tu C, Cong J, Zhang Q, et al. Bi-allelic mutations of *DNAH10* cause primary male infertility with asthenoteratozoospermia in humans and mice. *Am J Hum Genet*. 2021;108(8):1466-1477. doi:10.1016/j.ajhg.2021.06.010

76. Li K, Wang G, Lv M, et al. Bi-allelic variants in *DNAH10* cause asthenoteratozoospermia and male infertility. *J Assist Reprod Genet*. 2022;39(1):251-259. doi:10.1007/s10815-021-02306-x

77. Whitfield M, Thomas L, Bequignon E, et al. Mutations in *DNAH17*, encoding a sperm-specific axonemal outer dynein arm heavy chain, cause isolated male infertility due to asthenozoospermia. *Am J Hum Genet*. 2019;105(1):198-212. doi:10.1016/j.ajhg.2019.04.015

78. Zhang B, Ma H, Khan T, et al. A *DNAH17* missense variant causes flagella destabilization and asthenozoospermia. *J Exp Med*. 2020;217(2):e20182365. doi:10.1084/jem.20182365

79. Nosková A, Hiltpold M, Janett F, et al. Infertility due to defective sperm flagella caused by an intronic deletion in *DNAH17* that perturbs splicing. *Genetics*. 2021;217(2):iyaa033. doi:10.1093/genetics/iyaa033

80. Martinez G, Kherraf Z-E, Zouari R, et al. Whole-exome sequencing identifies mutations in *FSIP2* as a recurrent cause of multiple morphological abnormalities of the sperm flagella. *Hum Reprod*. 2018;33(10):1973-1984. doi:10.1093/humrep/dey264

81. Liu M, Sun Y, Li Y, Sun J, Yang Y, Shen Y. Novel mutations in *FSIP2* lead to multiple morphological abnormalities of the sperm flagella and poor ICSI prognosis. *Gene*. 2021;781:145536. doi:10.1016/j.gene.2021.145536

82. Omolaoye TS, Omolaoye VA, Kandasamy RK, Hachim MY, Du Plessis SS. Omics and male infertility: highlighting the application of transcriptomic data. *Life*. 2022;12(2):280. doi:10.3390/life12020280

83. Nayernia K, Adham IM, Burkhardt-Göttges E, et al. Asthenozoospermia in mice with targeted deletion of the sperm mitochondrion-associated cysteine-rich protein (*Smcp*) gene. *Mol Cell Biol*. 2002;22(9):3046-3052. doi:10.1128/MCB.22.9.3046-3052.2002

84. Tu C, Nie H, Meng L, et al. Novel mutations in *SPEF2* causing different defects between flagella and cilia bridge: the phenotypic link

between MMAF and PCD. *Hum Genet.* 2020;139(2):257-271. doi:10.1007/s00439-020-02110-0

85. Lehti MS, Zhang F-P, Kotaja N, Sironen A. SPEF2 functions in microtubule-mediated transport in elongating spermatids to ensure proper male germ cell differentiation. *Development.* 2017;144(14):2683-2693. doi:10.1242/dev.152108

86. Zhu F, Wang F, Yang X, et al. Biallelic *SUN5* mutations cause autosomal-recessive acephalic spermatozoa syndrome. *Am J Hum Genet.* 2016;99(6):1405. doi:10.1016/j.ajhg.2016.11.002

87. Xiang M, Wang Y, Wang K, et al. Novel mutation and deletion in *SUN5* cause male infertility with acephalic spermatozoa syndrome. *Reprod Sci.* 2022;29(2):646-651. doi:10.1007/s43032-021-00665-5

88. Lorès P, Dacheux D, Kherraf Z-E, et al. Mutations in *TTC29*, encoding an evolutionarily conserved axonemal protein, result in asthenozoospermia and male infertility. *Am J Hum Genet.* 2019;105(6):1148-1167. doi:10.1016/j.ajhg.2019.10.007

89. Xu Y, Han Q, Ma C, et al. Comparative proteomics and phosphoproteomics analysis reveal the possible breed difference in yorkshire and duroc boar spermatozoa. *Front Cell Dev Biol.* 2021;9:652809. doi:10.3389/fcell.2021.652809

90. Dias TR, Agarwal A, Pushparaj PN, Ahmad G, Sharma R. Reduced semen quality in patients with testicular cancer seminoma is associated with alterations in the expression of sperm proteins. *Asian J Androl.* 2020;22(1):88-93. doi:10.4103/aja.aja_17_19

91. Wang Z, Ding Z, Guan Y, et al. Altered gene expression in the testis of infertile patients with nonobstructive azoospermia. *Comput Math Methods Med.* 2021;2021:5533483. doi:10.1155/2021/5533483

92. Hu T, Luo S, Xi Y, et al. Integrative bioinformatics approaches for identifying potential biomarkers and pathways involved in nonobstructive azoospermia. *Transl Androl Urol.* 2021;10(1):243-257. doi:10.21037/tau-20-1029

93. Platts AE, Dix DJ, Chemes HE, et al. Success and failure in human spermatogenesis as revealed by teratozoospermic RNAs. *Hum Mol Genet.* 2007;16(7):763-773. doi:10.1093/hmg/ddm012

94. Kui F, Ye H, Chen X-L, Zhang J. Microarray meta-analysis identifies candidate genes for human spermatogenic arrest. *Andrologia.* 2019;51(8):e13301. doi:10.1111/and.13301

95. Ing NH, Konganti K, Ghaffari N, et al. Identification and quantification of coding and long non-coding RNAs in stallion spermatozoa separated by density. *Andrology.* 2020;8(5):1409-1418. doi:10.1111/andr.12791

96. Zhang J, Yan R, Wu C, et al. Spermatogenesis-associated 48 is essential for spermatogenesis in mice. *Andrologia.* 2018;50(6):e13027. doi:10.1111/and.13027

97. Alvarez-Rodriguez M, Martinez C, Wright D, Barranco I, Roca J, Rodriguez-Martinez H. The transcriptome of pig spermatozoa, and its role in fertility. *Int J Mol Sci.* 2020;21(5):1572. doi:10.3390/ijms21051572

98. Schumacher J, Ramljak S, Asif AR, Schaffrath M, Zischler H, Herlyn H. Evolutionary conservation of mammalian sperm proteins associates with overall, not tyrosine, phosphorylation in human spermatozoa. *J Proteome Res.* 2013;12(12):5370-5382. doi:10.1021/pr400228c

99. Xin M, You S, Xu Y, et al. Precision glycoproteomics reveals distinctive N-glycosylation in human spermatozoa. *Mol Cell Proteomics.* 2022;21(4):100214. doi:10.1016/j.mcpro.2022.100214

100. LaMunyon CW, Ward S. Evolution of larger sperm in response to experimentally increased sperm competition in *Caenorhabditis elegans. Proc Biol Sci.* 2002;269(1496):1125-1128. doi:10.1098/rspb.2002.1996

101. Tourmente M, Gomendio M, Roldan ERS. Sperm competition and the evolution of sperm design in mammals. *BMC Evol Biol.* 2011;11:12. doi:10.1186/1471-2148-11-12

102. Lüke L, Vicens A, Tourmente M, Roldan ERS. Evolution of protamine genes and changes in sperm head phenotype in rodents. *Biol Reprod.* 2014;90(3):67. doi:10.1095/biolreprod.113.115956

103. Cohen-Dayag A, Ralt D, Tur-Kaspa I, et al. Sequential acquisition of chemotactic responsiveness by human spermatozoa. *Biol Reprod.* 1994;50(4):786-790. doi:10.1095/biolreprod50.4.786

104. Primakoff P, Myles DG. Penetration, adhesion, and fusion in mammalian sperm-egg interaction. *Science.* 2002;296(5576):2183-2185. doi:10.1126/science.1072029

105. Chakradhar S. Puzzling over privilege: how the immune system protects-and fails-the testes. *Nat Med.* 2018;24(1):2-5. doi:10.1038/nm0118-2

106. Wang X, Jiang C, Dai S, Shen G, Yang Y, Shen Y. Identification of nonfunctional SPATA20 causing acephalic spermatozoa syndrome in humans. *Clin Genet.* 2023;103(3):310-319. doi:10.1111/cge.14268

107. Lin Y-N, Roy A, Yan W, Burns KH, Matzuk MM. Loss of zona pellucida binding proteins in the acrosomal matrix disrupts acrosome biogenesis and sperm morphogenesis. *Mol Cell Biol.* 2007;27(19):6794-6805. doi:10.1128/MCB.01029-07

108. Yeh C-H, Kuo P-L, Wang Y-Y, et al. SEPT12/SPAG4/LAMINB1 complexes are required for maintaining the integrity of the nuclear envelope in postmeiotic male germ cells. *PLoS One.* 2015;10(3):e0120722. doi:10.1371/journal.pone.0120722

109. Gilany K, Lakpour N, Vafakhah M, Sadeghi MR. The profile of human sperm proteome; a mini-review. *J Reprod Infertil.* 2011;12(3):193-199.

110. Balbach M, Beckert V, Hansen JN, Wachten D. Shedding light on the role of cAMP in mammalian sperm physiology. *Mol Cell Endocrinol.* 2018;468:111-120. doi:10.1016/j.mce.2017.11.008

111. Jimenez T, McDermott JP, Sánchez G, Blanco G. Na,K-ATPase alpha4 isoform is essential for sperm fertility. *Proc Natl Acad Sci U S A.* 2011;108(2):644-649. doi:10.1073/pnas.1016902108

112. Curi SM, Ariagno JI, Chenlo PH, et al. Asthenozoospermia: analysis of a large population. *Arch Androl.* 2003;49(5):343-349. doi:10.1080/01485010390219656

113. Oseguera-López I, Ruiz-Díaz S, Ramos-Ibeas P, Pérez-Cerezales S. Novel techniques of sperm selection for improving IVF and ICSI outcomes. *Front Cell Dev Biol.* 2019;7:298. doi:10.3389/fcell.2019.00298

114. Villani MT, Morini D, Spaggiari G, et al. Are sperm parameters able to predict the success of assisted reproductive technology? A retrospective analysis of over 22,000 assisted reproductive technology cycles. *Andrology.* 2022;10(2):310-321. doi:10.1111/andr.13123

115. Ferlin A, Raicu F, Gatta V, Zuccarello D, Palka G, Foresta C. Male infertility: role of genetic background. *Reprod Biomed Online.* 2007;14(6):734-745. doi:10.1016/s1472-6483(10)60677-3

116. Cox LJ, Larman MG, Saunders CM, Hashimoto K, Swann K, Lai FA. Sperm phospholipase Czeta from humans and cynomolgus monkeys triggers Ca2+ oscillations, activation and development of mouse oocytes. *Reproduction.* 2002;124(5):611-623. doi:10.1530/rep.0.1240611

117. Rogers NT, Hobson E, Pickering S, Lai FA, Braude P, Swann K. Phospholipase Czeta causes Ca2+ oscillations and parthenogenetic activation of human oocytes. *Reproduction.* 2004;128(6):697-702. doi:10.1530/rep.1.00484

118. Escoffier J, Lee HC, Yassine S, et al. Homozygous mutation of PLCZ1 leads to defective human oocyte activation and infertility that is not rescued by the WW-binding protein PAWP. *Hum Mol Genet.* 2016;25(5):878-891. doi:10.1093/hmg/ddv617

119. Govindaraju A, Dogan S, Rodriguez-Osorio N, Grant K, Kaya A, Memili E. Delivering value from sperm proteomics for fertility. *Cell Tissue Res.* 2012;349(3):783-793. doi:10.1007/s00441-012-1452-2

120. Hu Y, Benedict MA, Ding L, Núñez G. Role of cytochrome c and dATP/ATP hydrolysis in Apaf-1-mediated caspase-9 activation and apoptosis. *EMBO J.* 1999;18(13):3586-3595. doi:10.1093/emboj/18.13.3586

121. Lécuyer C, Dacheux JL, Hermand E, Mazeman E, Rousseaux J, Rousseaux-Prévost R. Actin-binding properties and colocalization with actin during spermiogenesis of mammalian sperm calicin. *Biol Reprod.* 2000;63(6):1801-1810. doi:10.1095/biolreprod63.6.1801

122. von Bülow M, Heid H, Hess H, Franke WW. Molecular nature of calicin, a major basic protein of the mammalian sperm head cytoskeleton. *Exp Cell Res.* 1995;219(2):407-413. doi:10.1006/excr.1995.1246

123. Ben Khelifa M, Coutton C, Zouari R, et al. Mutations in *DNAH1*, which encodes an inner arm heavy chain dynein, lead to male infertility from multiple morphological abnormalities of the sperm flagella. *Am J Hum Genet.* 2014;94(1):95-104. doi:10.1016/j.ajhg.2013.11.017

124. Levkova M, Radanova M, Angelova L. Potential role of dynein-related genes in the etiology of male infertility: a systematic review and a meta-analysis. *Andrology.* 2022;10(8):1484-1499. doi:10.1111/andr.13287

125. Davis FM, Goulding EH, D'Agostin DM, et al. Male infertility in mice lacking the store-operated Ca(2+) channel Orai1. *Cell Calcium.* 2016;59(4):189-197. doi:10.1016/j.ceca.2016.02.007

126. Albarran L, Lopez JJ, Jardin I, et al. EFHB is a novel cytosolic Ca2+ sensor that modulates STIM1–SARAF interaction. *Cell Physiol Biochem.* 2018;51(3):1164-1178. doi:10.1159/000495494

127. Di Gioia SA, Letteboer SJF, Kostic C, et al. FAM161A, associated with retinitis pigmentosa, is a component of the cilia–basal body complex and interacts with proteins involved in ciliopathies. *Hum Mol Genet.* 2012;21(23):5174-5184. doi:10.1093/hmg/dds368

128. Khanal S, Leung MR, Royfman A, et al. A dynamic basal complex modulates mammalian sperm movement. *Nat Commun.* 2021;12(1):3808. doi:10.1038/s41467-021-24011-0

129. Hess H, Heid H, Franke WW. Molecular characterization of mammalian cylicin, a basic protein of the sperm head cytoskeleton. *J Cell Biol.* 1993;122(5):1043-1052. doi:10.1083/jcb.122.5.1043

130. Panner Selvam MK, Agarwal A. Update on the proteomics of male infertility: a systematic review. *Arab J Urol.* 2018;16(1):103-112. doi:10.1016/j.aju.2017.11.016

131. Gilany K, Minai-Tehrani A, Savadi-Shiraz E, Rezadoost H, Lakpour N. Exploring the human seminal plasma proteome: an unexplored gold mine of biomarker for male infertility and male reproduction disorder. *J Reprod Infertil.* 2015;16(2):61-71.

132. Martins AD, Panner Selvam MK, Agarwal A, Alves MG, Baskaran S. Alterations in seminal plasma proteomic profile in men with primary and secondary infertility. *Sci Rep.* 2020;10(1):7539. doi:10.1038/s41598-020-64434-1

133. Agarwal A, Bertolla RP, Samanta L. Sperm proteomics: potential impact on male infertility treatment. *Expert Rev Proteomics.* 2016;13(3):285-296. doi:10.1586/14789450.2016.1151357

SUPPORTING INFORMATION

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