FULL TITLE: Novel and conventional approaches for the analysis of quantitative proteomic data are complementary towards the identification of seminal plasma alterations in infertile patients

AUTHORS: Ferran Barrachina1; Meritxell Jodar1; David Delgado-Dueñas1; Ada Soler-Ventura1; Josep Maria Estanyol2; Carme Mallofré3; Josep Lluís Ballescà4; Rafael Oliva1

AFFILIATIONS: 1Molecular Biology of Reproduction and Development Research Group, Institut d’Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Fundació Clinic per a la Recerca Biomèdica, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain and Biochemistry and Molecular Genetics Service, Hospital Clínic, Barcelona, Spain; 2Proteomics Unit, Scientific Technical Services, University of Barcelona, Barcelona, Spain; 3Department of Pathology, University of Barcelona, Hospital Clínic, Barcelona, Spain; 4Clinic Institute of Gynaecology, Obstetrics and Neonatology, Hospital Clínic, Barcelona, Spain.

* These authors contributed equally

CORRESPONDING AUTHOR: Rafael Oliva, Genetics Unit, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, University of Barcelona, Carrer Casanova 134, 08036 Barcelona, Spain. Tel.: +34 934021877; E-mail.: roliva@ub.edu

AUTHOR’S MAILING ADDRESSES: Ferran Barrachina (fbarrach@clinic.cat); Meritxell Jodar (meritxell.jodar@ub.edu); David Delgado-Dueñas (david.delgado82@gmail.com); Ada Soler-Ventura (ada.soler@ub.edu); Josep Maria Estanyol (jmestanyol@ub.edu); Carme Mallofré (mallofre@clinic.cat); Josep Lluís Ballescà (ballesca@clinic.cat); Rafael Oliva (roliva@ub.edu).

RUNNING TITLE: Seminal plasma proteome signatures

KEYWORDS: Male infertility, Seminal parameters, Seminal plasma, Quantitative proteomics, Stable-protein pairs, Biomarker: Diagnostic, Clinical proteomics, High Throughput Screening, Mass Spectrometry, Quantification.
ABBREVIATIONS

ANOVA, one-way analysis of variance;

BCA, bicinchoninic acid;

CASA, computer assisted semen analysis;

FDR, false discovery rate;

HCD, higher energy collision dissociation;

HPA, Human Protein Atlas;

IAA, iodoacetamide;

LC-MS/MS, liquid chromatography coupled with tandem mass spectrometry;

MS/MS, tandem mass spectrometry;

PAS, periodic acid and schiff’s reagent;

PSMs, peptide spectrum matches;

PTMs, post-translational modifications;

RT, room temperature;

SEMGs, semenogelins;

TBST, TBS with 0.1% (v/v) Tween 20;

TCEP, tris (2-carboxyethyl) phosphine;

TEAB, triethyl ammonium bicarbonate;

TMT, tandem mass tag;

WHO, World Health Organization.
SUMMARY

Our aim was to define seminal plasma proteome signatures of infertile patients categorized according to their seminal parameters using TMT-LC-MS/MS. To that extent, quantitative proteomic data was analyzed following two complementary strategies: (i) the conventional approach based on standard statistical analyses of relative protein quantification values; and (ii) a novel strategy focused on establishing stable-protein pairs. By conventional analyses, the abundance of some seminal plasma proteins was found to be positively correlated with sperm concentration. However, this correlation was not found for all the peptides within a specific protein, bringing to light the high heterogeneity existing in the seminal plasma proteome due to both the proteolytic fragments and/or the post-translational modifications. This issue was overcome by conducting the novel stable-protein pairs analysis proposed herein. A total of 182 correlations comprising 24 different proteins were identified in the normozoospermic-control population, while this proportion was drastically reduced in infertile patients with altered seminal parameters (18 in patients with reduced sperm motility, 0 in patients with low sperm concentration and 3 in patients with no sperm in the ejaculate). These results suggest the existence of multiple etiologies causing the same alteration in seminal parameters. Additionally, the repetition of the stable-protein pair analysis in the control group by adding the data from a single patient at a time enabled to identify alterations in the stable-protein pairs profile of individual patients with altered seminal parameters. These results suggest potential underlying pathogenic mechanisms in individual infertile patients, and might open up a window to its application in the personalized diagnostic of male infertility.
INTRODUCTION

Infertility is a worldwide frequent problem that affects approximately 15% of reproductive-aged couples. Around the 50% of the fertility problems is due to a male factor and, from those, the 40-60% of the infertile patients present some alterations in at least one of the seminal parameters assessed by a routine semen analysis (sperm concentration, motility and morphology) (1). According to these seminal parameters, infertile patients are categorized in: (i) patients with low or absent sperm concentration (oligozoospermia or azoospermia, respectively), (ii) patients with defective sperm motility (asthenozoospermia) and/or (iii) patients with abnormal sperm morphology (teratozoospermia) (2). The semen evaluation, together with a complete medical history and physical examination, will determine whether the initial assessment needs to be complemented with genetic and/or hormonal analyses, urinalysis or testicular biopsies. Unfortunately, the current available tools for the evaluation of male fertility are limited and insufficient, and the development of new methodologies to better discern the male factor etiology is required (3–6). Currently, the application of high-throughput proteomics for the study of the human sperm cell has resulted in the identification of 6871 proteins as well as some pathogenic mechanisms involved in male infertility (7–11). For instance, studies on the sperm proteome from asthenozoospermic patients revealed alterations mainly in proteins and pathways related with energy production and cytoskeleton (9, 12). However, since the semen is not just composed by sperm cells, the exploration of seminal plasma could help to better elucidate the causes of male infertility.

The human seminal plasma is a complex and protein-enriched biological fluid that constitutes 95% of the semen volume, whereas only the remaining 5% corresponds to spermatozoa (10). This fluid is composed by secretions from the testis (1-2%), the epididymis (2-4%) and the male accessory sex glands including the seminal vesicles (65-75%), the prostate gland (25-30%) and the bulbourethral glands (< 1%) (13–15). Seminal plasma contains a diversity of molecules including DNA, RNA, microRNAs, lipids, proteins and metabolites, together with a highly abundant population of extracellular vesicles, which are mainly secreted by the prostate (10, 16). So far, the compiled proteome profile of human seminal plasma
includes 2064 non-redundant proteins from 9 independent studies, revealing that this fluid is not only a simple medium to carry the spermatozoa through the female reproductive tract (10, 17). In fact, once the spermatogenesis is completed, the testicular spermatozoa, although morphologically differentiated, are immotile germ cells unable to fertilize the oocyte by its own (10). Seminal plasma plays a crucial role for the sperm maturation, motility and capacitation, prevention of premature acrosome reaction and sperm-to-zona pellucida recognition and fusion, thereby providing to the spermatozoa the capability to fertilize the oocyte (18–25). Additionally, the contact of the seminal plasma with the female reproductive tract provides an optimal environment that enhances embryo implantation and development as well as contributes promoting maternal immune tolerance of the semiallogenic fetus (26–29).

The analysis of the seminal plasma proteome is currently undergoing an intense study (14, 30). With the aim to shed light on the physiological role of the seminal plasma and to seek for specific protein biomarkers for male infertility diagnosis and/or prognosis, several groups have studied the seminal plasma proteome in certain subtypes of male infertility and/or alterations associated with oxidative stress and sperm functional traits, among others (31–49). In addition, there is a particular interest to identify seminal plasma protein biomarkers that could be predictive for the presence of sperm cells in the testis of azoospermic patients, which could avoid unnecessary testicular biopsies (34, 50–57).

Therefore, the analysis of the seminal plasma proteome mimicking the infertility classification according to seminal parameters is warranted, since it might help to decipher potential pathogenic mechanisms resulting in these sperm alterations. However, it is important to take into account that a wide range of factors may alter both sperm and seminal plasma compositions, leading to a high heterogeneity within patients sharing the same phenotype. This is challenging for the conventional analysis of the quantitative proteomic data based on the search of differential proteins between groups of patients with similar characteristics. Therefore, the main aim of the present study was to define the seminal plasma proteome signatures of infertile patients categorized according to their seminal parameters (normozoospermia, NZ; asthenozoospermia, AS; oligozoospermia, OZ; azoospermia, AZ) by applying
conventional and novel approaches for the analysis of quantitative proteomic data to try to better stratify the different subgroups of infertile patients. The results derived from this study suggest that the combination of conventional and novel analytical approaches may be useful towards the identification of pathogenic mechanisms of male infertility and, furthermore, to provide the bases for future studies to design new therapies to improve male fertility and move towards the application of an individual and personalized diagnostic of male infertility.

**EXPERIMENTAL PROCEDURES**

**Biological material and sample collection**

Semen samples - Human semen samples (n=34) were obtained from patients undergoing routine semen analysis at the Assisted Reproduction Unit from the Clinic Institute of Gynaecology, Obstetrics and Neonatology, from the Hospital Clinic (Barcelona, Spain), after signed informed consent. The ejaculates were collected by masturbation into sterile containers after 3-5 days of sexual abstinence. The evaluation of the seminal parameters was performed using the automatic semen analysis system CASA (Proiser, Paterna, Spain), and the sperm viability test was assessed using 0.5% (w/v) Eosin Y, following the WHO recommendations (2). The semen samples included in this study were classified according to semen parameters as normozoospermic (NZ, n=10), asthenozoospermic (AS, n=4), oligozoospermic (OZ, n=10) and azoospermic (AZ, n=10) (Supplemental Table S1).

Testis and epididymis tissues – Biopsies from normal testis (n=1) and normal epididymis (n=1) were provided by the Department of Pathology from the Hospital Clinic (Barcelona, Spain).

Ethics Statement – All samples were used in accordance to the appropriate ethical guidelines and Internal Review Board, and the biological material storing and processing was approved by the Clinical Research Ethics Committee of the Hospital Clinic of Barcelona. The written informed consent was obtained from all subjects in accordance with the Declaration of Helsinki.
Purification and isolation of proteins from seminal plasma

Liquefied semen samples were centrifuged at 500 g for 10 min and 1500 g for 10 min, to separate the sperm cells from the seminal plasma. The resulting seminal plasma was filtered (0.45 μm pore size) to remove any cellular leftovers. All cell-free seminal plasma samples were frozen at -80 ºC until further processing. After thawing, seminal plasma samples were centrifuged at 16000 g for 10 min at 4 ºC and the protein concentration from the supernatant of each sample was determined using the BCA protein Assay Kit (Pierce™ BCA protein Assay Kit, Thermo Fisher Scientific, Rockford, IL, USA), following manufacturer’s recommendations.

Seminal plasma peptide isotopic labeling (TMT 10-plex)

A total of 16 seminal plasma samples were selected for the proteomics study, including 4 NZ, 4 AS, 4 OZ and 4 AZ patients (Fig. 1). Differential peptide labeling was performed using TMT 10-plex isotopic label reagent set (TMT 10-plex Mass Tag Labeling; Thermo Fisher Scientific, Rockford, IL, USA), following manufacturer’s instructions. Briefly, 60 μg of protein from each seminal plasma sample was adjusted to a final volume of 60 μl with 100 mM TEAB, and protein quantification was repeated to ensure that all samples had the same concentration (1 μg/μl). Proteins were reduced in 9.5 mM TCEP for 1 h at 55 ºC, alkylated with 17 mM IAA for 30 min in the dark, and precipitated with 500 μl of cold 100% acetone at -20 ºC overnight. Samples were centrifuged at 17500 g for 10 min at 4 ºC and the acetone-precipitated protein pellets were resuspended in 60 μl of 100 mM TEAB. Trypsin was then added at 1:20 protein-to-protease ratio and incubated overnight at 37 ºC with constant shaking. Prior to peptide labeling, aliquots with the same volume and concentration were taken out from each of the 16 samples and combined together, in order to constitute the internal control. After that, 30 μg of peptides from each individual seminal plasma sample (n=16) and internal control (n=1) were labeled with TMT isobaric tags (reporter ions intensity from m/z 127.1 to m/z 131.1 (TMT-127N, -127C, -128C, -129N, -129C, -130N, -130C, -131) and m/z 126 (TMT-126), respectively; Fig. 1). Specifically, 19.5 μl of the TMT label reagents previously equilibrated at RT and dissolved in ACN (Sigma-Aldrich, St. Louis, MO, USA) were
added to the corresponding reduced and alkylated peptides. After 1 h of incubation at RT, the reaction was quenched with 4 μl of 5% hydroxylamine for 15 min. Labeled peptides from each sample were combined at equal amounts constituting two different multiplex pools (Pool A and Pool B; each one consisting of 8 different samples plus the same internal control; Fig. 1), which were dried in a speed-vacuum centrifuge and peptides were resuspended in 20 μl of 0.5% TFA (Sigma-Aldrich, St. Louis, MO, USA) in 5% ACN. Finally, the peptides were cleaned up via reversed-phase C18 spin columns (Pierce C18 Spin Columns, Thermo Scientific, Rockford, IL, USA), following manufacturer’s instructions.

<INSERT FIGURE 1>

**LC-MS/MS analysis**

Labeled peptides were analyzed by a nano-LC Ultra 2D Eksigent (AB Sciex, Brugg, Switzerland) attached to an LTQ-Orbitrap Velos (Thermo scientific, San Jose, CA, USA). For HPLC separation, peptides were injected onto a C18 trap column (L 0.5 cm, 300 μm ID, 5μm, 100 Å; Thermo Fisher Scientific, San Jose, CA, USA). Chromatographic analyses were performed using an analytical column (L 15 cm, 75 μm ID, 3 μm, 100 Å; Thermo scientific, San Jose, CA, USA). Two different buffer systems were used for the analysis: buffer A (97% H2O-3% ACN, 0.1% Formic acid) and buffer B (3% H2O-97% ACN, 0.1% Formic acid). The following gradient was applied for peptide separation on the analytical column: from 0-4 min 0% of B to 4% of B, from 4-300 min 4% of B to 35% of B, from 300-305 min 35% of B to 100% of B, at a flow rate of 400 nl/min, and from 305-320 min 100% of B at a flow rate of 400 nl/min. MS/MS analyses were performed using an LTQ-Orbitrap Velos (Thermo Fisher Scientific, Waltham, MA, USA) directly coupled to a nanoelectrospray ion source. The LTQ-Orbitrap Velos settings included one 30000 resolution at 400 m/z MS1 scan for precursor ions followed by MS2 scans of the 50 most intense precursor ions, at 30000 resolution at 400 m/z, in positive ion mode. The lock mass option was enabled, and polysiloxane (m/z 445.12003) was used for internal recalibration of the mass spectra. MS/MS data acquisition was completed using Xcalibur 2.1 (Thermo Fisher Scientific, Waltham, MA, USA). The normalized collision energy for HCD-MS2 was set to 40%.
Protein identification

LC-MS/MS data was analyzed using Proteome Discoverer 1.4.1.14 (Thermo Fisher Scientific, Waltham, MA, USA). For database searching, raw mass spectrometry files were submitted to the in-house Homo sapiens UniProtKB/Swiss-Prot database with Sus scrofa Trypsin added to it (HUMAN_Tryp_UP_SP_R_2016_03.fasta; released March 2016; 20155 protein entries) using SEQUEST HT version 28.0 (Thermo Fisher Scientific, Waltham, MA, USA). For re-scoring, percolator search node was used. Searches were performed using the following parameters: five maximum missed cleavage sites for trypsin, TMT-labeled lysine (+229.163 Da) and methionine oxidation (+15.995 Da) as dynamic modifications, cysteine carbamidomethylation (+57.021 Da) as a static modification, 20 ppm precursor mass tolerance, 0.6 Da fragment mass tolerance, 5 mmu peak integration tolerance, and most confident centroid peak integration method. Percolator was used for protein identification with the following identification criteria: at least one unique peptide per protein with a FDR of 1%.

Proteomic data analysis by conventional and novel approaches

Conventional relative protein quantification – Normalized TMT quantitative values for each identified spectrum derived from the ratio of the intensity of reporter ions from HCD MS2 spectra corresponding to each individual samples (TMT-127N to TMT-131) with the internal control (TMT-126), which were obtained using Proteome Discoverer software (Supplemental Fig. S1) (12, 58). Different isoforms of the same protein were treated as dissociated or “ungrouped” from their respective families, to avoid any possible ambiguity (12, 58). Only those proteins with at least 1 unique peptide quantified by ≥ 2 PSMs in all the samples and a coefficient of variation < 50% in at least 75% of the samples were considered for further statistical analyses. Significant statistical differences among the different subtypes of infertile patients were evaluated after normalizing the relative proteomic quantification values by log2 transformation using ANOVA combined with Tukey's multiple comparison test. Additionally, the correlations between sperm concentration and normalized relative proteomic quantification values were
assessed using Pearson correlation test followed by the adjustment of the p-values to FDR. An MS expert checked the spectra of all the differential proteins.

Establishment of stable-protein pairs profile – The intensity values from HCD MS2 spectra corresponding to each individual sample (TMT-127N to TMT-131), but not from the internal control (TMT-126), were used to establish the stable-protein pairs for each group of patients. This strategy has been previously applied to the study of stable-transcript pairs obtained from RNA-seq data (59, 60). In our case, only those proteins with at least 2 unique peptides quantified by $\geq 2$ PSMs for all samples with a coefficient of variation $< 50\%$ in at least $75\%$ of the samples were considered. Stable-protein pairs were determined by applying the following statistical principle: 2 proteins (with more than 1 peptide quantified for each one) were highly-correlated when $\geq 75\%$ of the possible peptide combinations had a Pearson correlation coefficient $\geq 0.9$ (Fig. 2). In order to determine alterations in individual samples, stable-protein pair analysis was repeated for the control group (NZ patients) by adding a patient with altered seminal parameters once at a time.

<INSERT FIGURE 2>

**Functional enrichment and expression analyses using public databases**

The seminal plasma proteomic datasets were uploaded to the Gene Ontology Consortium database (http://www.geneontology.org/) (61), based on PANTHER v13.1 database (Release date 2018-02-03), in order to predict the functional involvement of the seminal plasma proteins. The significance of enrichment analyses was calculated by a Fisher’s exact test. P-values $< 0.05$ after FDR adjustment were considered statistically significant.

The HPA Database (http://www.proteinatlas.org/) (62, 63) was used to assess the expression of specific proteins in different human male reproductive tissues.
Immunoblotting

Protein extracts from cell-free seminal plasma samples from an independent set of patients (n=18; 6 NZ, 6 OZ, 6 AZ; Supplemental Table S1) were used for Western blotting validation of ECM1 protein. A total of 40 μg of seminal plasma protein extracts from each sample were separated by SDS-PAGE and transferred onto Immobilon-P PVDF membranes (Merck Millipore, Tullagreen, Ireland) as described elsewhere (7). The membranes were blocked in TBST and 5% (w/v) skim milk for 1 h at RT. For immunostaining, anti-ECM1 antibody (polyclonal rabbit ECM1 antibody, #43263, SAB Signalway Antibody, Baltimore, MD, USA) diluted 1:500 in TBST was used. After washing in TBST, membranes were incubated with an ECL horseradish peroxidase-labeled donkey anti-rabbit IgG antibody (Amersham™, GE Healthcare, Little Chalfont, UK). Finally, the detection was done using WesternBright™ ECL Western blotting detection kit (Advantast, Menlo Park, CA, USA) and a LAS-4000 image analyzer (ImageQuant LAS 4000, GE Healthcare Life Sciences, Little Chalfont, UK). Optical densities from ECM1 bands were obtained and analyzed using Quantity One software (BioRad, Hercules, CA, USA). The correlation between sperm concentration and the optical densities values from ECM1 was evaluated using Pearson correlation test.

Immunohistochemistry

Cross-sections from normal testis (n=1) and a normal epididymis (n=1) were used to detect the expression pattern of ECM1 and NPC2. Bouin’s fixed, paraffin-embedded testicular and epididymal sections (4 μm) were deparaffinized in toluene (3×) and hydrated through graded series of ethanol (100%, 100%, 90%, 70%, H2O milliQ), with a 0.3% hydrogen peroxide (Sigma-Aldrich, St. Louis, MO, USA) intermediate incubation between the two 100% ethanol incubations. For antigen retrieval, sections were incubated with 10 mM sodium citrate (pH 6.0) at 99.5 ºC for 20 min. Sections were blocked with PBS-5% skim milk for 30 min at RT and, then, incubated with Avidin/Biotin Blocking Kit (Vector Laboratories, Burlingame, CA, USA). Afterwards, sections were incubated for 16 h at 4 ºC with the primary antibody of interest: anti-ECM1 (polyclonal rabbit ECM1 antibody, #EPP12545, Elabscience, Houston, TX, USA)
diluted 1:20 in PBS-1% skim milk, and anti-NPC2 (polyclonal rabbit NPC2 antibody, #CQA1207, Cohesion Biosciences, London, United Kingdom) diluted 1:100 in PBS-5% skim milk. Negative controls for non-specific binding of the primary antibodies were included using isotype rabbit IgG (Vector Laboratories, Burlingame, CA, USA). Then, sections were incubated with a biotinylated goat anti-rabbit secondary antibody (Vector Laboratories, Burlingame, CA, USA), and with an avidin-biotin-peroxidase detection kit (Vectastain ABC Elite Kit, Vector Laboratories, Burlingame, CA, USA). Finally, slides were incubated with ImmPACT™ DAB Peroxidase Substrate kit (Vector Laboratories, Burlingame, CA, USA). All slides were then washed and stained with haematoxylin, and testis slides were additionally stained with PAS (Sigma-Aldrich, St. Louis, MO, USA). Sections were dehydrated in 100% ethanol (3×), cleared in toluene (3×) and mounted in Eukitt Mounting Medium (Sigma-Aldrich, St. Louis, MO, USA) to be finally analyzed by a transmission light microscope (Olympus BX50, Olympus, Tokyo, Japan).

**Experimental Design and Statistical Rationale**

The experimental design of this study is shown in Fig. 1. Specifically, seminal plasma proteome from 16 infertile patients including individuals with normal seminal parameters (NZ; control group) and infertile patients with altered seminal parameters (AS, OZ and AZ patients; Supplemental Table S1) was characterized and compared. A total of 4 biological replicates per group were used for proteomic analysis. Significant statistical differences among the different subtypes of infertile patients were evaluated after the normalization of the relative proteomic quantification values by log2 transformation using ANOVA combined with Tukey's multiple comparison test. Since some differential proteins showed a gradual decreasing reliant on sperm concentration, we tested the correlation between sperm concentration and normalized values of the relative proteomic quantification using Pearson correlation test followed by the p-value adjustment to FDR. To further validate these results, immunoblotting analysis of one differential protein (ECM1) was performed in an independent set of infertile patients as biological replicates (6 NZ, 6 OZ, 6 AZ). The correlation between sperm concentration and the abundance of ECM1 protein was evaluated using Pearson correlation test.
The expression profiles of the differential proteins in human reproductive tissues were also assessed using the information available at the HPA, in order to discern whether the proteins correlated with sperm concentration were just reflecting a variation in the amount of sperm leftovers in the ejaculate or, in contrast, were the result of some proteomic alterations in the secretions from accessory sex glands. Additionally, immunohistochemistry was performed in testicular (n=1) and epididymal (n=1) biopsies to infer the tissue origin of 2 differentially expressed proteins, ECM1 and NPC2. Finally, the correlation between sperm concentration and normalized relative protein abundance was also explored at peptide level, since discrepancies within our findings and results published by others suggested differences on the specific peptides analyzed for each protein (50–52).

In order to assess the heterogeneity of the seminal plasma proteomic profile within the subgroups of infertile patients characterized according to seminal parameters as well as the specific protein alterations in individual patients, a new approach based on the analysis of the stable-proteins pairs was conducted.

Both, statistical analyses and the establishment of stable-protein pairs were carried out using R software version 3.4.4 (http://www.r-project.org) (64). P-values < 0.05 were considered statistically significant. All graphs were constructed using GraphPad Prism software version 5.01 (GraphPad Software Inc., San Diego, CA, USA).

RESULTS

Proteomic analysis of human seminal plasma

LC-MS/MS analysis resulted in the identification of a total of 349 proteins in the seminal plasma proteome from 16 infertile patients, with at least one unique peptide and 1% FDR (Supplemental Tables S2 and S3). However, just 60 of the 349 seminal plasma proteins fit our strict quantification criteria (at least 1 unique peptide quantified with ≥ 2 PSMs in all samples with a coefficient of variation < 50% in at least 75% of the samples) and, therefore, only these proteins were used for subsequent analyses. Detailed
information of the quantifiable peptides and corresponding proteins are presented in Supplemental Tables S4 and S5, respectively.

With the aim to assess the potential role of seminal plasma proteins in the functionality of spermatozoa, as well as the ability of seminal plasma proteome signatures to reflect disturbances in spermatogenesis and sperm maturation processes that could explain the alteration of seminal parameters, we compared the seminal plasma proteome from infertile patients categorized according to their seminal parameters (NZ, AS, OZ, AZ). Two different strategies were applied in our comparative quantitative proteomics study: (i) A conventional approach based on standard statistic analyses (ANOVA and Pearson correlation test) of relative protein quantification values; and (ii) a novel analysis method based on the establishment of the stable-protein pairs separately in groups of patients classified according seminal parameters, as well as the identification of protein alterations in individual samples based on the variations of the stable-protein pairs defined in NZ patients (Fig. 1). The results of each analysis are shown below.

Altered seminal plasma protein abundance in infertile patients

Comparison of the seminal plasma proteomes from the 4 different subtypes of infertile patients (NZ, AS, OZ, AZ) by conventional data analysis revealed a set of 6 differentially expressed proteins among the groups (p-value < 0.05; ANOVA with Tukey’s Post Hoc test; Fig. 3, Supplemental Table S6). Of interest, the Post Hoc test reflected that the protein abundance of 3 of these 6 differentially expressed proteins (CRISP1, NPC2 and SPINT3) was reduced in patients with low or absence of sperm cells in the ejaculate (OZ and AZ patients, respectively), whereas SCPEP1 protein abundance was increased in AZ patients (Fig. 3). Additionally, only the protein ANPEP displayed reduced protein abundance in patients with decreased sperm motility (AS patients).

<INSERT FIGURE 3>
Relationship between altered seminal plasma proteins and the sperm concentration

The correlation between the relative amount of seminal plasma proteins and the sperm concentration parameter was assessed in order to test whether the altered protein abundance detected in OZ and AZ patients relied in the number of sperm cells present in the ejaculate independently of sperm motility rate. Remarkably, the abundance of the proteins SPINT3, NPC2, ECM1, CRISP1 and IGHG2 increased with higher sperm concentration (Table I). To validate these results, a Western blotting analysis for ECM1 protein was performed in an independent set of samples not used for MS analysis (n=18; 6 NZ, 6 OZ, 6 AZ). As expected, higher protein abundance for ECM1 protein was found with an increased sperm count (p-value < 0.01; Fig. 4). The HPA database showed that 3 of the 5 seminal plasma proteins positively correlated with sperm concentration are mainly expressed in epididymis, but not detected in testis (Fig. 5A). We extended the protein patterns provided by the HPA database by immunohistochemical validation of ECM1 and NPC2 proteins in human testis (n=1) and human epididymis (n=1) biopsies (Fig. 5B). However, the analysis at peptide level showed that whereas the majority of peptides quantified for NPC2 protein (6 of 7 peptides) maintain the correlation with sperm concentration, only one peptide quantified for each of the remaining proteins was found correlated (1 of 1 peptide for SPINT3, 1 of 2 peptides for CRISP1 and IGHG2 and 1 of 3 peptides for ECM1; Table II).

Seminal plasma stable-protein pair profiles among patient groups and individuals

A total of 182 stable-protein pairs between 24 different proteins were identified for patients with normal semen parameters (NZ) (Table III). Of note, those 24 proteins are functionally involved in processes already ascribed to seminal plasma, such as the regulation of sperm function, semen
coagulation-liquefaction processes, immune system and lipid metabolism, among others. In contrast, very few stable-protein pairs were observed in the different subtypes of patients with altered seminal parameters: 18 stable-protein pairs comprising 16 proteins in AS, 0 in OZ and 3 comprising 5 proteins in AZ (Table III).

<INSERT TABLE III>

In order to assess alterations of stable-protein pairs in individual patients, the analysis of the stable-protein pairs for the NZ patients group was repeated by adding one single patient with an altered seminal parameter at a time (Table IV). This strategy revealed that the asthenozoospermic patient AS2 had a similar seminal plasma proteomic signature to that found in NZ men, since > 75% of the stable-protein pairs established for NZ population were maintained after performing the stable protein-pair analysis with proteomic data from NZ samples (n=4) and patient AS2 (Table IV). In contrast, when adding individually three of the four azoospermic patients (AZ1, AZ2 and AZ4) into the analysis of NZ stable-protein pairs, we detected < 50% stable-protein pairs determined in the NZ individuals, therefore reflecting huge differences in the seminal plasma proteome signature of those AZ patients (Table IV).

<INSERT TABLE IV>

DISCUSSION

The heterogeneous composition of the seminal plasma together with the rapid changes that occur in its molecular composition after ejaculation, such as the proteolytic cascade associated to the coagulation-liquefaction process, introduce further complexity to seminal plasma proteomic studies (10, 14). In the present study, a total of 349 proteins were identified in the 16 seminal plasma samples analyzed (Supplemental Table S2), which are functionally related to metabolism, response to stress, proteolysis, immune system and energy production. Also, with less extent, these identified human seminal plasma proteins seemed to be involved in processes related to fertilization and embryogenesis (8). This apparently low number of identified seminal plasma proteins could be explained by the detection of the
semenogelins I and II (SEMG1 and SEMG2) as the most abundant proteins of the seminal plasma. Specifically, around 40% of the PSMs identified in our proteomic study corresponded to SEMG1 and SEMG2, thus hindering the detection of low abundant proteins. This low number of protein identifications is also observed in other studies assessing the human seminal plasma proteome using MS methods and identification criteria comparable to ours (13, 31, 32, 40–42, 46, 47, 49, 65). For this reason, future studies should consider the incorporation of strategies to deplete SEMGs prior the proteomic characterization of seminal plasma, as for example the use of HPLC columns containing antibodies against SEMGs (10, 13).

Conventional approach to analyze quantitative proteomics data

Protein quantification of TMT-labeled peptides using conventional approaches is obtained from the average of relative ion abundance ratios for all peptides encompassing the same protein (12). The conventional statistical analyses conducted in this study showed: (i) The underexpression of the glycoprotein ANPEP in patients with altered sperm motility (Fig. 3), as previously reported by others (33); and (ii) a gradual decline of CRISP1, NPC2 and SPINT3 abundance in infertile patients, ranging from high to low sperm concentration (in decreasing order: NZ-AS, OZ and AZ) (Table I, Fig. 3). Of note, this gradual declined abundance was also observed for the protein levels of SPINT3, NPC2, ECM1 and CRISP1, independently of sperm motility parameter (Table I). The low abundance of those proteins in seminal plasma from patients with low or absence of sperm cells could reflect either proteomic alterations in the accessory sex glands secretions, or the presence of low amounts of male germ cells remnants coming from apoptotic sperm or from sperm cytoplasmic droplets. This specific question may be elucidated by deciphering the potential tissue origin of these altered seminal plasma proteins. According to the HPA Database, the testicular or extra-testicular origin of NPC2 could not be assessed, since, although NPC2 is a major component of epididymal secretions, it is also expressed in testis (43) (Fig. 5). In contrast, ECM1, SPINT3 and CRISP1 are mainly expressed in the epididymis while they are not detected in testis (Fig. 5), suggesting that epididymal secretions could be regulated by the presence of
sperm themselves in the epididymis. Interestingly, this potential cross-talk between the spermatozoa and epididymis has also been observed in rat and bovine species (66, 67). Drabovich and colleagues, in contrast, detected decreased levels of SPINT3, NPC2, ECM1 and CRISP1 only in obstructive azoospermic patients but not in patients with non-obstructive azoospermia (16). Other groups assessing potential seminal biomarkers to discern the different subtypes of azoospermia found other distinct differential proteins, such as LGALS3BP (57) and STAB2, CP135, GNRP and PIP (34). Altogether, it indicates the presence of some differences between seminal plasma proteomic data from different studies. Of note, a similar lack of concordance has also been observed in quantitative proteomic data of sperm samples from the same type of patients. This is exemplified by the detection of only 17 proteins out the 179 reported as differentially expressed in the sperm cells in at least 2 of the 7 comparative proteomic studies conducted so far for the study of asthenozoospermia (68). Differences in sample collection, handling and storage, proteomic strategies, and the biological intra- and inter-individual variance may be important causes contributing to this lack of reproducibility between studies (14, 17, 69). Moreover, the high and fast protease activity in seminal plasma after ejaculation could introduce even more heterogeneity to the results due to the presence of distinct proteolytic fragments, in addition to the protein PTMs not detected in standard proteomic procedures (70). In order to evaluate this putative heterogeneity of seminal plasma proteome we also assessed the correlation of peptides encompassing the proteins SPINT3, NPC2, ECM1, CRISP1 and IGHG2 with sperm concentration (Table II). Of note, NPC2 is the only protein showing the majority of its corresponding peptides correlated with the sperm concentration (Table II). In agreement with our findings, Giacomini and colleagues, by using nano LC-electrospray ionization-MS/MS, found decreased levels of NPC2 in seminal plasma from idiopathic oligoasthenozoospermic patients (43). As a summary, there is a need of novel approaches to analyze the results from quantitative shotgun proteomic studies, in order to overcome the limitations produced by the heterogeneity of seminal plasma in the proteolytic fragments, as well as by its conjunction with other variations such as the protein PTMs not detectable by standard proteomic procedures.
Novel approaches to analyze quantitative proteomics and identify patient-specific alterations

A set of seminal plasma strictly co-regulated proteins was established by following a new approach based on the correlation of the intensities of all unique peptides comprising one specific protein with all the unique peptides quantified for all the other detected proteins in the sample (Fig. 2). This strategy, called stable-protein pairs analysis herein, may contribute to reduce the heterogeneity observed in the seminal plasma proteomic data, since only those proteins displaying a consistent pattern in a specific phenotype are obtained. A total of 182 stable-protein pairs comprising 24 proteins were detected in patients with normal semen parameters (Table III), reflecting the strict co-regulation of these proteins in NZ individuals. These stable-protein pairs include gene products involved in: (i) Sperm function, such as CKB that plays a critical role in the demand of energy necessary for sperm motility (71), CST3, PAEP and CLU that regulate sperm capacitation (72–74) and CRISP1 and NPC2, which are necessary for sperm-oocyte binding and fertilization (75, 76); (ii) the regulation of semen clotting-liquefaction processes, such as ACCP, KLK3 and MME, which are directly involved in the proteolysis of the SEMGs or other proteins (77, 78), or as WFDC2, ALB and TGM4, which regulate other components required for clotting-liquefaction process such as proteases, zinc ions and polyamines (79–81); (iii) immunology, including proteins that could participate in the leukocyte-mediated immune response, such as IGKC, IGHG2, ANPEP, LGALSBP3, ECM1 and B2M (82, 83) or in antimicrobial activity, as for example CPE and QSOX1 (84, 85); and finally,(iv) other functions such as lipid metabolism (AZGP1, TF and IDH1) and matrix assembly (VWA1). In contrast to the high number of stable-protein pairs identified in NZ individuals, the stable correlations drastically decreased in the different groups of patients with altered semen parameters (Table III). Indeed, just 18, 3 and 0 stable-protein pairs were detected in AS, AZ and OZ patients, respectively. The low number of stable-protein pairs observed in AZ patients was not surprising, since this group contains patients indistinctly diagnosed with obstructive and non-obstructive azoospermia, which probably results in different semen protein profiles as previously reported by others (16). However, a high heterogeneity was also observed in the proteomic profile of seminal plasma from
AS and OZ patients. The few stable-protein pairs detected in infertile patients with altered seminal parameters indicate that alterations in different proteins may result or be a consequence of the same altered phenotype. Although some hints for the presence of protein-pair correlations was already reported in the sperm proteome using ancillary proteomic methods (86), the present study clearly demonstrates the potential of this approach using proteomics data at peptide level (Table III, Fig. 2).

To assess which protein pairs might be associated to the alterations of the seminal parameters in individual samples, we repeated the analysis of the stable-protein pairs in the NZ population but adding data from one patient with altered parameters at a time. First of all, we observed that the patient AS2 had a very similar seminal plasma protein signature to the NZ population, since it maintained more than 75% of the NZ stable-protein pairs (Table IV). Of note, one of the proteins that loses more correlations in patient AS2 is the TF, a protein involved in lipid metabolism and sperm protection against oxidative stress (87) (Table IV). It is interesting to note that TF loses more than 75% of the correlations in another asthenozoospermic patient (AS1) (Table IV), suggestive that oxidative stress may be related to the impairment of the sperm motility in both patients AS1 and AS2.

We also observed that proteins involved in the induction of proteolysis of SEMGs and other regulators of the semen clotting-liquefaction process lose > 75% of the correlations in OZ and AZ patients (KLK3 in OZ2 and AZ2; ACPP in OZ2, AZ1, AZ2 and AZ3; and MME in AZ1; Table IV), suggesting that the sperm also contribute with regulators for the semen clot proteolysis (88). Whereas some of the protein correlations were lost in the majority of the individual patients, as observed for CLU, PAEP and WFDC2, suggesting that the correlations established for these proteins in NZ samples are weak, some seminal plasma proteins seem to be altered only in one unique sample such as TGM4 in patient AS4, IDH1 in OZ1, CST3 and LGALS3BP in OZ4, CPE in Z1, B2M in Z3 and IGKC in Z4, although these alterations could not clearly explain the observed phenotype.
Quantitative proteomics as a tool to provide insights in seminal plasma proteome signatures of infertile patients

So far, proteomics biomarker discovery experiments have shown a relatively low concordance among different studies. In fact, we demonstrated that the results from relative quantitative proteomics are different if the analyses are performed at protein or at peptide level. These differences could explain the apparent lack of reproducibility of some of the findings, a fact that should be taken into account also when using antibody-based techniques recognizing specific peptides (such as Western blotting or ELISA) or targeted proteomic approaches to selected specific peptides. Here, we propose introducing a novel complementary approach for the analysis of quantitative proteomic data, which is based on the establishment of stable-protein pairs. This strategy has been previously applied to the study of RNA-seq data (59, 60), but to the best of our knowledge this is the first time it is applied to the proteomic data. The use of this new approach in our seminal plasma proteome dataset has allowed determining highly stable seminal plasma proteome signatures of men presenting normal seminal parameters (NZ). In contrast, we demonstrated that the current classification of infertile patients based on altered semen parameters resulted in a high heterogeneous seminal plasma proteomic profile, thereby suggesting that the current male infertility stratification performed in fertility clinics is not sufficient to obtain a good diagnosis. Moreover, the stable-protein pairs approach has the potential to pinpoint proteins potentially related to pathogenic mechanisms in individual samples when this strategy is applied for the evaluation of the NZ stable-protein pair alterations in individual infertile patients. Although our study has limitations, the novel data analysis approach proposed herein could be valuable towards the identification of altered proteins and pathogenic mechanisms of male infertility and might open a window to the personalized diagnosis of male infertility in future studies.

ACKNOWLEDGEMENTS

This work was supported by grants to R.O. from the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competitividad; fondos FEDER ‘una manera de hacer Europa’ PI13/00699,
PI16/00346), from Fundación Salud 2000 (SERONO 13-015), from EUGIN-UB (EUREP 2014), and from EU-FP7-PEOPLE-2011-ITN289880. F.B. is granted by Spanish Ministry of Education, Culture and Sports (Ministerio de Educación, Cultura y Deporte para la Formación de Profesorado Universitario, FPU15/02306). M.J. is granted by Government of Catalonia (Generalitat de Catalunya, pla estratègic de recerca i innovació en salut, PERIS 2016-2020, SLT002/16/00337).

We greatly thank Dr Judit Castillo for critical revision of the manuscript. The authors also recognize Raquel Ferreti and Alicia Diez for their assistance in the routine seminograms and sample collection.

DATA AVAILABILITY

The MS proteomics data generated have been deposited to the ProteomeXchange Consortium via the PRIDE (89) partner repository with the dataset identifier PXD010579. The data files can be downloaded at: https://www.ebi.ac.uk/pride/archive/ (username: reviewer97302@ebi.ac.uk, password: Ghw37Tda).

AUTHOR’S CONTRIBUTIONS

F.B., M.J., A.S-V., R.O. designed the study; F.B., M.J., A.S-V. performed research; JM.E. performed proteomic technology; F.B., M.J., D.D-D. performed the data analysis; F.B., M.J., R.O. interpreted the data and wrote the manuscript; C.M., JL.B. were involved in the clinical diagnosis of patients; F.B., M.J., D.D-D., A.S-V., JM.E., C.M., JL.B., R.O. critically reviewed and approved the final version of the manuscript.
REFERENCES

1. Dohle, G. R., Colpi, G. M., Hargreave, T. B., Papp, G. K., Jungwirth, A., and Weidner, W. (2005) EAU guidelines on male infertility. *Eur. Urol.* 48, 703–711
2. World Health Organization (2010) *WHO laboratory manual for the Examination and processing of human semen*, 5th Ed., p. 272, World Health Organization, Geneva, Switzerland
3. Öehninger, S., Franken, D. R., and Ombelet, W. (2014) Sperm functional tests. *Fertil. Steril.* 102, 1528–1533
4. Weber, R. F. A., Dohle, G. R., and Romijn, J. C. (2005) Clinical laboratory evaluation of male subfertility. *Adv. Clin. Chem.* 40, 317–364
5. Milardi, D., Grande, G., Vincenzoni, F., Castagnola, M., and Marana, R. (2013) Proteomics of human seminal plasma: Identification of biomarker candidates for fertility and infertility and the evolution of technology. *Mol. Reprod. Dev.* 80, 350–357
6. Bieniek, J. M., Drabovich, A. P., and Lo, K. C. (2016) Seminal biomarkers for the evaluation of male infertility. *Asian J. Androl.* 18, 426–433
7. Amaral, A., Castillo, J., Estanyol, J. M., Ballescà, J. L., Ramalho-Santos, J., and Oliva, R. (2013) Human sperm tail proteome suggests new endogenous metabolic pathways. *Mol. Cell. Proteomics* 12, 330–342
8. Castillo, J., Jodar, M., and Oliva, R. (2018) The contribution of human sperm proteins to the development and epigenome of the preimplantation embryo. *Hum. Reprod. Update* 24, 535–555
9. Amaral, A., Castillo, J., Ramalho-Santos, J., and Oliva, R. (2013) The combined human sperm proteome: cellular pathways and implications for basic and clinical science. *Hum. Reprod. Update* 20, 40–62
10. Jodar, M., Soler-Ventura, A., and Oliva, R. (2017) Semen proteomics and male infertility. *J. Proteomics* 162, 125–134
11. Netherton, J. K., Hetherington, L., Ogle, R. A., Velkov, T., and Baker, M. A. (2017) Proteomics analysis of good and poor quality human sperm demonstrates several proteins are routinely aberrantly regulated. *Biol. Reprod.* 99, 395–408
12. Amaral, A., Paiva, C., Attardo Parrinello, C., Estanyol, J. M., Ballescà, J. L., Ramalho-Santos, J., and Oliva, R. (2014) Identification of proteins involved in human sperm motility using high-throughput differential proteomics. *J. Proteome Res.* 13, 5670–5684
13. Rolland, A. D., Lavigne, R., Dauly, C., Calvel, P., Kervarrec, C., Frecour, T., Evrard, B., Rioux-Leclercq, N., Auger, J., and Pineau, C. (2013) Identification of genital tract markers in the human seminal plasma using an integrative genomics approach. *Hum. Reprod.* 28, 199–209
14. Camargo, M., Intasqui, P., and Bertolla, R. P. (2018) Understanding the seminal plasma proteome and its role in male fertility. *Basic Clin. Androl.* 28, 6
15. McDonough, C. E., Whittington, E., Pitnick, S., and Dorus, S. (2016) Proteomics of reproductive systems: Towards a molecular understanding of postmating, prezygotic reproductive barriers. *J. Proteomics* 135, 26–37
16. Drabovich, A. P., Saraon, P., Jarvi, K., and Diamandis, E. P. (2014) Seminal plasma as a diagnostic fluid for male reproductive system disorders. *Nat. Rev. Urol.* 11, 278–288
17. Gilany, K., Minai-Tehrani, A., Savadi-Shiraz, E., Rezadoost, H., and Lakpour, N. (2015) Exploring the human seminal plasma proteome: An unexplored gold mine of biomarker for male Infertility and male reproduction disorder. *J. Reprod. Infertil.* 16, 61–71
18. Lindholm, C. (1974) The Importance of Seminal Plasma for Human Sperm Motility. *Biol. Reprod.* 10, 533–542
19. Primakoff, P., and Myles, D. G. (2002) Penetration, adhesion, and fusion in mammalian sperm-egg interaction. *Science* 296, 2183–2185
20. Burden, H. P., Holmes, C. H., Persad, R., and Whittington, K. (2006) Prostasomes-their effects on human male reproduction and fertility. *Hum. Reprod. Update* 12, 283–292
21. Allegrucci, C., Ronquist, G., Ove Nilsson, B., Carlsson, L., Lundqvist, M., Minelli, A., and Larsson, A. (2001) Circulating human antisperm antibodies recognize prostasomes. *Am. J. Reprod. Immunol.* 46, 211–219

22. Kravets, F. G., Lee, J., Singh, B., Trochia, A., Pentyala, S. N., and Khan, S. A. (2000) Prostasomes: current concepts. *Prostate* 43, 169–174

23. Saez, F., Frenette, G., and Sullivan, R. (2003) Epididymosomes and Prostasomes: Their Roles in Posttesticular Maturation of the Sperm Cells. *J. Androl.* 24, 149–154

24. Arieti, G. (1999) The motility of human spermatozoa as influenced by prostasomes at various pH levels. *Biol. Cell* 91, 51–54

25. García-Rodríguez, A., de la Casa, M., Peinado, H., Gosálvez, J., and Roy, R. (2018) Human prostasomes from normozoospermic and non-normozoospermic men show a differential protein expression pattern. *Andrology* 6, 585–596

26. Aitken, R. J., and Baker, M. A. (2013) Oxidative stress, spermatozoa and leukocytic infiltration: relationships forged by the opposing forces of microbial invasion and the search for perfection. *J. Reprod. Immunol.* 100, 11–19

27. Rodríguez-Martínez, H., Kvist, U., Ernerudh, J., Sanz, L., and Calvete, J. J. (2011) Seminal plasma proteins: What role do they play? *Am. J. Reprod. Immunol.* 66, 11–22

28. Robertson, S. A. (2005) Seminal plasma and male factor signalling in the female reproductive tract. *Cell Tissue Res.* 322, 43–52

29. Dorus, S., Skerget, S., and Karr, T. L. (2012) Proteomic discovery of diverse immunity molecules in mammalian spermatozoa. *Syst. Biol. Reprod. Med.* 58, 218–228

30. Borzía, K., Álvarez-Fernández, A., L Karr, T., Pizzari, T., and Dorus, S. (2016) The Seminal fluid proteome of the polyandrous Red junglefowl offers insights into the molecular basis of fertility, reproductive ageing and domestication. *Sci. Rep.* 6, 35864

31. Intasqui, P., Camargo, M., Antoniassi, M. P., Cedenho, A. P., Carvalho, V. M., Cardozo, K. H. M., Zylbersztejn, D. S., and Bertolla, R. P. (2016) Association between the seminal plasma proteome and sperm functional traits. *Fertil. Steril.* 105, 617–628

32. Antoniassi, M. P., Intasqui, P., Camargo, M., Zylbersztejn, D. S., Carvalho, V. M., Cardozo, K. H. M., and Bertolla, R. P. (2016) Analysis of the functional aspects and seminal plasma proteomic profile of sperm from smokers. *BJU Int.* 118, 814–822

33. Wang, J., Wang, J., Zhang, H.-R., Shi, H.-J., Ma, D., Zhao, H.-X., Lin, B., and Li, R.-S. (2009) Proteomic analysis of seminal plasma from asthenozoospermia patients reveals proteins that affect oxidative stress responses and semen quality. *Asian J. Androl.* 11, 484–491

34. Yamakawa, K., Yoshida, K., Nishikawa, H., Kato, T., and Iwamoto, T. (2007) Comparative analysis of interindividual variations in the seminal plasma proteome of fertile men with identification of potential markers for azoospermia in infertile patients. *J. Androl.* 28, 858–865

35. Fung, K. Y. C. C., Glode, L. M., Green, S., and Duncan, M. W. (2004) A comprehensive characterization of the peptide and protein constituents of human seminal fluid. *Prostate* 61, 171–181

36. Davalieva, K., Kiprijanovska, S., Niveski, P., Plaseski, T., Kocevska, B., Broussard, C., and Plaseska-Karanfilska, D. (2012) Proteomic analysis of seminal plasma in men with different spermatogenic impairment. *Andrologia* 44, 256–264

37. Vilagran, I., Yeste, M., Sancho, S., Castillo, J., Oliva, R., and Bonet, S. (2015) Comparative analysis of boar seminal plasma proteome from different freezability ejaculates and identification of Fibronectin 1 as sperm freezability marker. *Andrology* 3, 345–356

38. da Silva, B. F., Meng, C., Helm, D., Pachl, F., Schiller, J., Ibrahim, E., Lynne, C. M., Brackett, N. L., Bertolla, R. P., and Kuster, B. (2016) Towards Understanding Male Infertility After Spinal Cord Injury Using Quantitative Proteomics. *Mol. Cell. Proteomics* 15, 1424–1434

39. Pilch, B., and Mann, M. (2006) Large-scale and high-confidence proteomic analysis of human seminal plasma. *Genome Biol.* 7, R40

40. Saraswat, M., Joenväärä, S., Jain, T., Tomar, A. K., Sinha, A., Singh, S., Yadav, S., and
Renkonen, R. (2017) Human Spermatozoa Quantitative Proteomic Signature Classifies Normo-
and Asthenozoospermia. *Mol. Cell. Proteomics* 16, 57–72

41. Herwig, R., Knoll, C., Planyavsky, M., Pourbiabany, A., Greilberger, J., and Bennett, K. L. (2013) Proteomic analysis of seminal plasma from infertile patients with oligoasthenoteratozoospermia due to oxidative stress and comparison with fertile volunteers. *Fertil. Steril.* 100, 355–66.e2

42. Intasqui, P., Antoniassi, M. P., Camargo, M., Nichi, M., Carvalho, V. M., Cardozo, K. H. M., Zylbersztejn, D. S., and Bertolla, R. P. (2015) Differences in the seminal plasma proteome are associated with oxidative stress levels in men with normal semen parameters. *Fertil. Steril.* 104, 292–301

43. Giacomini, E., Ura, B., Giolo, E., Luppi, S., Martinelli, M., Garcia, R. C., and Ricci, G. (2015) Comparative analysis of the seminal plasma proteomes of oligoasthenozoospermic and normozoospermic men. *Reprod. Biomed. Online* 30, 522–531

44. Agarwal, A., Ayaz, A., Samanta, L., Sharma, R., Assidi, M., Abuzenadah, A. M., and Sabanegh, E. (2015) Comparative proteomic network signatures in seminal plasma of infertile men as a function of reactive oxygen species. *Clin. Proteomics* 12, 23

45. Bianchi, L., Carnemolla, C., Viviani, V., Landi, C., Pavone, V., Luzzi, A., Piomboni, P., and Bini, L. (2018) Soluble protein fraction of human seminal plasma. *J. Proteomics* 174, 85–100

46. Sharma, R., Agarwal, A., Mohanty, G., Jesudasan, R., Gopalan, B., Willard, B., Yadav, S. P., and Sabanegh, E. (2013) Functional proteomic analysis of seminal plasma proteins in men with various semen parameters. *Reprod. Biol. Endocrinol.* 11, 38

47. Sharma, R., Agarwal, A., Mohanty, G., Du Plessis, S. S., Gopalan, B., Willard, B., Yadav, S. P., and Sabanegh, E. (2013) Proteomic analysis of seminal fluid from men exhibiting oxidative stress. *Reprod. Biol. Endocrinol.* 11, 85

48. Intasqui, P., Camargo, M., Del Giudice, P. T., Spaine, D. M., Carvalho, V. M., Cardozo, K. H. M., Zylbersztejn, D. S., and Bertolla, R. P. (2013) Sperm nuclear DNA fragmentation rate is associated with differential protein expression and enriched functions in human seminal plasma. *BJU Int.* 112, 835–843

49. Milardi, D., Grande, G., Vincenzoni, F., Messana, I., Pontecorvi, A., De Marinis, L., Castagnola, M., and Marana, R. (2012) Proteomic approach in the identification of fertility pattern in seminal plasma of fertile men. *Fertil. Steril.* 97, 67–73

50. Drabovich, A. P., Dimitromanolakis, A., Saraon, P., Soosaipillai, A., Batruch, I., Mullen, B., Jarvi, K., and Diamandis, E. P. (2013) Differential Diagnosis of Azoospermia with Proteomic Biomarkers ECM1 and TEX101 Quantified in Seminal Plasma. *Sci. Transl. Med.* 5, 212ra160

51. Batruch, I., Smith, C. R., Mullen, B. J., Grober, E., Lo, K. C., Diamandis, E. P., and Jarvi, K. A. (2012) Analysis of seminal plasma from patients with non-obstructive azoospermia and identification of candidate biomarkers of male infertility. *J. Proteome Res.* 11, 1503–1511

52. Drabovich, A. P., Jarvi, K., and Diamandis, E. P. (2011) Verification of male infertility biomarkers in seminal plasma by multiplex selected reaction monitoring assay. *Mol. Cell. Proteomics* 10, M110.004127

53. Starita-Geribaldi, M., Roux, F., Garin, J., Chevallier, D., Fénichel, P., and Pointis, G. (2003) Development of narrow immobilized pH gradients covering one pH unit for human seminal plasma proteomic analysis. *Proteomics* 3, 1611–1619

54. Starita-Geribaldi, M., Poggioli, S., Zucchini, M., Garin, J., Chevallier, D., Fénichel, P., and Pointis, G. (2001) Mapping of seminal plasma proteins by two-dimensional gel electrophoresis in men with normal and impaired spermatogenesis. *Mol. Hum. Reprod.* 7, 715–22

55. Korbakis, D., Schiza, C., Brinc, D., Soosaipillai, A., Karakosta, T. D., Légaré, C., Sullivan, R., Mullen, B., Jarvi, K., Diamandis, E. P., and Drabovich, A. P. (2017) Preclinical evaluation of a
TEX101 protein ELISA test for the differential diagnosis of male infertility. *BMC Med.* 15, 60

57. Freour, T., Com, E., Barriere, P., Bouchot, O., Jean, M., Masson, D., and Pineau, C. (2013) Comparative proteomic analysis coupled with conventional protein assay as a strategy to identify predictors of successful testicular sperm extraction in patients with non-obstructive azoospermia. *Andrology* 1, 414–420

58. Bogle, O. A., Kumar, K., Attardo-Parrinello, C., Lewis, S. E. M., Estanyol, J. M., Ballescà, J. L., and Oliva, R. (2017) Identification of protein changes in human spermatozoa throughout the cryopreservation process. *Andrology* 5, 10–22

59. Lalancette, C., Platts, A. E., Johnson, G. D., Emery, B. R., Carrell, D. T., and Krawetz, S. A. (2009) Identification of human sperm transcripts as candidate markers of male fertility. *J. Mol. Med.* 87, 735–748

60. Mao, S., Goodrich, R. J., Hauser, R., Schrader, S. M., Chen, Z., and Krawetz, S. A. (2013) Evaluation of the effectiveness of semen storage and sperm purification methods for spermatozoa transcript profiling. *Syst. Biol. Reprod. Med.* 59, 287–295

61. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matise, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000) Gene Ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29

62. Uhlen, M., Oksvold, P., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A.-K., Odeberg, J., Djureinovic, D., Takanen, J. O., Hofer, S., Alm, T., Edqvist, P.-H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J. M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J., and Ponten, F. (2015) Tissue-based map of the human proteome. *Science* 347, 1260419

63. Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hofer, S., Wernerus, H., Björling, L., and Ponten, F. (2010) Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* 28, 1248–1250

64. R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

65. Del Giudice, P. T., Belardin, L. B., Camargo, M., Zylbersztejn, D. S., Carvalho, V. M., Cardozo, K. H. M., Bertola, R. P., and Cedenho, A. P. (2016) Determination of testicular function in adolescents with varicocele - a proteomics approach. *Andrology* 4, 447–455

66. Reyes-Moreno, C., Laflamme, J., Frenette, G., Sirard, M. A., and Sullivan, R. (2008) Spermatozoa modulate epididymal cell proliferation and protein secretion in vitro. *Mol. Reprod. Dev.* 75, 512–520

67. Garrett, J. E., Garrett, S. H., and Douglass, J. (1990) A spermatozoa-associated factor regulates proenkephalin gene expression in the rat epididymis. *Mol Endocrinol.* 4, 108–118

68. Jodar, M., Barrachina, F., and Oliva, R. (2017) The use of sperm proteomics in the assisted reproduction laboratory. *A practical guide to sperm analysis*, Garrido N, Rivera R eds., pp. 233-244, CRC Press, Boca Raton, FL

69. Codina, M., Estanyol, J. M., Fidalgo, M. J., Ballescà, J. L., and Oliva, R. (2015) Advances in sperm proteomics: best-practise methodology and clinical potential. *Expert Rev. Proteomics* 12, 255–277

70. Laflamme, B. A., and Wolfner, M. F. (2013) Identification and function of proteolysis regulators in seminal fluid. *Mol. Reprod. Dev.* 80, 80–101

71. Hallak, J., Sharma, R. K., Pasqualotto, F. F., Ranganathan, P., Thomas, A. J., and Agarwal, A. (2001) Creatine kinase as an indicator of sperm quality and maturity in men with oligospermia. *Urology* 58, 446–451

72. Chiu, P. C. N., Chung, M. K., Tsang, H. Y., Koistinen, R., Koistinen, H., Seppala, M., Lee, K. F., and Yeung, W. S. B. (2005) Glycodelin-S in human seminal plasma reduces cholesterol efflux and inhibits capacitation of spermatozoa. *J. Biol. Chem.* 280, 25580–25589
73. Lee, R. K.-K., Tseng, H.-C., Hwu, Y.-M., Fan, C.-C., Lin, M.-H., Yu, J.-J., Yeh, L.-Y., and Li, S.-H. (2018) Expression of cystatin C in the female reproductive tract and its effect on human sperm capacitation. Reprod. Biol. Endocrinol. 16, 8
74. Saewu, A., Kadunganattil, S., Raghupathy, R., Kongmanas, K., Astudillo, P. D., Hermo, L., and Tanphaichitr, N. (2017) Clusterin in the mouse epididymis: possible roles in sperm maturation and capacitation. Reproduction 154, 867–880
75. Maldera, J. A., Muñoz, M. W., Chirinos, M., Busso, D., Raffo, F. G. E., Battistone, M. A., Blaquier, J. A., Larrea, F., and Cuasnicu, P. S. (2014) Human fertilization: Epididymal hCRISP1 mediates sperm-zona pellucida binding through its interaction with ZP3. Mol. Hum. Reprod. 20, 341–349
76. Busso, D., Oñate-Alvarado, M. J., Balboa, E., Castro, J., Lizama, C., Morales, G., Vargas, S., Härtel, S., Moreno, R. D., and Zanlungo, S. (2014) Spermatozoa from mice deficient in Niemann-Pick disease type C2 (NPC2) protein have defective cholesterol content and reduced in vitro fertilising ability. Reprod. Fertil. Dev. 26, 609–621
77. Brillard-Bourdet, M., Réhault, S., Juliano, L., Ferrer, M., Moreau, T., and Gauthier, F. (2002) Amidolytic activity of prostatic acid phosphatase on human semenogelins and semenogelin-derived synthetic substrates. Eur. J. Biochem. 269, 390–395
78. Lilja, H. (1985) A kallikrein-like serine protease in prostatic fluid cleaves the predominant seminal vesicle protein. J. Clin. Invest. 76, 1899–1903
79. Chhikara, N., Saraswat, M., Tomar, A. K., Dey, S., Singh, S., and Yadav, S. (2012) Human Epididymis Protein-4 (HE-4): A Novel Cross-Class Protease Inhibitor. PLoS One 7, e47672
80. Frenette, G., Tremblay, R. R., and Dubé, J. Y. (1989) Zinc binding to major human seminal coagulum proteins. Syst. Biol. Reprod. Med. 23, 155–163
81. Romijn, J. C. (1990) Polyamines and transglutaminase actions: Polyamine and transglutaminase- Wirkungen. Andrologia 22, 83–91
82. Le, T. M., Le, Q. V. C., Truong, D. M., Lee, H. J., Choi, M. K., Cho, H., Chung, H. J., Kim, J. H., Do, J. T., Song, H., and Park, C. (2017) β2-microglobulin gene duplication in cetartiodactyla remains intact only in pigs and possibly confers selective advantage to the species. PLoS One 12, e0182322
83. Hannan, N. J., and Salamonsen, L. A. (2008) CX3CL1 and CCL14 regulate extracellular matrix and adhesion molecules in the trophoblast: potential roles in human embryo implantation. Biol. Reprod. 79, 58–65
84. Kumar, S., Tomar, A. K., Singh, S., Gill, K., Dey, S., Singh, S., and Yadav, S. (2014) Heparin binding carboxypeptidase E protein exhibits antibacterial activity in human semen. Int. J. Biol. Macromol. 64, 319–327
85. Michael, C. O., and Kistler, W. S. (1980) Properties of a Flavoprotein Sulphydryl Oxidase from Rat Seminal Vesicle Secretion. Biochemistry 19, 2639–2645
86. de Mateo, S., Martínez-Heredia, J., Estanyol, J. M., Domínguez-Fandos, D., Domínguez-Fandos, D., Vidal-Taboada, J. M., Ballescà, J. L., and Oliva, R. (2007) Marked correlations in protein expression identified by proteome analysis of human spermatozoa. Proteomics 7, 4264–4277
87. Wakabayashi, H., Matsumoto, H., Hashimoto, K., Teraguchi, S., Takase, M., and Hayasawa, H. (1999) Inhibition of Iron / Ascorbate-Induced Lipid Peroxidation by an N-Terminal Peptide of Bovine Lactoferrin and Its Acylated Derivatives. Biosci. Biotechnol. Biochem. 63, 955–957
88. Jodar, M., Sendler, E., and Krawetz, S. A. (2016) The protein and transcript profiles of human semen. Cell Tissue Res. 363, 85–96
89. Vizcaino, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Terten, T., Xu, Q.-W., Wang, R., and Hermjakob, H. (2016) 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. 44, D447-56
FIGURE LEGENDS

FIG. 1. Overall strategy used for the identification of proteomic alterations in different subtypes of infertile patients using TMT 10-plex peptide labeling followed by LC-MS/MS.

FIG. 2. Example of ECM1 and NPC2 as stable-correlated proteins in patients with normal seminal parameters (NZ; n=4). Pearson correlation coefficient ($r$) for ECM1 peptides (3 peptides) and NPC2 peptides (7 peptides) is equal or higher than 0.9 for all the possible peptide-correlations.

FIG. 3. Seminal plasma proteins detected with significant different abundance in different subtypes of infertile patients according to seminal parameters. Statistical comparisons are performed by ANOVA with Tukey's Post Hoc test (* $p < 0.05$; ** $p < 0.01$). Horizontal lines represent mean protein abundance for each group. NZ: Normozoospermic, AS: Asthenozoospermic, OZ: Oligozoospermic, AZ: Azoospermic. The detailed information is found in supplemental Table S6.

FIG. 4. Correlation between sperm concentration and seminal plasma ECM1 protein abundance. A, Scatterplot of sperm concentration and ECM1 protein abundance values generated by LC-MS/MS (Table I). The figure describes the significant Pearson correlation coefficient ($r$) and its p-value ($p$). B, Validation of ECM1 protein biomarker by Western blotting in seminal plasma proteins obtained from an independent set of samples (n=18; 6 normozoospermic (NZ), 6 oligozoospermic (OZ) and 6 azoospermic (AZ) patients). One representative example is provided for the results of the Western blotting in a patient of each type. C, Quantification of the results in all independent set of samples analyzed. Optical Densities of the bands obtained in ECM1 Western blotting are plotted as a function of the sperm concentration. Correlation coefficient ($r$) and p-value ($p$) are from Pearson correlation analysis.

FIG. 5. Protein expression pattern in the male reproductive tract. A, Protein expression data obtained from the HPA database of the seminal plasma proteins identified in the present study correlating with sperm concentration (n=5, Table I). Intensity values are based on the immunohistochemical antibody staining score information available in the HPA database. B, Results from the subsequent
immunohistochemical analysis in human testis (n=1) and epididymis (n=1) sections using ECM1 and NPC2 antibodies assayed in the present study. The obtained protein patterns validate the HPA data. ECM1 is an epididymis-specific protein whereas NPC2 is found in epididymis but also in testes. A lower magnification (×100; images on top) and a higher magnification (×400; images below) from the boxed area are shown. Negative controls with rabbit IgG showed non-specific staining (data not shown). H: Haematoxylin stain, PAS: Periodic acid-Schiff stain.
# TABLES

Table I. Pearson correlation analysis between sperm concentration and seminal plasma protein abundance (LC-MS/MS data). Pearson correlation coefficient, p-value and FDR-adjusted p-value (q-value) are indicated in the table.

| Gene name | Uniprot Accession number | Pearson Correlation (r-value) | Significance (p-value) | Significance FDR corrected (q-value) |
|-----------|---------------------------|------------------------------|------------------------|-------------------------------------|
| SPINT3    | P49223                    | 0.746                        | 0.001                  | 0.03                                |
| NPC2      | P61916                    | 0.680                        | 0.004                  | 0.06                                |
| ECM1      | Q16610                    | 0.683                        | 0.004                  | 0.06                                |
| CRISP1    | P54107                    | 0.643                        | 0.007                  | 0.08                                |
| IGHG2     | P01859                    | 0.542                        | 0.030                  | 0.30                                |
Table II. Pearson correlation analysis between sperm concentration and seminal plasma peptide abundance (LC-MS/MS data) for the 5 proteins identified as correlated in Table I. Pearson correlation coefficient and \( p \)-value are indicated in the table.

| Peptide sequence and modifications | Uniprot Accession number | Gene name | Pearson Correlation (r-value) | Significance (p-value) |
|-----------------------------------|--------------------------|-----------|-------------------------------|------------------------|
| dLLPNVcAFPMk N-Term (TMT10Plex); C7(Carboxamidomethyl); K13(TMT10Plex) | P49223 | SPINT3 | 0.746 | **0.001** |
| sGiNpcPIQk N-Term (TMT10Plex); C5(Carboxamidomethyl); K9(TMT10Plex) |  |  | 0.778 | **0.0004** |
| eVNVSPcPTQcQLSk N-Term (TMT10Plex); C7(Carboxamidomethyl); C12(Carboxamidomethyl); K16(TMT10Plex) |  |  | 0.753 | **0.001** |
| aVVHGILMGVPVFPPIEPDGc N-Term (TMT10Plex); C22(Carboxamidomethyl); K23(TMT10Plex) | P61916 | NPC2 | 0.752 | **0.001** |
| deGSVDGVlk N-Term (TMT10Plex); C2(Carboxamidomethyl); K10(TMT10Plex) |  |  | 0.692 | **0.003** |
| lVVEWQLQDDk N-Term (TMT10Plex); K11(TMT10Plex) |  |  | 0.595 | **0.015** |
| nQLFcWEIIPQSHL N-Term (TMT10Plex); C6(Carboxamidomethyl) |  |  | 0.549 | **0.028** |
| cLPSLQHPNEQk N-Term (TMT10Plex); K12(TMT10Plex) | Q16610 | ECM1 | 0.377 | **0.149** |
| eVGPPLQEAEPk N-Term (TMT10Plex); K15(TMT10Plex) |  |  | 0.377 | **0.149** |
| lDGFPGRPSDNLNQcLPNR N-Term (TMT10Plex); C18(Carboxamidomethyl) | P54107 | CRISP1 | 0.699 | **0.003** |
| tGVPeAcPSNcEDk N-Term (TMT10Plex); C5(Carboxamidomethyl); C8(Carboxamidomethyl); C12(Carboxamidomethyl); K15(TMT10Plex) |  |  | 0.411 | **0.113** |
| lVTDLPNQEEIVNIHNLAR N-Term (TMT10Plex) |  |  | 0.411 | **0.113** |
| sTSESTAALGcLVk N-Term (TMT10Plex); C11(Carboxamidomethyl); K14(TMT10Plex) | P01859 | IGHG2 | 0.319 | 0.228 |
Table III. Proteins involved in the stable-protein pairs for the different subtypes of infertile patients according to their seminal parameters. The correlation of two proteins is counted as a single correlation.

Color code refers to protein function: orange, sperm function; green, semen clotting-liquefaction; blue, immune system; purple, other functions.

| Proteins      | Normozoospermic (NZ1-4) | Asthenozoospermic (AS1-4) | Oligozoospermic (OZ1-4) | Azoospermic (AZ1-4) |
|---------------|-------------------------|---------------------------|-------------------------|---------------------|
| CRISP1        | 21                      | 4                         | -                       | 1                   |
| NPC2          | 19                      | -                         | -                       | -                   |
| CST3          | 17                      | 3                         | -                       | -                   |
| PAEP          | 8                       | 5                         | -                       | 1                   |
| CKB           | 5                       | -                         | -                       | -                   |
| CLU           | 3                       | -                         | -                       | -                   |
| MME           | 19                      | 1                         | -                       | -                   |
| ACPP          | 17                      | 1                         | -                       | -                   |
| KLK3          | 17                      | 1                         | -                       | -                   |
| WFDC2         | 14                      | -                         | -                       | -                   |
| ALB           | 12                      | -                         | -                       | -                   |
| TGM4          | 8                       | 1                         | -                       | -                   |
| SERPINA5      | -                       | 1                         | -                       | -                   |
| SERPINA1      | -                       | -                         | -                       | 2                   |
| IGKC          | 21                      | 1                         | -                       | -                   |
| ANPEP         | 20                      | 1                         | -                       | -                   |
| IGHG2         | 20                      | 2                         | -                       | 1                   |
| ECM1          | 18                      | -                         | -                       | -                   |
| QSOX1         | 17                      | -                         | -                       | -                   |
| LGALS3BP      | 16                      | -                         | -                       | -                   |
| CPE           | 16                      | 1                         | -                       | -                   |
| B2M           | 16                      | 3                         | -                       | 1                   |
| AZGP1         | 20                      | 3                         | -                       | -                   |
| TF            | 14                      | -                         | -                       | -                   |
| IDH1          | 14                      | 5                         | -                       | -                   |
| VWA1          | 12                      | 3                         | -                       | -                   |
| Total Proteins involved | 24          | 16                        | 0                       | 5                   |
| Correlations  | 182                     | 18                        | 0                       | 3                   |
Table IV. Stable-protein correlations in control population (normal sperm parameters, NZ; n=4) altered in individual infertile patients. Color code refers to protein function: orange, sperm function; green, semen clotting-liquefaction; blue, immune system; purple, other functions.

| Samples                | Number of correlations | Correlations in common with NZ stable-protein pairs group | Proteins involved | Proteins that lost > 75% of correlations |
|------------------------|------------------------|----------------------------------------------------------|-------------------|----------------------------------------|
| **Stable-protein pairs for the NZ group (Control group; n=4)** |                        |                                                          |                   |                                        |
| NZ group               | 182                    | 24                                                       |                   |                                        |
| **Alterations in individual infertile patients** |                        |                                                          |                   |                                        |
| NZ group + AS1         | 118                    | 116                                                      | 22                | PAEP, CKB, CLU, WFDC2, ECM1, TF        |
| NZ group + AS2         | 142                    | 138                                                      | 24                | PAEP, CLU, WFDC2, ALB                  |
| NZ group + AS3         | 101                    | 99                                                       | 21                | PAEP, CLU, WFDC2, ALB                  |
| NZ group + AS4         | 95                     | 94                                                       | 22                | PAEP, CLU, WFDC2, TGM4, ANPEP, VWA1   |
| NZ group + OZ1         | 98                     | 96                                                       | 20                | PAEP, CKB, CLU, WFDC2, ECM1, QSOX1, IDH1 |
| NZ group + OZ2         | 98                     | 96                                                       | 22                | PAEP, CKB, ACPP, KLK3, WVA1           |
| NZ group + OZ3         | 88                     | 88                                                       | 22                | PAEP, CLU, WFDC2                      |
| NZ group + OZ4         | 115                    | 113                                                      | 22                | CST3, PAEP, CLU, WFDC2, LGALS3BP, VWA1 |
| NZ group + AZ1         | 89                     | 85                                                       | 22                | CKB, MME, ACPP, ANPEP, CPE            |
| NZ group + AZ2         | 92                     | 91                                                       | 19                | PAEP, CKB, CLU, ACPP, KLK3, QSOX1     |
| NZ group + AZ3         | 133                    | 127                                                      | 23                | CLU, ACPP, B2M, VWA1                  |
| NZ group + AZ4         | 87                     | 85                                                       | 23                | PAEP, CLU, WFDC2, ALB, IGKC, QSOX1, VWA1 |

NZ, normozoospermic; AS, asthenozoospermic; OZ, oligozoospermic; AZ, azoospermic
FIG. 1. Human semen samples (n=16) 
- 4 Normozoospermic (NZ)
- 4 Oligozoospermic (OZ)
- 4 Asthenozoospermic (AS)
- 4 Azospermic (AZ)

Centrifugation at 500 g, 1500 g and 0.45 μm membrane filtration

SEMINAL PLASMA

Protein quantification, denaturation, reduction, alkylation and tryptic digestion

Differential peptide labeling (TMT 10-plex) and pooling

Control  NZ1  AS1  OZ1  AZ1
TMT 126  TMT 127C  TMT 129N  TMT 128C  TMT 129C

N22  AS2  O22  AZ2
TMT 129N  TMT 130C  TMT 130N  TMT 131

Control  OZ3  AZ3  N23  AS3
TMT 126  TMT 127C  TMT 129N  TMT 128C  TMT 129C

O42  AZ4  N42  AS4
TMT 129N  TMT 130C  TMT 130N  TMT 131

POOL A

POOL B

LC-MS/MS

Protein identification and quantification

Statistical analyses of quantitative proteomic data

Conventional approaches on relative proteomic quantification data

New approaches

ANOVA + Tukey’s Post-Hoc test

Stable-protein pairs analysis

Pearson correlation (for sperm concentration)

Groups of samples (same phenotype)

Normozoospermic–control samples + individual patients

Proteins correlated at peptide level to sperm concentration by Pearson correlation

Homogeneity/heterogeneity of groups

Altered proteins or pathways in individual patients

Identification of stable conserved proteins/pathways in normal seminal plasma and altered proteins and proteome signatures in infertile patients
FIG. 2.

ECM1 protein

NPC2 protein
FIG. 3.

CRISP1

Protein abundance

NPC2

Protein abundance

IGHG2

Protein abundance

SPINT3

Protein abundance

SCPEP1

Protein abundance

ANPEP

Protein abundance

FIG. 4.

A

ECM1

Protein abundance

Sperm concentration

\[ r = 0.683 \]

\[ p = 0.004 \]

B

ECM1

(76kDa)

C

ECM1 Optical Density

Sperm concentration

\[ r = 0.620 \]

\[ p = 0.0061 \]
FIG. 5.

A

| Gene name | Uniprot Accession number | Protein expression overview (HPA) |
|-----------|--------------------------|-----------------------------------|
|           |                          | Testis   | Epididymis | Seminal vesicles | Prostate |
| **Extra-testicular proteins (n=3)** | | | | | |
| ECM1      | Q16610                   | -        | High       | Low             | -        |
| CRISP1    | P54107                   | -        | High       | -               | -        |
| SPINT3    | P49223                   | -        | High       | -               | -        |
| **Testicular/ Unknown origin (n=2)** | | | | | |
| NPC2      | P61916                   | High     | High       | High            | High     |
| IGHG2     | P01859                   | NA       | NA         | NA              | NA       |

*: not detected, NA: not available

B

Validation of the HPA data

| ECM1 | NPC2 |
|------|------|
| ![Epididymis](H) | ![Epididymis](H) |
| ![Testis](H + PAS) | ![Testis](H + PAS) |