Alterations in the abundance of protamine proteoforms related to sperm chromatin packaging, obesity and age in normozoospermic men.

Castillo J, Gay M, De La Iglesia A, Arauz-Garofalo G, Vilanova M, Leiva M, Corral JM, Guimerà M, Manau D, Vilaseca M, Jodar M, Oliva R.

Mol Hum Reprod. 2025 May 24:gaaf019.

doi: 10.1093/molehr/gaaf019. Online ahead of print. PMID: 40411759

Alterations in the abundance of protamine proteoforms related to sperm chromatin packaging, obesity and age in normozoospermic men

Running title: Protamine proteoform alterations in human sperm

Judit Castillo[1,†,*], Marina Gay[2,†], Alberto de la Iglesia[1,3,†], Gianluca Arauz-Garofalo[2], Mar Vilanova[2], Marina Leiva[1], Juan Manuel Corral[4], Marta Guimerà[5], Dolors Manau[5,6], Marta Vilaseca[2], Meritxell Jodar[1,7,*], Rafael Oliva[1,7,*]

[1] Molecular Biology of Reproduction and Development Research Group, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Fundació de Recerca Clínic Barcelona (FRCB), Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona (UB), Barcelona, Spain.

[2] Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology (BIST), Barcelona, Spain.

[3] Current address: Université Paris Cité, INSERM, CNRS, Institut Cochin, Paris, France.

[4] Urology Service, Andrology Unit, Nephrology and Urology Clinical Institute (ICNU), Hospital Clínic Barcelona, Barcelona, Spain.

[5] Gynecology, Obstetrics and Neonatology Clinical Institute (ICGON), Hospital Clínic Barcelona, Barcelona, Spain.

[6] Gynecology, Human Reproduction and Women´s Health Research Group, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain.

[7] Biochemistry and Molecular Genetics Service, Hospital Clínic Barcelona, Barcelona, Spain.

[†] These authors contributed equally to this work.

[*] Corresponding authors: Judit Castillo (ORCID 0000-0002-9407-2675; juditcastillo@ub.edu), Rafael Oliva (ORCID 0000-0003-4876-2410; roliva@ub.edu), and Meritxell Jodar (ORCID 0000-0002-3272-0163; meritxell.jodar@ub.edu).

Molecular Biology of Reproduction and Development Research Group, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Fundació de Recerca Clínic Barcelona (FRCB), Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona (UB), Casanova 143, 08036, Barcelona, Spain.

ABSTRACT

Protamines are considered among the most relevant sperm proteins because of their functional implications on paternal genome packaging and protection. Although the proteomic evaluation of protamines is technically challenging, mass spectrometry-based studies have shown a complex population of protamine proteoforms in the human sperm. This included intact, truncated and modified forms for protamine 1 (P1) and mature and immature components of protamine 2 family (P2). However, it is still unknown whether global or specific protamine proteoforms levels may be unbalanced under conditions that may impair paternal chromatin maturity and epigenetic information. In this study, protamines from normozoospermic men stratified according to body mass index, age and chromatin maturity (assessed through the P1/P2 ratio derived from acid-urea electrophoresis) were evaluated using a refined top-down mass spectrometry protocol for protamine proteoform quantification and comparative analysis. Accumulation of the P2 immature forms HPS1 and HPI2 was significantly associated with abnormally high P1/P2 ratios, suggesting either impaired eviction of P2 immature forms or defective P2 processing during spermatogenesis in these men clinically classified as normozoospermic. When considering weight and age as factors, P1 was the only affected protamine. Sperm from obese men, which were found to be exposed to high levels of oxidative damage derived from lipid peroxidation, showed mass shift(s) of +61 Da from the unmodified P1 protein sequence. Men of advanced age showed a specific loss of diphosphorylated P1, mainly on Ser 11 and 22. Our results allow the hypothesis that protamine proteoforms in the male gamete act as additional layers of epigenetic information, the alteration of which might be related to some cases of impaired sperm function.

Key words: human sperm / protamines / normozoospermia / age / obesity / chromatin / mass spectrometry / proteoforms / protein modifications / epigenetics

INTRODUCTION

Protamines are considered among the most relevant mature sperm nuclear proteins due to their highly specific function of chromatin compaction into semi-crystalline structures. This enables paternal genome protection from external damage and nuclear streamlining for a proper sperm functionality (Oliva, 2006; Balhorn, 2007; Oliva and Castillo, 2011; Barrachina *et al.*, 2018). Protamines have short and characteristic amino acid sequences, which are extremely arginine-rich and contain cysteine residues, resulting in highly basic (positively charged) proteins with the potential to form inter- and intra-molecular disulfide bonds. These unique physicochemical properties lead to an extremely tight genome packaging into structural units that can wrap ~250-fold more DNA than nucleosomes (Oliva and Dixon, 1991; Oliva, 2006; Balhorn, 2007; Castillo *et al.*, 2014).

Although protamines have a highly specific function in sperm genome packaging that is well conserved, their protein sequence and isoform diversity are quite variable among mammals (Queralt *et al.*, 1995; Oliva, 2006; Kasinsky *et al.*, 2011; Jodar and Oliva, 2014; Luke *et al.*, 2016; de la Iglesia *et al.*, 2023). In fact, a complex protamine proteoforms landscape has been described in human and mouse sperm by our group and others (Brunner *et al.*, 2014; Castillo *et al.*, 2015; Soler-Ventura *et al.*, 2020; Arauz-Garofalo *et al.*, 2021; Schon *et al.*, 2023). In humans, sperm contain two

types of protamines: protamine 1 (P1), translated as mature protein from a single gene (*PRM1*), and protamine 2 family (P2), translated from *PRM2* gene as a precursor (pre-P2) that is then proteolyzed into three mature components: HP2, HP3 and HP4, with HP2 being the most abundant. Additionally, truncated forms of P1 and pre-P2, as well as additional P2 immature proteoforms (HPI1, HPI2, HPS1, HPS2), with still unknown function in the mature sperm, have been identified (de Yebra *et al.*, 1998; de Mateo *et al.*, 2011; Soler-Ventura *et al.*, 2020; Arauz-Garofalo *et al.*, 2021). This highly diverse profile of protamine proteoforms raises the question of whether they have a redundant function in chromatin compaction, or whether there are some proteoform-specific functional implications, the deregulation of which would impact the information delivered to the oocyte.

Moreover, an additional level of complexity in human protamine proteoform profile is incorporated, considering that protamines also bear post-translational modifications (PTMs), such as phosphorylation (Pruslin *et al.*, 1987; Chirat *et al.*, 1993; Pirhonen *et al.*, 1994; Brunner *et al.*, 2014; Castillo *et al.*, 2015; Soler-Ventura *et al.*, 2020; Arauz-Garofalo *et al.*, 2021; Moritz *et al.*, 2023; Schon *et al.*, 2023). By adapting mass spectrometry (MS) approaches and proteoform identification algorithms to the challenging characteristics of protamines, our group has described a complete protamine phosphorylation profile in the normal and mature human sperm (Soler-Ventura *et al.*, 2020). Besides the well-established need of phosphate groups for the correct incorporation of protamines into the chromatin of testicular spermatids during spermiogenesis (Willmitzer *et al.*, 1977; Papoutsopoulou *et al.*, 1999; Wu *et al.*, 2000; Gou *et al.*, 2020), the presence of residual phosphorylated amino acids in the ejaculated mature sperm suggest implications towards sperm function, oocyte fertilization and beyond (de la Iglesia *et al.*, 2023).

Indeed, a role for protamine phosphorylation has been demonstrated in paternal nuclear reprogramming in the zygote (Gou *et al.*, 2020). Although it must be noted that residue-specific identification of protamine phosphorylation is a challenging task, some studies have managed to characterize phosphorylation sites in human sperm protamines, such as S9 and S11 in P1, and S37 and S59 in P2 (Chirat *et al.*, 1993; Hornbeck *et al.*, 2015; Soler-Ventura *et al.*, 2020). P1S9 and P2S59 have been also identified in mouse homolog sequences, which raises interest in deciphering a functional role in male gamete (Brunner *et al.*, 2014).

Relative protein amounts of P1 and P2 (expressed as P1/P2 ratio calculated from the optical density of stained bands obtained by acid-urea polyacrylamide gel electrophoresis, PAGE) have been traditionally used as marker of chromatin maturity. While similar abundances of P1 and P2 at protein level (P1/P2 ~1, range 0.8–1.2) are related to normality, an altered P1/P2 ratio has been associated with impaired sperm count, motility and morphology, as well as compromised chromatin integrity and lower assisted reproduction technologies (ART) success rate (Balhorn *et al.*, 1988; de Yebra *et al.*, 1993; Mengual *et al.*, 2003; Hammoud *et al.*, 2009; Nanassy *et al.*, 2011). Detailed implications of protamine content in human male fertility have been reviewed elsewhere (Jodar and Oliva, 2014; Ni *et al.*, 2016; Soler-Ventura *et al.*, 2018; de la Iglesia *et al.*, 2023). Apart from chromatin maturity defects, other potential factors that may affect the male gamete chromatin content and, thus, sperm contribution to embryo development, are under investigation. This is the case for sociodemographic factors, including increased obesity rates measured through the body mass index (BMI) or advanced paternal age at the time of conception (Mascarenhas *et al.*, 2012; Tiegs *et al.*, 2018; Agarwal *et al.*, 2021). Obesity has been described to affect spermatogenesis at many levels, with impacts on the sperm epigenome that can be transmitted to

subsequent generations and affect newborn's health (Fullston *et al.*, 2012, 2013; Soubry *et al.*, 2013, 2016; McPherson *et al.*, 2014; Craig *et al.*, 2017; Pepin *et al.*, 2022). Despite presenting normal seminal parameters, sperm from obese men carry altered DNA methylation and small-noncoding RNAs (Donkin *et al.*, 2016; Soubry *et al.*, 2016). Also, obese mice show impaired histone marks, histone-modifying enzymes, and histone positioning within the genome (Palmer *et al.*, 2011; Terashima *et al.*, 2015; Deshpande *et al.*, 2021). In turn, advanced paternal age is associated with rise in de novo mutations, telomeric elongation, increased DNA fragmentation, and an altered methylation pattern (Jenkins *et al.*, 2014, 2018; Laurentino *et al.*, 2020; Pohl *et al.*, 2021; Wood and Goriely, 2022). Transgenerational effects of age-associated DNA methylation impairments have been also described in mouse (Milekic *et al.*, 2015).

Despite all the above knowledge, there is a current lack of quantitative data of protamine proteoforms levels using mass spectrometry in males with normal semen parameters (normozoospermic) and differing characteristics regarding protamine content, weight and age. This information would provide insights on the sperm contribution to the oocyte and propose specific chromatin defects to be validated as potential causes of unexplained infertility. For that reason, in this study, we have refined the top-down MS and the density-based spatial clustering of applications with noise (DBSCAN) algorithm developed by our group (Arauz-Garofalo *et al.*, 2021), to identify quantitative alterations in the protamine proteoform profile of normozoospermic men from different cohorts. Furthermore, we used the derived proteomic data to assign phosphorylated residues. With the aim of overcoming potential bias from associated co-morbidities, a strict inclusion criterion was followed resulting in highly homogeneous groups of men differing in a) acid-urea PAGE-based P1/P2 ratio, b) BMI and c) male age, from the control group. Therefore, we report

herein altered levels of specific protamine proteoforms essentially attributed to impaired chromatin status, obesity and advanced paternal age. These hypothesis-generating results open a window to the evaluation of changes in sperm protamine proteoforms abundance in men with unknown infertility, as epigenetic alterations with potential impacts on oocyte fertilization and beyond.

MATERIALS AND METHODS

Biological materials and groups of study

A schematic summary of the study workflow is shown in Figure 1. Human semen samples from men undergoing routine semen analysis were obtained at the Assisted Reproduction Unit from the Clinic Institute of Gynecology, Obstetrics and Neonatology, at the hospital Clínic Barcelona, Spain. The ejaculates were collected by masturbation into sterile containers after 3–5 days of sexual abstinence. Patients reporting a known infertility factor and associated diseases or toxic habits, such as smoking, were excluded. After evaluation of seminal parameters using the automatic semen analysis system CASA (Proiser, Paterna, Spain), only samples classified as normozoospermic according to the latest update of World Health Organization guidelines (World Health Organization, 2021) were selected for the study (n = 28). Samples were classified according to chromatin quality (in terms of P1/P2 ratio calculated by acid-urea PAGE), BMI and age to establish the following groups of study: a “Control” group formed by males with sperm showing a normal P1/P2 ratio <1.2, no obesity nor overweight (BMI < 25), and less than 45 years old (n = 11); an “Altered P1/P2” group formed by males whose spermatozoa show a P1/P2 ratio > 1.2, and with

BMI < 25 and age < 45 years old (n = 3); an "Obesity" group formed by males whose spermatozoa show normal P1/P2 ratio <1.2 and with BMI > 30 and age < 45 years old (n = 10); and an "Advanced age" group formed by males whose spermatozoa show normal P1/P2 ratio <1.2, and with BMI < 25 and age > 45 years old (n = 4). The males included in the "Control" group have fathered at least one child either by natural conception or through ART after sample collection. Using a one-way ANOVA test and Tukey's multiple comparison correction, it was ensured that the groups of study significantly differed from each other in the single parameter that characterized the cohort.

Sperm samples were purified from any contaminating round cell through 50% density gradient using Puresperm® (NidaCon International AB, Gothenburg, Sweden), according to manufacturer's recommendations.

Ethical approval for the use of human samples

All individuals included in this study signed informed consent in accordance with the Declaration of Helsinki. Demographic and clinical data were collected, cleaned and de-identified under the frame of the hospital Clínic Barcelona. Subsequently, a codified study identification number was assigned to each sample for research purposes. Samples were used following the appropriate ethical guidelines and Internal Review Board, and the Clinical Research Ethics Committee of the hospital Clínic Barcelona approved the biological material storing and processing (Spain; International Review Board study number R121031-096-2012/7942, granted on 22nd November 2012).

Purification of protamines from human semen samples and quantification

The protamine-rich fraction of sperm nuclear proteins was extracted as previously described by our group (Soler-Ventura *et al.*, 2018, 2020). Briefly, histones and other basic proteins loosely attached to DNA were removed from sperm nuclei incubating the cells in 0.5 M HCl at 37°C and centrifuged three times at 2000 g for 20 min at 4°C. The histone-enriched supernatant was kept as quality control of the protocol and the pellet was washed in 0.5 % Triton X-100, 20 mM Tris-HCl (pH 8), and 2 mM $MgCl_2$ and centrifuged at 8940 g for 5 min at 4°C. Cells were subsequently resuspended in miliQ H_2O with 1 mM PMSF and centrifuged at 8940 g for 5 min at 4°C. Chromatin was then denatured with 575 mM DTT in 6 M GuHCl and precipitated by incubation with cold ethanol at –20°C and centrifuged at 12880 g for 15 min at 4°C. Basic nuclear proteins were extracted from DNA incubating with 0.5 M HCl at 37°C and centrifuged at 17530 g for 10 min at 4°C. Protamine precipitation was carried out with 20 % TCA on ice and centrifuged at 17530 g for 10 min at 4°C. Protamines were washed twice with 1 % β-mercaptoethanol in acetone, and centrifuged at 17530 g for 5 min at 4°C, then dried out at room temperature.

Protamine extracts were quantified according to published procedures of our group (Soler-Ventura *et al.*, 2018). Dried purified extracts were resuspended in loading buffer (5.5 M urea, 20% β-mercaptoethanol, 5% acetic acid) and separated using acid-urea PAGE together with increasing amounts of protamine standard (Mengual *et al.*, 2003). Gels were stained with EZBlue™ Gel Staining Reagent (Sigma-Aldrich) and intensity of the two main bands, corresponding to P1 and P2, was quantified using Quantity One 1-D analysis software (BioRad, Hercules, CA, USA). A regression curve, obtained from protein standard intensity determination, allowed calculating the quantity of P1 and P2 for each sample. Final quantification of each protamine sample was obtained by

205 calculating the mean of at least three technical replicates conducted in different acid-urea PAGE.
206 Acid-urea PAGE P1/P2 ratios were calculated with the purpose of classifying samples in the
207 different groups of study.

208 Protein extracts corresponding to histones and other basic proteins loosely attached to DNA
209 that were removed during the protamine purification protocol were visualized in an acid-urea
210 PAGE. Global optical density of histone bands was quantified using Quantity One 1-D analysis
211 software and related to global optical density of main protamine bands (histone/protamine ratio).
212

213 **Characterisation of protamine proteoforms by top-down MS-based proteomics**

214 Purified protamine extracts from 16 samples were evaluated by top-down MS (5 “Control”,
215 3 “Altered P1/P2”, 4 “Obesity” and 4 “Advance age”). Protamine-enriched fractions were
216 reconstituted in 50 mM NH_4HCO_3, reduced with 2 mM DTT for 1h at room temperature (RT) and
217 carbamidomethylated for 30 min in the dark at RT with 5 mM iodoacetamide. DTT was added to a
218 final concentration of 2mM to consume any unreacted iodoacetamide. Protamines were further
219 desalted using PolyLC C18 filter tips (PolyLC Inc., Columbia, MD, USA), eluted with 80% acetonitrile
220 and 1% trifluoroacetic acid, evaporated to dryness, and reconstituted in aqueous solution of 3 %
221 acetonitrile and 1 % formic acid for nanoLC-MS/MS system injection. Samples were loaded to a 300
222 µm × 5 mm PepMap100, 5 µm, 100 Å, C18 µ-precolumn (Thermo Scientific) at a flow rate of 15
223 µl/min using a Dionex Ultimate 3000 chromatographic system (Thermo Scientific). Proteins were
224 separated using a C18 analytical column Acclaim PepMap® RSLC (75 µm × 50 cm, nanoViper, C18,
225 2µm, 100 Å) (Thermo Scientific) with a 60 min run, comprising two consecutive steps with linear
226 gradients from 3 to 15% B in 30 min and from 15 to 85% B in 5 min, followed by isocratic elution at

85% B in 5 min and stabilization to initial conditions (A = 0.1% FA in water, B = 0.1% FA in CH_3CN). The column outlet was directly connected to an Advion TriVersa NanoMate (Advion) fitted on an Orbitrap Fusion Lumos™ Tribrid (Thermo Scientific). The mass spectrometer was operated in a data-dependent acquisition (DDA) mode. Survey MS scans were acquired in the Orbitrap with the resolution (defined at 200 m/z) set to 120,000. The lock mass was user-defined at 445.12 m/z in each Orbitrap scan. The top speed (most intense) ions per scan were fragmented by ETD and detected in the orbitrap. The ion count target value was 400,000 and 1,000,000 for the survey scan and for the MS/MS scan respectively. Target ions already selected for MS/MS were dynamically excluded for 30s. Spray voltage in the NanoMate source was set to 1.60 kV. The RF Lenses were tuned to 30%. The minimal signal required to trigger MS to MS/MS switch was set to 50,000. The spectrometer was working in positive polarity mode and singly charge state precursors were rejected for fragmentation. Between two and six technical replicates were analyzed per biological sample.

Protamine proteoform search, data analysis, and proteoform annotation

Top-down proteomics workflow was based on a twin search strategy, encompassing the bioinformatics tools ProSight PD (LeDuc *et al.*, 2004) (PS; Thermo Scientific) and TopPIC suite (Kou *et al.*, 2016) (TP; Indianapolis, IN, USA), to identify proteoform spectrum matches (PrSMs). Proteome Discoverer v2.5.0.400 (Thermo Fisher Scientific) with PS v4.0 and TP v1.4.8 was used for protamine identification and PTM assignment. Database preparation for PS searches began from an xml-annotated database from SwissProt containing P1 and P2, P2-2 and P2-3 entries. This xml database was then translated to psdb format, through the PS database tool included in Proteome Discoverer,

to finally enable the PS search. The annotated database was modified to include phosphorylation modification to all STY residues, and oxidation in M. A 3-tier search (Absolute, Biomarker and Absolute) was performed with precursor mass tolerances of 2.2 Da / 10 ppm / 500 Da, respectively, and fragment mass tolerances of 10 ppm. Searches in TP were done using TopFD tool for top-down spectral deconvolution and a FASTA format database, which contained P1 and P2 sequences. Common modifications were included (oxidations in M, phosphorylation in STY, acetylation in K, methylation in KR, dimethylation in KR and trimethylation in K). Precursor and fragment mass tolerances were set to 15 ppm and maximum mass shift was set to 500 Da. Proteoform spectrum matches with FDR < 1% were considered for further data integration and clustering analysis. PS and TP top-down search outputs were combined into a single harmonized dataset. Identified PrSMs were filtered by defining a 0 to 48 minutes retention time window to get rid of these identifications coming from the column wash. The remaining PrSMs were subsequently processed by clustering deconvoluted masses values with the DBSCAN algorithm previously developed by our group (Arauz-Garofalo *et al.*, 2021) with ε = 0.0012 and nmin = 9 as DBSCAN clustering hyperparameters (Ester *et al.*, 1996).

Statistical analysis for comparisons of protamine proteoform abundance

For the statistical analysis of comparisons of protamine proteoform abundance among sample groups, data from both biological and technical replicates were considered. Proteoforms were grouped according to the native protein (P1, mature P2 and immature P2) and the type of modification (unmodified, phosphorylated and +61Da) to evaluate the relative abundance of proteoforms groups. In this phase, we treated proteoforms that share the same mass (such as

those with a single phosphorylation occurring at distinct residues) as identical, to simplify the analysis. Ratios were computed with linear intensity values of the involved PrSMs at the Node pair level and compared among groups with pairwise t-test using R statistical software R 4.0.1 with rstatix library (R Core Team, 2020). Differential analysis of protamine proteoforms abundance was conducted with lme4 package using R 4. PS and TP intensities were $\log_2$-transformed and fitted to a linear model according to lmer function from lme4 package or lm from stats. “Altered P1/P2”, “Obesity” and “Advanced age” groups were compared against the “Control” group, and multiple comparisons were corrected by Benjamini & Hochberg. The significant protein established threshold was $|FC| > 1.5$ and p-value < 0.05.

Assignment of phosphorylated residues in P1+2ph proteoform

Hands-on verification of the phosphorylation site locations (p-sites) across all PrSMs of the P1+2Ph proteoform using ProSight Lite was conducted in the “Control” and “Advance age” groups (Fellers *et al.*, 2015). To assess the impact of each phosphorylated proteoform within the overall P1+2ph proteoforms landscape, we counted the PrSMs that were confirmed in both groups of study.

Separation of protamine 1- and protamine 2-enriched fractions by liquid chromatography

Liquid chromatography separation of total protamine extracts was performed on the Scientific and Technological Centres (CCiTUB), Universitat de Barcelona. Purified protamine extracts from a pool of normozoospermic semen samples were dissolved in 1% formic acid at a concentration of 2 µg/µL for subsequent liquid chromatography separation. A total of 80 µg

293 protamine were injected into a Waters Symmetry (3,9 x 150 mm, 5µm) C18 analytic column,
294 coupled to a Waters 2795 Separations Module and a Waters 2996 Photodiode Array Detector
295 (Waters, Milford, MA, USA). Separation was achieved using a mobile phase consisting of 0.1% FA
296 in water (solvent A) and 0.1% FA in acetonitrile (solvent B), applying the following program with a
297 flow rate of 0.8 mL/min: linear gradient from 97%A/3%B to 68%A/32%B in 10 min, followed by an
298 increase to 10%A/90%B in 1 min, which was kept for additional 6 min. The column was then re-
299 stabilized by decreasing to initial 97%A/3%B in 1 min, keeping this condition for another 6 min.
300 Afterwards, a total of 400 µg protamine were injected in five consecutive purification cycles.
301 Collected fractions were dried using speed-vacuum. The equivalent to 5% of each fraction was
302 visualized by acid-urea PAGE to identify the most relevant P1- and P2-enriched fractions for
303 subsequent top-down MS proteoform identification.
304
305 **Determination of oxidative stress levels produced by lipid peroxidation**
306 Sperm intracellular state of lipid peroxidation, as indicative of oxidative stress, was
307 measured in purified spermatozoa from obese patients (n=6) in comparison to controls (n=6).
308 Malondialdehyde (MDA) levels were quantified using the OxiSelect™ TBARS Assay Kit (Cell Biolabs,
309 Inc, CA, USA) and following the manufacturer's instructions. Briefly, between 20 and 50 million
310 purified sperm of each sample and increasing amounts of an MDA standard were lysed and
311 incubated with TBA reagent at 95 °C for 1 hour. Subsequently, samples were centrifuged at 825 g
312 for 15 min at room temperature, and the supernatants were transferred to a 96-well microplate.
313 Up to three technical replicates were included per sample and standard. The MDA levels were
314 quantified by reading absorbance at 532 nm with the Infinite® MNANO+ 200 PRO microplate reader

(Tecan, Switzerland). Values were normalized per number of cells and expressed as ng MDA/million sperm. Differences in oxidative stress levels between obese patients and controls were determined by applying a Shapiro Wilk/Kolmogorov Smirnov test, to evaluate data normal distribution, followed by a Student's t test. A p-value < 0.05 was considered significant.

RESULTS

Acid-urea PAGE-based ratios and classification of study groups

P1/P2 ratios, determined from quantification of the optical densities of main P1 and P2 bands detected by acid-urea PAGE, and histone/protamine ratios, relating the optical densities of main histone and protamine bands in the acid-urea PAGE, are shown in Supplementary Table S1 and Supplementary Table S2, respectively.

Sperm quality parameters, acid-urea PAGE-derived P1/P2 values, BMI and male age were considered to classify samples among groups of study before Top-Down MS analysis (Figure 1, Table 1 and Supplementary Table S1). Samples with a P1/P2 ratio above the literature-based normality threshold of 1.2 were classified as "Altered P1/P2" (Figure 1 and Supplementary Table S1). The rest of the samples included in this study showed acid-urea PAGE-based P1/P2 ratios below 1.2 (Supplementary Table S1).

As observed in Table 1, the "Altered P1/P2" group only differed from the "Control" group by a higher P1/P2 ratio (1.4 ± 0.1 vs 1.0 ± 0.2, respectively, p-value < 0.05), the "Obesity" group only differed from the "Control" group by a higher BMI (37.6 ± 4.4 vs 24.2 ± 1.8, respectively, p-

336 value < 0.05), and the “Advance age” group only differed from the “Control” group by a higher age
337 at the moment of sample collection (50.3 ± 3.2 y.o. vs 35.2 ± 0.4 y.o., respectively, p-value < 0.05).
338
339 **Global protamine proteoform profiles among study groups**
340 Protamine proteoform identification was performed using the DBSCAN-based approach
341 and the two identifying nodes Prosight PD (PS) and TopPIC suite (TP), as previously described by
342 our group (Arauz-Garofalo *et al.*, 2021) (Figure 1). Proteoform intensities from both software
343 evidenced a good Pearson correlation (R = 0.91, Figure 2A) and results from both PS and TP were
344 considered for further analyses. From the total 36,554 proteoform spectrum matches (PrSMs)
345 identified by top-down MS, 95% passed quality filters, 88% were successfully clustered, and 84%
346 were successfully classified (Table 2). The resulting 30,689 classified PrSMs yielded 28 protamine
347 proteoforms after collapsing artifact PTMs (carbamidomethylation, oxidation, mesityloxidation,
348 and pyroglutamic acid from glutamine). From those, 13 proteoforms were identified for P1, 6 for
349 P2 mature forms (HP2, HP3 and HP4) and 9 for P2 immature forms (Pre-P2, HPS1, HPS2 and HPI2)
350 (Figure 2B). Identified proteoforms included unmodified molecules as well as proteoforms bearing
351 phosphate groups and a +61Da modification, either as a single modification in one molecule or in
352 combination in the same protamine sequence (Figure 2B). P1 proteoforms were detected with
353 higher intensity than P2 proteoforms due to their more favourable ionization, as shown in
354 Supplementary Figure S1.
355
356 **Relative abundance of protamine proteoforms groups in stratified normozoospermic males**

The relative abundance of proteoform groups was evaluated to detect global deregulations of the protamine proteoform profile according to native protein and modification status. Intensities for all proteoforms identified for P1 and P2 mature (HP2, HP3 and HP4) and immature (pre-P2, HPS1, HPS2 and HPI2) components were used to establish the MS-based ratios P2-immature/P2 mature (relating the global amounts of P2 immature and mature forms), P1-ph/P1 (relating the global amount of P1 phosphorylated forms over total P1 amount), and P2-mature-ph/P2 mature (relating the global amount of mature P2 phosphorylated forms over total P2 amount; Figure 3). No significant differences were found on the relative levels of global P2 immature forms over P2 mature forms (P2-immature/P2-mature) among groups (Figure 3). Also, stable levels of global protamine phosphorylation were detected in "Altered P1/P2", "Obesity", and "Advanced age" groups in comparison with the control group (p-value > 0.05; Figure 3).

Normozoospermic men with abnormally high P1/P2 ratio calculated by acid-urea PAGE accumulate specific P2 immature proteoforms

Quantitative proteoform analysis revealed a specific increase in the levels of the P2 immature form HPS1 in "Altered P1/P2" samples compared to "Control" group (FC_{PS} = 82.69, p-$value_{PS}$ = 0.008; FC_{TP} = 434.93, p-$value_{TP}$ = 0.022) (Figure 4, Supplementary Table S3). Moreover, the P2 immature form HPI2 was exclusively identified in "Altered P1/P2" patients and was not observed in any other group of study (Figure 2 and 4). These results suggest an accumulation of P2 immature forms in men with an abnormally high acid-urea PAGE-based P1/P2 ratio. The proteoform P1+2(61)+2ph was found to be significantly less abundant in the "Altered P1/P2" group, but only when the PS node is considered (Figure 4, Supplementary Table S3).

Obesity and advance age are associated with reduced abundance of specific modified P1 proteoforms

The evaluation of the "Obesity" group revealed a remarkable reduction of P1 proteoforms bearing a +61 Da mass shift, either coupled or not to a monophosphorylation (Figure 5A, Supplementary Table S3). In particular, the levels of the P1+1(61)+1ph proteoform were significantly lower compared to controls (FC_{PS} = 0.35, *p*-$value_{PS}$ = 0.021; FC_{TP} = 0.40, p-$value_{TP}$ = 0.033). Quantitative levels of P1+3(61) were also found to be significantly reduced, although only according to TopPIC (FC_{TP} = 0.003, *p*-$value_{TP}$ = 0.024). The MS data obtained for the +61Da modification did not allow the assignment to any described modification. Of note, sperm from obese men showed a significant increase in intracellular levels of MDA in comparison to controls (Figure 5B), which is related to higher levels of oxidative damage derived from lipid peroxidation that may affect the chromatin state and integrity of the male gametes.

In turn, diphosphorylated P1 (P1+2ph), was found with significant lower abundance in "Advanced age" group, compared to controls, according to TP (FC_{TP} = 0.01, p-$value_{TP}$ = 0.012) (Figure 6A, Supplementary Table S3). The alteration on this P1 proteoform, not bearing any other modification but phosphorylation, lead us to look for the potential residues contributing to a reduction on P1+2ph abundance. We validated phospho-site localization of 42 out of 64 PrSMs. The remaining 22 PrSMs were ambiguous. From all possible combinations, Ser 11 and 22 showed the most pronounced phosphorylation loss (13 counts in "Control" group vs 3 counts in "Advanced age" group), being likely responsible to the observed alteration (Figure 6B). Diphosphorylation in Ser 11 and 29 were reduced to half counts in the "Advance age" group (4 vs 2 counts in "Control"

and “Advanced age” groups, respectively) which, despite not being a relevant decrease, highlights Ser 11 as the main contributor of this change (17 vs 5 counts in “Control” and “Advanced age” groups, respectively; Figure 6B).

Assignment of protamine proteoforms to acid urea-PAGE electrophoretic bands

The proteomic evaluation of P1- and P2-enriched protein extracts fractions obtained by liquid chromatography purification confirmed the accumulation of P1 and mature P2 (HP2, HP3 and HP4) proteoforms in the main electrophoretic bands observed in the acid-urea PAGE (Supplementary Figure S2). Specifically, MS evaluation demonstrated the identification of unmodified forms of P1 and mature P2, and some modified forms of P1 and HP2. The immature P2 forms Pre-P2 and HPS1 have been assigned to upper bands in the electrophoretic gel (Supplementary Figure S2). However, the specific assignment of low abundant modified forms to specific electrophoretic bands was challenging with the current approach.

DISCUSSION

The present work reports the presence of protamine proteoform quantitative alterations in human spermatozoa from normozoospermic men under different biological- and sociodemographic-associated conditions related with impaired chromatin status. Here, we show a refined MS approach comprising top-down proteomics and data analysis with the clustering algorithm DBSCAN for the quantitative comparison of protamine proteoforms, which cannot be detected nor quantified using standard MS strategies. This design has increased the reliability and

robustness of the identification of proteoforms, leading to the identification of a smaller number of proteoforms than that reported in our previous descriptive work (Soler-Ventura *et al.*, 2020). The detection of protamine proteoforms presenting differential abundance among the different study groups suggests the presence of additional layers of sperm chromatin epigenetic alterations so far unexplored in the male gamete.

Acidic electrophoresis-based estimation of the P1/P2 ratio has been applied in this work to classify samples among the study groups, since up to now it has been considered the most reliable strategy to evaluate protamine levels in sperm (Lescoat *et al.*, 1988; Bach *et al.*, 1990; Khara, 1997; Carrell and Liu, 2001; Nasr-Esfahani *et al.*, 2004; Aoki *et al.*, 2005, 2006; Torregrosa *et al.*, 2006; de Mateo *et al.*, 2009; Hammadeh *et al.*, 2010; Castillo *et al.*, 2011; Simon *et al.*, 2014; Soler-Ventura *et al.*, 2018). However, the detection of specific protamine proteoforms with significantly different abundance in all study groups, compared to the control, shows that our top-down proteomics pipeline precisely identifies and quantifies proteoforms that are undetectable by the electrophoretic approach. Furthermore, we have shown that in-gel protamine intensity quantification of the main P1 and P2 electrophoretic bands mainly relies on unmodified forms of P1 and mature P2 (HP2, HP3 and HP4) and some modified forms. However, the contribution of other protamine proteoforms to the density of these main electrophoretic bands is still unknown. The limitation of the acid-urea PAGE strategy evidenced herein might be overcome by using MS data. However, this might need to be taken carefully since P1 shows a more favourable ionization than P2, which results in a predominant MS identification of P1 proteoforms over P2 proteoforms and biased derived P1/P2 ratios. Furthermore, some low abundant protamine proteoforms could have not pass the quality cut-offs set herein for MS identification and were, thus, not included in

the present report. It is therefore required to develop further studies similar and complementary to this one in independent patient cohorts, to increase the number of samples evaluated by MS and to collect enough evidence to determine the normal range for MS-derived P1/P2 ratios.

A slightly higher relative proportion of the total of immature P2 forms over mature P2 forms has been observed in the "Altered P1/P2" group compared to the "Control" group, although no statistically significant. Only when individual abundance of protamine proteoforms was evaluated, an accumulation of two specific P2 immature proteoforms, HPS1 and HPI2, was revealed. Recent findings have shown that mice with impaired pre-P2 processing were infertile (Schneider *et al.*, 2016, 2020; Arévalo *et al.*, 2022). In fact, the loss of the characteristic pre-P2 N-terminal domain in mice (named as "cleaved PRM2" or "cP2") results in a remarkable impairment of chromatin remodeling during spermiogenesis. In particular, it was reported that there is an abnormal retention of transition proteins, and an incorrect accumulation of unprocessed P2 in the nucleus and of mature P2 in spermatid cytoplasm and residual bodies (Arévalo *et al.,* 2022). All these data suggest that the accumulation of HPS1 and HPI2 in the "Altered P1/P2" group might be the result of either failed transport of immature P2 outside the nucleus during chromatin condensation, as observed in mice, or a compensatory mechanism through which Pre-P2 expression and processing is increased to achieve normal P2 levels in these men, generating an excess of immature forms as subproducts. Relative levels of histone and protamines, quantified from the acid-urea PAGE, suggest an accumulation of histones in the altered P1/P2 compared to control. However, this accumulation was not exclusive for this cohort, being observed also in the other groups. Further studies evaluating the cellular location of HPS1 and HPI2 as well as transition protein retention will allow deciphering of the mechanism behind immature P2 accumulation in these patients.

Interestingly, the detrimental impact on proteoform content observed in this study was different depending on the presence of either molecular (P1/P2 levels) or non-molecular (obesity and advance age) alterations in the normozoospermic men. While a direct protamine-related impairment, potentially caused during testicular male germ cells development, suggests P2 cleavage or eviction defects, life-style conditions, such as weight and male age at conception, are preferentially related with alterations in P1 PTMs. Two main mass shifts have been detected in P1, one corresponding to incorporation of 80 Da (or 80 mass units), assigned to phosphorylation, and another corresponding to incorporation of 61 Da, which showed MS characteristics that did not match with any described PTM. A recent study evaluating sperm from men with altered seminogram using MS have also reported the presence of methylated amino acids in protamines (Schon *et al.*, 2023). However, this modification has not been identified with our top-down strategy and the cut-offs established in our study, and neither matches a +61Da mass shift.

Despite the unknown correspondence of the +61-mass shift to a specific modification, it is worth mentioning that it was detected with a significantly different abundance in all P1 proteoforms described in obese men, either alone or coupled to a monophosphorylation. Our initial hypothesis for this modification was that it may correspond to a Zn^{2+} molecule, which is known to be crucial to stabilize the nucleoprotamine structure (Björndahl and Kvist, 2011; Soler-Ventura *et al.*, 2020). However, the PrSM isotopic pattern of the +61 Da modification detected herein does not match with that of Zn^{2+}. We then conducted a deeper evaluation of the ejaculated sperm from the obese normozoospermic men, which revealed the presence of significantly higher intracellular levels of MDA, in comparison to controls. MDA is a cytotoxic product derived from cell membrane lipid peroxidation due to the degradation and reaction of polyunsaturated fatty acids

and radical species (Jové *et al.*, 2020). Therefore, our results agree with the relation observed between high fat dietary patterns and increased sperm membrane lipid peroxidation (McPherson *et al.*, 2014; McPherson and Lane, 2015; Akbarian *et al.*, 2021). In the cell, MDA may react with DNA, amine headgroups of phospholipids, and proteins (Ishii *et al.*, 2006; Weisser *et al.*, 2017; Kompella and Vasquez, 2019). Of note, the mass of a N-(2-propenal)-lysine modification derived from the lipid peroxidation product MDA might coincide with that increment observed in protamines in our study (~60 Da). However, before establishing this relationship, it should be considered that under physiological conditions MDA mainly modifies lysines (Uchida *et al.*, 1997; Voitkun and Zhitkovich, 1999; Weisser *et al.*, 2017; Jové *et al.*, 2020), which are not part of the sequence of human P1. Under in vitro conditions, in contrast, additional reactions with histidine, tyrosine and arginine were observed (Jové *et al.*, 2020). Another critical point is the fact that obese men show significant lower levels of P1 (+61) than controls, which, in principle, would be opposite to the higher levels of lipid peroxidation in the cells. In any case, the MDA-DNA reaction has mutagenic effects and produces endogenous DNA-protein crosslinks that may impede the ability of the cells to turnover chromatin damage (Voitkun and Zhitkovich, 1999; Galligan and Marnett, 2017; Kompella and Vasquez, 2019). Therefore, sperm from obese normozoospermic men are under intracellular stress conditions affecting chromatin stability, which makes it essential to develop specific studies focused on deciphering the unknown effect observed in this study on P1 proteoforms related to obesity.

Regarding protamine proteoform alterations related to aging, a reduction of diphosphorylated forms of P1 was observed, likely happening due to loss of phosphate groups in Ser11 in combination with Ser22. Previous results in mouse models evidenced that, while loss of a

single protamine phosphorylation site has no effect on fertilization potential, likely due to a redundancy of monophosphorylated protamine proteoforms in the sperm cell, the simultaneous depletion of two phosphorylation sites in P1 ends up in drastically impaired fertility (Gou *et al.*, 2020). This evidence reinforces the relevance of the decrease of this specific proteoform in men with advance age found in the present study. Besides, protamine phosphorylation status may have a role on zygotic male pronuclear reprogramming, knowing that protamine phosphorylation is an essential mechanism to unwrap paternal DNA after fertilization (Gou *et al.*, 2020). Studies in men alongside ageing would be useful to unravel whether alterations on protamine proteoforms worsen over time.

Altogether, the evaluation of the protamine proteoform profile in normozoospemic men has suggested that alterations of protein modifying mechanisms, rather than protein expression, are critical factors for the proper maturity of the sperm chromatin. In addition, the refined method and the results reported herein open a window to functionally validate the role of protamine proteoforms as an additional layer of epigenetic information critical for sperm function, preimplantation embryo development, and offspring health with animal studies.

DATA AVAILABILITY

The MS proteomics data underlying this article have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository [75] at https://www.ebi.ac.uk/pride/, and can be accessed with the dataset identifier PXD049367.

ACKNOWLEDGMENTS

Mass spectrometry/Proteomics was performed at the IRB Barcelona Mass Spectrometry and Proteomics Core Facility, which was a member of ProteoRed, PRB3-ISCIII, supported by grant PRB3 (IPT17/0019 - ISCIII-SGEFI / ERDF) and is granted in the framework of the 2014-2020 ERDF Operational Programme in Catalonia, co-financed by the European Regional Development Fund (ERDF), Reference: IU16-015983. We thank the Scientific and Technological Centres (CCiTUB), Universitat de Barcelona, and Esther Miralles, for their support and advice on liquid chromatography protein separation.

AUTHORS' ROLES

JC, MJ and RO designed the study; JC, MGa, AI, GAG, MVilanova, ML, JMC, MGu and DM acquired samples and data and perform methodology; JC, MGa, AI, GAG, MVilaseca, MJ and RO analyzed and interpreted data; JC, AI, MJ and RO drafted the article and all authors revised it critically for important intellectual content; all authors approved the final version of the manuscript and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work was appropriately investigated and resolved.

FUNDING

This study was funded by grants from the "Plataforma en Red de Proteómica Carlos III (PROTEORED)" to R.O., and by Instituto de Salud Carlos III (ISCIII) through the projects PI20/00936 to R.O., and PI23/00986 to J.C, and the grants CD17/00109 to J.C., and FI17/00224 to A.I., and co-funded by the European Union. The work has the support of "Departament de Recerca i Universitats de la Generalitat de Catalunya" to the research group 2021-SGR-01327. J.C. is a Serra Húnter fellow (Universitat de Barcelona, Generalitat de Catalunya). This article is based on work

from COST Action Andronet CA20119, supported by COST (European Cooperation in Science and Technology, www.cost.eu).

CONFLICT OF INTERESTS

The authors have no competing interest to disclose.

REFERENCES

Agarwal A, Baskaran S, Parekh N, Cho C-L, Henkel R, Vij S, Arafa M, Panner Selvam MK, Shah R. Male infertility. Lancet 2021;**397**:319–333.

Akbarian F, Rahmani M, Tavalaee M, Abedpoor N, Taki M, Ghaedi K, Nasr-Esfahani MH. Effect of Different High-Fat and Advanced Glycation End-Products Diets in Obesity and Diabetes-Prone C57BL/6 Mice on Sperm Function. Int J Fertil Steril 2021;**15**:226–233.

Aoki VW, Liu L, Carrell DT. Identification and evaluation of a novel sperm protamine abnormality in a population of infertile males. Hum Reprod 2005;**20**:1298–1306.

Aoki VW, Liu L, Jones KP, Hatasaka HH, Gibson M, Peterson CM, Carrell DT. Sperm protamine 1/protamine 2 ratios are related to in vitro fertilization pregnancy rates and predictive of fertilization ability. Fertil Steril 2006;**86**:1408–1415.

Arauz-Garofalo G, Jodar M, Vilanova M, de la Iglesia A, Castillo J, Soler-Ventura A, Oliva R, Vilaseca M, Gay M. Protamine Characterization by Top-Down Proteomics: Boosting Proteoform Identification with DBSCAN. Proteomes 2021;**9**:21.

Arévalo L, Merges GE, Schneider S, Oben FE, Neumann IS, Schorle H. Loss of the cleaved-protamine 2 domain leads to incomplete histone-to-protamine exchange and infertility in mice. PLoS Genet 2022;**18**:e1010272.

575 Bach O, Glander HJ, Scholz G, Schwarz J. Electrophoretic patterns of spermatozoal nucleoproteins
576 (NP) in fertile men and infertility patients and comparison with NP of somatic cells. Andrologia
577 1990;**22**:217–224.
578 Balhorn R. The protamine family of sperm nuclear proteins. Genome Biol 2007;**8**:227.
579 Balhorn R, Reed S, Tanphaichitr N. Aberrant protamine 1/protamine 2 ratios in sperm of infertile
580 human males. Experientia 1988;**44**:52–55.
581 Barrachina F, Soler-Ventura A, Oliva R, Jodar M. Sperm Nucleoproteins (Histones and Protamines).
582 In Zini A, Agarwal A, editors. A Clinician's Guide to Sperm DNA and Chromatin Damage 2018; p.
583 31–51.
584 Björndahl L, Kvist U. A model for the importance of zinc in the dynamics of human sperm chromatin
585 stabilization after ejaculation in relation to sperm DNA vulnerability. Syst Biol Reprod Med
586 2011;**57**:86–92.
587 Brunner AM, Nanni P, Mansuy IM. Epigenetic marking of sperm by post-translational modification
588 of histones and protamines. Epigenetics Chromatin 2014;**7**:2.
589 Carrell DT, Liu L. Altered protamine 2 expression is uncommon in donors of known fertility, but
590 common among men with poor fertilizing capacity, and may reflect other abnormalities of
591 spermiogenesis. J Androl 2001;**22**:604–610.
592 Castillo J, Amaral A, Oliva R. Sperm nuclear proteome and its epigenetic potential. Andrology
593 2014;**2**:326–338.
594 Castillo J, Estanyol J, Ballescà J, Oliva R. Human sperm chromatin epigenetic potential: genomics,
595 proteomics, and male infertility. Asian J Androl 2015;**17**:601.

596 Castillo J, Simon L, Mateo S de, Lewis S, Oliva R. Protamine/DNA ratios and DNA damage in native
597 and density gradient centrifuged sperm from infertile patients. J Androl 2011;**32**:324–332.
598 Chirat F, Arkhis A, Martinage A, Jaquinod M, Chevaillier P, Sautière P. Phosphorylation of human
599 sperm protamines HP1 and HP2: identification of phosphorylation sites. Biochim Biophys Acta
600 1993;**1203**:109–114.
601 Craig JR, Jenkins TG, Carrell DT, Hotaling JM. Obesity, male infertility, and the sperm epigenome.
602 Fertil Steril 2017;**107**:848–859.
603 de la Iglesia A, Jodar M, Oliva R, Castillo J. Insights into the sperm chromatin and implications for
604 male infertility from a protein perspective. WIREs Mechanisms of Disease 2023;**15**:e1588.
605 de Mateo S, Gazquez C, Guimera M, Balasch J, Meistrich ML, Ballesca JL, Oliva R. Protamine 2
606 precursors (Pre-P2), protamine 1 to protamine 2 ratio (P1/P2), and assisted reproduction
607 outcome. Fertil Steril 2009;**91**:715–722.
608 de Mateo S, Ramos L, Boer P de, Meistrich M, Oliva R. Protamine 2 Precursors and Processing.
609 Protein Pept Lett 2011;**18**:778–785.
610 de Yebra L, Ballescà JL, Vanrell JA, Bassas L, Oliva R. Complete selective absence of protamine P2 in
611 humans. J Biol Chem 1993;**268**:10553–10557.
612 de Yebra L, Ballescá JL, Vanrell JA, Corzett M, Balhorn R, Oliva R. Detection of P2 precursors in the
613 sperm cells of infertile patients who have reduced protamine P2 levels. Fertil Steril 1998;**69**:755–
614 759.
615 Deshpande SSS, Nemani H, Balasinor NH. Diet-induced- and genetic-obesity differentially alters
616 male germline histones. Reproduction 2021;**162**:411–425.

617 Donkin I, Versteyhe S, Ingerslev LR, Qian K, Mechta M, Nordkap L, Mortensen B, Appel EVR,
618 Jørgensen N, Kristiansen VB, et al. Obesity and Bariatric Surgery Drive Epigenetic Variation of
619 Spermatozoa in Humans. Cell Metab 2016;**23**:369–378.
620 Ester M, Kriegel H-P, Sander J, Xu X. A density-based algorithm for discovering clusters in large
621 spatial databases with noise. KDD-96 Proceedings 1996;**96**:226–231.
622 Fellers RT, Greer JB, Early BP, Yu X, LeDuc RD, Kelleher NL, Thomas PM. ProSight Lite: graphical
623 software to analyze top-down mass spectrometry data. Proteomics 2015;**15**:1235–1238.
624 Fullston T, Ohlsson Teague EMC, Palmer NO, DeBlasio MJ, Mitchell M, Corbett M, Print CG, Owens
625 JA, Lane M. Paternal obesity initiates metabolic disturbances in two generations of mice with
626 incomplete penetrance to the F2 generation and alters the transcriptional profile of testis and
627 sperm microRNA content. FASEB J 2013;**27**:4226–4243.
628 Fullston T, Palmer NO, Owens JA, Mitchell M, Bakos HW, Lane M. Diet-induced paternal obesity in
629 the absence of diabetes diminishes the reproductive health of two subsequent generations of
630 mice. Hum Reprod 2012;**27**:1391–1400.
631 Galligan JJ, Marnett LJ. Histone Adduction and Its Functional Impact on Epigenetics. Chem Res
632 Toxicol 2017;**30**:376–387.
633 Gou L-T, Lim D-H, Ma W, Aubol BE, Hao Y, Wang X, Zhao J, Liang Z, Shao C, Zhang X, et al. Initiation
634 of Parental Genome Reprogramming in Fertilized Oocyte by Splicing Kinase SRPK1-Catalyzed
635 Protamine Phosphorylation. Cell 2020;**180**:1212-1227.
636 Hammadeh ME, Hamad MF, Montenarh M, Fischer-Hammadeh C. Protamine contents and P1/P2
637 ratio in human spermatozoa from smokers and non-smokers. Human Reproduction
638 2010;**25**:2708–2720.

639 Hammoud S, Liu L, Carrell DT. Protamine ratio and the level of histone retention in sperm selected
640 from a density gradient preparation. Andrologia 2009;**41**:88–94.
641 Hornbeck P v, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014:
642 mutations, PTMs and recalibrations. Nucleic Acids Res 2015;**43**:D512-20.
643 Ishii T, Kumazawa S, Sakurai T, Nakayama T, Uchida K. Mass Spectroscopic Characterization of
644 Protein Modification by Malondialdehyde. Chem Res Toxicol 2006;**19**:122–129.
645 Jenkins TG, Aston KI, Carrell DT. Sperm epigenetics and aging. Transl Androl Urol 2018;**7**:S328–
646 S335.
647 Jenkins TG, Aston KI, Pflueger C, Cairns BR, Carrell DT. Age-associated sperm DNA methylation
648 alterations: possible implications in offspring disease susceptibility. PLoS Genet
649 2014;**10**:e1004458.
650 Jodar M, Oliva R. Protamine alterations in human spermatozoa. Adv Exp Med Biol 2014;**791**:83–
651 102.
652 Jové M, Mota-Martorell N, Pradas I, Martín-Gari M, Ayala V, Pamplona R. The Advanced
653 Lipoxidation End-Product Malondialdehyde-Lysine in Aging and Longevity. Antioxidants
654 2020;**9**:1132.
655 Kasinsky HE, Eirín-López JM, Ausió J. Protamines: structural complexity, evolution and chromatin
656 patterning. Protein Pept Lett 2011;**18**:755–771.
657 Khara KK. Human protamines and male infertility. J Assist Reprod Genet 1997;**14**:282–290.
658 Kompella P, Vasquez KM. Obesity and cancer: A mechanistic overview of metabolic changes in
659 obesity that impact genetic instability. Mol Carcinog 2019;**58**:1531–1550.

660 Kou Q, Xun L, Liu X. TopPIC: a software tool for top-down mass spectrometry-based proteoform
661 identification and characterization. Bioinformatics 2016; **32**:3495-3497.
662 Laurentino S, Cremers J, Horsthemke B, Tüttelmann F, Czeloth K, Zitzmann M, Pohl E, Rahmann S,
663 Schröder C, Berres S, et al. A germ cell-specific ageing pattern in otherwise healthy men. Aging
664 Cell 2020;**19**:e13242.
665 LeDuc RD, Taylor GK, Kim Y-B, Januszyk TE, Bynum LH, Sola J v., Garavelli JS, Kelleher NL. ProSight
666 PTM: an integrated environment for protein identification and characterization by top-down
667 mass spectrometry. Nucleic Acids Res 2004;**32**:W340–W345.
668 Lescoat D, Colleu D, Boujard D, Lannou D Le. Electrophoretic characteristics of nuclear proteins
669 from human spermatozoa. Arch Androl 1988;**20**:35–40.
670 Luke L, Tourmente M, Roldan ERS. Sexual selection of protamine 1 in mammals. Mol Biol Evol
671 2016;**33**:174–184.
672 Mascarenhas MN, Flaxman SR, Boerma T, Vanderpoel S, Stevens GA. National, Regional, and Global
673 Trends in Infertility Prevalence Since 1990: A Systematic Analysis of 277 Health Surveys. PLoS
674 Med 2012;**9**:e1001356.
675 McPherson NO, Fullston T, Bakos HW, Setchell BP, Lane M. Obese father's metabolic state,
676 adiposity, and reproductive capacity indicate son's reproductive health. Fertil Steril
677 2014;**101**:865–873.
678 McPherson NO, Lane M. Male obesity and subfertility, is it really about increased adiposity? Asian
679 J Androl 2015;**17**:450–458.
680 Mengual L, Ballescá JL, Ascaso C, Oliva R. Marked differences in protamine content and P1/P2 ratios
681 in sperm cells from percoll fractions between patients and controls. J Androl 2003;**24**:438–447.

682 Milekic MH, Xin Y, O'Donnell A, Kumar KK, Bradley-Moore M, Malaspina D, Moore H, Brunner D,
683 Ge Y, Edwards J, et al. Age-related sperm DNA methylation changes are transmitted to offspring
684 and associated with abnormal behavior and dysregulated gene expression. Mol Psychiatry
685 2015;**20**:995–1001.
686 Moritz L, Schon SB, Rabbani M, Sheng Y, Agrawal R, Glass-Klaiber J, Sultan C, Camarillo JM, Clements
687 J, Baldwin MR, et al. Sperm chromatin structure and reproductive fitness are altered by
688 substitution of a single amino acid in mouse protamine 1. Nat Struct Mol Biol 2023;**30**:1077–
689 1091.
690 Nanassy L, Liu L, Griffin J, Carrell D. The clinical utility of the protamine 1/protamine 2 ratio in
691 sperm. Protein Pept Lett 2011;**18**:772–777.
692 Nasr-Esfahani MH, Razavi S, Mozdarani H, Mardani M, Azvagi H. Relationship between protamine
693 deficiency with fertilization rate and incidence of sperm premature chromosomal condensation
694 post-ICSI. Andrologia 2004;**36**:95–100.
695 Ni K, Spiess A-N, Schuppe H-C, Steger K. The impact of sperm protamine deficiency and sperm DNA
696 damage on human male fertility: a systematic review and meta-analysis. Andrology 2016;**4**:789–
697 799.
698 Oliva R. Protamines and male infertility. Hum Reprod Update 2006;**12**:417–435.
699 Oliva R, Castillo J. Proteomics and the genetics of sperm chromatin condensation. Asian J Androl
700 2011;**13**:24–30.
701 Oliva R, Dixon GH. Vertebrate Protamine Genes and the Histone-to-Protamine Replacement
702 Reaction. Prog Nucleic Acid Res Mol Biol 1991;**40**:25–94.

703 Palmer NO, Fullston T, Mitchell M, Setchell BP, Lane M. SIRT6 in mouse spermatogenesis is
704 modulated by diet-induced obesity. Reprod Fertil Dev 2011;**23**:929–939.
705 Papoutsopoulou S, Nikolakaki E, Chalepakis G, Kruft V, Chevaillier P, Giannakouros T. SR protein-
706 specific kinase 1 is highly expressed in testis and phosphorylates protamine 1. Nucleic Acids Res
707 1999;**27**:2972–2980.
708 Pepin A-S, Lafleur C, Lambrot R, Dumeaux V, Kimmins S. Sperm histone H3 lysine 4 tri-methylation
709 serves as a metabolic sensor of paternal obesity and is associated with the inheritance of
710 metabolic dysfunction. Mol Metab 2022;**59**:101463.
711 Perez-Riverol Y, Bai J, Bandla C, García-Seisdedos D, Hewapathirana S, Kamatchinathan S, Kundu
712 DJ, Prakash A, Frericks-Zipper A, Eisenacher M, et al. The PRIDE database resources in 2022: a
713 hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res 2022;**50**:D543–
714 D552.
715 Pirhonen A, Linnala-Kankkunen A, Mäenpää PH. Identification of phosphoseryl residues in
716 protamines from mature mammalian spermatozoa. Biol Reprod 1994;**50**:981–986.
717 Pohl E, Gromoll J, Wistuba J, Laurentino S. Healthy ageing and spermatogenesis. Reproduction
718 2021;**161**:R89–R101.
719 Pruslin FH, Imesch E, Winston R, Rodman TC. Phosphorylation state of protamines 1 and 2 in human
720 spermatids and spermatozoa. Gamete Res 1987;**18**:179–190.
721 Queralt R, Adroer R, Oliva R, Winkfein RJ, Retief JD, Dixon GH. Evolution of protamine P1 genes in
722 mammals. J Mol Evol 1995;**40**:601–607.

723 R Core Team. R: A Language and Environment for Statistical Computing [Internet]. 2020; R
724 Foundation for Statistical Computing: Vienna, Austria. Available from: https://www.R-
725 project.org.
726 Schneider S, Balbach M, Jan F Jikeli, Fietz D, Nettersheim D, Jostes S, Schmidt R, Kressin M,
727 Bergmann M, Wachten D, et al. Re-visiting the Protamine-2 locus: deletion, but not
728 haploinsufficiency, renders male mice infertile. Sci Rep 2016;**6**:36764.
729 Schneider S, Shakeri F, Trötschel C, Arévalo L, Kruse A, Buness A, Poetsch A, Steger K, Schorle H.
730 Protamine-2 Deficiency Initiates a Reactive Oxygen Species (ROS)-Mediated Destruction
731 Cascade during Epididymal Sperm Maturation in Mice. Cells 2020;**9**: 1789.
732 Schon SB, Moritz L, Rabbani M, Meguid J, Juliano BR, Ruotolo BT, Aston K, Hammoud SS. Proteomic
733 analysis of human sperm reveals changes in protamine 1 phosphorylation in men with infertility.
734 F S Sci 2024; **5**:121-129
735 Simon L, Liu L, Murphy K, Ge S, Hotaling J, Aston KI, Emery B, Carrell DT. Comparative analysis of
736 three sperm DNA damage assays and sperm nuclear protein content in couples undergoing
737 assisted reproduction treatment. Human Reproduction 2014;**29**:904–917.
738 Soler-Ventura A, Castillo J, de la Iglesia A, Jodar M, Barrachina F, Ballesca JL, Oliva R. Mammalian
739 Sperm Protamine Extraction and Analysis: A Step-By-Step Detailed Protocol and Brief Review of
740 Protamine Alterations. Protein Pept Lett [Internet] 2018;**25**:424–433.
741 Soler-Ventura A, Gay M, Jodar M, Vilanova M, Castillo J, Arauz-Garofalo G, Villarreal L, Ballescà JL,
742 Vilaseca M, Oliva R. Characterization of Human Sperm Protamine Proteoforms through a
743 Combination of Top-Down and Bottom-Up Mass Spectrometry Approaches. J Proteome Res
744 2020;**19**:221–237.

745 Soubry A, Guo L, Huang Z, Hoyo C, Romanus S, Price T, Murphy SK. Obesity-related DNA
746 methylation at imprinted genes in human sperm: Results from the TIEGER study. Clin Epigenetics
747 2016;**8**:51.
748 Soubry A, Schildkraut JM, Murtha A, Wang F, Huang Z, Bernal A, Kurtzberg J, Jirtle RL, Murphy SK,
749 Hoyo C. Paternal obesity is associated with IGF2 hypomethylation in newborns: results from a
750 Newborn Epigenetics Study (NEST) cohort. BMC Med 2013;**11**:29.
751 Terashima M, Barbour S, Ren J, Yu W, Han Y, Muegge K. Effect of high fat diet on paternal sperm
752 histone distribution and male offspring liver gene expression. Epigenetics 2015;**10**:861–871.
753 Tiegs AW, Landis J, Garrido N, Scott R, Hotaling J. Total motile sperm count trend over time across
754 two continents: evaluation of semen analyses from 119,972 infertile men. Fertil Steril
755 2018;**110**:e27.
756 Torregrosa N, Dominguez-Fandos D, Camejo MI, Shirley CR, Meistrich ML, Ballesca JL, Oliva R.
757 Protamine 2 precursors, protamine 1/protamine 2 ratio, DNA integrity and other sperm
758 parameters in infertile patients. Hum Reprod 2006;**21**:2084–2089.
759 Uchida K, Sakai K, Itakura K, Osawa T, Toyokuni S. Protein Modification by Lipid Peroxidation
760 Products: Formation of Malondialdehyde-DerivedNϵ-(2-Propenal)lysine in Proteins. Arch
761 Biochem Biophys 1997;**346**:45–52.
762 Voitkun V, Zhitkovich A. Analysis of DNA–protein crosslinking activity of malondialdehyde in vitro.
763 Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 1999;**424**:97–
764 106.

765 Weisser J, Ctortecka C, Busch CJ, Austin SR, Nowikovsky K, Uchida K, Binder CJ, Bennett KL. A
766 Comprehensive Analytical Strategy To Identify Malondialdehyde-Modified Proteins and
767 Peptides. Anal Chem 2017;**89**:3847–3852.
768 Willmitzer L, Bode J, Wagner KG. Phosphorylated protamines. II. Circular dichroism of complexes
769 with DNA, dependency on ionic strength. Nucleic Acids Res 1977;**4**:163–176.
770 Wood KA, Goriely A. The impact of paternal age on new mutations and disease in the next
771 generation. Fertil Steril 2022;**118**:1001–1012.
772 World Health Organization. WHO laboratory manual for the examination and processing of human
773 semen, sixth edition. 2021; World Health Organization; 2021. License: CC BY-NC-SA 3.0 IGO.:
774 Geneva (Switzerland).
775 Wu JY, Ribar TJ, Cummings DE, Burton KA, McKnight GS, Means AR. Spermiogenesis and exchange
776 of basic nuclear proteins are impaired in male germ cells lacking Camk4. Nat Genet
777 2000;**25**:448–452.

778

779 **FIGURE LEGENDS**

780

781 **Figure 1. Schematic summary of the study workflow.** From top to bottom, semen samples from
782 normozoospermic men were collected and spermatozoa were purified by density gradient
783 centrifugation prior acidic protamine extraction. Protamine levels and P1/P2 ratios were calculated
784 per individual subjects by determination of optical densities of the main electrophoretic bands
785 corresponding to P1 and P2. Individual samples were sorted per group of study (Control, Altered
786 P1/P2, Obesity, and Advanced age). A set of samples from the Control and the Obesity groups were

subjected to the evaluation of oxidative stress produced by lipid peroxidation by quantifying malondialdehyde (MDA) levels. Protamine proteoform identification and quantification were conducted by top-down mass spectrometry, applying the density-based spatial clustering of applications with noise (DBSCAN) algorithm and two search nodes, Prosight PD and TopPIC suite. Mass shifts from the unmodified protamine proteoforms masses were identified and correlated to post-translational modifications. In the bottom section, an example of incorporation of phosphate groups (mass shifts of 80 mass units) to protamine 1 is shown. Quantitative differences among groups were considered significant when a p-value <0.05 and FC >1.5 were obtained.

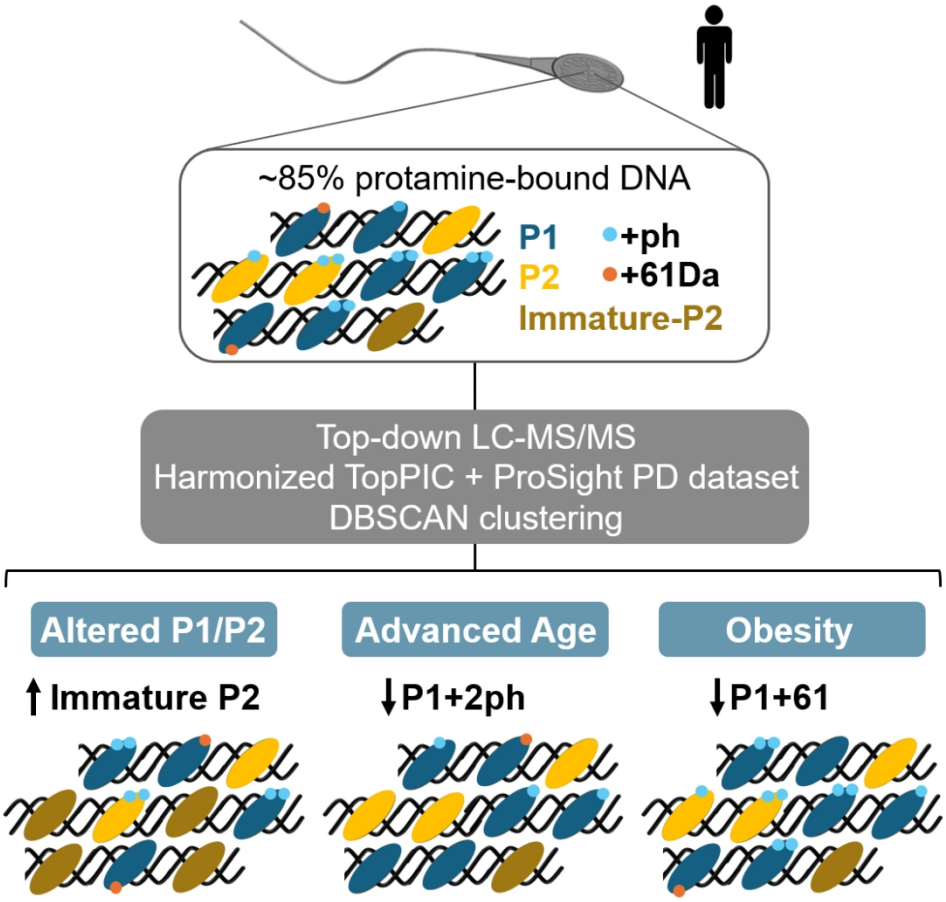
Figure 2. Human protamine proteoform identification by top-down proteomics using DBSCAN algorithm. (**A**) Scatter plot showing the correlation between proteoform $\log_2$ Intensity according to the nodes ProSight PD (PS; y-axis) and TopPIC suite (TP; x-axis). Each datapoint represents a protamine proteoform identification at the technical replicate level. Missing values were imputed at 10 only for visualization purposes. Pearson correlation coefficient without considering missing values was R = 0.91. (**B**) Proteoform landscape of the different groups of study after collapsing artifact modifications. “ph” stands for Phosphorylation, “(61)” represents a not assigned +61 Da mass shift, “PrSMs” are proteoform spectrum matches.

Figure 3. Global MS-based relative abundances of protamines among the different groups. Relation of the abundance of P2 immature forms to P2 mature forms, P1 phosphorylated forms to total P1, and P2 phosphorylated forms to total P2, calculated by ratios. The facet grid breaks down the ratios defined (columns) and the searching nodes Prosight PD and TopPIC suite (rows). When

809 considering global abundances, no significant differences were detected among study groups (p-
810 value > 0.05). P2-immature, all proteoforms corresponding to the P2 immature forms Pre-P2, HPS1,
811 HPS2 and HPI2; P2-mature, all proteoforms corresponding to the P2 mature forms HP2, HP3 and
812 HP4; P1-ph, phosphorylated P1 proteoforms; P1, all P1 proteoforms; P2-mature-ph,
813 phosphorylated proteoforms corresponding to the P2 mature forms HP2, HP3 and HP4.
814
815 **Figure 4. Individual abundance of protamine proteoforms identified in the "Altered P1/P2" group**
816 **in comparison with "Control" males.** Plot showing the comparison of the $\log_2$ Fold Change (FC)
817 values for "Altered P1/P2" (y-axis) and "Control" (x-axis) groups, using data obtained with ProSight
818 PD (PS, left) and TopPIC suite (TP, right). Dot size represents $-\log(p)$. Red and blue dots indicate
819 individual proteoforms with significant higher and lower abundance, respectively ($|FC| > 1.5$, $p <$
820 0.05). The proteoform identified in all the biological and technical replicates of only one of the
821 compared groups (b/w identification) is shown in orange.
822
823 **Figure 5. Differential abundance of individual protamine proteoforms in normozoospermic males**
824 **with obesity. (A)** Plots showing the comparison of $\log_2$ Fold Change (FC) values for "Obesity" (y-
825 axis) versus "Control" (x-axis) groups obtained with ProSight PD (PS, left) and TopPIC Suite (TP,
826 right). Dot size represents -log(*p*). Blue dots indicate individual proteoforms with significant lower
827 abundance in obese males ($|FC| > 1.5$, $p < 0.05$). **(B)** Intracellular levels of malondialdehyde (MDA),
828 as marker of oxidative stress produced by lipid peroxidation, in purified sperm from "Control" and
829 "Obesity" samples. The results are expressed in ng MDA/ million of sperm. The asterisk indicates
830 differences with p-value < 0.05.

Figure 6. Individual protamine proteoforms with significant differential abundance in males with advanced age. (**A**) Plots showing the comparison of $\log_2$ Fold Change (FC) values between "Advanced age" (y-axis) and "Control" (x-axis) groups obtained with ProSight PD (PS, left) and TopPIC Suite (TP, right). Dot size represents -log(*p*). Blue dots indicate significant lower abundance in the "Advanced age" group ($|FC| > 1.5$, $p < 0.05$). (**B**) Quantitative prediction of simultaneously phosphorylated P1 residues in "Advance age" and "Control" males to identify the sites most likely contributing to the significant reduced abundance of diphosphorylated P1 detected in males with advanced age.

Graphical abstract. Alterations in the abundance of specific protamine proteoforms were identified in normozoospermic men with altered P1/P2 ratios, obesity and advanced age, using an optimized proteomics strategy.



199x208mm (150 x 150 DPI)

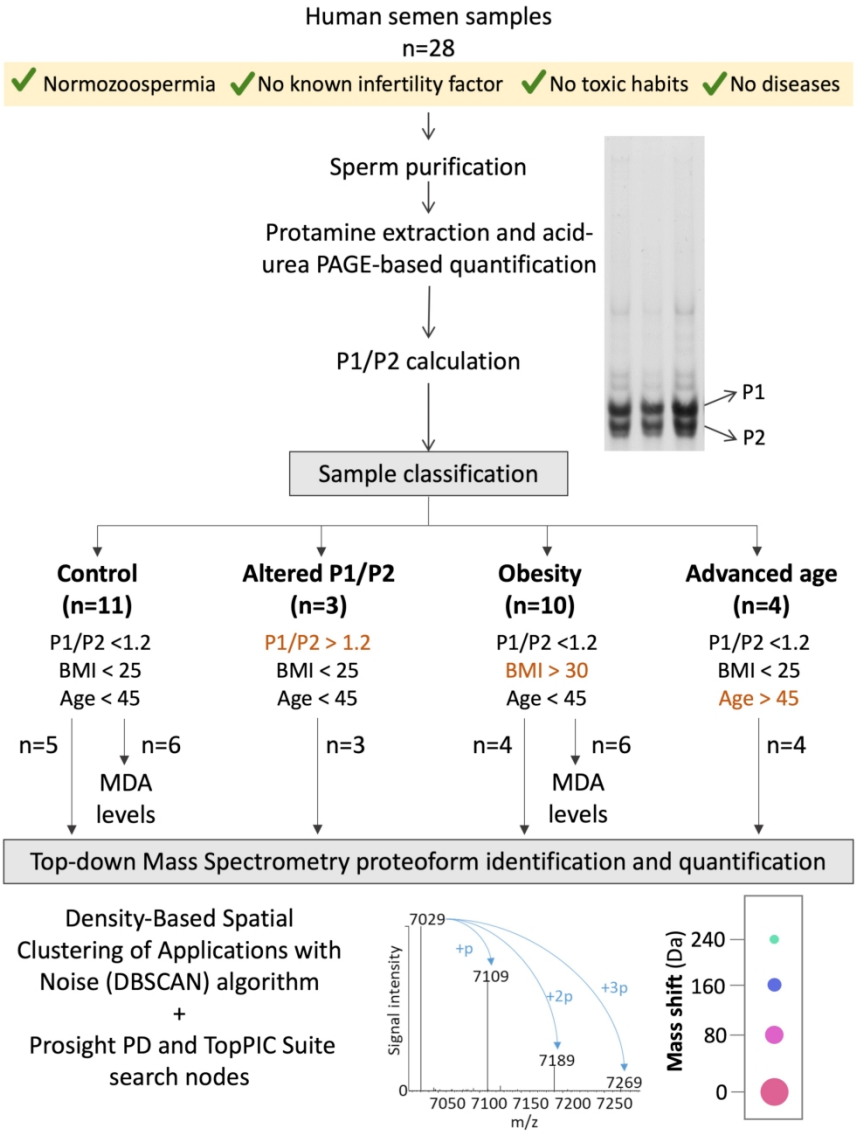


Figure 1. Schematic summary of the study workflow. From top to bottom, semen samples from normozoospermic men were collected and spermatozoa were purified by density gradient centrifugation prior acidic protamine extraction. Protamine levels and P1/P2 ratios were calculated per individual subjects by determination of optical densities of the main electrophoretic bands corresponding to P1 and P2. Individual samples were sorted per group of study (Control, Altered P1/P2, Obesity, and Advanced age). A set of samples from the Control and the Obesity groups were subjected to the evaluation of oxidative stress produced by lipid peroxidation by quantifying malondialdehyde (MDA) levels. Protamine proteoform identification and quantification were conducted by Top-Down Mass Spectrometry, applying the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm and two search nodes, Prosight PD and TopPIC suite. Mass shifts from the unmodified protamine proteoforms masses were identified and correlated to post-translational modifications. At the bottom part, an example of incorporation of phosphate groups (mass shifts of 80 mass units) to protamine 1 is shown. Quantitative differences among groups were considered significant when p-value <0.05 and FC >1.5 were obtained.

157x202mm (600 x 600 DPI)

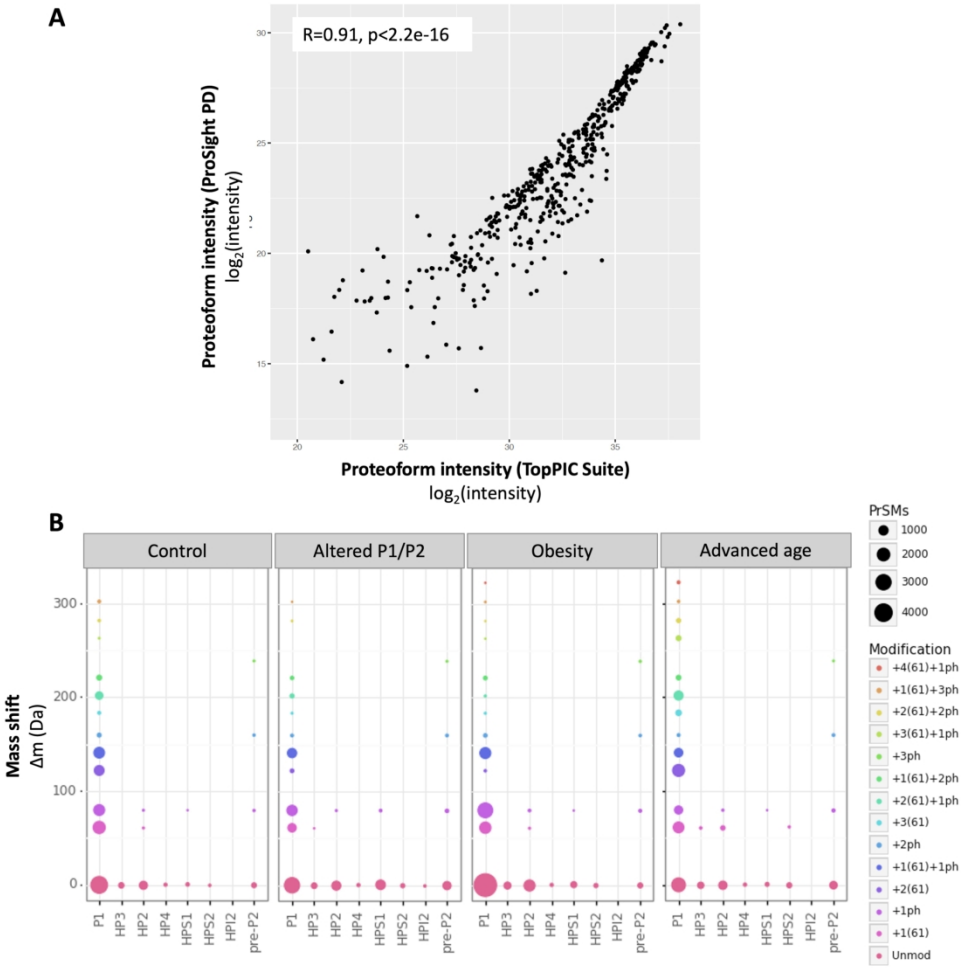


Figure 2. Human protamine proteoform identification by Top-Down proteomics using DBSCAN algorithm. (A) Scatter plot showing the correlation between proteoform log2 Intensity according to the nodes ProSight PD (PS; y-axis) and TopPIC suite (TP; x-axis). Each datapoint represents a protamine proteoform identification at the technical replicate level. Missing values were imputed at 10 only for visualization purposes. Pearson correlation coefficient without considering missing values was R = 0.91. (B) Proteoform landscape of the different groups of study after collapsing artifact modifications. "ph" stands for Phosphorylation, "(61)" represents a not assigned +61 Da mass shift, "PrSMs" are proteoform spectrum matches.

245x251mm (600 x 600 DPI)

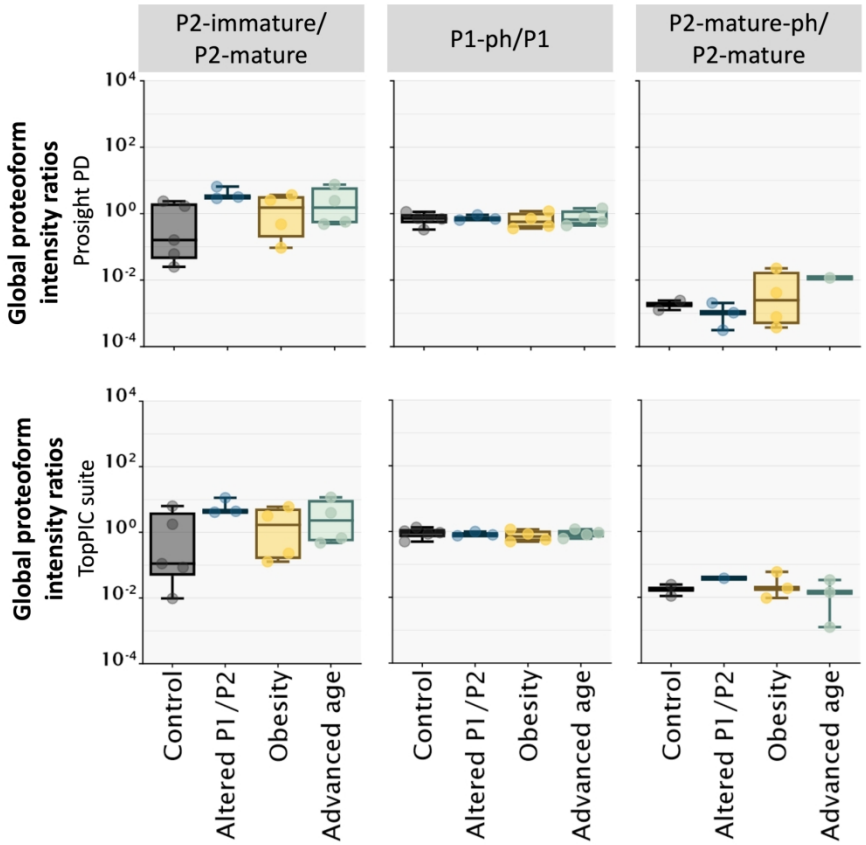


Figure 3. Global MS-based protamine relative abundances among the experimental groups. Relation of the abundances of P2 immature forms over P2 mature forms, P1 phosphorylated forms over total P1, and P2 phosphorylated forms over total P2, calculated by ratios. The facet grid breaks down the ratios defined (columns) and the searching nodes Prosight PD and TopPIC suite (rows). When considering global abundances, no significant differences were detected among study groups (p-value > 0.05). P2-immature: all proteoforms corresponding to the P2 immature forms Pre-P2, HPS1, HPS2 and HPI2; P2-mature: all proteoforms corresponding to the P2 mature forms HP2, HP3 and HP4; P1-ph: phosphorylated P1 proteoforms; P1: all P1 proteoforms; P2-mature-ph: phosphorylated proteoforms corresponding to the P2 mature forms HP2, HP3 and HP4.

210x193mm (600 x 600 DPI)

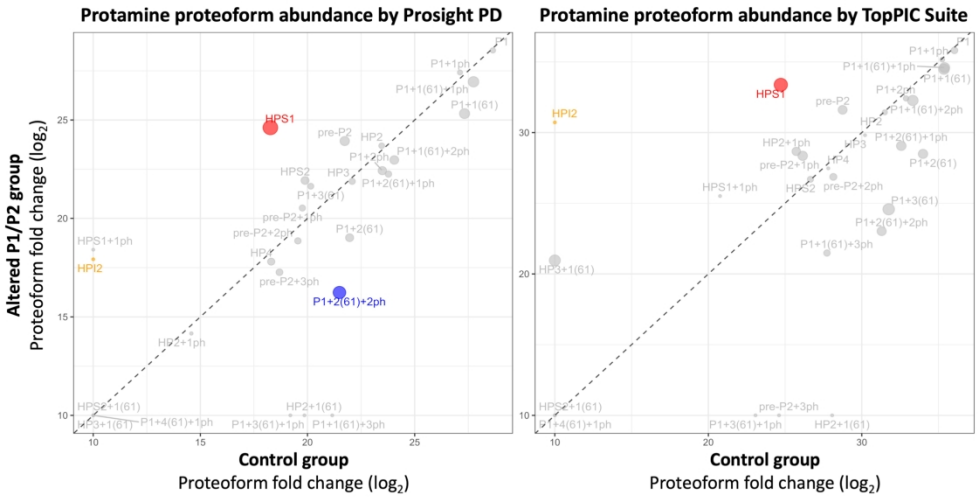


Figure 4. Individual abundance of protamine proteoforms identified in the "Altered P1/P2" group in comparison with "Control" males. Plot showing the comparison of the log2 Fold Change (FC) values for "Altered P1/P2" (y-axis) and "Control" (x-axis) groups, using data obtained with ProSight PD (PS, left) and TopPIC suite (TP, right). Dot size represents -log(p). Red and blue dots indicate individual proteoforms with significant higher and lower abundance, respectively (|FC| > 1.5, p < 0.05). Proteoform identified in all the biological and technical replicates of only one of the compared groups (b/w identification) is shown in orange.

201x111mm (600 x 600 DPI)

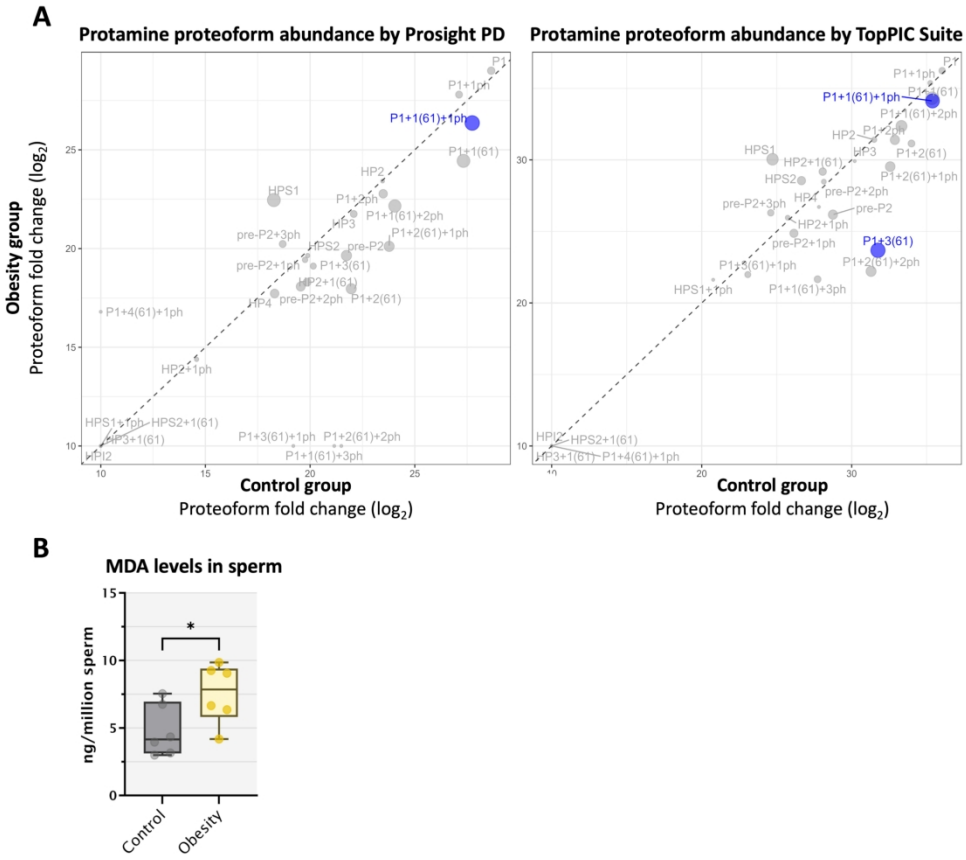


Figure 5. Differential abundance of individual protamine proteoforms in normozoospermic males with obesity. (A) Plots showing the comparison of log2 Fold Change (FC) values for "Obesity" (y-axis) vs "Control" (x-axis) groups obtained with ProSight PD (PS, left) and TopPIC Suite (TP, right). Dot size represents -log(p). Blue dots indicate individual proteoforms with significant lower abundance in obese males (|FC| > 1.5, p < 0.05). (B) Intracellular levels of malondialdehyde (MDA), as marker of oxidative stress produced by lipid peroxidation, in purified sperm from "Control" and "Obesity" samples. The results are expressed in ng MDA/ million of sperm. The asterisk indicates differences with p-value < 0.05.

199x177mm (600 x 600 DPI)

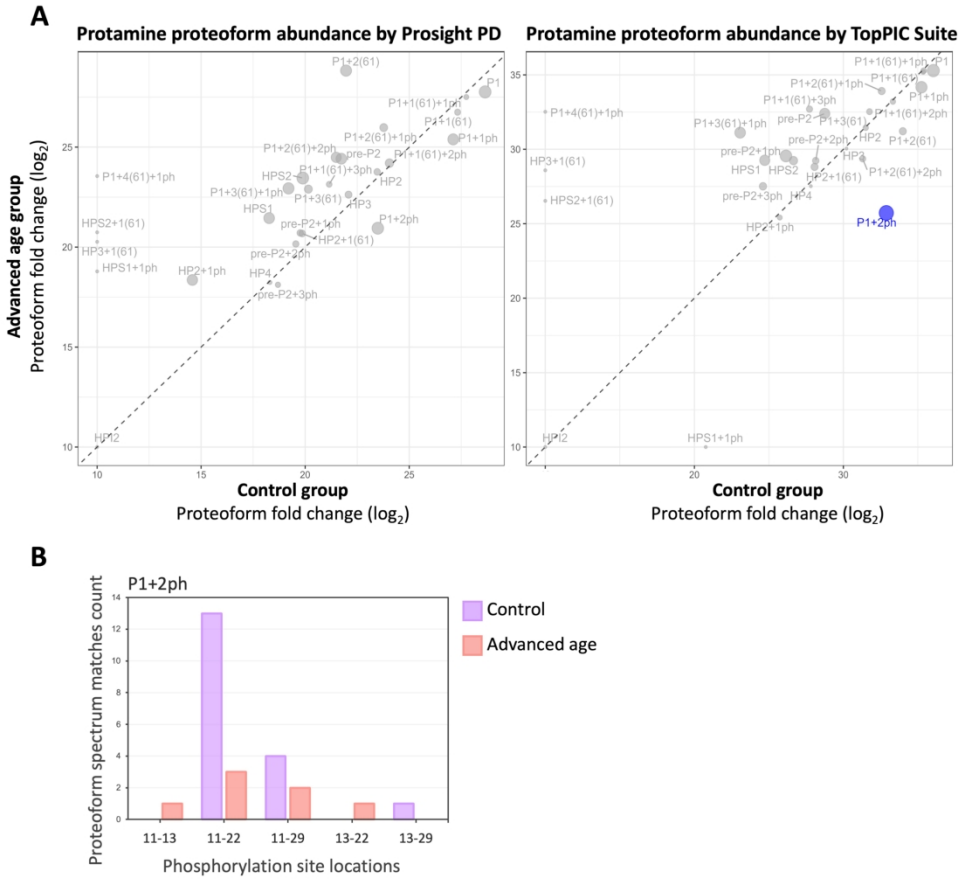


Figure 6. Individual protamine proteoforms with significant differential abundance in males with advanced age. (A) Plots showing the comparison of log2 Fold Change (FC) values between "Advanced age" (y-axis) and "Control" (x-axis) groups obtained with ProSight PD (PS, left) and TopPIC Suite (TP, right). Dot size represents -log(p). Blue dots indicate significant lower abundance in the "Advanced age" group (|FC| > 1.5, $p < 0.05$). (B) Quantitative prediction of simultaneously phosphorylated P1 residues in "Advance age" and "Control" males, to identify the sites most likely contributing to the significant reduced abundance of diphosphorylated P1 detected in males with advanced age.

250x230mm (600 x 600 DPI)

Table 1. Characteristics of the study groups included in the protamine proteoform Top-Down Mass Spectrometry-based quantitative comparison.

Sample group	Conc. (Mz/mL)	A +B (%)	Vol. (mL)	NF (%)	P1/P2 ratio	BMI	Age
Control (n=5)	155.4 ± 117.6	59.3 ± 16.8	2.4 ± 1.2	16.8 ± 4.3	1.0 ± 0.1	24.2 ± 1.8	35.2 ± 0.4
Altered P1/P2 (> 1.2) (n=3)	74.8 ± 5.7	48.3 ± 6.9	2.6 ± 1.2	15.0 ± 5.3	**1.4 ± 0.1[a]**	26.0 ± 1.7	33.7 ± 2.1
Obesity (BMI > 30) (n=4)	125.5 ± 50.3	54.8 ± 19.7	3.0 ± 1.1	24.5 ± 10.5	1.0 ± 0.0	**37.6 ± 4.4[b]**	35.8 ± 4.4
Advanced age (> 45 y.o.) (n=4)	129.3 ± 38.3	51.4 ± 18.9	2.2 ± 1.1	16.3 ± 10.6	1.0 ± 0.1	25.0 ± 1.1	**50.3 ± 3.2[c]**

Values corresponding to relevant seminal parameters indicating normality (Conc.: sperm concentration; A + B: sperm with progressive motility; Vol.: semen volume; NF: sperm with normal forms), acid-urea PAGE-based P1/P2 ratio, body mass index (BMI) and age are shown as mean value ± standard deviation. Differential values used to classify the samples among groups of study are highlighted in bold. a: “Control” vs “Altered P1/P2”, p-value < 0.05; b: “Control” vs “Obesity”, p-value < 0.0001; c: “Control” vs “Advanced age”, p-value < 0.01.

Table 2. Proteoform spectrum matches (PrSMs) identified by Top-Down proteomics of purified human protamine extracts.

Node	PrSMs (before filters)	PrSMs (after filters)	PrSMs Clustered	PrSMs Classified
PS	19 235	18 358	17 965	17 461
PS (%)	100.0	95.4	93.4	90.8
TP	17 319	16 505	14 268	13 228
TP (%)	100.0	95.3	82.4	76.4
Total	36 554	34 863	32 233	30 689
Total (%)	100.0	95.4	88.2	84.0

PS: ProSight PD, TP: TopPIC suite