Document downloaded from:

[\[http://redivia.gva.es/handle/20.500.11939/6374\]](http://redivia.gva.es/handle/20.500.11939/6374)

This paper must be cited as:

[Casares-Crespo, L., Fernández-Serrano, P., & Viudes-de-Castro, M. P. (2019). Proteomic characterization of rabbit (*Oryctolagus cuniculus*) sperm from two different genotypes. *Theriogenology*, 128, 140-148.]

ivia
Institut Valencià
d'Investigacions Agràries

The final publication is available at

[\[https://doi.org/10.1016/j.theriogenology.2019.01.026\]](https://doi.org/10.1016/j.theriogenology.2019.01.026)

Copyright [Elsevier]

**PROTEOMIC CHARACTERIZATION OF RABBIT (*Oryctolagus cuniculus*) SPERM
FROM TWO DIFFERENT GENOTYPES**

Lucía Casares-Crespo, Paula Fernández-Serrano and María P. Viudes-de-Castro^{*}

Animal Technology and Research Center (CITA), Instituto Valenciano de Investigaciones
Agrarias (IVIA), Polígono La Esperanza nº 100, 12400 Segorbe (Castellón), Spain.

^{*}Corresponding author. Tel.: +34-964-712-166.

E-mail address: viudes_mar@gva.es

Abstract

The present study was conducted to characterise rabbit sperm proteins focusing on the influence of the genetic origin. Six samples were recovered during two months from five males from genotype A (New Zealand White origin) and five from genotype R (California origin). Sperm proteins were extracted and subjected to in-gel digestion nano LC-MS/MS and bioinformatics analysis. The resulting library included 487 identified proteins validated with $\geq 95\%$ Confidence (unused Score ≥ 1.3). All the identified proteins belonged to *Oryctolagus cuniculus* taxonomy. These data are available via ProteomeXchange with identifier PXD007989. Only 7 proteins were specifically implicated in reproductive processes according to Gene Ontology annotation. Regarding the comparison of the sperm proteins abundance between genotypes, forty proteins were differentially expressed. Among them, 25 proteins were over-expressed in genotype A, while 15 proteins were over-expressed in genotype R. In conclusion, this study characterizes for the first time rabbit sperm proteins and provides evidence that genotype is related to a specific abundance of spermatozoa proteins.

Keywords: rabbit, spermatozoa, proteome, genotype, LC-MS/MS.

1. Introduction

New advances in proteomics are having a major impact on our understanding of how spermatozoa acquire their capacity for fertilization [1]. Spermatozoa cell is one of the most highly differentiated cells and is composed of a head with a highly compacted chromatin structure and a large flagellum with midpiece that contains the required machinery for movement and therefore to deliver the paternal genetic and epigenetic content to the oocyte [2]. By being so highly differentiated, spermatozoa are advantageous cells to study proteomics of specific compartments such as the membrane, which basically is the area of major importance for its role in interacting with the surroundings and the oocyte [3]. The fusion of a sperm and an oocyte is a sophisticated process that must be preceded by suitable changes in the sperm's membrane composition [4]. Recent studies of spermatozoa from the proteomic point of view have allowed the identification of different proteins in spermatozoa that are responsible for the regulation of normal/defective sperm functions [5].

While several techniques are available in proteomics, LC-MS based analysis of complex protein/peptide mixtures has turned out to be a mainstream analytical technique for quantitative proteomics [6]. Using this method, detailed proteomic data are now available for human [7], macaque [8,9], mouse [10], rat [11], bull [12,13,14], stallion [15], fruit fly [16], *Caenorhabditis elegans* [17], carp [18], rainbow trout [19], mussel [20], ram [21], honeybee [22], rooster [23] and sika deer [9] sperm proteins.

Rabbit (*Oryctolagus cuniculus*) is an important mammalian species worldwide, being at the same time of commercial interest and a research model animal. In a previous work, we identified and quantified rabbit seminal plasma proteins between two different genotypes [24], concluding the clear effect of genotype in the abundance of certain seminal plasma proteins. However, it is unknown at present whether these differences also exist at sperm

proteome level. Therefore, the aim of the present study was to characterise rabbit sperm proteins through nano LC-MS/MS analysis focusing on the influence of the genetic origin.

2. Materials and methods

Unless stated otherwise, all chemicals in this study were purchased from Sigma-Aldrich Química S.A (Madrid, Spain). All the experimental procedures used in this study were performed in accordance with Directive 2010/63/EU EEC for animal experiments.

2.1. Localization and animals

The experiment was carried out with 20 males from two Spanish commercial rabbit genetic lines (genotypes A and R) from March to April 2017. Line A is based on New Zealand White rabbits selected since 1980 by a family index for litter size at weaning over 45 generations. Line R comes from the fusion of two lines, one founded in 1976 with Californian rabbits reared by Valencian farmers and another founded in 1981 with rabbits belonging to specialised paternal lines. All bucks were of proven fertility and subjected to a weekly pattern of ejaculate collection. All animals were housed at the Animal Technology and Research Centre (CITA-IVIA, Segorbe, Castellón, Spain) experimental farm in flat deck indoor cages (75×50×30 cm), with free access to water and commercial pelleted diets. The photoperiod was set to provide 16 h of light and 8 h of dark, and the room temperature was regulated to keep temperatures between 14°C and 28°C.

2.2. Semen collection, evaluation and sperm samples preparation

Semen samples were obtained by artificial vagina and collected into a sterile tube. One ejaculate was collected per male and week. Collections were performed on the same day of the week during two months. For each genotype, one pooled sample each week was selected

to develop the experiment (Figure 1). A total of 6 pools (3 for each genetic line) were analysed. Quality semen parameters (concentration, sperm abnormality, acrosome integrity and sperm motility) were performed using the methodology described by Viudes de Castro et al [25] to assess the initial seminal quality. Only ejaculates exhibiting a white colour and possessing motility rate higher than 70% were used in the experiment. Then ejaculates from the same genotype were pooled each day as a single sample. Each sperm sample was centrifuged at 7,400 x g for 10 min at 22 °C. The resulting pellets were washed twice by centrifugation at 900 x g for 10 min in PBS. Sperm proteins were extracted according to Casares-Crespo et al. [26] protocol with few modifications. Briefly, sperm pellets were resuspended in 1% SDS (w/v) in TCG (Tris-citrate-glucose supplemented with a 1% v/v protease inhibitor cocktail, P2714) and sonicated on ice 6 times for 5 s at 30% amplitude using an Ultrasonic Lab Homogenizer UP 100 H (Hielscher Ultrasonics GmbH). After sonication, the solution was kept in ice for 15 min and centrifuged for 10 min at 15,000 x g at 4°C. Protein lysates were stored at -80°C until analysis. Before the proteomic analysis, total protein concentration was quantified in triplicate by the bicinchoninic acid method (BCA) using BSA as standard protein [27] and seminal samples were adjusted to 5 µg/µL in saline solution.

2.3. In-gel digestion

The proteomic analysis was performed in the proteomics facility of SCSIE University of Valencia that belongs to ProteoRed, PRB2-ISCI, supported by grant PT13/0001. Thirty µg of total protein was loaded onto a 1-D SDS PAGE gel but not resolved [28]. The entire sample was cut and analyzed as a single band. Samples were digested with sequencing grade trypsin (Promega) as described elsewhere [29]. Six hundred ng of trypsin in 100µL of ammonium bicarbonate (ABC) solution was used for each sample. The digestion was stopped

with trifluoroacetic acid (Fisher Scientific; 1% final concentration); a double extraction with acetonitrile (ACN) (Fisher Scientific) was done and all the peptide solutions were dried in a rotatory evaporator. The final mixture was resuspended with 30 μ L of 2% ACN; 0.1% trifluoroacetic acid (TFA).

2.4. Nano LC-MS/MS analysis

Three μ L of each sample were loaded onto a trap column (nano LC Column, 3 μ m particles size C18-CL, 350 μ m diameter x 0.5mm long; Eksigent Technologies) and desalted with 0.1% TFA at 3 μ L/min during 5 min. The peptides were then loaded onto an analytical column (LC Column, 3 μ m particles size C18-CL, 75 μ m diameter x12cm long, Nikkyo) equilibrated in 5% acetonitrile (ACN) 0.1% formic acid (FA). Peptide elution was carried out with a linear gradient of 5% to 35% of solvent B in A for 60 min. (A: 0.1% FA; B: ACN, 0.1% FA) at a flow rate of 300nL/min. Peptides were analysed in a mass spectrometer nano ESIqQTOF (5600 TripleTOF, ABSCIEX).

Eluted peptides were ionized applying 2.8 kV to the spray emitter. The mass spectrometric analysis was carried out in a data-dependent mode. Survey MS1 scans were acquired from 350–1250 m/z for 250 ms. The quadrupole resolution was set to ‘UNIT’ for MS2 experiments, which were acquired from 100–1500 m/z for 25 ms in ‘high sensitivity’ mode. Following switch criteria were used: charge: 2+ to 5+; minimum intensity; 70 counts per second (cps). Up to 25 ions were selected for fragmentation after each survey scan. Dynamic exclusion was set to 15 s. The system sensitivity was controlled with 2 fmol of 6 proteins mixture (LC Packings). Samples were injected in a random order.

The proteomics data and result-files from the analysis have been deposited to the ProteomeXchange Consortium via the PRIDE [30] partner repository, with the dataset identifier PXD007989 and 10.6019/PXD007989.

2.5. Protein identification

The SCIEX.wiff data-files were processed using Protein Pilot v5.0 search engine (AB SCIEX). ProteinPilot default parameters were used to generate peak list directly from 5600 TripleTofwiff files. The Paragon algorithm of Protein Pilot v5.0 was used to search UniProt_mammals' protein database (22/06/2017) with the following parameters: trypsin specificity, cys-alkylation, taxonomy restricted, FDR (False Discovery Rate) calculation and the search effort set to through. For building the spectra library all the files combined were searched with the parameters previously used.

To avoid using the same spectral evidence in more than one protein, the identified proteins are grouped based on MS/MS spectra by the Protein-Pilot Progroup algorithm. A protein group in a Pro Group Report is a set of proteins that share some physical evidence. Unlike sequence alignment analyses where full length theoretical sequences are compared, the formation of protein groups in Pro Group is guided entirely by observed peptides only. Since the observed peptides are actually determined from experimentally acquired spectra, the grouping can be considered to be guided by usage of spectra. Then, unobserved regions of protein sequence play no role in explaining the data. Only peptide and protein identifications with $\geq 95\%$ Confidence (unused Score ≥ 1.3) were validated. Protein identifications were accepted if they contained at least two identified peptides.

2.6. Label-free protein quantification using Chromatographic Areas

For quantification, the group file generated by Protein Pilot was used. The ions areas were extracted from the wiff files obtained from LC-MS/MS experiment by Peak View® v1.1. Only peptides assigned with confidence $\geq 95\%$, among those without modifications or

shared by different proteins were extracted. A total of 6 samples were analysed and 487 proteins were quantified.

2.7. Bioinformatics analysis

Gene names of the proteins were obtained from UniProt database by running a Retrieve/ID mapping tool of the protein accession numbers. Gene ontology terms for biological process, molecular function and cellular component were obtained using PANTHER v12.0 (<http://www.pantherdb.org/> accessed on 16/08/2017) [31]. Gene names of all sperm proteins were used to search against the panther database with *Homo sapiens* as the organism to maximise classifications, and a PANTHER™ Go Slim analysis was performed for each category. KEGG pathway enrichment analysis was undertaken using DAVID [32]. Protein-protein interactions and functional enrichment of the differentially expressed proteins were predicted using Search Tool for the Retrieval of Interacting Genes/Proteins [33]. Network analysis was set at medium stringency (STRING score=0.4). Proteins were linked based on seven criteria: neighbourhood, gene fusion, co-occurrence, co-expression, experiments, databases and text mining.

2.8. Statistical analysis

The quantitative data obtained by PeakView® were analysed by Marker View®v1.3 (AB Sciex). First, areas were normalized by total areas summa, and then a Discriminant Analysis (DA) and a T-Test were done. Proteins were considered differentially expressed between genotypes if the adjusted p-value < 0.05. Mean quantity of proteins were calculated and the fold-changes between the two groups were estimated. No multiple corrections were performed. The standard deviation was pooled out by calculating a separate t value for each

peak. Finally, differentially expressed proteins between genotypes were represented in a heat map using RStudio v.3.5.

3. Results

The seminal pools presented an average sperm concentration of 447×10^6 permatozoa/mL. Total motility rate, progressive motility rate, percentage of spermatozoa with a normal apical ridge and viability rate are shown in table 1.

3.1 Rabbit sperm proteome

The Proteomics System Performance Evaluation Pipeline (PSPEP) Software was used to perform a false discovery rate analysis on ParagonTM algorithm results. The complete spectral library included 487 proteins validated with $\geq 95\%$ Confidence (unused Score ≥ 1.3) when using at least 2 peptides for identification (Table S1). All the identified proteins belonged to *Oryctolagus cuniculus* taxonomy. These 487 proteins were quantified based on their chromatographic or peak areas (Table S2).

The complete rabbit sperm proteome was classified under different categories based on their molecular function, biological process and cellular components (PANTHER analysis). The results are shown in Figure 2. For molecular function (Fig. 2a), a total of 343 hits were found. The catalytic activity was the predominant function (49%), followed by binding (26%) and structural molecule activity (14%). Regarding biological process (Fig. 2b), a total of 669 hits were found. The cellular process (29%) and the metabolic (25%) were the most abundant categories, but it is worth mentioning that 8 hits (1%) were classified in the category of reproduction, 4 in the fertilization and the other 4 in gamete generation process. Finally, a total of 336 hits were found for cellular component category (Fig. 2c). Cell part (47%),

organelle (27%), macromolecular complex (15%) and membrane (6%) were the most abundant cellular components of the studied proteins.

3.2 Effect of genetic origin on spermatozoa proteome

The results of the sperm proteome comparison between both genotypes (A and R) are shown in Fig. 3. Discriminant Analysis (DA) showed a clear separation between samples from different genetic origin, classifying the six sperm samples into two different main clusters corresponding to both genotypes. Given that the proteome between both genotypes presented high variability, a t-test was done. Results showed a total of 40 differentially expressed proteins ($p < 0.05$) between genetic lines. Of the differentially expressed proteins, 25 proteins were over-expressed in genetic line A and 15 proteins over-expressed in line R (Table 2). The hierarchical clustering of differentially expressed sperm proteins separated the six samples into two different main clusters, differentiating between genotypes (Fig. 4).

The most significant metabolic pathway which were enriched in the differentially expressed sperm proteins between genotypes were glycolysis/gluconeogenesis, carbon metabolism, lysosome and hypertrophic cardiomyopathy. The protein-protein interactions between the differentially expressed proteins were performed by using STRING. The resulting protein network is shown in Fig. 5. Twenty-nine proteins were found to be linked either directly or indirectly with a total of fifty-six edges between them. Proteins are represented as nodes. Small nodes indicate proteins of unknown 3D structure and coloured nodes show query proteins and first shell of interactors. The number of coloured edges between proteins indicates the strength of data support, thereby; more lines between nodes represent stronger associations. Light blue and pink lines mean known interactions from curated databases and experimentally determined, respectively. Predicted interactions by gene neighbourhood are shown in dark green, by gene fusions in red and by gene co-occurrence in dark blue.

Interactions obtained by textmining are shown in light green, by co-expression in black and by protein homology in lilac.

Functional enrichment of this string network showed several biological processes (GO) related to reproduction: binding of sperm to zona pellucida, fertilization, reproduction, single organism reproductive process and single fertilization.

4. Discussion

To the best of our knowledge this is the first study in which rabbit sperm proteins are characterised. As a consequence of the lack of knowledge of rabbit proteomics in comparison with other mammalian species, among the 487 identified proteins 325 were catalogued as uncharacterized proteins. In UniProtKB database the 98% of the protein sequences have been derived from cDNA or genomic sequencing, thus most of the available protein sequences are reliant on the quantity and quality of DNA or RNA-derived information for that species [34]. As a result, studies of the reproductive proteome to date have been limited to model species such as human, mice and fruit fly [10,16,35] supported by extensive, high quality genomic information or with dedicated genome projects such as honeybee [36] among which rabbit species is not included. In addition, Bayram et al. [34] studied the species origin of database matches using mammalian proteome database search in UniProtKB and the dominant species were few: human, mouse, rat, sheep and cattle, which account for 87.5% of the sequences entries. Therefore, the fact that rabbit is a non-model species and lacks a fully annotated genome explains the high number of uncharacterised proteins found in rabbit spermatozoa in the present work. To solve this drawback, all protein accession numbers were translated into gene names in order to analyse the results.

Bioinformatics analysis of rabbit spermatozoa proteome revealed that 49% of identified proteins were related to catalytic activity and the second dominant group of proteins were

assigned to a binding function (26%). These proportions agree with previous proteomic studies of carp and honeybee spermatozoa [22,18]. Regarding biological process, the cellular process (29%) was the most abundant category in rabbit spermatozoa followed by metabolic process (25%), which coincides with honeybee sperm proteins [22]. It is also noticeable that only seven of the 487 proteins identified in rabbit spermatozoa are to date recorded in GO as being directly associated with reproductive processes such as fertilization and gamete generation. These proteins related with reproductive processes are the following: acrosin-binding protein (ACRBP), zonadhesin (ZAN), sperm equatorial segment protein 1 (SPESP1), serine protease inhibitor Kazal type 8 (SPINK8), dual specificity testis-specific protein kinase 2 (TESK2), dynein heavy chain 9 (DNAL1) and four and a half LIM domains protein 1 (FHL1). It is surprising that so few spermatozoa proteins are found in the category of reproduction and that many of the proteins studied in this work have not been assigned a specific role in GO. In previous studies in boar and in rabbit seminal plasma, authors encountered the same situation [37,25]. Despite of this, we are sure that the identified sperm proteins are directly or indirectly related to reproduction processes. Finally, GO analysis revealed that the majority of rabbit spermatozoa proteins were extracted from cell parts, organelles and membranes.

Among the twenty-five more abundant proteins in genotype A, we found proteins related to different biological functions such as: protein folding, glycolysis, protein transport, metabolism, ion transport and fertilization.

CCT8, CCT3, CCT2 and CCT6B are chaperones that contain the TCP1 complex and whose function is to assist the folding of other proteins such as tubulin and actin upon ATP hydrolysis [<http://www.uniprot.org/uniprot/G1SHZ8>]. In addition, arylsulfatase A protein (ARSA), has been localized in the acrosomal region of human spermatozoa and it is known to increase its surface expression significantly during capacitation [38]. These proteins have

been related to sperm zona pellucida interaction role in mouse and human spermatozoa by several authors [38,39,40]. On the other hand, zonadhesin protein' (ZAN) role is to mediate species-specific zona pellucida adhesion [41] and has been localized in the anterior acrosome of rabbit spermatozoa [42]. In a previous work [25], zonadhesin, was also found more abundant in seminal plasma of rabbit A genotype compared to genotype R. Furthermore, other proteins over-expressed in genotype A are: cullin 3 (CUL3), which has been recently found to have an important role in the human sperm flagellum [43], phosphoglycerate kinase 2 (PGK2), that has shown to be essential for sperm motility and male infertility in mice [44] and SERPINE2, which may play a role as a decapacitation factor [45]. All of these findings, especially the increased amount of these proteins observed in genotype A in comparison with genotype R could explain in part the enhanced fertility and prolificacy previously described in genotype A [46,47].

On the other hand, among the fifteen more abundant proteins in genotype R, we found proteins related to different biological functions such as: antioxidant activity, binding, catalytic activity, transporter activity and structural molecule activity.

For instance, the protein HSPB1 belongs to a superfamily of mammalian small stress proteins and its function has been suggested to be related to the cytoskeleton [48]. Just like ACTB protein, which is an actin involved in cell motility and cytokinesis [49]. On the other hand, peroxiredoxins such as PRDX1 are highly sensible to oxidative stress proteins which are involved in the antioxidant protection of the mammalian spermatozoa [50]. Finally, BASP1 protein, which has been located in rat mature spermatozoa, may be participating in biochemical processes through the activation of calcium [51].

5. Conclusions

In summary, our data provide evidence that genotype has a huge impact on protein abundance in rabbit sperm. The present work together with the previous study of rabbit seminal plasma, leads to the complete characterization of rabbit semen. Further studies are needed in order to elucidate the reproductive role of these identified proteins and the different evolution of these genotypes that gives rise to the intraspecies variation.

Author contributions

LCC was responsible for acquisition, analysis and drafting the article as part of her doctoral thesis; PFS assisted in the acquisition, as well as in analysis of data. MPVC conceived, designed and supervised the study and also participated in the interpretation of data and manuscript development. All authors read and approved the final manuscript.

Competing interest

The author declares that they have not competing interest.

Acknowledgements

This research was supported in part by the RTA2013-00058-00-00 from INIA, the European Social Fund and the European FEDER Funds. L. Casares-Crespo is supported by a scholarship from Instituto Valenciano de Investigaciones Agrarias (IVIA) and the European Social Fund. P. Fernández-Serrano is supported by Spanish funds from IVIA and Ministerio de Empleo y Seguridad Social (Youth Guarantee Program).

References

1. Aitken RJ, Baker MA. The role of proteomics in understanding sperm cell biology. *Int J Androl* 2008; 31: 295-302.

- 325 2. Codina M, Estanyol JM, Fidalgo MJ, Ballescà JL, Oliva R. Advances in sperm
326 proteomics: best-practice methodology and clinical potential. *Expert Rev Proteomics*
327 2015; 12: 255-277.
- 328 3. Rodríguez-Martínez H, Kvist U, Ernerudh J, Sanz L, Calvete JJ. Seminal plasma proteins:
329 what role do they play? *Am J Reprod Immunol* 2011; 66: 11-22.
- 330 4. Nowicka-Bauer K, Kurpisz M. Current knowledge of the human sperm proteome. *Expert*
331 *Rev Proteomics* 2013; 10, 591-605.
- 332 5. Rahman M S, Lee JS, Kwon WS, Pang MG. Sperm proteomics: road to male fertility and
333 contraception. *Int J Endocrinol* 2013; 360986.
- 334 6. Tuli L, Ressom HW. LC-MS based detection of differential protein expression. *J*
335 *Proteomics Bioinform* 2009; 2:416-38.
- 336 7. Baker MA, Reeves G, Hetherington L, Müller J, BaurI, Aitke, RJ. Identification of gene
337 products present in Triton X100 soluble and insoluble fractions of human spermatozoa
338 lysates using LC-MS/MS analysis. *Proteom Clin Appl* 2007; 1:524-532.
- 339 8. Skerget S, Rosenow M, Polpitiya A, Petritis K, Dorus S, Karr TL. The Rhesus macaque
340 (*Macaca mulatta*) sperm proteome. *Mol. Cell. Proteomics* 2013; 12: 3052-3067.
- 341 9. Kawase O, Cao S, Xuan X. Sperm membrane proteome in wild Japanese macaque
342 (*Macaca fuscata*) and Sika deer (*Cervus nippon*). *Theriogenology* 2015; 83: 95-102.
- 343 10. Baker MA, Hetherington L, Reeves GM, Aitken RJ. The mouse sperm proteome
344 characterized via IPG strip prefractionation and LC-MS/MS identification. *Proteomics*
345 2008; 8:1720-1730.
- 346 11. Baker MA, Hetherington L, Reeves G, Müller J, Aitken RJ. The rat sperm proteome
347 characterized via IPG strip prefractionation and LC-MS/MS identification. *Proteomics*
348 2008; 8:2312-2321.

349 12. Byrne K, Leahy T, McCulloch R, Colgrave ML, Holland MK. Comprehensive mapping
350 of the bull sperm surface proteome. *Proteomics* 2012; 12:3559-3579.

351 13. Ashrafzadeh A, Nathan S, Karsani SA. Comparative analysis of Mafriwal (*Bos Taurus* ×
352 *Bos indicus*) and Kedah Kelantan (*Bos indicus*) sperm proteome identifies sperm proteins
353 potentially responsible for higher fertility in a tropical climate. *Int J Mol Sci* 2013;
354 14:15860-15877.

355 14. Somashekar L, Selvaraju S, Parthipan S, Ravindra JP. Profiling of sperm proteins and
356 association of sperm PDC-109 with bull fertility. *Syst biol reprod med* 2015; 61: 376-387.

357 15. Swegen A, Curry BJ, Gibb Z, Lambourne SR, Smith ND, Aitken RJ. Investigation of the
358 stallion sperm proteome by mass spectrometry. *Reproduction* 2015; 149: 235-244.

359 16. Wasbrough ER, Dorus S, Hester S, Howard-Murkin J, Lilley K, Wilkin E, Polpitiya A,
360 Petritis K, Karr TL. The *Drosophila melanogaster* sperm proteome-II (DmSP-II). *J*
361 *Proteomics* 2010; 73: 2171-2185.

362 17. Ma X, Zhu Y, Li C, Xue P, Zhao Y, Chen S, Yang F, Miao L. Characterisation of
363 *Caenorhabditis elegans* sperm transcriptome and proteome. *BMC Genomics* 2014;
364 15:168.

365 18. Dietrich MA, Dietrich GJ, Mostek A, Ciereszko A. Motility of carp spermatozoa is
366 associated with profound changes in the sperm proteome. *J Proteomics* 2016; 138:124-
367 135.

368 19. Nynca J, Arnold GJ, Fröhlich T, Otte K, Ciereszko A. Proteomic identification of rainbow
369 trout sperm proteins. *Proteomics* 2014; 14:1569-1573.

370 20. Zhang Y, Mu H, Lau SC, Zhang Z, Qiu J. Sperm proteome of *Mytilus gallo provincialis*:
371 Insights into the evolution of fertilization proteins in marine mussels. *Proteomics* 2015;
372 15: 4175-4179.

- 373 21. Pini T, Leahy T, Soleilhavoup C, Tsikis G, Labas V, Combes-Soia L, Harichaux G,
374 Rickard JP, Druart X, de Graaf SP. Proteomic Investigation of Ram Spermatozoa and the
375 Proteins Conferred by Seminal Plasma. *J Proteome Res* 2016;15:3700-3711.
- 376 22. Zareie R, Eubel H, Millar AH, Baer B. Long-term survival of high quality sperm: insights
377 into the sperm proteome of the honeybee *Apis mellifera*. *J Proteome Res* 2013; 12: 5180-
378 5188.
- 379 23. Labas V, Grasseau I, Cahier K, Gargaros A, Harichaux G, Teixeira-Gomes A, Alves S,
380 Bourin M, Gérard N, Blesbois E. Qualitative and quantitative peptidomic and proteomic
381 approaches to phenotyping chicken semen. *J Proteomics* 2015; 112:313-335.
- 382 24. Casares-Crespo L, Fernández-Serrano P, Vicente JS, Marco-Jiménez F, Viudes-de-Castro
383 MP. Rabbit seminal plasma proteome: the importance of the genetic origin. *Anim.*
384 *Reprod. Sci.* 2018; 189:30-42.
- 385 25. Viudes-de-Castro MP, Casares-Crespo L, Monserrat-Martínez A, Vicente JS.
386 Determination of enzyme activity in rabbit seminal plasma and its relationship with
387 quality semen parameters. *World Rabbit Sci.* 2015; 23: 247-253.
- 388 26. Casares-Crespo L, Fernández-Serrano P, Viudes-de-Castro MP. Effect of the detergent
389 used on the resulting rabbit sperm membrane proteome. XVII Jornadas Sobre Producción
390 Animal, Zaragoza, España, AIDA 2017; 383-385.
- 391 27. Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto
392 EK, Goeke NM, Olson BJ, Klenk DC. Measurement of protein using bicinchoninic acid.
393 *Anal. Biochem.* 1985; 150:76-85.
- 394 28. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass
395 spectrometric characterization of proteins and proteomes. *Nat. Protoc.* 2006;1:2856–2860.
- 396 29. Shevchenko A, Jensen ON, Podtelejnikov AV, Sagliocco F, Wilm M, Vorm O, Mortensen
397 P, Shevchenko A, Boucherie H, Mann M. Linking genome and proteome by mass

spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc. Natl. Acad. Sci. 1996;93(25):14440-14445.

30. Vizcaíno JA, Csordas A, del-Toro N, Dianes JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, Xu QW, Wang R, Hermjakob H. 2016 update of the PRIDE database and related tools. Nucleic Acids Res 2016; 44(D1): D447-D456.

31. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Res 2017; 45: D183-D189.

32. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57.

33. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017; 45:D362-368.

34. Bayram HL, Claydon AJ, Brownridge PJ, Hurst JL, Mileham A, Stockley P, Beynon RJ, Hammond DE. Cross-species proteomics in analysis of mammalian sperm proteins. J Proteomics 2016; 135: 38-50.

35. Amaral A, Castillo J, Ramalho-Santos J, Oliva R. The combined human sperm proteome: cellular pathways and implications for basic and clinical science. Hum Reprod Update 2014; 20: 40-62.

36. Poland V, Eubel H, King M, Solheim C, Harvey Millar A, Baer B. Stored sperm differs from ejaculated sperm by proteome alterations associated with energy metabolism in the honeybee *Apis mellifera*. Mol Ecol 2011; 20:2643-2654.

37. Pérez-Patiño C, Barranco I, Parrilla I, Valero ML, Martínez EA, Rodríguez-Martínez H, Roca J. Characterization of the porcine seminal plasma proteome comparing ejaculate portions. *J Proteomics* 2016; 142:15-23.
38. Redgrove KA, Nixon B, Baker MA, Hetherington L, Baker G, Liu DY, Aitken RJ. The molecular chaperone HSPA2 plays a key role in regulating the expression of sperm surface receptors that mediate sperm-egg recognition. *PLOS ONE* 2012; 7(11): e50851.
39. Dun MD, Smith ND, Baker MA, Lin M, Aitken RJ, Nixon B. The chaperonin containing TCP1 complex (CCT/TRiC) is involved in mediating sperm-oocyte interaction. *J Biol Chem* 2011; 286:36875-36887.
40. Redgrove KA, Anderson AL, Dun MD, McLaughlin EA, O'Bryan MK, Aitken RJ, Nixon B. Involvement of multimeric protein complexes in mediating the capacitation-dependent binding of human spermatozoa to homologous zonae pellucidae. *Dev Biol* 2011; 356:460-474.
41. Tardif S, Wilson MD, Wagner R, Hunt P, Gertsenstein M, Nagy A, Lobe C, Koop BF, Hardy DM. Zonadhesin is essential for species specificity of sperm adhesion to the egg zona pellucida. *J Biol Chem* 2010; 285:24863-24870.
42. Lea IA, Sivashanmugam P, O'Rand MG. Zonadhesin: characterization, localization, and zona pellucida binding. *Biol Reprod* 2001; 65:1691-1700.
43. Jumeau F, Chalmel F, Fernandez-Gomez FJ, Carpentier C, Obriot H, Tardivel M, Caillet-Boudin ML, Rigot JM, Rives N, Buee L, Sergeant N, Mitchell V. Defining the human sperm microtubulome: an integrated genomics approach. *Biol Reprod* 2017; 96:93-106.
44. Danshina PV, Geyer CB, Dai Q, Goulding EH, Willis WD, Kitto GB, McCarrey JR, Eddy EM, O'Brien DA. Phosphoglycerate kinase 2 (PGK2) is essential for sperm function and male fertility in mice. *Biol Reprod* 2010; 82:136-145.

445 45. Lu C, Lee RK, Hwu Y, Chu S, Chen Y, Chang W, Lin S, Li S. SERPINE2, a serine
446 protease inhibitor extensively expressed in adult male mouse reproductive tissues, may
447 serve as a murine sperm decapacitation factor. *Biol Reprod* 2011; 84:514-525.

448 46. Lavara R, Mocé E, Lavara F, Viudes-de-Castro MP, Vicente JS. Do parameters of seminal
449 quality correlate with the results of on-farm inseminations in rabbits? *Theriogenology*
450 2005; 64:1130-1141.

451 47. Safaa HM, Vicente JS, Lavara R, Viudes-de-Castro MP. Semen evaluation of two selected
452 lines of rabbit bucks. *World Rabbit Sci* 2008; 16:141-148.

453 48. Fontaine JM, Rest JS, Welsh MJ, Benndorf R. The sperm outer dense fiber protein is the
454 10th member of the superfamily of mammalian small stress proteins. *Cell Stress Chaperon*
455 2003;8: 62-69.

456 49. Zeng C, He L, Peng W, Ding L, Tang K, Fang D, Zhang Y. Selection of optimal reference
457 genes for quantitative RT-PCR studies of boar spermatozoa cryopreservation.
458 *Cryobiology* 2014; 68:113-121.

459 50. Liu Y, O'Flaherty C. In vivo oxidative stress alters thiol redox status of peroxiredoxin 1
460 and 6 and impairs rat sperm quality. *Asian J Androl* 2017; 19:73-79.

461 51. Mosevitsky MI, Snigirevskaya ES, Komissarchik YY. Immunoelectron microscopic study
462 of BASP1 and MARCKS location in the early and late rat spermatids. *Acta Histochem*
463 2012; 114:237-243.

Figure Legends

Figure 1. Experimental design scheme.

Figure 2. Pie charts showing the distribution of rabbit sperm proteins based on their a) molecular function, b) biological process and c) cellular component, using UniProt database in combination with PANTHER.

Figure 3. Discriminant Analysis (DA) showing the classification of spermatozoa protein samples from genotypes A and R, based on relative protein amount.

Figure 4. Heat map representing levels of differentially expressed sperm proteins between genetic origins A and R and hierarchical clustering, showing two main clusters comprising genotype A and R.

Figure 5. Connectivity of differentially expressed proteins determined by STRING.

481 **Figure 1**

482

483

484 **Figure 2**

509

510 **Figure 3**

511

512

513

Figure 4

Supporting Information

Supporting Information Table S1 contains the complete list of the 487 proteins identified in rabbit spermatozoa with a cut off of two unique peptides and validated with $\geq 95\%$ Confidence (unused Score ≥ 1.3).

Supporting Information Table S2 contains the complete list of the chromatographic areas of the 487 proteins identified in the two rabbit genotypes (3 replicates per sample).

Supporting Information Table S3 shows the results of the protein quantity T-test comparison between genotypes, including mean protein quantity, t-value, p-value, fold change and log (fold change) of the 487 quantified proteins.

545 **Table 1.** Rabbit sperm characteristics (LSM±S.E.).

Genotype	N	Acrosome integrity (%)	Sperm Viability (%)	Total sperm motility (%)	Progressive sperm motility (%)
A	9	84.84±2.8	70.45±3.4	67.56±3.4	37.78 ±3.8
R	9	88.63± 2.8	70.35±3.4	62.33±3.4	36.44 ±3.8

546 N: number of pools

547
548
549
550
551
552
553
554
555
556
557
558

559 **Table 2.** List of differentially expressed rabbit spermatozoa proteins, included in *Oryctolagus cuniculus* taxonomy, between genotypes A and R.

Biological process	Gene ID	Gene name	Mean protein amount		Log (FoldChange)	p-value
			Line A	Line R		
	BASP1	Brainacid soluble protein 1	24770,351	88227,532	-0.552	0.00256
	A0A0G2JH23	Uncharacterizedprotein	14742,829	39547,488	-0.429	0.00648
	USP5	Ubiquitincarboxyl-terminal hydrolase 5	19155,099	7991,209	0.380	0.00775
Protein folding, protein complex assembly	CCT8	T-complex protein 1 subunit theta	289998,762	197226,072	0.167	0.00824
	AKR1B1	Aldosereductase	2042126,661	684791,411	0.475	0.00885
Glycolysis	PGK2	Phosphoglyceratekinase 2	1853797,274	1018697,679	0.260	0.00950
Cellular component morphogenesis	FLNB	Filamin-B	80934,073	142157,789	-0.245	0.01146
	PSMA1	Proteasomesubunitalpha type- 1	43475,764	19905,528	0.339	0.01313
Intracellular protein transport, lipid metabolic process, lipid transport, receptor-mediated endocytosis	SORT1	Sortilin	34590,467	16905,899	0.311	0.01353
Anion transport, catabolic and cellular process	ATP11A	Probable phospholipid- transporting ATPase IH	7586,766	13297,619	-0.244	0.01413
DNA replication, cell cycle, macrophage activation	S100A8	Protein S100-A8	3106,240	42277,814	-1.134	0.01579

Protein folding, immune system process, response to stress	HSPB1	Heat shock protein beta-1	122006,550	270497,552	-0.346	0.01613
Protein folding, protein complex assembly	CCT3	T-complex protein 1 subunit gamma	190302,157	113244,540	0.225	0.01738
	PRDX1	Peroxisomal oxidoreductin-1	20463,159	48178,975	-0.372	0.01975
Cell cycle, endocytosis, exocytosis, intracellular protein transport	ACTC1	Actin, alpha-cardiac muscle 1	17064660,120	19815038,340	-0.065	0.02015
	RAB14	Ras-related protein Rab-14	11035,514	7647,710	0.159	0.02028
Generation of precursor metabolites and energy, carbohydrate metabolic process, cellular amino acid catabolic process	ME1	NADP-dependent malic enzyme	63596,915	24116,024	0.421	0.02191
	G1T1S4	Uncharacterized protein	7527,446	24318,855	-0.509	0.02279
Protein folding, protein complex assembly	CCT2	T-complex protein 1 subunit beta	139552,440	70079,970	0.299	0.02652
Mitosis, cellular component movement and organization, cytokinesis, intracellular protein transport, intracellular signal transduction	MYO1C	Unconventional myosin-Ic	9273,656	27610,368	-0.474	0.02678
Cellular process, nuclear transport, protein localization and targeting	CSE1L	Exportin-2	43538,784	31223,895	0.144	0.02712
Protein folding, protein complex assembly	CCT6B	T-complex protein 1 subunit zeta-2	46030,918	30228,978	0.183	0.02933

Ion transport, response to toxic substance	ARSA	ATPase ASNA1	235936,189	118217,646	0.300	0.02996
Protein folding, protein complex assembly	TCP1	T-complex protein 1 subunit alpha	280703,715	159641,857	0.245	0.03036
	PSMB5	Proteasomesubunit beta type-5	38988,095	20945,469	0.270	0.03086
Anion transport, immune system process, response to toxic substance	MPST	3-mercaptopyruvate sulfurtransferase	17474,356	7178,149	0.386	0.03302
Ferredoxin metabolic process, nitrogen compound metabolic process, respiratory electron transport chain	TXNRD1	Thioredoxinreductase 1, cytoplasmic	15003,186	4118,362	0.561	0.03578
Intracellular protein transport, receptor-mediated endocytosis	AP1G1	AP-1 complexsubunit gamma-1	4014,558	9804,403	-0.388	0.03596
Sulfurcompound metabolic process	SULT1C4	Sulfotransferase 1C4	207417,507	46024,524	0.654	0.03816
Muscle contraction	GSTP1	Glutathione S-transferase P	6118,846	24472,019	-0.602	0.03941
Cellcyle, endocytosis, exocytosis, intracelular protein transport	ACTB	Actin, cytoplasmic 1	16089029,870	17666884,880	-0.041	0.04093
	RSPH1	Radial spoke head 1 homolog	37959,409	25088,770	0.180	0.04116
Cell adhesion, cellular process, fertilization	ZAN	Zonadhesin	851340,621	657973,081	0.112	0.0424
	LY6K	Lymphocyteantigen 6K	17476,014	10019,657	0.242	0.04276
Gluconeogenesis, glycolysis	GPI	Glucose-6-phosphate	357966,171	251913,213	0.153	0.04322

isomerase						
Catabolic process, cellular process, cellular protein modification process, proteolysis	CUL3	Cullin-3	43018,851	27406,064	0.196	0.04441
	NT5C3A	Cytosolic 5'-nucleotidase 3A	23230,678	44228,290	-0.280	0.04754
	YWHAE	14-3-3 protein epsilon	187532,157	243472,884	-0.113	0.04769
Regulation of biological process	SERPINE2	Glia-derived nexin	5889,179	1960,672	0.478	0.04867
Cellular process, cellular protein modification process	KLHL10	Kelch-like protein 10	14523,219	9150,116	0.201	0.04907

560 Log Fold Change > 0 indicates protein overexpression in A genotype

561 Log Fold Change < 0 indicates protein overexpression in R genotype