Data in Brief 45 (2022) 108578

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

# Dataset of the sperm proteome of stallions with different motility

Gemma Gaitskell-Phillips[a], Francisco E. Martín-Cano[a], José M Ortiz-Rodríguez[a], Antonio Silva-Rodríguez[b], Eva da Silva-Álvarez[a], Maria C. Gil[a], Cristina Ortega-Ferrusola[a], Fernando J. Peña[a,*]

[a] *Laboratory of Equine Reproduction and Spermatology, Department of Animal Medicine, Veterinary Teaching Hospital, University of Extremadura, Cáceres, Spain*
[b] *SIPA University of Extremadura, Cáceres, Spain*

ARTICLE INFO

*Article history:*
Received 25 September 2021
Revised 29 July 2022
Accepted 30 August 2022
Available online 5 September 2022

Dataset link: Stallion sperm proteins in 10 stallions with different motility (Original data)

*Keywords:*
Spermatozoa
Proteomics
Horses
Motility
CASA

ABSTRACT

This paper provides a detailed set of data on how the stallion sperm proteome differs among stallions with different sperm motilities, although within normal ranges. Findings distinguish proteins that may help to identify stallions of superior sperm motility. Sperm proteins were analyzed using a UHPLC/MS/MS system comprising of an Agilent 1290 infinity series UHPLC coupled to an Agilent 6550 Q-TOF mass spectrometer (Agilent Technologies, Santa Clara, CA, USA). These data can be used to disclose potential targets to identify good sperm samples and to study specific pathways involved in the regulation of sperm motility. This data article is linked to the paper "Proteins involved in mitochondrial metabolic functions and fertilization predominate in stallions with better motility Journal of Proteomics 247:104335 doi: 10.1016/j.jprot.2021.104335".



DOI of original article: 10.1016/j.jprot.2021.104335
* Corresponding author.
*E-mail address:* fjuanpvega@unex.es (F.J. Peña).
*Social media:* @Fjpvega (F.J. Peña)

https://doi.org/10.1016/j.dib.2022.108578

## Specifications Table

| | |
|---|---|
| Subject | Veterinary Sciences; Reproductive Biology. |
| Specific subject area | Proteomics of stallion spermatozoa. |
| Type of data | Raw data |
| How data were acquired | Samples were analyzed using a UHPLC/MS/MS system consisting of an Agilent 1290 infinity series UHPLC coupled to an Agilent 6550 Q-TOF mass spectrometer (Agilent Technologies, Santa Clara, CA, USA) |
| Data format | Processed<br>Analyzed |
| Parameters for data collection | Equine spermatozoa from 10 fertile stallions with a median age of 10.8 years with different values of motility estimated using computer assisted sperm analysis (CASA). Proteins were extracted from the spermatozoa of these stallions and analyzed using UHPLC/MS/MS. |
| Description of data collection | Comparative mass spectrometry proteomic profiling of the sperm proteome of fertile stallions with different motility. |
| Data source location | Cáceres, Extremadura, Spain. |
| Data accessibility | ProteomeXchange Consortium *via* the PRIDE partner repository with the dataset identifier PXD025807<br>ProteomeXchange title: Stallion sperm proteins in 10 stallions with different motility<br>ProteomeXchange accession: PXD025807<br>PubMed ID: 34298182<br>ProjectWebpage: https://www.ebi.ac.uk/pride/archive/projects/PXD025807<br>FTP download:<br>http://ftp.pride.ebi.ac.uk/pride/data/archive/2021/07/PXD025807/ |
| Related research article | Gaitskell-Phillips G, Martín Cano FE, Ortiz Rosdríguez JM, Silva Rodríguez A, Rojo-Domínguez P, Tapia JA, Gil MC, Ortega Ferrusola C, Peña FJ (2021). Proteins involved in metabolic functions and fertilization predominate in stallions with better motility Journal of Proteomics 247:104335<br>https://doi.org/10.1016/j.jprot.2021.104335 |

## Value of the Data

- The proteomic data reported here show that stallions with different sperm motility present notable differences in their proteome.
- Re-analysis of the data set may help to characterize traits related to sperm motility in stallions and the influence of factors such as breed or age in the motility.

## 1. Data Description

A data set of 903 proteins obtained from equine spermatozoa is provided (Supplementary Table 1). Proteins were extracted from spermatozoa from 10 fertile stallions, of a median age of 10.8 years old, of the following breeds; Arab, Andalusian, Andalusian-Arab, Spanish Sport Horse, Lusitanian and Anglo-Arabian. Relative abundance of proteins is also provided. These proteins were obtained from 10 stallions with different motility and classified as having good ($>$ 90% total motility), average (between 80 and 90% total motility), and poor $<$ 80% total motility), although all of them have sperm parameters within normal ranges [1] . Motility parameters were studied using Computer Assisted Sperm Analysis (CASA), including motility defined as the percentage of motile spermatozoa, linear motility, as the percentage of spermatozoa showing a linear pattern of movement, VCL defined as the actual velocity along the true trajectory in $\mu$m/s, VAP defined as the average velocity in $\mu$m/s, and the VSL defined as the straight-line velocity in $\mu$m/s. This information, after re-analysis, may be used to improve our understanding of the regulation of stallion sperm motility, and to develop on farm tests to check sperm quality. The data set has been deposited to the ProteomeXchange Con-

sortium [2] *via* the PRIDE partner repository with the dataset identifier PRIDE PXD025807, https://www.ebi.ac.uk/pride/archive/projects/PXD025807

## 2. Experimental Design, Materials and Methods

### *2.1. Sample Preparation*

The same protocol as described in [3–9] was used. Stallion ejaculates were washed in PBS (600gx 10'), then the pellet formed by the spermatozoa was frozen immediately in liquid nitrogen and kept frozen at -80 °C. Once the pellets were thawed the absence of debris and potential contaminating cells was assessed using phase contrast microscopy. For the analysis, proteins were extracted from 200 $\times$ $10^6$ spermatozoa. The protein extraction was initiated with the solubilization of the sperm proteins in the lysis buffer which composition follows, (C7:C7Bz0 [3-(4-heptyl) phenyl-(3-hydroxypropyl) dimethylammoniopropanesulfonate], 7M urea, 2M thiourea and 40 mM Tris (pH 10.4) as described in [1]. For the quantification of proteins we used the 2-D Quant Kit (GE Healthcare, Sevilla Spain) as indicated by the manufacturer, to normalize samples to 100 $\mu$g of protein. In the next step 200 µL of the solution obtained containing 100 $\mu$g of protein was mixed with 100 $\mu$l of 25 mM ammonium bicarbonate buffer pH 8.5 (100 $\mu$g of protein in 300 $\mu$L of solution). Then, proteins were reduced by adding 30 µL of 10 mM DTT and incubated at 56 °C for 20 min. In the next step, proteins were alkylated by adding 30 µL of 20 mM IAA solution, and incubated at room temperature in the dark for 30 min. The last step was digestion of the proteins, adding 2 $\mu$L of Trypsin Proteomics Grade (Sigma) (Trypsin solution: 1 g/L in 1 mM HCl) and incubating for at least 3 h at 37 °C. After incubation, the reaction was stopped with 10 $\mu$L of 0.1% formic acid, and the peptide solution was filtered through 0.2 $\mu$m (hydrophilic PTFE) into 2 mL dark glass vials. Finally, the protein samples were dried using a nitrogen current with the vial in a heating block at 35 °C. The dried samples were diluted in 20 $\mu$l of buffer A, (water/acetonitrile/formic acid (94.9:5:0.1)).

### *2.2. UHPLC-MS/MS Analysis*

The separation and analysis were conducted in an Agilent 1290 Infinity II Series UHPLC (Agilent Technologies, Santa Clara, CA, USA) equipped with an automated multisampler module and a High Speed Binary Pump, coupled to a Mass Spectrometer (Agilent 6550 Q-TOF, Agilent Technologies, Santa Clara, CA, USA) using an Agilent Jet Stream Dual electrospray (AJS-Dual ESI) interface, controlled by the MassHunter Workstation Data Acquisition software (Agilent Technologies, Rev. B.06.01), following previously published protocols [2,4,5,9]. The digested proteins were injected onto an Agilent Advance Bio Peptide Mapping HPLC column (2.7 $\mu$m, 150 $\times$ 2.1 mm, Agilent technologies), at 55 °C, at a flow rate of 0.4 ml/min. The gradient program was initiated with 2% of buffer B (water/acetonitrile/formic acid, 10:89.9:0.1) remaining in isocratic mode for 5 min, after this time, linearly was increased up to 45 % B for 40 min, and up to 95 % B over additional 15 min and finally remained constant for 5 min. After this initial 70 min run, 5 min followed, for the conditioning of the column for the next run using the initial condition. The mass spectrometer operated in positive mode. The nebulizer gas pressure was set to 35 psi, the drying gas flow was set to 10 l/min at a temperature of 250 °C, and the sheath gas flow was set to 12 l/min at 300 °C. The voltages of the capillary spray, fragmentor and octopole RF Vpp were, respectively, 3500 V, 340 V and 750 V. Profile data were acquired for both MS and MS/MS scans in extended dynamic range mode. MS and MS/MS mass ranges were 50–1700 m/z and scan rates were 8 spectra/sec for MS and 3 spectra/sec for MS/MS. Auto MS/MS mode was used with precursor selection by abundance and a maximum of 20 precursors selected per cycle. A ramped collision energy was used with a slope of 3.6 and an offset of -4.8. The same ion was rejected after two consecutive scans.

### 2.3. Data Processing

Data were analyzed using the Spectrum Mill MS Proteomics Workbench (Rev B.04.01, Agilent Technologies, Santa Clara, CA, USA) as described in previous publications from our laboratory [5,6]. The raw data were extracted under default conditions: selection of non fixed or variable modifications; [MH]+ 50–10,000 m/z; maximum precursor charge +5; retention time and m/z tolerance ± 60 s; minimum signal-to-noise MS (S/N) 25; finding $^{12}C$ signals. The MS/MS search against the updated protein database (Uniprot/Horse) was done as follows: non fixed modifications were selected and as variable modifications carbamidomethylated cysteines, tryptic digestion with 5 maximum missed cleavages, ESI-Q-TOF instrument, minimum matched peak intensity 50%, maximum ambiguous precursor charge +5, monoisotopic masses, peptide precursor mass tolerance 20 ppm, product ion mass tolerance 50 ppm and calculation of reversed database scores. The auto-validation strategy used was auto-threshold, in which the peptide score is automatically optimized for a target % FDR (1.2%), and immediately protein polishing validation is then performed to increase the sequence coverage of validated results with the restriction of a new maximum target protein FDR (0 %).

## 3. Ethics Statement

Stallions were maintained following specific institutional and European regulations for animal care (Law 6/2913 June 11th and European Directive 2010/63/EU). The ethical committee of the University approved this study REF IB-20008.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

## Data Availability

Stallion sperm proteins in 10 stallions with different motility (Original data) (ProteomeXchange).

## CRediT Author Statement

**Gemma Gaitskell-Phillips:** Methodology, Writing – review & editing; **Francisco E. Martín-Cano:** Methodology; **José M Ortiz-Rodríguez:** Methodology; **Antonio Silva-Rodríguez:** Methodology; **Eva da Silva-Álvarez:** Methodology; **Maria C. Gil:** Supervision, Formal analysis; **Cristina Ortega-Ferrusola:** Supervision, Formal analysis; **Fernando J. Peña:** Formal analysis, Funding acquisition, Writing – review & editing, Supervision, Validation.

## Acknowledgments

The authors received financial support for this study from the Ministerio de Economía y Competitividad-FEDER, Madrid, Spain, grants, AGL2017-83149-R Junta de Extremadura-FEDER (IB 20008 and GR 21060). JMOR holds a PhD grant from the Junta de Extremadura- FEDER (PD 18005), Mérida, Spain, GG-P holds a PhD grant from the Ministry of Science, Madrid, Spain (PRE2018-083354).

## Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2022.108578.

## References

[1] Gaitskell-Phillips G., Martín Cano F.E., Ortiz Rosdríguez J.M., Silva Rodríguez A., Rojo-Domínguez P., Tapia J.A., Gil M.C., Ortega Ferrusola C., Peña F.J. (2021). Stallion sperm proteins in 10 stallions with different motility. https://www.ebi.ac.uk/pride/archive/projects/PXD025807.

[2] G. Gaitskell-Phillips, F.E. Martin-Cano, J.M. Ortiz-Rodriguez, A. Silva-Rodriguez, E. da Silva-Alvarez, P. Rojo-Dominguez, et al., Proteins involved in mitochondrial metabolic functions and fertilization predominate in stallions with better motility, J. Proteom. 247 (2021) 104335.

[3] Y. Perez-Riverol, J. Bai, C. Bandla, D. Garcia-Seisdedos, S. Hewapathirana, S. Kamatchinathan, et al., The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences, Nucleic Acids Res. 50 (2022) D543–DD52.

[4] G. Gaitskell-Phillips, F.E. Martin-Cano, J.M. Ortiz-Rodriguez, A. Silva-Rodriguez, E. da Silva-Alvarez, M.C. Gil, et al., The seminal plasma proteins peptidyl arginine deaminase 2, rRNA adenine N (6)-methyltransferase and KIAA0825 are linked to better motility post thaw in stallions, Theriogenology 177 (2022) 94–102.

[5] G. Gaitskell-Phillips, F.E. Martin-Cano, J.M. Ortiz-Rodriguez, E. da Silva-Alvarez, J. Masot, E. Redondo, et al., Seminal plasma proteins as potential biomarkers for sperm motility and velocities, Theriogenology 177 (2022) 34–41.

[6] G Gaitskell-Phillips, FE Martin-Cano, JM Ortiz-Rodriguez, A Silva-Rodriguez, MC Gil, C Ortega-Ferrusola, et al., In stallion spermatozoa, superoxide dismutase (Cu-Zn) (SOD1) and the aldo-keto-reductase family 1 member b (AKR1B1) are the proteins most significantly reduced by cryopreservation, J. Proteome Res. 20 (2021) 2435–2446.

[7] G. Gaitskell-Phillips, F.E. Martin-Cano, J.M. Ortiz-Rodriguez, A. Silva-Rodriguez, M.C. Gil, C. Ortega-Ferrusola, et al., Differences in the proteome of stallion spermatozoa explain stallion-to-stallion variability in sperm quality post-thaw, Biol. Reprod. 104 (2021) 1097–1113.

[8] G. Gaitskell-Phillips, F.E. Martin-Cano, J.M. Ortiz-Rodriguez, A. Silva-Rodriguez, H. Rodriguez-Martinez, M.C. Gil, et al., Seminal plasma AnnexinA2 protein is a relevant biomarker for stallions which require removal of seminal plasma for sperm survival upon refrigerationd, Biol. Reprod. 103 (2020) 1275–1288.

[9] F.E. Martin-Cano, G. Gaitskell-Phillips, J.M. Ortiz-Rodriguez, A. Silva-Rodriguez, A. Roman, P. Rojo-Dominguez, et al., Proteomic profiling of stallion spermatozoa suggests changes in sperm metabolism and compromised redox regulation after cryopreservation, J. Proteom. 221 (2020) 103765.