

1 **Seminal plasma proteins as potential biomarkers for**  
2 **sperm motility and velocities**

3

4 Gemma Gaitskell-Phillips, Francisco E. Martín-Cano, José M Ortiz-Rodríguez, Eva da  
5 Silva-Álvarez, Javier Massot, Eloy Redondo, Maria C. Gil, Cristina Ortega-Ferrusola, \*  
6 Fernando J. Peña

7

8 *Laboratory of Equine Reproduction and Equine Spermatology, Veterinary Teaching*  
9 *Hospital, University of Extremadura, Cáceres, Spain.*

10

11

12

13 \*Correspondence to Dr. FJ Peña, Veterinary Teaching Hospital, Laboratory of Equine  
14 Spermatology and Reproduction, Faculty of Veterinary Medicine University of  
15 Extremadura Avd. de la Universidad s/n 10003 Cáceres Spain. E-mail  
16 [fjuanpvega@unex.es](mailto:fjuanpvega@unex.es)

17 phone + 34 927-257167

18 fax +34 927257102

19

20 **Acknowledgements**

21

22 The authors received financial support for this study from the Ministry of Science-  
23 FEDER, Madrid, Spain, grants, [AGL2017-83149-R](#) Junta de Extremadura-FEDER (GR  
24 18008 and IB 20008). JMOR holds a PhD grant from the Junta de Extremadura- FEDER  
25 (PD 18005), Mérida, Spain. GG-P holds a PhD grant from the Ministry of Science,  
26 Madrid, Spain (PRE2018-083354).

27

28

29

30

31 **Abstract**

32

33 Seminal plasma proteins have important roles in sperm functionality, and different  
34 mechanisms including micro-vesicle transport of proteins are involved in the regulation  
35 of sperm biology. Due to the role of seminal plasma, we hypothesized that specific  
36 proteins present in seminal plasma may be used as discriminant variables with potential  
37 to identify stallions producing different quality ejaculates; 10 fertile stallions, with  
38 different motility and velocity values (although within normal ranges) were used in this  
39 study. Motilities and velocities were studied using computer assisted sperm analysis  
40 (CASA), while protein composition of the seminal plasma was studied using UHPLC-  
41 MS/MS. Specific proteins were more abundant in samples with poorer percentages of  
42 total motility, average path velocity and circular velocity, and were: *Secreted*  
43 *phosphoprotein 1*, *Fructose-bisphosphate aldolase* ( $p=1,95E-09$ ;  $q=0.0005$ ) and *Malate*  
44 *dehydrogenase 1* ( $p=1,41E-11$ ;  $q=0.002$ ), to the contrary samples with better straight-  
45 line velocity values were enriched in *Glutathione peroxidase* ( $p=0.00013$ ;  $q=0.04$ ) and  
46 *Triosephosphate isomerase* ( $p=0.00015$ ;  $q=0.04$ ).

47

48 **Key words:** stallion, seminal plasma, proteomics, CASA

49

## 50 1.- Introduction

51

52 The stallion ejaculate is composed of a mixture of spermatozoa and the secretions from  
53 the accessory sex glands [1]. Numerous proteins are vehiculated in the seminal plasma,  
54 mainly in micro vesicles such as prostasomes; microRNAs are also vehiculated in this  
55 way [2, 3]. Seminal plasma influences sperm functionality in different ways, for example  
56 the presence of high amounts of Annexin A2 may impair the ability of the ejaculates to  
57 sustain prolonged conservation periods under refrigeration [4]. However seminal plasma  
58 also contains proteins that support sperm metabolism, probably through vesicle mediated  
59 transport [4]; this group of proteins support sperm metabolism and also help in the redox  
60 regulation of these cells. [2, 3, 5, 6]. Seminal plasma proteins may interact with the  
61 spermatozoa vehiculated in micro-vesicles; thus, these proteins may influence and  
62 regulate sperm functionality [6]. Nowadays, the use of mass spectrometry allows for the  
63 identification of numerous proteins in the spermatozoa and the seminal plasma, in  
64 addition to the increased availability of bioinformatic software and on line platforms  
65 which allow curation of the information gathered from large data sets [7-10]. We  
66 hypothesized that bioinformatic analysis of seminal plasma proteins may provide

67 candidates for biomarkers of sperm quality in stallions, with potential to be used as  
68 discriminant variables to forecast the quality of a particular ejaculate.

69

## 70 **2.- Material and methods**

71

### 72 ***2.1.- Reagents and media***

73

74 All chemicals were purchased from Sigma-Aldrich (Madrid, Spain), unless otherwise  
75 stated.

76

### 77 ***2.2.- Semen collection and processing***

78

79 Semen was collected from 10 stallions (three ejaculates per stallion) of various breeds  
80 maintained as indicated in institutional and European regulations for animal care (Law  
81 6/2913 June 11th and European Directive 2010/63/EU). Data on the reproductive record  
82 of the stallions used are presented in supplementary table 1. All stallions were of proven  
83 fertility, with a median age of 10.8 years old. Semen was collected from all stallions on a  
84 regular basis following the standard protocol in our center in which collections are  
85 performed in the morning between 10 and 12 h. Ejaculates used in this study were  
86 collected after depletion of the extragonadal sperm reserves, during the 2019 breeding  
87 season. The University ethics committee approved this study. Ejaculates were collected  
88 using a warmed, lubricated Missouri model artificial vagina, and the gel was removed  
89 with an inline filter. Semen was transported immediately to the laboratory after collection  
90 for evaluation and processing. Upon arrival at the laboratory, the seminal plasma (SP)  
91 was removed by serial centrifugation (2 x 1500g 10') and stored at -80° C until proteomic  
92 analysis. An aliquot of the ejaculate was processed through colloidal centrifugation [[11](#),  
93 [12](#)] to remove dead spermatozoa, and contaminating cells, and then was re-suspended in  
94 Tyrode's medium (20mM HEPES, 5mM Glucose, 96mM NaCl, 15mM NaHCO<sub>3</sub>, 1mM  
95 Na-Pyruvate, 21.6 mM Na-Lactate, 2mM CaCl<sub>2</sub>\*2H<sub>2</sub>O, 3.1mM KCl, 0.4mM  
96 MgSO<sub>4</sub>\*7H<sub>2</sub>O, 0.3mM NaH<sub>2</sub>PO<sub>4</sub>\*H<sub>2</sub>O, 0.3% BSA) 315 mOsm/kg and pH 7.4 [[13](#)],  
97 for assessment of motility and sperm velocities.

98

### 99 ***2.3.- Sample preparation***

100

101 Samples were processed immediately after collection. Aliquots of isolated SP were kept  
102 frozen at -80°C until further analysis. Phase contrast microscopy was used to control the  
103 absence of spermatozoa, moreover SP was filtered (0.22µM) before snap freezing and  
104 further processing.

105

#### 106 ***2.4.- Protein solubilization and quantification***

107

108 Aliquots of SP were solubilized in lysis buffer and incubated under constant rotation at  
109 4°C for 1 hour as described in previous studies [2, 14]. The amount of protein was then  
110 normalized to obtain a final concentration of 100 µg of protein per sample.

111

#### 112 ***2.5.- In-solution trypsin digestion.***

113

114 Trypsin digestion was performed as described in preceding studies [2, 14, 15]. In brief  
115 the proteins were mixed with a bicarbonate buffer, reduced with DTT and lastly alkylated.  
116 They were then digested by trypsin overnight.

117

#### 118 ***2.6.- UHPLC-MS/MS analysis.***

119

120 The separation and analysis of the samples were performed as previously described [2,  
121 14, 16] using an Agilent 1290 Infinity II Series UHPLC (Agilent Technologies, Santa  
122 Clara, CA, USA) equipped with an automated multisampler module and a high-speed  
123 binary pump coupled to an Agilent 6550 Q-TOF Mass Spectrometer (Agilent  
124 Technologies, Santa Clara, CA, USA) using an Agilent Jet Stream Dual electrospray  
125 (AJS-Dual ESI) interface. Control of the UHPLC and Q-TOF was using Mass Hunter  
126 Workstation Data Acquisition software (Agilent Technologies, Rev. B.06.01).

127

#### 128 ***2.7.- Data processing***

129

130 Data processing and analysis was performed using the Spectrum Mill MS Proteomics  
131 Workbench (Rev B.04.01, Agilent Technologies, Santa Clara, CA, USA) as previously  
132 described [14].

133

134 **2.8.- *Bioinformatic Analysis***

135

136 **2.8.1.- *Variance filtering and PCA***

137

138 Data were normalized and log<sub>2</sub> transformed using Qlucore Omics Explorer  
139 (<https://qlucore.com>) as described in previous publications [14]. Principal Component  
140 Analysis (PCA) was used to visualize the data set in a three-dimensional space, after  
141 filtering out variables with low overall variance to reduce the impact of noise and  
142 centering and scaling the remaining variables to zero mean and unit variance. The  
143 projection score [17] was used to determine the optimal filtering threshold.

144

145 **2.8.2.- *Identifying discriminating variables***

146

147 Qlucore Omics Explorer Ver. 3.7 (<https://qlucore.com>) was used to identify the  
148 discriminating variables able to find differences in motility, and velocities (VCL, VAP  
149 and VSL) among the stallions. This software works by fitting a linear model for each  
150 variable with condition proteins of the seminal plasma from stallions showing different  
151 values of the above parameters as predictors, including the stallion, breed, and age as  
152 nuisance covariates. P-values were adjusted for multiple testing using the Benjamini-  
153 Hochberg method [18, 19] and variables with adjusted p-values (q values, equivalent to  
154 false discovery rate FDR) below 0.1 were considered significant. Then, to further validate  
155 the potential biomarkers previously identified (seminal plasma proteins) we also used the  
156 Biomarker Workbench in Qlucore version 3.7.21; this functionality allows the  
157 simultaneous analysis of multiple variables. A model including, age, breed, individual  
158 stallion, and the computer assisted sperm analysis (CASA) derived parameters:  
159 percentage of total motile spermatozoa, percentage of linear spermatozoa, circular  
160 velocity (VCL), average path velocity (VAP) and straight-line velocity (VSL) was used.  
161 In this analysis variables were considered potential biomarkers only if q values were  
162 below 0.05.

163

164 **2.9.- *Computer-assisted sperm analysis (CASA)***

165

166 Sperm motility and velocities were assessed with a computer-assisted sperm analysis  
167 (CASA) system (ISAS Proiser, Valencia, Spain) [12, 20]. Samples were loaded into

168 Leja® chambers with a depth of 20 µm (Leja, Amsterdam, The Netherlands) and placed  
169 on a warmed stage at 38 °C. Sixty consecutive digitized images obtained using a 10x  
170 negative phase-contrast objective (Olympus CX 41), and 500 spermatozoa per sample  
171 were analyzed in random fields. Spermatozoa with an average path velocity (VAP) > 35  
172 µm/s were considered motile. Spermatozoa deviating < 45% from a straight line were  
173 classified as linearly motile.

174

## 175 **2.10.- Statistical analysis**

176

177 The following parameters were measured: percentages of total and linear motile  
178 spermatozoa, circular (VCL), straight line (VSL) and average (VAP) velocities in µm/s.  
179 The normality of the data was assessed using the Kolmogorov-Smirnoff test. Paired t-  
180 tests and one-way ANOVA followed by Dunnett's multiple comparisons test were  
181 performed using GraphPad Prism version 8.00 for Mac, La Jolla California USA,  
182 ([www.graphpad.com](http://www.graphpad.com)).

183

## 184 **3.- Results**

185

### 186 **3.1.- Identification of stallions with different motility and velocities**

187

188 The percentage of motile sperm ranged from  $74.3 \pm 6.7$  to  $92.0 \pm 1.1$  % ( $P < 0.05$ ; Figure  
189 1). Stallions showing ejaculates with median total motility over 90% were classified as  
190 “good”, those with total motility over 80% were classified as “average” and those with  
191 percentages below 80% were classified as “poor” (Figure 1A). The values for good,  
192 average, and poor percentages of linearly motile spermatozoa were 70% 60% and 40%  
193 respectively (Fig 1B). Sperm velocities also showed significant variation, in terms of  
194 circular velocity, good stallions had VCL values of  $212.3 \pm 6.7$  µm/s, for stallions  
195 classified as average VCL was  $181.1 \pm 6.17$  µm/s and for stallions classified as poor VCL  
196 was  $141.7 \pm 10.8$  µm/s (Fig 1C). Average path velocity in the good, average, and poor  
197 groups was  $117.3 \pm 1.8$ ,  $99.9 \pm 3.0$  and  $70.3 \pm 7.6$  respectively (Fig 1D). Finally, VSL was  
198  $81.5 \pm 3.7$  µm/s in good stallions,  $65.5 \pm 2.6$  µm/s in average and  $51.3 \pm 2.4$  µm/s in the  
199 poor group (Fig 1 E)

200

201 **3.2.- Seminal plasma proteins differ in stallions with good and poor motility and**  
202 **velocities**

203

204 Raw data were uploaded to the ProteomeXchange Consortium via the PRIDE partner  
205 repository with the dataset identifier PXD026375 [21] and the **Username:**  
206 reviewer\_pxd026375@ebi.ac.uk **Password:** CobS4ffZ. Volcano plots were constructed  
207 to obtain an overview of seminal plasma proteins differentially expressed in stallions  
208 showing poorer motility (Figure 2A), lower VCL and VAP (Figure 2 B and C) and better  
209 (faster) VSL (Figure 2D). Next, discriminant variables ( $p= 9.6e-04$ ;  $q=0.069$ ) were  
210 identified for % total motility for VCL and VAP ( $p= 8.7x e-4$ ,  $q=0.0069$ ) and for VSL  
211 ( $p= 0.001$ ;  $q=0.08$ ). We identified 6 proteins in seminal plasma with discriminant power  
212 for stallions with lower percentages of total motility in their ejaculates, then these proteins  
213 were further curated ( $q<0.05$ ) using the biomarker workbench (Table 1) to find the  
214 proteins with the highest potential as biomarkers, and *fructose-bisphosphate aldolase*  
215 ( $p=2.56e-5$ ;  $q=0.0070$ ), *secreted phosphoprotein 1* ( $p=2.58e-5$ ;  $q=0.0070$ ) and *malate*  
216 *dehydrogenase 1* ( $p=0.00015$ ;  $q= 0.028$ ) were identified as strong biomarkers of  
217 ejaculates with poor motility (Figure 3). Discriminant variables were also identified for  
218 VCL and VAP (Figures 4 and 5) and in both cases were found to be the same proteins as  
219 in the case of the percentages of total motile spermatozoa. However, in the case of  
220 straight-line velocity two proteins were identified as being discriminant variables for  
221 stallions with higher VSL (Table 1). Higher VSL values were better explained by two  
222 variables, *glutathione peroxidase* ( $p=0.00013$ ;  $q=0.04$ ) and *triosephosphate isomerase*  
223 ( $p=0.00015$ ;  $q=0.04$ ).

224

225 **4.- DISCUSSION**

226

227 Stallion ejaculates presenting percentages of total motility and velocities which were  
228 different, but within normal ranges, differed in the amounts of specific seminal plasma  
229 proteins, present in ejaculates. These findings should be considered in the context of in  
230 vitro situation. Many proteins present in seminal plasma are vehiculated in exosomes  
231 (epididysomes, prostasomes and vesiculosomes), in fact it has already been discovered  
232 that many seminal plasma proteins attach to stallions sperm membranes [22]. Although  
233 further research is needed, presence of these proteins in the seminal plasma may reflect  
234 the secretory activity of the male genital tract and thus its functionality [3, 6]. Specific

235 proteins were associated with samples showing lower percentages of total motility, VCL  
236 and VAP. To the contrary, two specific seminal plasma proteins were present in higher  
237 amounts in samples with higher VSL. Interestingly, the same discriminant proteins were  
238 able to predict samples which had poor motility, VCL and VAP. Fructose bisphosphate  
239 aldolase catalyzes the reversible reaction that splits fructose 1, 6 bisphosphate into  
240 dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3 -phosphate (GP3). These are  
241 intermediate metabolites in glycolysis, and in a subsequent step phosphate groups are  
242 eliminated. Irreversible elimination of the phosphate groups from DAHP and 3GP forms  
243 methylglyoxal (MG) as a by product; this is a 2-oxoaldehyde that is a strong electrophile  
244 due to adjacent carbonyl groups that rapidly and spontaneously react with nucleophiles  
245 from proteins, lipids and DNA **forming** advanced glycation end products (AGEs) [23].  
246 These compounds are potentially cytotoxic and mutagenic, although they may also be  
247 involved in regulatory functions. Besides, MG can form adducts with superoxide  
248 dismutase 1 (SOD1) impairing the antioxidant action of this enzyme and promoting  
249 oxidative stress [24]. In relation with this fact recent research from our laboratory  
250 identifies SOD1 as one of the most important antioxidant systems in the spermatozoa [14,  
251 25]. Furthermore we have recently described the toxic nature of MG for the stallion  
252 spermatozoa [26]. Extenders containing high amounts of glucose, produce high amounts  
253 of MG causing sperm malfunction including drops in motility and sperm velocities [26].  
254 This provides an explanation to our findings linking high amounts of this enzyme in  
255 seminal plasma and poor motility, VCL and VAP. While relevant potential biomarkers  
256 were found for poor motility, average and circular velocity, to the contrary biomarkers  
257 for good straight-line velocity were evidenced. *Glutathione peroxidase* and  
258 *triosephosphatase isomerase* were biomarkers of good VSL. A BLAST analysis showed  
259 that the glutathione peroxidase found in our study had an 89% homology to the  
260 *epididymal secretory glutathione peroxidase* from *Sus scrofa*, 85.5% to the same protein  
261 in *Canis lupus familiaris* and 79.3% to the same protein in *Homo sapiens*. This enzyme  
262 catalyzes the reduction of hydrogen peroxide, lipid peroxides and organic hydroperoxide  
263 by glutathione. This protein constitutes a glutathione peroxidase-like protective system  
264 against peroxide damage in sperm membrane lipids [27]. It is not surprising that a protein  
265 with roles in lipid peroxide detoxification relates to better characteristics of sperm  
266 kinematics. Spermatozoa are cells that need a tight redox regulation, and the loss of redox  
267 equilibrium rapidly leads to sperm malfunction [5, 14]. Interestingly, our findings may  
268 also suggest a different regulation of different aspects of sperm kinematics, suggesting



269 that motility, circular velocity, and average path velocities, may have similar regulation,  
270 while straight line velocity may have distinctive particularities.

271

272 In sum, seminal plasma proteins may have a major impact of sperm functionality, and  
273 specific seminal plasma proteins may be used as discriminant variables for poor motility,  
274 VCL and VAP, while discriminant variables for good VSL were also identified.  
275 Validation of these data and further research may help to develop potent biomarkers of  
276 sperm functionality which are rapidly applicable in clinical settings. These findings also  
277 underpin the role of seminal plasma in sperm functionality.

278

## 279 **5.- REFERENCES**

280

281 [1] Topfer-Petersen E, Ekhlasi-Hundrieser M, Kirchhoff C, Leeb T, Sieme H. The role  
282 of stallion seminal proteins in fertilisation. *Anim Reprod Sci.* 2005;89:159-70.

283 [2] Gaitskell-Phillips G, Martin-Cano FE, Ortiz-Rodriguez JM, Silva-Rodriguez A,  
284 Rodriguez-Martinez H, Gil MC, et al. Seminal plasma AnnexinA2 protein is a relevant  
285 biomarker for stallions which require removal of seminal plasma for sperm survival  
286 upon refrigeration. *Biol Reprod.* 2020.

287 [3] Barranco I, Padilla L, Parrilla I, Alvarez-Barrientos A, Perez-Patino C, Pena FJ, et  
288 al. Extracellular vesicles isolated from porcine seminal plasma exhibit different  
289 tetraspanin expression profiles. *Sci Rep.* 2019;9:11584.

290 [4] Gaitskell-Phillips G, Martin-Cano FE, Ortiz-Rodriguez JM, Silva-Rodriguez A,  
291 Rodriguez-Martinez H, Gil MC, et al. Seminal plasma AnnexinA2 protein is a relevant  
292 biomarker for stallions which require removal of seminal plasma for sperm survival  
293 upon refrigeration. *Biol Reprod.* 2020;103:1275-88.

294 [5] Pena FJ, O'Flaherty C, Ortiz Rodriguez JM, Martin Cano FE, Gaitskell-Phillips GL,  
295 Gil MC, et al. Redox Regulation and Oxidative Stress: The Particular Case of the  
296 Stallion Spermatozoa. *Antioxidants (Basel).* 2019;8.

297 [6] Rodriguez-Martinez H, Martinez EA, Calvete JJ, Pena Vega FJ, Roca J. Seminal  
298 Plasma: Relevant for Fertility? *Int J Mol Sci.* 2021;22.

299 [7] Griffin RA, Baker M, Aitken RJ, Swegen A, Gibb Z. What makes a fertile sperm?  
300 Unique molecular attributes of stallion fertility. *Reproduction.* 2019;158:R125-R37.

301 [8] Perez-Patino C, Parrilla I, Li J, Barranco I, Martinez EA, Rodriguez-Martinez H, et  
302 al. The Proteome of Pig Spermatozoa Is Remodeled During Ejaculation. *Mol Cell*  
303 *Proteomics*. 2019;18:41-50.

304 [9] Mlecnik B, Galon J, Bindea G. Comprehensive functional analysis of large lists of  
305 genes and proteins. *J Proteomics*. 2018;171:2-10.

306 [10] Codina M, Estanyol JM, Fidalgo MJ, Balleca JL, Oliva R. Advances in sperm  
307 proteomics: best-practise methodology and clinical potential. *Expert Rev Proteomics*.  
308 2015;12:255-77.

309 [11] Morrell JM, Garcia BM, Pena FJ, Johannisson A. Processing stored stallion semen  
310 doses by Single Layer Centrifugation. *Theriogenology*. 2011;76:1424-32.

311 [12] Ortega-Ferrusola C, Garcia BM, Gallardo-Bolanos JM, Gonzalez-Fernandez L,  
312 Rodriguez-Martinez H, Tapia JA, et al. Apoptotic markers can be used to forecast the  
313 freezeability of stallion spermatozoa. *Anim Reprod Sci*. 2009;114:393-403.

314 [13] Hall SE, Aitken RJ, Nixon B, Smith ND, Gibb Z. Electrophilic aldehyde products  
315 of lipid peroxidation selectively adduct to heat shock protein 90 and arylsulfatase A in  
316 stallion spermatozoa. *Biol Reprod*. 2017;96:107-21.

317 [14] Gaitskell-Phillips G, Martin-Cano FE, Ortiz-Rodriguez JM, Silva-Rodriguez A, Gil  
318 MC, Ortega-Ferrusola C, et al. In Stallion Spermatozoa, Superoxide Dismutase (Cu-Zn)  
319 (SOD1) and the Aldo-Keto-Reductase Family 1 Member b (AKR1B1) Are the Proteins  
320 Most Significantly Reduced by Cryopreservation. *J Proteome Res*. 2021.

321 [15] Martin-Cano FE, Gaitskell-Phillips G, Ortiz-Rodriguez JM, Silva A, Gil C, Ortega-  
322 Ferrusola C, et al. Data set of the proteome of fresh and frozen thawed stallion  
323 spermatozoa. *Data Brief*. 2020;31:105887.

324 [16] Martin-Cano FE, Gaitskell-Phillips G, Ortiz-Rodriguez JM, Silva-Rodriguez A,  
325 Roman A, Rojo-Dominguez P, et al. Proteomic profiling of stallion spermatozoa  
326 suggests changes in sperm metabolism and compromised redox regulation after  
327 cryopreservation. *J Proteomics*. 2020;221:103765.

328 [17] Fontes M, Soneson C. The projection score--an evaluation criterion for variable  
329 subset selection in PCA visualization. *BMC Bioinformatics*. 2011;12:307.

330 [18] Tamhane AC, Hochberg Y, Dunnett CW. Multiple test procedures for dose finding.  
331 *Biometrics*. 1996;52:21-37.

332 [19] Viskoper RJ, Laszt A, Oren S, Hochberg Y, Villa Y, Drexler I, et al. The  
333 antihypertensive effect of atenolol and bopindolol in the elderly. *Neth J Med*.  
334 1989;35:185-91.

335 [20] Ortega-Ferrusola C, Macias Garcia B, Suarez Rama V, Gallardo-Bolanos JM,  
336 Gonzalez-Fernandez L, Tapia JA, et al. Identification of sperm subpopulations in  
337 stallion ejaculates: changes after cryopreservation and comparison with traditional  
338 statistics. *Reprod Domest Anim.* 2009;44:419-23.

339 [21] Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu  
340 DJ, et al. The PRIDE database and related tools and resources in 2019: improving  
341 support for quantification data. *Nucleic Acids Res.* 2019;47:D442-D50.

342 [22] Ekhlesi-Hundrieser M, Schafer B, Kirchhoff C, Hess O, Bellair S, Muller P, et al.  
343 Structural and molecular characterization of equine sperm-binding fibronectin-II  
344 module proteins. *Mol Reprod Dev.* 2005;70:45-57.

345 [23] Nevin C, McNeil L, Ahmed N, Murgatroyd C, Brison D, Carroll M. Investigating  
346 the Glycating Effects of Glucose, Glyoxal and Methylglyoxal on Human Sperm. *Sci*  
347 *Rep.* 2018;8:9002.

348 [24] Polykretis P, Luchinat E, Boscaro F, Banci L. Methylglyoxal interaction with  
349 superoxide dismutase 1. *Redox Biol.* 2020;30:101421.

350 [25] Gaitskell-Phillips G, Martin-Cano FE, Ortiz-Rodriguez JM, Silva-Rodriguez A, Gil  
351 MC, Ortega-Ferrusola C, et al. Differences in the proteome of stallion spermatozoa  
352 explain stallion-to-stallion variability in sperm quality post thawdagger. *Biol Reprod.*  
353 2021.

354 [26] Ortiz-Rodriguez JM, Martin-Cano FE, Gaitskell-Phillips G, Silva A, Ortega-  
355 Ferrusola C, Gil MC, et al. Low glucose and high pyruvate reduce the production of 2-  
356 oxoaldehydes, improving mitochondrial efficiency, redox regulation and stallion sperm  
357 function. *Biol Reprod.* 2021.

358 [27] Taylor A, Robson A, Houghton BC, Jepson CA, Ford WC, Frayne J. Epididymal  
359 specific, selenium-independent GPX5 protects cells from oxidative stress-induced lipid  
360 peroxidation and DNA mutation. *Hum Reprod.* 2013;28:2332-42.

361  
362  
363  
364  
365  
366  
367  
368

369 Table 1.- The Biomarker Workbench in the bioinformatic software Qlucore Omics  
 370 Explorer 3.7. (<https://qlucore.com>) was used to identify seminal plasma proteins able to  
 371 identify poor and good samples in terms of motility and velocity. The number of variables  
 372 identified specifically for each trait with a q value <0.05 are given. Variables (seminal  
 373 plasma proteins SPP) able to discriminate between stallions were found, and after  
 374 correction for FDR 3 were able to differentiate between stallions with poorer motility,  
 375 VCL and VAP and 2 proteins were able to identify stallions with better VSL.  
 376

Explanatory Variable	Explanatory Variable	Eliminated Factors	SPP p<0.05	SPP q<0.05
Type	Details			
Multi Group Comparison	STALLION		19	0
Two Group Comparison	motility [POOR >All]		17	3
Two Group Comparison	linear motility [POOR >All]		11	0
Two Group Comparison	VAP [POOR >All]		17	3
Two Group Comparison	VCL [POOR >All]		17	3
Two Group Comparison	VSL [GOOD>All]		84	2

377

378

379 Figure legends

380

381 Figure 1.- Computer assisted sperm analysis in 10 stallions (3 ejaculates each, n=30).  
 382 Ejaculates were obtained as described in the material and methods and diluted to  $50 \times 10^6$   
 383 spermatozoa in Tyrode's media. Aliquots were then loaded into a Leja chamber and at  
 384 least 500 spermatozoa were analyzed. Stallions were classified into three categories  
 385 according to their performance in total and linear motility, with stallions showing  
 386 ejaculates with median percentages of total motility over 90% classified as "good", those  
 387 with total motility over 80% classified as average and those with percentages below 80%  
 388 classified as poor. The values for good, average, and poor percentages of linear motile  
 389 spermatozoa were 70% 60% and 40% respectively. Violin plots were used showing the  
 390 data distribution in each group. A) Percentage of total motile spermatozoa B) Percentage

391 of linear motile spermatozoa C) Circular velocity (VCL;  $\mu\text{m/s}$ ) D) Average path velocity  
392 (VAP;  $\mu\text{m/s}$ ) E) Straight line velocity (VSL;  $\mu\text{m/s}$ ) \*  $P < 0.05$ \*\*  $P < 0.01$ , ns=non-  
393 significant.

394

395 Figure 2.- Volcano plots showing differentially expressed proteins in the ejaculates of 10  
396 stallions, different colors show different values of each protein represented by the z-  
397 scores represented in the heat map. The difference of protein content ( $\log_2$  fold change)  
398 is plotted against the significance of the difference  $-\log_{10}(p)$  between the two conditions  
399 (poor motility vs good and average for % total motility, VAP and VCL, or good versus  
400 poor and average in the case of VSL. Three independent ejaculates from 10 different  
401 stallions, in addition to two technical replicates ( $n = 60$  samples) were used to derive  
402 results from.- A) Differential amounts of proteins in stallions showing poor values of %  
403 motility; these stallions showed increased amounts of proteins as seen in the upper right  
404 quadrant of the plot (in red). B) Stallions showing poorer values of VCL showing  
405 increased amounts of proteins depicted in the upper right quadrant of the plot (in red) C)  
406 Stallions showing poorer values of VAP showing increased amounts of proteins in the  
407 upper right quadrant (in red). D) Stallions showing better values of VSL had higher  
408 amounts of proteins depicted in the upper right quadrant (in green).

409

410 Figure 3.- Discriminant variables retrieved after bioinformatic analysis able to identify  
411 stallions with poor total motility, defined as stallions with motility within normal ranges,  
412 but below 80% total motility. Qlucore omics explorer bioinformatic software (Lund  
413 Sweden) was used to identify these variables, through the comparison of relative amounts  
414 of proteins based on spectral counts among stallions classified as good average or poor in  
415 terms of total motility. Proteins were  $\text{Log}_2$  transformed and normalized, then variables  
416 with a corrected p value,  $q < 0.05$  were considered as biomarkers; 3 independent  
417 ejaculates from 10 different stallions in addition to two technical replicates ( $n = 60$   
418 samples) were used to derive results from.

419

420

421 Figure 4.- Discriminant variables for stallions with poor VCL, these were normal fertile  
422 stallions but showed a VCL in the lower rank in our group of stallions with an average  
423 VCL of  $141.7 \pm 10.8 \mu\text{m/s}$ . Qlucore Omics Explorer bioinformatic software (Lund  
424 Sweden) was used to identify these variables through the comparison of the relative

425 amounts of proteins based on spectral counts among stallions classified as good, poor and  
426 average in terms of VCL. Proteins were Log<sub>2</sub> transformed and then variables with a  
427 corrected p value,  $q \leq 0.05$  were considered as biomarkers; 3 independent ejaculates from  
428 10 different stallions in addition to two technical replicates (n = 60 samples) were used  
429 to derive results from.

430

431

432 Figure 5.- Discriminant variables for stallions with poor VAP, consisting of those  
433 stallions with VAP of  $70.3 \pm 7.68 \mu\text{m/s}$ . Qlucore Omics Explorer bioinformatic software  
434 (Lund Sweden) was used to identify discriminant variables measuring the relative  
435 amounts of proteins based on spectral counts among stallions classified as good, poor and  
436 average. Proteins were Log<sub>2</sub> transformed and normalized and then variables with a  
437 corrected p value,  $q \leq 0.05$  were considered as biomarkers; 3 independent ejaculates from  
438 10 different stallions in addition to two technical replicates n = 60 samples were used to  
439 derive results from.

440

441

442 Figure 6.- Discriminant variables identifying stallions with good VSL, (mean value of  
443  $81.5 \pm 3.7 \mu\text{m/s}$ ). Qlucore Omics Explorer (Lund Sweden) bioinformatic software was  
444 used to compare relative amounts of proteins based on spectral counts among stallions  
445 classified as good, poor and average for VSL. Proteins were Log<sub>2</sub> transformed and  
446 normalized and then variables with a corrected p value,  $q \leq 0.05$  were considered as  
447 biomarkers; 3 independent ejaculates from 10 different stallions in addition to two  
448 technical replicates (n = 60 samples) were used to derive results from.

449

450

451

452

453

454

455