

Biomarkers of *in vivo* fertility in sperm and seminal plasma of fertile stallions

S. Novak^{a,*}, T.A. Smith^a, F. Paradis^a, L. Burwash^b, M.K. Dyck^a, G.R. Foxcroft^a,
W.T. Dixon^a

^a Department of Agricultural, Food and Nutritional Science, 4-10 Ag/For Centre, University of Alberta, T6G 2P5, Canada

^b Alberta Agriculture and Rural Development, 97 East Lake Ramp NE, Airdrie, Alberta, T4A 0C3, Canada

Received 6 October 2009; received in revised form 30 March 2010; accepted 22 April 2010

Abstract

The global proteome of sperm and seminal plasma of fertile stallions was investigated to determine whether associations with relative *in vivo* fertility exist. Seven stallions at stud in a commercial breeding station were collected throughout the breeding season and bred to a total of 164 mares to determine conception rates. On three occasions during the breeding season, raw semen was obtained from a regular collection for proteomic analysis using two-dimensional electrophoresis and also assessed for routine semen quality end points. First cycle conception rate was negatively related to ejaculate volume ($r = -0.43$, $P = 0.05$) and total IGF1 content (ng) per ejaculate ($r = -0.58$, $P = 0.006$), whereas overall pregnancy rate was positively related to sperm concentration ($r = 0.56$, $P = 0.01$). The abundance of three proteins known to be involved in carbohydrate metabolism in sperm was positively related to fertility. Furthermore, the abundance of four seminal plasma proteins were identified as being negatively related to fertility; these were identified as kallikrein-1E2 (KLK2), clusterin, and seminal plasma proteins 1 (SP1) and 2 (SP2). Abundance of cysteine-rich secretory protein 3 (CRISP3) was positively related to first cycle conception rate ($r = 0.495$, $P = 0.027$) and may provide a good marker of fertility. Based on stepwise regression analysis, clusterin and SP1 in seminal plasma together with sperm citrate synthase were predictive of fertility ($r = 0.77$, $P < 0.0001$). This study identified proteins within sperm and seminal plasma that could serve as biomarkers of semen quality and fertility in stallions.

© 2010 Elsevier Inc. All rights reserved.

Keywords: Protein; Seminal plasma; Insulin-like growth factor 1; Semen quality; Fertility; Horse

1. Introduction

The horse is unique in that selection of stallions used for breeding is based on genetics and desired performance and conformational characteristics. However, with increased use of reproductive technologies and transported semen, stallion fertility is of increasing im-

portance. Even if subfertile stallions can be identified through conventional semen quality assessments, there is substantial variance in relative *in vivo* fertility in stallions considered fertile (60% per cycle conception rate) [1]. Although a stallion may be fertile over repeated breedings, there is economic value for both the mare and stallion breeder if the stallion has a high first cycle conception rate. The ability to select these fertile stallions or predict fertility using biomarkers is a promising goal. In recent years, research efforts have focused on identifying reliable markers of fertility at

* Corresponding author. Tel.: (780) 248-1159; fax: (780) 492-4265.

Email address: susan.novak@ualberta.ca (S. Novak).

either the genomic or proteomic level. Polymorphisms in the gene encoding CRISP3, an abundant seminal plasma protein, have been associated with a reduction in fertility [2], whereas inhibin beta A (INHBA) polymorphisms were associated with pregnancy rates [1], in a large-scale association study with Hanoverian stallions.

Proteins in seminal plasma are involved in various processes to preserve viability of sperm, interactions with the female reproductive tract, and the process of fertilization [3] and are good candidates for markers of fertility. The seminal plasma proteome has been investigated in the stallion [4], bull [5], and human [6], using 2D gel electrophoresis. The proteins in the sperm membrane are also interesting candidates, as they are important for sperm adhesion to the sperm reservoir in the female tract and sperm-egg binding. Proteomics has identified proteins that are differentially expressed in human sperm from normal versus asthenozoospermic samples [7]. In the boar, recent studies have shown negative relationships between seminal plasma proteins and fertility [8]. In the bull, specific proteins in seminal plasma were associated with fertility [9,10], whereas in equine seminal plasma, one protein was positively (72 kDa, pI 5.6, osteopontin) and three were negatively associated with fertility [11]. More recently, IGF1 concentrations in stallion seminal plasma were also shown to have a relationship to fertility [12]. Three proteins on the stallion sperm membrane (SNARE proteins, caveolin-1 and NSF) have also recently been identified and implicated in fertility [13].

The objective of this study was to investigate the global proteome of both sperm and seminal plasma of stallions known to be fertile, to determine if any of the observed proteins were associated with conventional semen quality measurements and relative *in vivo* fertility of stallions at a commercial breeding station. Relationships between IGF1 concentrations in seminal plasma of these stallions and their fertility were also explored.

2. Materials and methods

2.1. Stallion semen collection

Seven fertile stallions with an average age of 12.5 y (range, 6–23 y) from a commercial stallion breeding station (Bassano, AB, Canada) were collected every second day throughout the breeding season and bred to a total of 164 mares from May to August (on average 25.3 mares per stallion). On three occasions during the breeding season, each 2 wk apart, an ejaculate was

assessed for routine semen quality parameters such as ejaculate volume, sperm concentration and percent progressively motile sperm by the same trained individual. One of the stallions was only collected on two occasions, as it was no longer needed for breeding. The ejaculate was then diluted with extender to at least 500×10^6 progressively motile sperm, and immediately used to inseminate mares at the breeding station. An aliquot (10 mL) of fresh, raw semen was obtained from these collections for protein analysis, and sperm and seminal plasma were immediately separated by centrifugation at $4,000 \times g$ for 10 min at room temperature, and seminal plasma was frozen at -20°C . The sperm pellet was washed three times with 5 mL PBS and centrifuged each time at $4,000 \times g$ for 10 min. The washed sperm pellet was immediately frozen at -20°C and transported back to the laboratory.

Stallions were assessed for fertility, based on first cycle conception rate for three 4-wk intervals throughout the breeding season, and the final pregnancy rate and average number of cycles bred was determined after breeding mares for a total of three cycles. Pregnancy was confirmed in the mares by ultrasonography approximately 30 d after insemination.

2.2. IGF1 radioimmunoassay

Concentrations of IGF1 in seminal plasma in all ejaculates were quantified using an established radioimmunoassay [14]. Samples were extracted singly and assayed in triplicate. The following modifications were made: 0.1 mL of seminal plasma was extracted and 0.3 mL of neutralized sample was taken to assay. Diluted neutralized control pools of boar and horse seminal plasma each showed parallelism to the standard curve. The product used for iodination and the standard curve was GroPep Human IGF-I Receptor Grade (GroPep, # CU100, GroPep Ltd., Adelaide, SA, Australia). Assay sensitivity defined as 86% bound was 3.5 pg/tube (equivalent to 1.81 ng/mL of raw plasma in the current assay). All samples had concentrations higher than sensitivity. Apparent radioinert % recovery was $44.24 \pm 1.92\%$ for horse seminal plasma. Results were not corrected for recovery. The intra-assay CV for the single assay run was 3.05%.

2.3. Sperm and seminal plasma protein preparation

Protein was extracted from the sperm pellet in rehydration lysis buffer (7M Urea, 2M thiourea, 4% (wt/vol) CHAPS, 20 mM DTT) with repeated homogenization for 1 h on ice. The sample was centrifuged at $10,000 \times g$ for 10 min to remove debris and the

supernatant containing the solubilized sperm proteins was collected and an aliquot was taken for determination of protein content using the 2D Quant Kit (GE Healthcare, Baie d'Urfé, QC, Canada). Seminal plasma proteins were quantified directly using the BCA protein assay (Pierce, Rockford, IL, USA).

2.4. 2D gel electrophoresis

The seminal plasma and sperm proteins (100 μg) were solubilized in rehydration buffer (7M Urea, 2M Thiourea, 20 mM DTT, 0.8% (vol/vol) IPG Buffer pH 3–10, Bromophenol Blue) and loaded by cup-loading onto rehydrated 24 cm Immobiline DryStrips IPG strips pH 3–10 Linear (GE Healthcare). There were 20 gels in total of seminal plasma and 20 gels of sperm proteins, as one stallion was not collected at one time point. The first dimension isoelectric focusing protocol included a gradual increase in voltage to a total of 70,000 Volt-h over a 24 h interval (IPGPhor, GE Healthcare) and immediately frozen at -80°C after isoelectric focusing. Prior to the second dimension separation, strips were equilibrated in 50 mM Tris-HCl, 6 M urea, 30% (vol/vol) glycerol, 2% (wt/vol) SDS and bromophenol blue, for 15 min with 1% (wt/vol) DTT, and then alkylated using 2.5% (wt/vol) iodoacetamide for 15 min. The 24 cm IPG strips were placed on 12% (wt/vol) polyacrylamide 26 cm-wide slab gels and initially separated at 5W/gel for 30 min, followed by 17W/gel for 3.5 to 4 h with buffer temperature set to 20°C in a ETTAN DALT-six gel electrophoresis system (GE Healthcare). To reduce technical variation, six gels were cast at once in the multi-gel caster and six gels were run at once in the second dimension system when possible.

The proteins were visualized by silver staining [15], and scanned using a 14-bit scanner (Imagescanner, GE Healthcare). Silver stain was chosen as a method for these large format gels, because it is very sensitive to detect proteins in lower abundance and it is cost-effective for the number of samples analyzed. Detection and matching of spots across gels and quantification were performed using Progenesis PG240 software (Nonlinear Dynamics, UK). Briefly, the program first performed a quality control for each image, verifying appropriate resolution and flagging saturated spots; none of the images used had saturated spots. A reference gel for each gel type (sperm or seminal plasma) was chosen and all gels were matched with the reference gel through visual examination and selection of user seeds using the computer aided-alignment in the software, then the software completed the matching analysis. A

total of 108 spots in seminal plasma and 182 spots in sperm were matched and analyzed across all gels. All matches were visually inspected for accuracy and any spots that needed editing (merging, splitting, or re-matching) were redone. Non-spots were eliminated from the analysis and the spot volume for each spot was normalized against the total spot volume for that gel, thus applying a correction factor for gels that differed in staining intensity. The individual normalized spot volumes were extracted from the software and statistically analyzed in SAS (SAS Institute, Cary, NC, USA).

2.5. Statistics

The quantified sperm and seminal plasma proteins, semen characteristics and IGF1 concentrations were analyzed to determine if there were differences across stallions with each of the three ejaculates for each stallion as replicates, using the general linear models procedure of SAS.

The quantified proteins were also correlated to visual semen quality characteristics (semen volume, sperm concentration, total sperm number per ejaculate, % motility, and total progressively motile sperm), IGF1 concentrations, protein concentrations, total IGF1 content and total protein content in seminal plasma and *in vivo* fertility (first cycle conception rate, and overall pregnancy rate).

Correlation analyses and stepwise regression analysis were conducted to determine the specific relationships between fertility and sperm and seminal plasma characteristics and proteins. A principal components analysis (PCA) was also conducted to confirm these relationships, all within SAS.

2.6. LC-MS/MS identification of sperm and seminal plasma proteins

A preparative 2D gel was run for each of the sperm (1 mg) and seminal plasma proteins (500 μg) on 24 cm Immobiline DryStrip Gels (GE Healthcare), in the linear pH 3–10 range, using an extended isoelectric focusing protocol for 85,000 Vh with rehydration loading, and loaded onto a 12 (wt/vol) % SDS-PAGE slab gel. The resulting gel was fixed overnight in 50% (vol/vol) methanol and stained using Bio-Safe Coomassie Blue (Bio-Rad Labs, Hercules, CA). The protein spots of interest were manually excised and sent to a mass spectrometry facility for further processing and identification (Centre Genomique du Quebec, Sainte-Foy, QC, Canada). All procedures for mass spectrometry and database searching for protein identification have been previously described [8].

Table 1

Reproductive and semen characteristics of the seven stallions. Fertility values were based on the entire breeding season from late May to early August and the semen characteristics are an average from three ejaculates, each taken 2 wk apart, from the end of May to mid July.

Stallion	First cycle conception rate (%)	Overall pregnancy rate (%)	No. cycles per pregnancy	Ejaculate volume (mL)	Sperm concentration ($\times 10^6$ /mL)	Progressive motility (%)	Total no. sperm ($\times 10^6$)	Total no. progressively motile sperm ($\times 10^6$)
1	63.9 ^{bc}	100	1.36	31.3 ^c	246.0 ^a	58.3	7609	4691
2	100 ^a	100	1.0	26.0 ^c	197.0 ^{ab}	70.0	4613	3127
3	63.3 ^{bc}	93.1	1.46	53.3 ^c	47.7 ^c	56.6	3020	2057
4	62.9 ^{bc}	87.5	1.35	93.7 ^a	118.3 ^{bc}	73.3	11049	8265
5	77.7 ^c	91.2	1.26	37.5 ^c	79.5 ^{bc}	62.5	2947	1850
6	52.9 ^c	75	1.36	31.0 ^c	79.3 ^{bc}	58.3	2459	1449
7	100 ^a	100	1.0	21.7 ^c	199.7 ^{ab}	71.3	2459	2048
Pooled	4.9	9.35	0.17	11.5	36.5	7.5	1722	1442
SEM								
Prob.*	< 0.0001	0.385	0.776	0.009	0.015	0.538	0.030	0.059

^{a–c} Within a column, means without a common superscript differed ($P < 0.05$).

* Probability for the main effect of stallion.

3. Results

3.1. Stallion fertility and semen quality

First cycle conception rate of the stallions was 50–100%; overall pregnancy rates ranged from 75–100% (Table 1), and all stallions were considered fertile. There were differences ($P < 0.05$) among stallions in sperm concentration, total number of sperm, and ejaculate volume (Table 1). Stallion 4 had a higher ejaculate volume than the other six stallions ($P = 0.0009$), whereas sperm concentration varied greatly among stallions. There were no differences in progressively motile sperm (PMS), or the total number of progressively motile sperm ($P > 0.05$). All stallions were considered to have acceptable sperm morphology ($>70\%$ normal sperm; data not available). First cycle pregnancy rate

was negatively related to ejaculate volume ($r = -0.436$, $P = 0.05$) and overall pregnancy rate was positively related to sperm concentration ($r = 0.564$, $P = 0.01$).

There were differences in total protein concentration ($P = 0.01$), but not total protein content ($P > 0.05$), in the seminal plasma across stallions (Table 2). However, total protein content in the ejaculate was negatively related to first cycle conception rate ($r = -0.496$, $P = 0.026$), but not protein concentration. Seminal plasma IGF1 concentrations and total IGF1 content were different among stallions ($P = 0.007$). Concentrations of IGF1 in seminal plasma were not related to fertility, however lower total seminal plasma IGF1 content was associated with the most fertile stallions, with an overall negative relationship to first cycle

Table 2

Total protein and IGF1 concentrations and total content in seminal plasma of stallions.

Stallion	Ejaculate volume (mL)	[IGF1] (ng/mL)	IGF1 content (ng)	Total protein concentration (mg/mL)	Total protein content (mg)
1	31.3 ^c	28.0 ^a	808.6 ^a	20.63 ^a	631.2
2	26.0 ^c	12.0 ^{bcd}	234.7 ^{bc}	14.08 ^{abc}	366.8
3	53.3 ^c	13.2 ^{bcd}	688.7 ^a	11.6 ^{bc}	550.7
4	93.7 ^a	10.5 ^{cd}	1000.7 ^a	6.27 ^c	600.6
5	37.5 ^c	22.4 ^{ab}	857.2 ^a	26.89 ^{ab}	642.0
6	31.0 ^c	15.3 ^{abc}	493.0 ^{ab}	12.99 ^{abc}	404.0
7	21.7 ^c	7.5 ^d	139.8 ^c	7.63 ^c	127.2
Pooled SEM	11.5	3.94	172.0	2.29	142.7
P-value	0.009	0.009	0.004	0.010	0.211

Semen values are expressed as means.

^{a–c} Within a column, means without a common superscript differed ($P < 0.05$).

*Probability for the main effect of stallion.

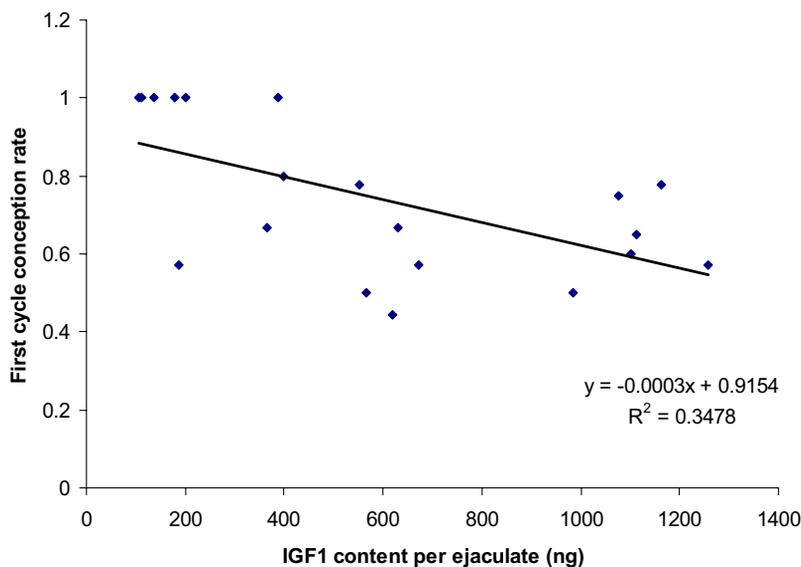


Fig. 1. Relationship between first cycle conception rate and IGF1 content in equine ejaculates ($P = 0.007$).

conception rate ($r = -0.589$, $P = 0.006$; Fig. 1). Total IGF1 content was also strongly positively correlated with ejaculate volume ($r = 0.724$, $P = 0.0003$), demonstrating a relationship between ejaculate volume and IGF1 content that may be indicative of fertility as both were related to first cycle conception rate.

3.2. Proteomics analysis of sperm and seminal plasma proteins

The software initially detected >3000 sperm proteins ($n = 20$ gels) and >4000 seminal plasma proteins ($n = 20$ gels). After all detected proteins were edited and correctly matched, 108 proteins total were submitted for analysis on the seminal plasma gels; of these 108 proteins, the 20 gels had on average 92 spots (range, 75–108). For the sperm proteins, 182 spots were considered for analysis after correct matches were made (average 141 spots; range 99–182 spots across 20 gels). All protein spots were numbered and those that were associated with fertility are identified by their number on Fig. 2 and listed in Table 3. For the proteins found in sperm, four proteins were positively related to fertility; Spot 1692, identified as malate dehydrogenase was positively related to first cycle conception rate ($r = 0.46$, $P = 0.04$), and Spot 1182 identified as fumarate hydratase was also associated with first cycle conception rate ($r = 0.466$, $P = 0.038$). Protein Spot 1213 was strongly related to first cycle conception rate ($r = 0.55$, $P = 0.011$) and overall pregnancy rate ($r = 0.54$, $P = 0.014$), and was identified as citrate synthase. Protein Spot 1145, identified as α -enolase, had a negative re-

lationship with seminal plasma IGF1 concentrations ($r = -0.448$, $P = 0.047$) and was positively associated with first cycle conception rate ($r = 0.517$, $P = 0.019$). Protein Spot 899, identified as dihydrolipoamide dehydrogenase, was negatively related to ejaculate volume ($r = -0.459$, $P = 0.042$), as well as positively related to first cycle conception rate ($r = 0.589$, $P = 0.007$).

For the proteins found in seminal plasma, six were related to fertility or semen characteristics (Table 3). Protein Spots 329, 322, 344, and 331 were all clustered in a similar region of the gel (Fig. 2) and were identified as a protein similar to palate, lung, and nasal epithelium protein (PLUNC). These spots represent different isoforms of the protein, were all negatively related to IGF1 concentrations ($r = -0.45$, $P = 0.04$), and had a positive relationship with first cycle conception rate ($r = 0.435$, $P = 0.05$) in seminal plasma. Spot 302 was negatively related to overall fertility ($r = -0.558$, $P = 0.01$), first cycle conception rate ($r = -0.721$, $P = 0.0003$; Fig. 3), and tended to be positively associated with IGF1 concentrations in seminal plasma ($r = 0.432$, $P = 0.06$). This protein was identified as clusterin, an abundant protein in the seminal plasma of horses. Furthermore, KLK2 was represented on a 2D gel as a cluster of protein isoforms of which we identified Spots 3851 and 966 (Fig. 2). Spot 3851 was negatively related to overall pregnancy rate ($r = -0.60$, $P = 0.028$), and positively related to ejaculate volume ($r = 0.46$, $P = 0.04$). Spot 966 was negatively related to overall pregnancy rate ($r = -0.62$, $P = 0.004$) and negatively

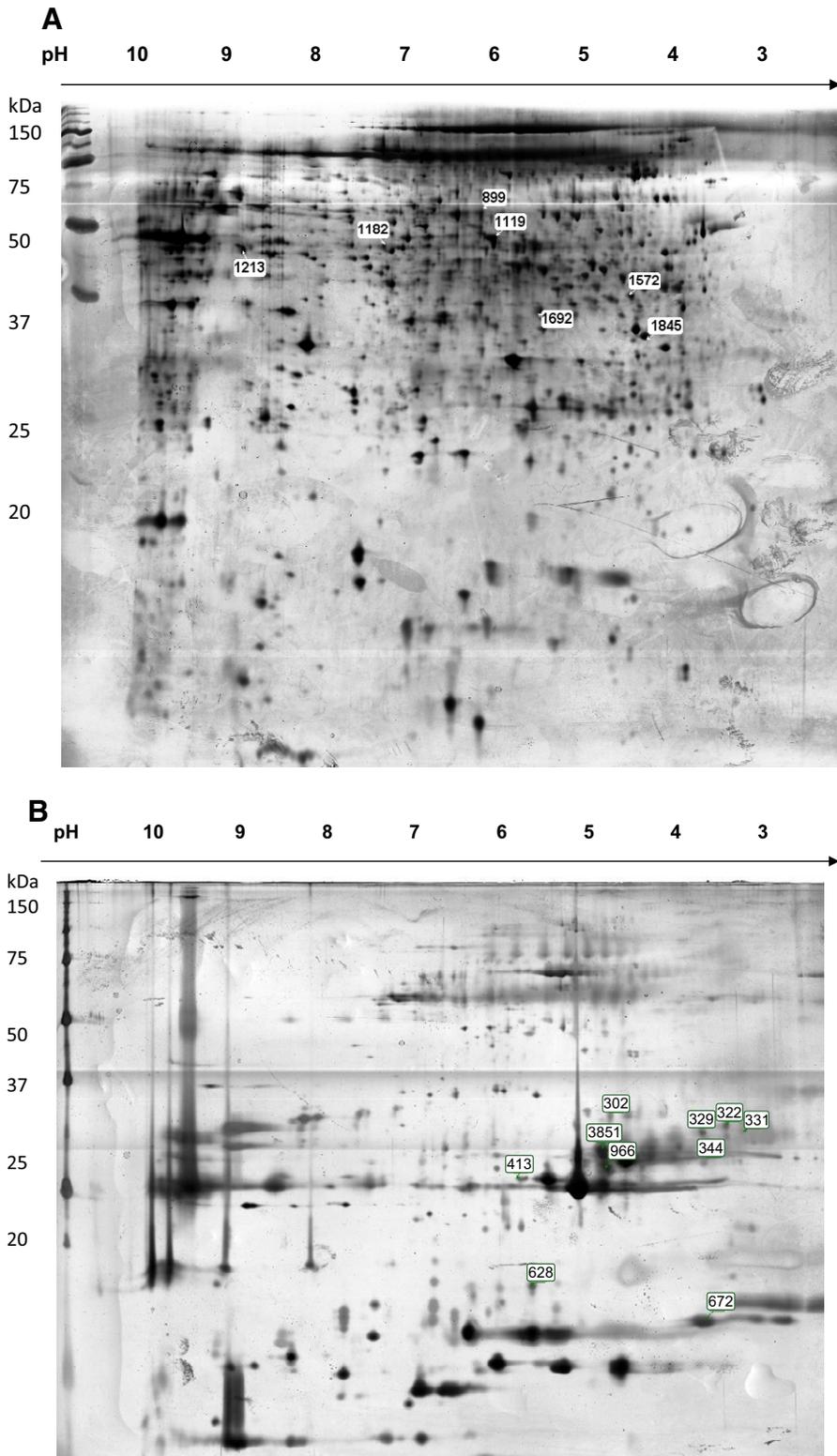


Fig. 2. Two-dimensional gels (pH 3–10 in the first dimension and 12% SDS-PAGE in the second dimension) of equine sperm (A) and seminal plasma (B). Proteins that were identified by LC-MS/MS (Table 3) are labeled with their corresponding number.

Table 3

Identification of proteins of interest in seminal plasma and sperm of stallions and their association with fertility (first cycle conception rate) and semen characteristics.

Spot #	Source	Association	Protein name	Accession number (Uni-Prot)	Theoretical MW (kDa)	No. unique peptides (% Coverage)
1213	Sperm	Fertility +	Citrate Synthase	UPI0001561064	52	11 (36%)
1182	Sperm	Fertility +	Fumarate hydratase	UPI0001560C16	50.9	20 (66%)
1692	Sperm	Fertility +	Malate dehydrogenase	UPI000155F7AE	36	12 (39%)
899	Sperm	Fertility + Volume –	Dihydrolipamide dehydrogenase	UPI000155E1EA	54	29 (62%)
1119	Sperm	Fertility + [IGF1] –*	α -enolase	UPI000155DDF2	41	41 (83 %)
413	Seminal Plasma	Fertility +	Cysteine-rich secretory protein 3 (CRISP3)	O19010	29.3	5 (36%)
966	Seminal Plasma	Fertility – [Sperm]–	Kallikrein-1E2 (KLK2)	Q6H321	29	8 (31%)
3851	Seminal Plasma	Fertility – Volume +	Kallikrein-1E2 (KLK2)	Q6H321	29	12 (49%)
302	Seminal Plasma	Fertility – [IGF1] +*	Clusterin (alpha-chain)	Q29482	52	21 (32%)
672	Seminal Plasma	Fertility – [IGF1] +	Seminal Plasma Protein 2 (SP2)	UPI000155EF95	19.2	13 (43%)
628	Seminal Plasma	Fertility – [sperm] – Volume +	Seminal Plasma Protein 1 (SP1)	P81121	14	5 (39%)
329	Seminal Plasma	Fertility + [IGF1] –	Similar to PLUNC	UPI00000156000D	26.9	9 (47%)
322	Seminal Plasma	Fertility + [IGF1] –	Similar to PLUNC	UPI00000156000D	26.9	8 (37%)
344	Seminal Plasma	Fertility + [IGF1] –	Similar to PLUNC	UPI00000156000D	26.9	6 (28%)
331	Seminal Plasma	Fertility + [IGF1] –	Similar to PLUNC	UPI00000156000D	26.9	7 (32%)

* Correlation ($P < 0.10$).

associated with sperm concentration ($r = -0.47$, $P = 0.036$). Two more proteins were negatively related to fertility, Spots 628 and 672. Spot 628 was identified as seminal plasma protein 1 (SP1) and was negatively associated with first cycle conception rate ($r = -0.59$, $P = 0.006$) and overall pregnancy rate ($r = -0.56$, $P = 0.01$), as well as negatively associated with sperm concentration ($r = -0.52$, $P = 0.02$) and positively associated with ejaculate volume ($r = 0.47$, $P = 0.037$). Spot 672 was also identified as a major component of seminal plasma, seminal plasma protein 2 (SP2) and was also related with first cycle conception rate ($r = -0.55$, $P = 0.012$), and positively related to IGF1 concentrations ($r = 0.67$, $P = 0.001$). The only protein identified in this analysis to be positively related to fertility was Spot 413, cysteine-rich secretory protein 3 (CRISP3), which was positively associated with first cycle conception rate ($r = 0.495$, $P = 0.027$).

3.3. Stepwise regression analysis

Stepwise regression analysis was performed to determine which variables contributed to the determination of first cycle conception rate (%). Although overall pregnancy rate was also used in the stepwise regression, each of the three normalized spot volume values for each stallion was regressed against single value obtained for overall pregnancy rate. First cycle conception rate was more accurate to predict the relative fertility of these stallions, and semen characteristics were represented within the three time points chosen throughout the breeding season. All variables were entered into the stepwise regression analysis including: ejaculate volume, sperm concentration, motility, total sperm, total progressively motile sperm, IGF1 concentrations, IGF1 content, seminal plasma protein concentrations and content, as well as all identified seminal plasma and sperm proteins (clusterin, CRISP3, PLUNC, SP1,

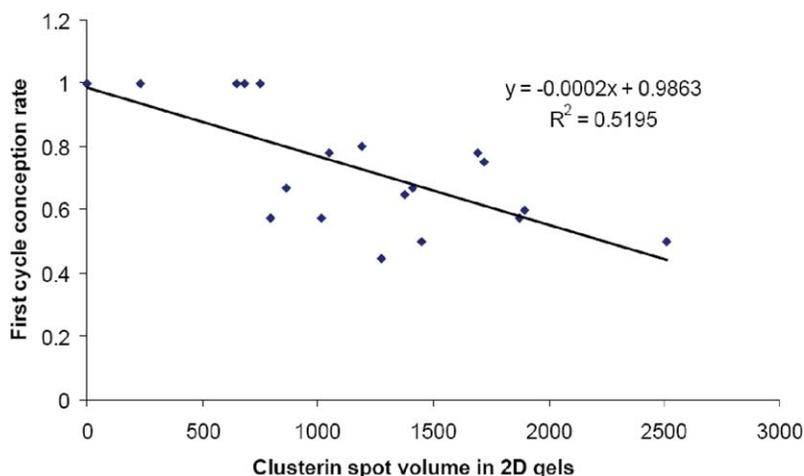


Fig. 3. The relationship of clusterin protein in equine seminal plasma and first cycle conception rate ($P = 0.0003$).

SP2, kallikrein, malate dehydrogenase, citrate synthase, fumarate hydratase, α -enolase, and dihydrolipoamide dehydrogenase).

3.3.1. Stepwise regression analysis for first cycle conception rate

The regression analysis retained variables that were significant at the $P < 0.15$ level in the model. The resulting model included clusterin (partial $R^2 = 0.52$, $P = 0.0003$), SP1 (partial $R^2 = 0.16$, $P = 0.01$), and citrate synthase (partial $R^2 = 0.096$, $P = 0.019$). The overall model had an $R^2 = 0.77$, $P < 0.0001$. The equation was as follows:

First cycle conception rate (%) = $0.90765 - 14.45$ (clusterin) $- 5.99$ (SP1) $+ 0.10$ (citrate synthase).

3.3.2. Stepwise regression analysis for overall pregnancy rate

The regression analysis retained variables that were significant at the $P < 0.15$ level in the model. The resulting model included kallikrein (Spot 966, partial $R^2 = 0.385$, $P = 0.0035$), citrate synthase (partial $R^2 = 0.1231$, $P = 0.0548$), and sperm concentration (partial $R^2 = 0.1171$, $P = 0.04$). The overall model had an $R^2 = 0.6251$, $P = 0.001$. The equation is as follows:

Overall pregnancy rate (%) = $0.8977 + 38.8$ (sperm concentration) $+ 0.0558$ (citrate synthase) $+ 3.55$ (kallikrein).

4. Discussion

This study demonstrated the identification of semen biomarkers associated with stallion *in vivo* fertility. First cycle conception rate is a good indicator of relative fertility, as it becomes an economically important

trait in breeding programs to reduce matings over multiple cycles. When first cycle conception rate was considered in this study, it was negatively related to some abundant proteins in semen and also semen volume, demonstrating that more dilute semen may not be beneficial for relative fertility in the stallion. Both the negative relationships with IGF1 content and clusterin in semen could be used independently as good predictors of fertility. In contrast, all of the sperm proteins shown to vary with relative fertility were positively related to first cycle conception rate, and there was also a strong positive relationship between overall pregnancy rate and sperm concentration.

Based on multiple regression analysis of all variables in the ejaculates, including semen characteristics and abundance of all sperm and seminal plasma proteins, clusterin, SP1, and citrate synthase, were all significant factors in an equation to predict fertility. The contributions of the sperm proteins to the model were approximately 10%, compared to the 90% contribution of seminal plasma proteins. Even though there were no significant differences in overall pregnancy rate among stallions, a stepwise regression analysis reconfirmed the relationship of citrate synthase with fertility. With pregnancy rate considered, other proteins included in the model were kallikrein. Interestingly, sperm concentration as a semen characteristic was included in this model, further supporting the concept that increased semen volumes were negatively associated with fertility.

In the present study, the volume of the raw ejaculate was negatively related to first cycle conception rate and higher sperm concentration in semen was positively associated with overall pregnancy rate. These ejacu-

lates tended to be diluted to a lesser extent in order to maintain at least 500×10^6 sperm for each insemination, and the increased amount of seminal plasma per ejaculate could have negatively affected the fertilizing ability of the sperm for immediate breeding. In support of this, the major seminal plasma proteins that exhibited a relationship with fertility in this study, such as kallikrein-1E2, SP1 and SP2, had a negative relationship with fertility, suggesting that these proteins may serve to bind to spermatozoa in the ejaculate to protect them and possibly reduce their ability to fertilize eggs, when artificial insemination of reduced sperm doses was done. We have recently shown similar results in the boar, demonstrating that the boar spermadhesins, AWN1, PSP1 and osteopontin in seminal plasma could have a negative effect on fertility in a low sperm dose situation [8]. During natural breeding, the entire raw ejaculate would contain a much higher concentration of sperm without dilution of seminal plasma, which could buffer the ratio of sperm to seminal plasma, thereby potentially improving the fertilizing ability of sperm in natural breeding compared to artificial insemination.

The concentrations of IGF1 in seminal plasma have previously been found to be positively associated with fertility in the stallion [12]. Although there were only seven stallions in the study, no relationships between seminal plasma IGF1 concentrations and either first-cycle conception rate or overall pregnancy rate were detected. Similar to the previous study, our stallions were part of a regular breeding program and were collected every second day throughout the breeding season. The IGF1 measurements were also an average of three samples for each stallion, and the variance between samples was small. Mouse IGF1 knockouts produced males with reduced testosterone concentrations and markedly lowered sperm production [16], and Macpherson et al [12] also reported a relationship between IGF1 concentrations and sperm motility in stallions. One of the possible reasons why we did not see any relationship between IGF1 concentrations and fertility may be inherent with the use of normal, relatively fertile stallions in the current study. However, although there was no relationship between fertility and IGF1 concentrations in the current study, total IGF1 content in the ejaculate was highly negatively related to first cycle conception rate and could prove to be a simple and effective tool to measure fertility. The measurement of total IGF1 content in the ejaculate is largely dependent on the ejaculate volume, which was independently shown to have a negative effect on first cycle conception rate in this study. Taken together, the mea-

surement of total IGF1 content may be an effective predictive measure of relative fertility that merits further investigation.

Seminal plasma proteins have been investigated as markers of fertility using two-dimensional electrophoresis in the bull [9], stallion [11], boar [8], and human [17]. The human seminal plasma proteome has considerable diversity in the proteins present, whereas the boar and bull contain a large proportion of spermadhesins and heavily glycosylated proteins. The stallion appears to be intermediate between these extremes, exhibiting large families of proteins, as well as many unique spots. One of the two proteins in seminal plasma that was identified as being positively related to fertility was cysteine-rich secretory protein 3 (CRISP3) (spot 413). This protein was present in large amounts in stallion seminal plasma [18]. As this appears to be a characteristic unique to this species, Hamann *et al.*, [1] postulated that a special role for this protein may exist in the horse, and also reported that certain polymorphisms in the CRISP3 gene were positively related to fertility using SNP analysis techniques. Interestingly, the polymorphism that was found to be negatively related to fertility in that study was an amino acid substitution from glutamic acid to lysine at position 208. Consistent with this finding, the amino acid at that position in the CRISP3 in our study was identified as glutamic acid, not lysine. In the mouse, the CRISP1 protein bound to sperm and was involved in sperm-egg fusion at specific binding sites [19], and may help explain the positive association with fertility and CRISP3 in the present study.

The abundance of a cluster of seminal plasma proteins (Spots 329, 322, 344, 331) in our study was also identified to be negatively associated with IGF1 concentrations and positively associated with fertility. These proteins were identified as being similar to palate, lung, and nasal epithelium protein (PLUNC), which is thought to be involved in the inflammatory response to irritants in the respiratory pathways. However this is first study to report the presence of this family of proteins in the male reproductive tract, specifically in seminal plasma. Fouchécourt et al [4] also identified proteins in this same area, based on their appearance in two-dimensional gels in equine seminal plasma, however they identified these proteins as isoforms belonging to the clusterin family. Although clusterin is a possible identification for these proteins in our analyses, the number of unique peptides matched, and sequence coverage, is much higher in all cases to the predicted PLUNC protein. The molecular weight re-

ported for PLUNC at 26.9 kDa is also very similar to the molecular weight observed on our gels, rather than the 52 kDa molecular weight reported for clusterin. As the study by Fouchécourt [4] was published in the year 2000, and the entry for the *Equus caballus* PLUNC protein was only published in 2008, this would have precluded their ability to identify this protein. Also, in our study, the clusterin protein was found to be negatively associated with fertility, whereas the PLUNC protein was positively associated with fertility, suggesting different roles for these proteins in seminal plasma. Functionally, the PLUNC protein is a lipopolysaccharide-binding protein and a bacteriocidal permeability-increasing protein, and its positive association with fertility suggests that it may play a role after insemination, possibly by aiding the mare's own immune response in clearing the female reproductive tract of bacteria in the horse.

In seminal plasma, the proteins that were identified as being negatively associated with fertility, were horse seminal plasma protein 1 (SP1) and seminal plasma protein 2 (SP2), clusterin (alpha-chain), and the kallikrein -1E2 protein, which is a homologue for prostate-specific antigen [20]. Two spots (966 and 3851) were identified as kallikrein 1-E2 (KLK2), and were also negatively related to overall pregnancy rate, with Spot 966 included in the equation to predict overall pregnancy rate. One of the kallikrein spots was positively related to ejaculate volume, and the other negatively related to sperm concentration. The kallikreins are serine proteases and the glandular kallikrein is regulated by androgens and secreted by the prostate. The amount of prostate-specific antigen present in serum is currently the best marker for the development of prostate tumors [21]. Perhaps the volume of seminal fluid and prostate secretions could have an inverse relationship with fertility; stallions producing more concentrated semen with higher sperm concentrations and lower amounts of seminal plasma would have higher fertilizing ability. To further support this concept, we also detected an inverse relationship with ejaculate volume and first cycle conception rate and, similar to the kallikrein protein, many of the abundant proteins found in seminal plasma, the clusterins and seminal plasma proteins 1 (SP1) and 2 (SP2), also exhibited a negative relationship with fertility. Additionally, clusterin and SP1 both had positive associations with ejaculate volume and were the two seminal plasma proteins included in the regression model equation used to predict first cycle conception rate. The immediate protective effect of these proteins on sperm, which may be inhibitory to

sperm activity and fertilization, would be diluted out with the increased extender volumes in these stallions. Support for this hypothesis comes from reports of similar relationships between seminal plasma proteins and relative fertility in the stallion [11] and in the boar [8].

The functions of the seminal plasma proteins, SP1 and SP2 are not yet fully elucidated, but they bind to sperm and are likely involved in capacitation [3]. The quantity of these proteins in seminal plasma, but not the amount bound to sperm, were negatively associated with fertility, but this may merely be a reflection of the volume of seminal plasma in the ejaculate, rather than an observed functional relationship between these proteins and fertility. Likewise, the clusterins were also negatively associated with fertility and are also an abundant protein family found in seminal plasma [4]. We have already established a strong negative relationship between first cycle conception rate and total IGF1 content in seminal plasma, however when all of the proteins were also included in a multiple regression analysis with IGF1, the abundance of clusterin protein and SP1 together replaced IGF1 content as a predictive measure of fertility. The abundance of these two proteins in seminal plasma explains the majority (82%) of the differences in fertility between stallions in the regression model, whereas IGF1 content only explained 50% of this relationship. However, IGF1 content is easier to measure and a more accessible end point than two seminal plasma proteins at the present time.

When the proteins in spermatozoa are considered, all five identified proteins had a positive relationship with fertility, which was opposite to the relationship predominantly seen in the seminal plasma. Three of these proteins (citrate synthase, fumarate hydratase, and malate dehydrogenase) are involved in carbohydrate metabolism, and are enzymes directly involved in the TCA cycle. It is likely that increased sperm metabolism and the ability to use carbohydrates as energy may have a positive effect on the fertilizing ability of sperm. Another protein, dihydrolipoamide dehydrogenase, was negatively related to ejaculate volume and strongly positively correlated to first cycle conception rate. Interestingly, Martinez-Heredia et al [7] also reported higher levels of both the precursor forms of fumarate hydratase and dihydrolipoamide dehydrogenase in asthenozoospermic (lower motility) sperm samples from men, and suggested that there was an impairment in these infertile sperm in their ability to process these enzymes to their functional mature form. Another enzyme, α -enolase was identified as having a positive relationship with first cycle conception rate and a neg-

ative tendency to be associated with IGF1 concentrations in seminal plasma. In infertile men, sperm-specific enolase enzyme activity was elevated in normal sperm compared to abnormal sperm [22], however in the present study we were unable to distinguish the enolase identified as the sperm-specific enolase in humans. There is not enough information available in the literature or in the available gene information banks to distinguish between these two isoforms at a proteomic level using mass spectrometry. It is interesting to consider that previous studies have focused on enzyme activities in sperm as potential markers of fertility, and the current study identified seven sperm proteins as markers of fertility, all of which were enzymes.

It is important to note that the relationships established in the present study were not always indicative of a functional relationship; however, the proteins may serve as markers of that attribute or characteristic. A good example of this concept is the negative relationship reported for the seminal plasma proteins, where it may not be reflective of function, but rather the relative abundance of seminal plasma or secretory activity of the specific glands contributing to the seminal fluid in the ejaculate. In addition, fertility is a complex multi-gene trait that is considered to be specific to the environment in which it is measured, whether *in vivo* using natural breeding or artificial insemination, whether the sperm are used fresh or frozen, or whether *in vitro* for use in assisted reproductive technologies; presumably, this contributed to so many reported markers of fertility in numerous species. A panel of biomarkers will likely be necessary to establish screening tests for breeding stock for specific purposes.

In conclusion, the current study demonstrated strong positive relationships between proteins involved in carbohydrate metabolism in the sperm and fertility, as well as negative relationships with some seminal plasma proteins, such as SP1, SP2, and clusterin. Furthermore, the negative relationship established between total IGF1 content in seminal plasma and fertility, or the negative relationship between clusterin and SP1 with fertility could also be predictive of fertility in stallions. As many of the characteristics that were negatively related to fertility were associated with ejaculate volume, and overall pregnancy rate was positively associated with sperm concentration, we inferred that the higher secretory activity of the accessory glands to produce larger volumes of semen may not be beneficial. A panel of biomarkers was identified to predict first cycle conception rate; these included clusterin and SP1 in seminal plasma, as well as citrate synthase in sperm.

A stepwise regression analysis with overall pregnancy rate also re-confirmed these relationships between fertility and citrate synthase and included kallikrein and sperm concentration. This study provides insight on the proteins within sperm and seminal plasma that could serve as biomarkers for semen quality in stallions. Further validation of these markers across a large population of stallions is necessary.

Acknowledgements

The authors express kind thanks to Shirley Shostak for her help in IGF1 RIAs, and Joan Turchinsky in the molecular laboratory. Also, the authors appreciate the efforts of Michael Vinsky in his help with the Progenesis software and analysis. This research study was generously funded with help from Alberta Agriculture and Rural Development, Horse Racing Alberta, the Horse Industry Association of Alberta and the University of Alberta. In-kind contributions of semen were kindly made by Sandy Ridge Stallion Station, Bassano, AB. The authors gratefully acknowledge A.F. Parlow, Harbor-UCLA Medical Center, Torrance, CA, and NIDDK's National Hormone and Peptide program for antisera to IGF1.

References

- [1] Giesecke K, Hamann H, Sieme H, Distl O. INHBA-associated markers as candidates for stallion fertility. *Reprod Domest Anim* 2009. Jan 8 [Epub ahead of print].
- [2] Hamann H, Jude R, Sieme H, Mertens U, Topfer-Petersen E, Distl O, Leeb T. A polymorphism within the equine CRISP3 gene is associated with stallion fertility in Hanoverian warm-blood horses. *Anim Genet* 2007;38:259–64.
- [3] Topfer-Petersen E, Ekhlas-Hundrieser M, Kirchhoff C, Leeb T, Sieme H. The role of stallion seminal proteins in fertilisation. *Anim Reprod Sci* 2005;89:159–70.
- [4] Fouchecourt S, Metayer S, Locatelli A, Dacheux F, Dacheux JL. Stallion epididymal fluid proteome: qualitative and quantitative characterization; secretion and dynamic changes of major proteins. *Biol Reprod* 2000;62:1790–803.
- [5] Moura AA, Chapman DA, Koc H, Killian GJ. A comprehensive proteomic analysis of the accessory sex gland fluid from mature Holstein bulls. *Anim Reprod Sci* 2007;98:169–88.
- [6] Dacheux JL, Belghazi M, Lanson Y, Dacheux F. Human epididymal secretome and proteome. *Mol Cell Endocrinol* 2006; 250:36–42.
- [7] Martinez-Heredia J, de Mateo S, Vidal-Taboada JM, Balleca JL, Oliva R. Identification of proteomic differences in asthenozoospermic sperm samples. *Hum Reprod* 2008;23:783–91.
- [8] Novak S, Ruiz-Sanchez A, Dixon WT, Foxcroft GR, Dyck MK. Seminal plasma proteins as potential markers of relative fertility in boars. *J Androl* 2010;31:188–200.

- [9] Killian GJ, Chapman DA, Rogowski LA. Fertility-associated proteins in Holstein bull seminal plasma. *Biol Reprod* 1993;49:1202–7.
- [10] Moura AA, Chapman DA, Koc H, Killian GJ. Proteins of the cauda epididymal fluid associated with fertility of mature dairy bulls. *J Androl* 2006;27:534–41.
- [11] Brandon CI, Heusner GL, Caudle AB, Fayrer-Hosken RA. Two-dimensional polyacrylamide gel electrophoresis of equine seminal plasma proteins and their correlation with fertility. *Theriogenology* 1999;52:863–73.
- [12] Macpherson ML, Simmen RC, Simmen FA, Hernandez J, Sheerin BR, Varner DD, Loomis P, Cadario ME, Miller CD, Brinsko SP, Rigby S, Blanchard TL. Insulin-like growth factor-I and insulin-like growth factor binding protein-2 and -5 in equine seminal plasma: association with sperm characteristics and fertility. *Biol Reprod* 2002;67:648–54.
- [13] Gamboa S, Ramalho-Santos J. SNARE proteins and caveolin-1 in stallion spermatozoa: possible implications for fertility. *Theriogenology* 2005;64:275–91.
- [14] Novak S, Treacy BK, Almeida FR, Mao J, Buhi WC, Dixon WT, Foxcroft GR. Regulation of IGF-I and porcine oviductal secretory protein (pOSP) secretion into the pig oviduct in the peri-ovulatory period, and effects of previous nutrition. *Reprod Nutr Dev* 2002;42:355–72.
- [15] Eschenbruch M, Burk RR. Experimentally improved reliability of ultrasensitive silver staining of protein in polyacrylamide gels. *Anal Biochem* 1982;125:96–9.
- [16] Baker J, Hardy MP, Zhou J, Bondy C, Lupu F, Bellve AR, Efstratiadis A. Effects of an Igf1 gene null mutation on mouse reproduction. *Mol Endocrinol* 1996;10:903–18.
- [17] Yamakawa K, Yoshida K, Nishikawa H, Kato T, Iwamoto T. Comparative analysis of interindividual variations in the seminal plasma proteome of fertile men with identification of potential markers for azoospermia in infertile patients. *J Androl* 2007;28:858–65.
- [18] Schambony A, Gentzel M, Wolfes H, Raida M, Neumann U, Topfer-Petersen E. Equine CRISP-3: primary structure and expression in the male genital tract. *Biochim Biophys Acta* 1998;1387:206–16.
- [19] Da Ros VG, Maldera JA, Willis WD, Cohen DJ, Goulding EH, Gelman DM, Rubinstein M, Eddy EM, Cuasnicu PS. Impaired sperm fertilizing ability in mice lacking Cysteine-Rich Secretory Protein 1 (CRISP1). *Dev Biol* 2008;320:12–8.
- [20] Carvalho AL, Sanz L, Baretino D, Romero A, Calvete JJ, Romao MJ. Crystal structure of a prostate kallikrein isolated from stallion seminal plasma: a homologue of human PSA. *J Mol Biol* 2002;322:325–37.
- [21] Ulmert D, O'Brien MF, Bjartell AS, Lilja H. Prostate kallikrein markers in diagnosis, risk stratification and prognosis. *Nat Rev Urol* 2009;6:384–91.
- [22] Force A, Viillard JL, Grizard G, Boucher D. Enolase isoforms activities in spermatozoa from men with normospermia and abnormospermia. *J Androl* 2002;23:202–10.