Implication of transcriptome profiling of spermatozoa for stallion fertility

Yara Suliman\textsuperscript{A}, Frank Becker\textsuperscript{A,C} and Klaus Wimmers\textsuperscript{B}

\textsuperscript{A}Institute for Reproductive Biology, Leibniz Institute for Farm Animal Biology Dummerstorf, D-18196 Dummerstorf, Wilhelm-Stahl-Allee 2, Germany.

\textsuperscript{B}Institute for Genome Biology, Leibniz Institute for Farm Animal Biology Dummerstorf, D-18196 Dummerstorf, Wilhelm-Stahl-Allee 2, Germany.

\textsuperscript{C}Corresponding author. Email: becker@fbn-dummerstorf.de

Abstract. Poor fertility of breeding stallions is a recognised problem in the equine industry. The aim of the present study was to detect molecular pathways using two groups of stallions that differed in pregnancy rates as well as in the proportion of normal and motile spermatozoa. RNA was isolated from spermatozoa of each stallion and microarray data were analysed to obtain a list of genes for which transcript abundance differed between the groups ($P \leq 0.05$, fold change $\geq 1.2$). In all, there were 437 differentially expressed (DE) genes between the two groups ($P \leq 0.05$, fold change $\geq 1.2$). Next, the DE genes were analysed using Database for Annotation, Visualisation, and Integrated Discovery (DAVID). Finally, ingenuity pathways analysis (IPA) was used to identify top biological functions and significant canonical pathways associated with the DE genes. Analysis using the DAVID database showed significant enrichment in the gene ontology (GO) term ‘RNA binding’ ($P = 0.05$) and in the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway cytokine–cytokine receptor interaction ($P = 0.02$). Furthermore, IPA analysis showed interconnected biological functions and canonical pathways involved in the regulation of spermatogenesis and male fertility. In addition, significantly enriched metabolic pathways were identified. In conclusion, the present study has identified, for the first time, molecular processes in stallion spermatozoa that could be associated with stallion fertility.

Additional keywords: microarray, sperm RNA, spermatogenesis.

Received 9 July 2016, accepted 6 December 2017, published online 14 March 2018

Introduction

Stallion fertility is of high economic importance for the horse industry. Among farm animals, horses have the lowest reproductive rate. This is reflected in lower per cycle pregnancy rates, which range from 43% to 60% (Morris and Allen 2002), compared with 85–90% in boars (Colenbrander et al. 1993) and 80–90% in rams (Menzies 1999). Unlike other domestic species, the selection of equines for breeding is based on their pedigree, athletic prowess and conformation characteristics, with no consideration of fertility potential during selection (Colenbrander et al. 2003). Breeding values for different breeds are calculated for important traits such as sportive results, movement, conformation and health. Stallion fertility can be determined in different ways, including considering parameters reflecting breeding or reproduction success rates, such as pregnancy rate, foaling rate and non-return rate, which show immense individual variation. Male fertility may also be assessed using sperm characteristics, which are also highly variable (Hamann et al. 2005). Attempts to find a relationship between stallion fertility and sperm characteristics have reported inconsistent results. For example, Jasko et al. (1990a) and Hirano et al. (2001) found a significant correlation between sperm features and fertility in stallion, but Dowsett and Pattie (1982) and Voss et al. (1981) reported no correlation between sperm characteristics and fertility in the stallion. Moreover, animal-based intrinsic, genetic and physiological data sets associated with male fertility are largely unknown. In recent years, increased attention has focused on the analysis and characterisation of sperm RNAs and their role in the regulation of spermatogenesis, fertilisation and early embryo development (Boerke et al. 2007). Spermatozoa contain a large population of RNAs, including mRNA, interference RNA, antisense RNA and micro-RNA (Dadoune 2009; Hosken and Hodgson 2014). In the study of Zhao et al. (2006) a large number of RNAs was detected in human spermatozoa. The quality and quantity of these RNAs, as well as their potential roles in male fertility, are still largely unknown. It was first believed that RNAs in spermatozoa are remnants and that their sole purpose was to transfer the paternal genome to the oocyte during fertilisation (Krawetz 2005). Recently, Das et al. (2013) characterised the global transcriptome in the spermatozoa of fertile stallions and explored the important role of these transcripts in male fertility. These...
findings provided a better concept of the biological importance of sperm RNAs, allowing the identification of biomarkers of stallion fertility. The sperm transcriptome has been studied in humans (Ostermeier et al. 2002), boars (Kempisty et al. 2008; Yang et al. 2009) and bulls (Gilbert et al. 2007; Bissonnette et al. 2009; Feugang et al. 2010), and the utility of sperm RNA as a marker of infertility has been explored (Miller 2000; Steger 2001), whereby differences in transcript levels in spermatozoa with different motilities (Bissonnette et al. 2009; Lambard et al. 2004), as well as between normal and abnormal sperm samples (Platts et al. 2007; Steger et al. 2003), have been reported. Krawetz (2005) demonstrated that some of the sperm RNAs can be delivered to the oocyte during fertilisation. Martins and Krawetz (2005) identified protamine2 and clusterin in zygotes using zona-free hamster egg and human spermatozoa, which were incubated together in culture medium. These transcripts were detected in human spermatozoa and zygotes, but not in unfertilised oocytes. These transcripts could have a role in the oocyte at the time of fertilisation and during early embryonic development (Boerke et al. 2007; Carrell 2008; Lalancette et al. 2008; Bukowska et al. 2013). Despite the advances achieved in transcriptional analysis of spermatozoa in several species, genetic studies of stallion spermatozoa are still limited. Therefore, the aim of the present study was to identify molecular processes that are associated with male fertility by comparing microarray-derived sperm transcriptomes of stallions that were clearly assigned to either a fertile or subfertile group based on reproductive success and sperm characteristics. The results of the present study may help predict stallion fertility, as well as in the selection of future breeding stallions.

Materials and methods

Animals and semen collection

All animal procedures were performed in accordance with the modified Council Directive 92/65/EC (https://ec.europa.eu/food/animals/semen/equine_en, assessed 15 December 2017). No particular ethical issues had to be considered because the semen samples used in the present study had been collected as part of routine breeding procedures. The stallions were properly housed and fed, and their surroundings were kept in sanitary conditions in accordance with the requirements to operate an AI station within the EU. Six warmblood stallions (Hanoverian, Mecklenburg and Oldenburg Warmblood) were used in the present study. The age of the stallions ranged from 7 to 12 years. Stallions were selected from more than 150 active stallions with a minimum covering number of 10 mares per season. Available seasonal breeding reports were analysed by the stud manager examining results after AI with extended fresh semen. All stallions were part of a regular semen collection regimen (four times from April until July) during the breeding period. In all, 24 ejaculates were collected (four per animal) for the present study using an artificial vagina (Hannover model; Minitüb). After semen collection, the volume of each gel-free ejaculate was measured using a graduated cylinder, and the sperm concentration was determined using a photometer (Minitüb). The semen samples were then diluted using an appropriate extender (Minitüb) to a concentration of 100 × 10^6 spermatozoa mL⁻¹. The extended semen was divided into 10-mL tubes, loaded into a transport container and transported at 4°C to the laboratory.

Semen quality assessment

Stallion ejaculates were assessed individually for routine semen quality parameters such as progressive motility and normal morphology. Progressive motility was estimated using computer-aided sperm analysis (CASA; Minitüb). First, the diluted semen was diluted further in the laboratory to a concentration of 25–50 × 10^6 spermatozoa mL⁻¹ and kept at room temperature for 10–15 min. Samples were analysed in a 10-μm Leja chamber at 37°C. The percentage of spermatozoa with normal morphology was determined using eosin–nigrosin staining (Nidadavolu). Semen smears were prepared by mixing 20 μL semen with the same volume of stain. The mixture was then smeared on a glass slide and air dried. A total of 100 spermatozoa in each ejaculates was examined and classified under phase contrast microscopy oil immersion (×1000). After assessing sperm quality, 1 mL fresh semen from each ejaculate (containing 100 × 10^6 spermatozoa) was washed three times at 1200g for 10 min at 4°C with phosphate-buffered saline (PBS), which was prepared by dissolving 1 tablet of PBS (Sigma Aldrich) in 200 μL diethylpyrocarbonate (DEPC)-treated water. After 30 min to allow the table to dissolve, the PBS medium was filtered through a Syringe-filter (SARSTEDT) with a pore size of 0.20 μm and stored at 4°C until use. Washed sperm pellets were then stored at −80°C until RNA extraction.

RNA isolation and purity assessment

An RNA isolation protocol was developed to extract pure sperm RNA without contamination by somatic cell RNAs. In this protocol frozen sperm samples were placed on ice for 5–10 min to thaw, after which they were centrifuged at 5000 g for 5 min at 4°C. After the supernatant had been removed, 3 mL hypotonic solution (99 mL DEPC water supplemented with 0.5 mL of 0.1% sodium dodecyl sulfate (SDS) and 0.5 mL of 0.5% Triton X-100) was added to the sperm pellet and the mixture was pipetted up and down five to 10 times and then placed on ice for 15 min to lyse the somatic cells. After centrifugation of samples at 5000g for 15 min at 4°C, the hypotonic solution was removed and the pellet was washed three times with PBS. Then, 1 mL TRizol (1 mL for 10^6 cells in suspension) was added to each sperm pellet to isolate total RNA from the spermatozoa. The pellet was lysed by aspirating the solution and releasing it 15–20 times with a 27-gauge needle connected to a 5-mL syringe. After 15 min incubation at room temperature (with vortexing every 5 min), the homogenised samples were centrifuged at 12 000g for 10 min at 4°C. The resulting pellet represents insoluble material, such as polysaccharides, extracellular membranes and high molecular weight DNA. The supernatant, containing the RNA, was transferred to a fresh tube with 1 mL TRizol. Then, 200 μL chloroform was added to each sample and the samples were homogenised by vortexing for 30 s, followed by incubation at room temperature for 5 min and then centrifugation at 12 000g for 15 min at 4°C. During centrifugation, the mixture separates into an organic phase, an interphase and an upper aqueous phase that contains the RNA. The aqueous phase was transferred to a
fresh tube and mixed with 1 volume isopropanol (100%) and 1 μL glycogen. The samples were then incubated for 30 min at room temperature to precipitate the RNA and then centrifuged at 20 000g for 30 min at 4°C. The supernatant was removed and the pellets washed three times with 75% ethanol. The RNA was sedimented by centrifugation at 12 000g for 5 min at 4°C, the ethanol was removed and the pellets dried for 5–10 min at room temperature. The RNA pellet was then dissolved in 20 μL water at 68°C until use.

Microarray experiments and data analysis

For the microarray experiments, individual RNA samples of each of the six stallions were pooled and used to prepare microarray probes. Labelled cRNA was prepared from 500 ng total RNA using the GeneChip WT PLUS Reagent Kit (Affymetrix) according to the manufacturer’s instructions. The cRNA targets were fragmented and individually hybridised onto a catalogue genome-wide Equine Gene 1.0 ST Array (Affymetrix) according to the manufacturer’s instructions. After staining and washing, the arrays were scanned and raw data were obtained using Affymetrix GCOS 1.1 software. First, the raw microarray data were subjected to quality control using Microarray Suite 5 (MASS; present and absent) and quantitative expression levels of the transcripts were estimated using Probe Logarithmic Intensity Error (PLIER) for normalisation using Affymetrix GCOS 1.1.1 software. Then, one-way analysis of variance (ANOVA) was used to detect transcripts with significant differences in abundance between fertile and subfertile stallions. According to the Minimum Information About a Microarray Experiment (MIAME) standards (Brazma et al. 2001) the microarray data have been deposited in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo, accessed 5 January 2017) under the Accession number GSE75350.

Functional annotation and pathway analysis to identify significantly overrepresented functions

The list of genes with significantly different abundance (P ≤ 0.05 and fold change (FC) ≥ 1.2 based on Affymetrix microarray technology sensitivity; corresponding to a false discovery rate (FDR) of <0.4) in the stallion groups was submitted to the Database for Annotation, Visualisation and Integrated Discovery (DAVID; http://david.abcc.ncifcrf.gov/, accessed 16 January 2017) and the web-based ingenuity pathways analysis (IPA; www.Ingentity.com, accessed 13 January 2017) to identify gene ontology (GO) terms, canonical pathways and biological functions associated with these genes. The significance of gene enrichment was determined by Fisher’s exact test, with a cut-off set at 0.05. P-values from Fisher’s exact test were adjusted for multiple testing using Benjamini–Hochberg multiple testing corrections.

Synthesis of cDNA

In preparation for first-strand synthesis, 1 μL total RNA was incubated with 1 μL oligo (dT) 13 primer (0.5 μg μL⁻¹), 1 μL RNase inhibitor (40 U μL⁻¹; Promega), 2.5 μL random hexamer primer (0.2 μg μL⁻¹; Fermentas) and 7 μL RNase-free water at 68°C for 5 min. The samples were then placed on ice for 5 min. The reaction mix for first strand synthesis consisted of 4 μL of 5× First Strand buffer (Invitrogen), 1 μL dithiothreitol (DTT; 0.1 M; Invitrogen), 1 μL dNTPs Mix (10 mM; Invitrogen), 0.5 μL RNase-free water and 1 μL SuperScript III Reverse Transcriptase (200 units μL⁻¹; Invitrogen). Samples were incubated at 25°C for 5 min, 50°C for 1 h and then heated to 75°C for 15 min to inactivate the enzyme. Finally, samples were cooled at 4°C. The cDNA was then stored at −20°C until use.

Quantitative real-time PCR

Genomic DNA contamination of the samples was evaluated using real-time PCR (Light cycler 480; Roche) using a set of
differentially expressed genes and Functional annotation and KEGG pathway analysis of gene expression were performed in the subfertile group. Comparison of spermatozoa with normal morphology) compared with stallions with a better semen profile (progressive motility and percentage of Fertility and Development) were assessed using GO and the KEGG pathway analysis. These genes were chosen according to their biological function and, in all cases, were validated as having high transcript abundance in the fertile stallion group (Table 4).

Results
Stallion fertility and semen quality
Four ejaculates were collected from each stallion during the breeding season. Motility and morphology measurements for each stallion, as well as seasonal pregnancy rate for each group, are presented in Table 2. The stallions in the fertile group had a better semen profile (progressive motility and percentage of spermatozoa with normal morphology) compared with stallions in the subfertile group.

Functional annotation and KEGG pathway analysis of differentially expressed genes
In all, 437 genes were found to be significantly differentially expressed (DE) between the two groups of stallions \( (P \leq 0.05, \text{FDR} \leq 1.2) \). To explore another set of functionally related genes together, the list of DE genes was evaluated using DAVID, and DE genes were assessed using GO and the KEGG pathway database. DAVID assigned 216 transcripts on the gene list to GO terms and KEGG pathways. Of these, three genes \( (OAS1, OAS2, RPL5) \) belonged to GO term ‘RNA binding’ and had a \( P \) value of 0.05. Another eight genes from the gene list, namely colony-stimulating factor 3 receptor (granulocyte) \( (CSF3R) \), granulocyte–macrophage colony-stimulating factor \( (GM-CSF) \), interleukin (IL)-17A \( (IL17A) \), IL13, IL22RA1, tumour necrosis factor receptor superfamily, member 4 \( (CD134) \), CD27 and tumour necrosis factor receptor superfamily, member 9 \( (SF9) \), were found to be significantly enriched in the KEGG pathway cytokine–cytokine receptor interaction \( (P = 0.02) \), as summarised in Table 3.

Biological functions and canonical pathways of DE genes
IPA was used to select the main biological functions and canonical pathways represented in the gene list generated in the present study. The top 20 significantly enriched IPA biological functions are shown in Fig. 1. Most of the DE genes were associated with biological functions related to cellular processes, including cellular assembly and organisation, cellular function and maintenance, cellular development and cellular growth and proliferation. IPA also identified canonical pathways from the DE genes, which included G1/S checkpoint regulation, cyclin and cell cycle regulation, ATM (ataxia-telangiectasia mutated) signalling and netrin signalling (Fig. 2). Significantly enriched metabolic pathways were related to phosphatase and D-myoinositol phosphate metabolism. The enriched signalling pathways are shown in Fig. 3.

Quantitative real-time PCR validation of microarray results
To validate the microarray results, four genes from the gene list \( (OAS1, OAS2, IL13 \text{ and } IL22RA1) \) that exhibited differential expression in the microarray were selected for real-time PCR analysis. These genes were chosen according to their biological function and, in all cases, were validated as having high transcript abundance in the fertile stallion group (Table 4).

Discussion
Research into human fertility, as well as that of other mammals, has advanced considerably in recent years, but research into the molecular mechanisms affecting stallion fertility remains in its infancy. Genetic markers associated with stallion fertility may be useful in selection and breeding management. Analysis of stallion sperm transcriptome by microarray analysis in the present study allowed comparison of transcript abundance between fertile and subfertile stallions. Analysis of the sperm transcriptome by gene expression microarray is typically affected by the RNA isolation technique used. RNA degradation is one of the most crucial problems in the RNA-extraction process. The concentration of RNA in spermatozoa is lower than in somatic cells; therefore, it is essential to avoid any RNA degradation during the extraction process. Furthermore, global microarray analysis of the low quantity of RNA in samples obtained from spermatozoa compared with the quantity of RNA obtained from somatic cells has traditionally required either one or more amplification step(s) or mixing of samples depending on the RNA concentration to obtain sufficient nucleic acid.
material for subsequent detection. In the present study, the RNA samples for each stallion were pooled, glycogen was used as a coprecipitant for the small amounts of RNA and the purity of the RNA samples was checked using real-time PCR and a set of primers specific to protamine2.

In several previous studies, many genes and proteins were analysed for their association with fertility in the stallion (Schambony et al. 1998; Hamann et al. 2007; Zaabal and Ahmed 2010). For example, equine cysteine-rich secretory protein (CRISP) genes were found to be significantly associated

<table>
<thead>
<tr>
<th>Stallion Ejaculate</th>
<th>Progressive sperm motility (%)</th>
<th>Normal sperm morphology (%)</th>
<th>Seasonal pregnancy rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Individual values</td>
<td>Mean for each stallion</td>
<td>Group mean</td>
</tr>
<tr>
<td><strong>Fertile group</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>67.58</td>
<td>72.02 ± 5.35</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>81.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>70.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>68.93</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>79.98</td>
<td>72.59 ± 4.87</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>67.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>68.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>73.86</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>80.92</td>
<td>66.38 ± 9.32</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>61.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>55.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>67.13</td>
<td></td>
</tr>
<tr>
<td><strong>Subfertile group</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>40.13</td>
<td>48.74 ± 8.62</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>57.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>56.84</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>40.13</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>39.84</td>
<td>38.65 ± 3.03</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>33.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>40.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>41.12</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td>39.84</td>
<td>42.72 ± 3.22</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>47.83</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>40.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>43.12</td>
<td></td>
</tr>
</tbody>
</table>

**P-value** 0.001 0.012 0.007

Table 3. Significant gene function according to the Database for Annotation, Visualisation, and Integrated Discovery (DAVID)
KEGG, Kyoto Encyclopedia of Genes and Genomes

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>No. genes</th>
<th>Gene symbol</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOTERM-MF-ALL</td>
<td>RNA binding</td>
<td>3</td>
<td>OAS1</td>
<td>2',5'-Oligoadenylate synthetase 1, 40/46 kDa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OAS2</td>
<td>2',5'-Oligoadenylate synthetase 2, 69/71 kDa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RPL5</td>
<td>Ribosomal protein L5</td>
</tr>
<tr>
<td>KEGG pathway</td>
<td>Cytokine–cytokine receptor interaction</td>
<td>8</td>
<td>CSF3R</td>
<td>Colony stimulating factor 3 receptor (granulocyte)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GM-CSF</td>
<td>Granulocyte–macrophage colony-stimulating factor</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IL17A</td>
<td>Interleukin-17A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IL22RA1</td>
<td>Interleukin-22 receptor, α1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IL13</td>
<td>Interleukin-13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CD134 homologue</td>
<td>Tumour necrosis factor receptor superfamily, member 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CD27 ligand</td>
<td>CD70 molecule</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SF9</td>
<td>Tumour necrosis factor receptor superfamily, member 9</td>
</tr>
</tbody>
</table>
with male fertility (Schambony et al. 1998). The CRISP proteins represent the major equine fraction of seminal plasma proteins and were found to be involved in sperm–oocyte fusion (Töpfer-Petersen et al. 2005). Furthermore, Giesecke et al. (2011) chose three genes, namely angiotensin-converting enzyme (ACE), sperm autoantigenic protein 17 (SPA17) and FSH beta subunit (FSHB), to test as candidates for determining Hanoverian stallion fertility, finding that haplotypes of all three genes significantly contributed to the paternal and embryonic fertility components of the pregnancy rate per oestrus. In addition, Gamboa and Ramalho-Santos (2005) investigated the presence of soluble N-ethylmaleimide-sensitive factor activating protein receptor (SNARE) proteins in equine spermatozoa and found that, stallions with fertility problems had the worst sperm and acrosomal membrane quality, as well as fewer sperm cells that stained positive for SNAREs and caveolins compared with spermatozoa from fertile donors.

In the present study, we chose six stallions with differing fertility, as evaluated in terms of semen profile (progressive motility and percentage of spermatozoa with normal morphology) and percentage of pregnancies per season, calculated by dividing the number of all pregnant mares (for each stallion individually) by the number of mares bred during the year. Recently, many developments have been achieved in the field of gene expression analysis, providing new molecular detection tools to aid our understanding of stallion fertility (Bright et al. 2009; Novak et al. 2010). RNA in the spermatozoa could be used in genomic analysis to evaluate male fertility. Currently, several studies have emerged focusing on an analysis of transcripts or proteins in spermatozoa that could be associated with male fertility. Feugang et al. (2010) identified 415 DE transcripts from high- and low-fertility Holstein bulls. These transcripts were associated with different cellular functions and biological processes. Kropp et al. (2017) focused in their study on bull fertility and demonstrated that bulls with different fertility status delivered embryos with different transcriptome profiles despite similar morphology and adequate development to blastocyst stage. In addition, the identification of paternal components, which are delivered into the oocyte during fertilisation, could lead to a better understanding of male fertility and suitable selection of breeding sires. Several genes and molecular pathways involved in spermatogenesis and male fertility were identified by Bansal et al. (2015). In that study, analysis of 2081 transcripts that were differentially expressed between three groups of men with known fertility,
**Fig. 2.** Significant canonical pathways associated with differentially expressed genes as generated by ingenuity pathways analysis. ‘Ratio’ refers to the ratio of genes from the submitted gene list divided by the total number of genes that make up the pathway. The orange horizontal line denotes the threshold for significance ($P = 0.05$). ATM, ataxia-telangiectasia mutated; CHK, choline/ethanolamine kinase; ESC, embryonic stem cell; NFAT, nuclear factor of activated T cells; MAPK, mitogen-activated protein kinase; NK, natural killer; Oct4, octamer-binding transcription factor 4; PAK, p21-activated kinase.

**Fig. 3.** Significantly enriched metabolic pathways associated with differentially expressed genes as generated by ingenuity pathways analysis. ‘Ratio’ refers to the ratio of genes from the submitted gene list divided by the total number of genes that make up the pathway. The orange horizontal line denotes the threshold for significance ($P = 0.05$).
RNA-binding proteins

RNA-binding proteins (RBPs) are a class of proteins that are expressed in the nucleus and are defined by their ability to bind RNA. RBPs play important roles in the organism. They are expressed in the nucleus and are defined by their ability to bind RNA. RBPs play important roles in the organism. They are involved in many cellular processes that occur during tissue development. Moreover, these proteins regulate immune activities in addition to their original functions in the immune system (Miller et al. 1999; Idler and Yan 2012). Cytokines act as growth and differentiation factors during pathological states, as well as under normal physiological conditions (Hales et al. 1999). Various immunological factors, cytokines, chemokines and growth factors have been documented in human semen (Poltich et al. 2007), as well as that from rodents (Ingman and Rebecca 2008) and other livestock species (Paulus et al. 2010). These molecules in semen affect vaginal immunology and female fertility following insemination and play an important role in regulating the proliferation, viability and differentiation of blastomeres in embryos (Kane et al. 1997). However, in the present study we investigated the presence of cytokines in stallion spermatozoa whose quantitative expression levels were higher in fertile than subfertile stallions. It has been reported that cytokines and growth factors play an essential role in both the implantation and development of embryos (Sharkey 1998). Robertson et al. (2001) reported the important role of GM-CSF in murine preimplantation embryos by promoting blastocyst formation and increasing the number of viable blastomeres by reducing the incidence of highly condensed and the sperm nucleus is transcriptionally inert and does not contain sufficient RNA to support translation (Miller et al. 1999). In post-meiotic spermatozoa, de novo transcription is silenced and the synthesis and storage of sufficient mRNA is very important to compensate for the lack of mRNAs during the subsequent transcriptionally inactive stage of spermatogenesis (Sassone-Corsi 2002). Germ cells have been shown to express high levels of RBPs throughout spermatogenesis, which are very essential to post-transcriptional events during all stages of spermatogenesis (Idler and Yan 2012). According to DAVID, in the present study three genes from the DE genes were involved in GO term ‘RNA binding’ (with \( P = 0.05 \); Table 3). This result highlights the importance of RBPs in the regulation of spermatogenesis and male fertility. Bettegowda and Wilkinson Miles (2010) reported the role of four RBPs, namely gonadotrophin-regulated testicular RNA helicase (GRTH), Src-associated in mitosis 68 kDa (SAM68), mouse Y-box protein 2 (MSY2) and deleted in azoospermia-associated protein-1 (DAZAPI), that were shown to be critical for spermatogenesis by promoting the translation of a subset of germ cell mRNAs, indicating the importance of translation control for normal spermatogenesis. Furthermore, Idler and Yan (2012) reported that in mice, the inactivation of many of the RBPs studied caused infertility and spermatogenic arrest at various steps of germ cell differentiation.

Cytokine–cytokine receptor interaction

The results of the present study also indicated that eight genes from the gene list created (i.e. CSF3R, GM-CSF, IL17A, IL13, IL22RA1, CD134, CD27 and SF9) were significantly enriched in the KEGG pathway cytokine–cytokine receptor interaction (\( P = 0.02 \); Fig. 4). Cytokines are a broadly defined group of regulatory proteins that have a wide range of biological activities in addition to their original functions in the immune system (Hales et al. 1999; Ihsan 2014). Several cytokines have direct effects on testicular cell function, and several of these are involved in normal reproductive physiology and fertility regulation (Hales et al. 1999; Hedger and Meinhardt 2003). Cytokines act as growth and differentiation factors during pathological states, as well as under normal physiological conditions (Hales et al. 1999). Various immunological factors, cytokines, chemokines and growth factors have been documented in human semen (Poltich et al. 2007), as well as that from rodents (Ingman and Rebecca 2008) and other livestock species (Paulus et al. 2010). These molecules in semen affect vaginal immunology and female fertility following insemination and play an important role in regulating the proliferation, viability and differentiation of blastomeres in embryos (Kane et al. 1997). However, in the present study we investigated the presence of cytokines in stallion spermatozoa whose quantitative expression levels were higher in fertile than subfertile stallions. It has been reported that cytokines and growth factors play an essential role in both the implantation and development of embryos (Sharkey 1998). Robertson et al. (2001) reported the important role of GM-CSF in murine preimplantation embryos by promoting blastocyst formation and increasing the number of viable blastomeres by reducing the incidence of

### Table 4. Validation of microarray results using real-time polymerase chain reaction (PCR) for selected genes

<table>
<thead>
<tr>
<th>Stallion</th>
<th>( \Delta Ct ) OAS1</th>
<th>( \Delta Ct ) OAS2</th>
<th>( \Delta Ct ) IL13</th>
<th>( \Delta Ct ) IL22RA1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fertile group A</td>
<td>-5.5</td>
<td>-3.9</td>
<td>-4.7</td>
<td>-2.0</td>
</tr>
<tr>
<td>B</td>
<td>-5.5</td>
<td>-3.2</td>
<td>-4.8</td>
<td>-4.5</td>
</tr>
<tr>
<td>C</td>
<td>-6.6</td>
<td>-3.0</td>
<td>-4.9</td>
<td>-1.4</td>
</tr>
<tr>
<td>Subfertile group D</td>
<td>-2.5</td>
<td>-0.2</td>
<td>-1.0</td>
<td>1.1</td>
</tr>
<tr>
<td>E</td>
<td>-4.7</td>
<td>-0.4</td>
<td>-3.7</td>
<td>-0.1</td>
</tr>
<tr>
<td>F</td>
<td>-3.5</td>
<td>-1.9</td>
<td>-2.0</td>
<td>2.1</td>
</tr>
</tbody>
</table>

**P-values**
- real-time PCR: 0.035, 0.017, 0.031, 0.033
- Microarray: 0.023, 0.0071, 0.004, 0.0026

shown that some of these transcripts were related to heat shock proteins, testis-specific genes and Y chromosome genes, which suggested that sperm RNA has considerable potential as a marker in the evaluation of fertility. Proteomics has also been used to identify putative biomarkers associated with male fertility. In particular, Ashrafzadeh et al. (2013) identified several proteins that differed in abundance between two groups of fertile and subfertile zebu cattle using frozen semen samples. These proteins were involved in energy metabolism, glycolysis, sperm motility and male fertility. Similarly, in the present study we compared the RNA profile of spermatozoa obtained from fertile and subfertile stallions using Affymetrix microarray technology. The analysis identified 437 transcripts that differed significantly between fertile and subfertile stallions. These transcripts were associated with different biological functions and molecular processes known to be fundamental for spermatogenesis and male fertility.
apoptosis and increasing the uptake of glucose. In addition, the in vitro culture of human embryos in GM-CSF demonstrated that GM-CSF has a positive effect on the development of embryos to the blastocyst stage by increasing the number of cells in the inner cell mass and in the trophectoderm (Sjöblom et al. 1999). Many other cytokines have been studied extensively in mice and other mammalian species, including IL-1, IL-2, IL-6, IL-8, IL-10, colony-stimulating factor 1 (CSF1) and anti-inflammatory cytokines of the transforming growth factor β family, because there is strong evidence that some of them are directly involved in the production of male and female gametes, embryo implantation and development (Sharkey 1998; Simón et al. 1998; Hedger and Meinhardt 2003; Robertson et al. 2007).

**Biological functions and canonical and metabolic pathways of DE genes**

IPA was used to obtain a deeper understanding of the significantly altered biological functions and canonical pathways associated with the DE genes that were identified in the microarray analysis. IPA of DE genes showed several different associated biological functions. The target genes were significantly enriched for diverse cellular functions and biological processes known to be fundamental for spermatogenesis and male fertility, including cellular assembly and organisation, cellular function and maintenance, cellular development, cellular growth and proliferation, tissue development, cell morphology and organisational injury and abnormalities. Further analysis was also conducted in IPA to understand the enriched

---

**Fig. 4.** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway for cytokine–cytokine receptor interaction (in the horse) as generated by Database for Annotation, Visualisation, and Integrated Discovery (DAVID). Genes that were significantly different abundant in the pathway are denoted by a red star.
and significant canonical pathways of DE genes identified in this dataset. In all, 17 significantly enriched canonical pathways were identified, using threshold values of $P < 0.05$ and FC $> 1.2$ (corresponding to an FDR $< 0.4$). Of these pathways, significance was higher (based on $P$-values) for pathways associated with G1/S checkpoint regulation, cyclin and cell cycle regulation, ATM signalling, netrin signalling and natural killer cell signalling. Furthermore, IPA identified significantly enriched metabolic pathways that were related to phosphatase and D-myoinositol phosphate metabolism. Sperm formation refers to the differentiation of diploid spermatagonia into mature spermatozoa. During this transformation, spermatagonia grow in the first gap phase (G1), synthesise DNA in the synthesis phase (S), prepare for mitosis in the second gap phase (G2) and undergo mitosis and meiosis in the mitosis phase (M) (Ruwanpura et al. 2010). The G1 phase is the longest phase in the cell cycle, during which the cell grows continuously and undergoes most of its physiological activity. The next phase of the cell cycle is the S phase, during which DNA replication occurs. Therefore, passing the G1 phase is tightly controlled by both cell cycle machinery and checkpoint pathways (Neganova and Lako 2008). The G1/S checkpoint is essential to ensure that all conditions of cell division are successfully met and to guarantee the possibility of replicating the DNA and producing healthy daughter cells. This transition is signalled and controlled by cell cycle regulatory proteins, and have been identified in simple as well as higher-order organisms (Wolgemuth et al. 2013). Cyclins are the principal cell cycle regulatory proteins, and have been identified in simple as well as higher-order organisms (Wolgemuth et al. 2002). Cyclins can be divided into eight classes (cyclin A–H) according to their amino acid similarity and the timing of their appearance during the cell cycle (Wolgemuth et al. 2002). Although the function of cyclins during mitosis and meiosis has been studied in somatic cells, their function during mitosis and meiosis in the germ line is poorly understood. Wolgemuth et al. (2013) reported that cyclin A2 is highly expressed in differentiated Type A and B spermatogonia and plays an important role during both meiosis and mitosis in the germ cell line. Based on the IPA in the present study, G1/S checkpoint regulation and cyclin and cell cycle regulation are the most highly significantly enriched canonical pathway identified for the DE genes. These pathways are known to be the principle regulatory elements involved in cell cycle regulation and therefore play an important role during sperm proliferation and male fertility. Using IPA of DE genes, several significantly enriched metabolic pathways related to phosphatase and D-myoinositol phosphate metabolism signalling pathways were identified. Inositol is a cyclic polyl that is considered as a member of vitamin B complex. Myoinositol is the most available form of inositol in nature. Myoinositol is distributed in the tissues of many organisms and plays an essential role in the signal transduction system in cells (Condorelli et al. 2011). In addition, inositol is involved in several activities and pathways, such as mRNA transcription, cytoskeleton and chromatin remodelling and P53 activity (Bizzarri et al. 2016). In conclusion, myoinositol plays a key role in the male reproductive system by regulating the osmolality of the seminal plasma, as well as sperm motility and acrosome reaction (Condorelli et al. 2011). Condorelli et al. (2012) suggested that myoinositol significantly increased the percentage of spermatozoa with progressive motility based on 2 h incubation of human spermatozoa with 2 mg mL$^{-1}$ myoinositol. The increase in progressive motility of spermatozoa was associated with a significant increase in the percentage of spermatozoa with high mitochondrial membrane potential. In addition, inositol has been reported to affect different processes involved in oocyte fertilisation, which improve the penetration of the ovum cumulus oophorus, binding with the zona pellucida and the acrosome reaction (Calogero et al. 2015). Furthermore, high concentrations of myoinositol have been found in the fluid of the seminiferous tubule in many mammals (Setchell et al. 1968; Hinton et al. 1980; Chauvin and Griswold 2004). Myoinositol is produced by two enzymatic steps from glucose-6-phosphate, and its synthesis is regulated by myoinositol-1-phosphate synthase and myoinositol monophosphatase-1. These activity of these two enzymes is high in the testes (Chauvin and Griswold 2004). These findings support our identification of myoinositol within the top metabolic pathways in the present study.

**Conclusion**

In summary, prediction of stallion fertility is one of the most important problems in horse breeding and the analysis of semen characteristics has limited capacity for predicting fertility. In addition, there are only a few genetic studies on stallion fertility, which may provide possibilities and new tools that would help in predicting male fertility. In the present study, comparison of transcript abundance between fertile and subfertile stallions provided a glimpse into the various functional pathways involved in stallion fertility. Furthermore, this analysis confirmed the involvement of the immune response in stallion fertility. These results identify molecular processes in spermatozoa that could be associated with stallion fertility.

**Conflicts of interest**

The authors declare no conflicts of interest.

**References**


