From the
Department of Veterinary Sciences
Faculty of Veterinary Medicine
Ludwig-Maximilians-Universität München
Chair for Molecular Animal Breeding and Biotechnology
Prof. Dr. E. Wolf

Qualitative and quantitative analysis of porcine sperm transcripts and characterization of a normalized cDNA library

Thesis for the attainment of the title Doctor in Veterinary Medicine
from the Faculty of Veterinary Medicine of the
Ludwig-Maximilians-Universität München

by
Lucio Pereira Rauber
from Santa Cruz do Sul - Brazil

Munich 2008

Gedruckt mit Genehmigung der Tierärztlichen Fakultät
der Ludwig-Maximilians-Universität München

Dekan:	Univ.-Prof. Dr. Braun
Berichterstatter:	Univ.-Prof. Dr. Wolf
Korreferent:	Univ.-Prof. Dr. Heinritzi

Tag der Promotion: 18. Juli 2008

Aus dem
Department für Veterinärwissenschaften
Tierärztliche Fakultät
Ludwig-Maximilians-Universität München
Lehrstuhl für Molekulare Tierzucht und Biotechnologie
Prof. Dr. E. Wolf

Qualitative und quantitative Analyse porciner Spermientranskripte und Charakterisierung der normalisierten cDNA-Bank

Inaugural Dissertation
zur Erlangung der tiermedizinischen Doktorwürde
der Tierärztlichen Fakultät der Ludwig-Maximilians-Universität München

von
Lucio Pereira Rauber
aus Santa Cruz do Sul – Brasilien

München 2008

Gedruckt mit Unterstützung des Deutschen Akademischen Austauchdienstes

Index

Abbreviations III
List of figures V
List of tables VI
1. Introduction 1
2. Literature 3
2.1 Spermatogenesis 3
2.1.1 Proliferative Phase 3
2.1.2 Meiotic Phase 5
2.1.3 Differentiation 5
2.1.4 Spermiogenesis 6
2.2 Transcripts, Proteins and their Regulation in Sperms 6
2.2.1 Nucleus Proteins 6
2.2.2 Transcription 7
2.2.3 Transcripts 9
2.2.4 Posttranscriptional Regulation 10
2.2.5 Alternative Transcripts 10
2.2.6 Non-coding RNA 11
2.2.7 Micro RNA and RNA Interference 12
2.2.8 PIWI-Interacting RNA 13
2.2.9 Translational Regulation 14
2.3 Chromatoid Body 16
2.4 Residual Bodies 16
2.5 Ribosomal Proteins and Translation Factors 17
2.6 Apoptosis 18
2.7 Cytoplasmic Bridges 19
2.8 Mitochondrion 20
2.9 Polyspermy 21
2.10 Classical Techniques for Semen Evaluation 21
2.11 Molecular Biology 22
2.11.1 Overview 22
2.11.2 Microarray techniques 25
2.12 Damaged DNA and paternal contribution 25
3. Material and Methods 27
3.1 Material 27

3.2 Equipment....27
3.2.1 Equipment for Gel Electrophoresis....27
3.2.2 PCR Equipment....28
3.2.3 Centrifuges....28
3.2.4 Chemicals....28
3.2.5 Solutions and Reagents....30
3.2.6 Water for PCR and RNA....32
3.2.7 Marker....32
3.2.8 Reagents and Enzymes....32
3.2.9 Other Materials....33
3.2.10 Oligonucleotides....33
3.3 Methods....34
3.3.1 Agarose Gel Electrophoresis....34
3.3.2 Sperm Recovery....34
3.3.3 Handling RNA....35
3.3.4 RNA Extraction....35
3.3.5 Generation of a Normalized cDNA Library....37
3.3.6 cDNA array Hybridization....41
4. Results....44
4.1 RNA Extraction....44
4.2 Normalized cDNA Library....54
4.3 Comparative cDNA Array Hybridization of Sperm RNA from a Polyspermy Boar and a Normal Control Boar....62
5. Discussion....67
5.1 Extraction and Characterization of RNA Isolated from Porcine Spermatozoa....67
5.2 Characterization of a Normalized cDNA Library Produced from Porcine Spermatozoa....70
5.3 Comparative Microarray Hybridization with Samples Derived from a Boar with Polyspermy and a Control Boar....75
5.4 Molecular Biology Tools for the Investigation of Infertility....76
6. Summary....78
7. Zusammenfassung....80
8. References....82
9. Acknowledgments....96

Abbreviations

°C	degree Celsius
µg	Microgram
µJ	Micro Joule (Energy, work, heat)
µl	Microlitre
µm	Micrometre
µM	Micromolar (µmol/l)
AGPC	Acid Guanidinium Thiocyanate/Phenol-Chloroform
AI	Artificial insemination
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
bp	Base pair
cDNA	Complementary DNA
cfu	Colony-formating unit
cm	Centimetre
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside (5')-triphosphate
ddNTP	2',3'-Dideoxynucleoside triphosphate
(ds)cDNA	Double-strand cDNA
(ds)RNA	Double-strand RNA
DTT	1,4-Dithio-DL-threitol
E. coli	*Escherichia coli*
EDTA	Ethylene diaminetetraacetate
e.g.	*exempli gratia* (for example)
EST	Expressed Sequence Tag
EtBr	Ethidium bromide
EtOH	Ethanol
g	Gram
g	relative centrifugal force (RCF)
G	Gauge
h	Hour
HeLa	Henrietta Lacks, cell line
i.e.	*i est* (that is)
IVF	*In vitro* Fertilisation
IVP	*In vitro* Production
kb	Kilobase
kg	Kilogram
l	Litre
LB	Luria-Bertani broth or lysogeny broth
M	Molar (mol/l)
mg	Milligram

min	Minute
ml	Millilitre
mm	Millimeter
mM	Millimolar (mmol/l)
miRNA	microRNA
mRNA	messenger RNA
mRNP	messenger Ribonucleoprotein
NaOAc	Sodium acetate
ncRNA	non-coding RNA
ng	Nanogram
nm	Nanometre
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
pg	Picogram
pH	*potentia Hydrogenii*
pmol	Picomol
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
rpm	Revolutions per minute
rRNA	ribosomal RNA
RT	Room temperature
RT-PCR	Reverse Transcription-PCR
sec	Second
SDS	sodium dodecyl sulfate
siRNA	small interfering RNA
(ss)cDNA	single-strand cDNA
TBE	Tris/borate/EDTA-Buffer
TE	Tris/EDTA-Buffer
tRNA	transfer RNA
Tris	Tris-(hydroxymethyl)-aminomethane
TU	Transcriptional unit
UTR	Untranslated region
UV	Ultra-Violett
V	Volt

List of figures

Figure 1: The spermatogenesis....4
Figure 2: Control of gene expression during spermatogenesis....8
Figure 3: N0 cDNA and N1 normalized cDNA....37
Figure 4: Multiprobe II Automated Liquid Handling System....40
Figure 5: Omnigrid Accent microarrayer....42
Figure 6**:** Spermatozoa after different treatments before and after treatment with Triton X-100, TRIzol and guanidinium thiocyanate....44
Figure 7**:** RNA extraction from the pellet after TRIzol treatment....45
Figure 8**:** Agarose gel analysis of sperm RNA after TRIzol extraction and AGPC extraction of the pellet after TRIzol homogenization....46
Figure 9: Comparison between standard TRIzol (T) protocol and heated TRIzol (HT) extraction....47
Figure 10: Comparison of 300 ng total RNA extracted from spermatozoa and RNA from HeLa cells....48
Figure 11: Bioanalyzer-electrophoresis profile of spermatozoa RNA....49
Figure 12: Assessment of RNA quality using spectrophotometric absorbance....50
Figure 13: Absorbance curve of RNA extracted from cells COLO 357 precipitated with or without glycogen as co-precipitant....51
Figure 14: Sperm RNA extracted with TRIzol and precipitated with: **1**- 5 M Na-citrate with 10% isopropanol, **2**- 5 M Na-citrate with 15% isopropanol, **3**- 5 M Na-citrate with 20% isopropanol....52
Figure 15: Distribution of amplified RNA derived from spermatozoa and HeLa cells after one-step amplification (for 12 h) with the MessageAmp II aRNA Amplification Kit....53
Figure 16: Agarose gel electrophoresis of amplified cDNA clones....55
Figure 17: Overview of DNA sequencing results....56
Figure 18: Illustration of the cellular functions of the genes found in the cDNA library....61
Figure 19: Array evaluation using AIDA Image Analyzer software....63
Figure 20: Significance analysis of microarray (SAM) plot for the data derived from one microarray membrane....64
Figure 21: Hybridization of RNA from two boars with the microarrays derived from the normalized cDNA library....66

List of tables

Table 1: Different regulated transcripts according to the literature. 23
Table 2: 260/280 and 260/230 ratios from RNA precipitated with 5 M Na-citrate and different concentrations of isopropanol. 53
Table 3: Amplification of total RNA from spermatozoa and HeLa cells pure or mixed. 54
Table 4. The twenty most prominent functional groups of genes in the normalized cDNA library for boar spermatozoa using the DAVID software suite. 57
Table 5. List of the 30 most frequent transcripts found in the normalized cDNA library for ejaculated spermatozoa from boar. 58
Table 6: Genes with higher RNA concentrations in samples derived from the boar without polyspermy. 65

1. Introduction

Male germ cells are highly specialized cells, adapted to the functions of transport and delivery of the male's genetic material to the female gamete. In the spermatozoon development, the male germ cell passes through three distinct phases. First, a proliferative phase, second a meiotic and, finally, a differentiation phase. Until beginning of the differentiation, the nuclear proteins (histones) are the same for somatic and spermatogenic cells. Then the histones are gradually replaced by protamines. This allows a six-fold higher condensation of the chromatin as with histones and enables the potential of transport and delivery of male genetic information to the oocyte. This nuclear packaging is accompanied by morphological changes like acrosome and tail formation and loss of part of the cytoplasm. Furthermore the replacement of histones by protamines in the late spermiogenesis terminates the transcription of messenger RNA. But transcribed RNA from former stages are stored in ribonucleoproteins (RNP) for later translation e.g. in the late spermatogenesis or during fertilization. These events are under stringent translational control to ensure appropriate temporal and spatial expression. Ejaculated spermatozoa contain high levels of repetitive transcripts, which are at least partly adenylated and play different functions in the cell, from basal metabolism to capacitation and fertilization. However, the origin of the sperm RNA is not completely understood.

The *in vitro* evaluation of semen is an important tool for pathologic investigation, but is typically used when the fertility is already affected. It is complementary to the clinical investigation and represents only a moment in the reproductive life of the male. Under routine production conditions only ejaculates with low quality are identified by semen analysis. Furthermore, it is difficult to differentiate samples of good quality from the intermediate ones, especially in pigs that are sensitive to temperature and humidity variations. The most valid fertility assessment is the pregnancy rate, but it is subjected to variations like female and insemination conditions. The development of new investigation techniques, like RNA profiling in sperm cells, could be a helpful tool for the complementary diagnosis and for the prognosis of the fertilization capacity.

The transcriptome is the set of all RNAs produced in a given cell population. The study of transcriptome, often using techniques based on microarray technology, examines the RNA profile at any given time. In the case of sperm cells it represents the past (spermatogenesis) and the future, represented by stored transcripts related to capacitation and fertilization. The characterization of the boar spermatozoa transcriptome may help us to understand the function of these molecules in the sperm cells and relate them to further events like fertilization and early embryo development.

The aim of this project was the characterization of the boar spermatozoa transcriptome and the production of a cDNA library, as basis for further investigations of infertility or other phenomena, using the microarrays technology. The knowledge of the mechanisms underlying successful fertilization could help to reduce losses related to semen deficiency, e.g. by supplementation of the missing factor. The combination of a normalized cDNA library and microarray technology was applied to identify transcript differences between boars with normal or pathological states, in this case polyspermy.

2. Literature

2.1 Spermatogenesis

Spermatogenesis is the process of cell division and differentiation that starts in the seminiferous tubules with undifferentiated spermatogonia and results in the production of spermatozoa (Hendriksen 1999). No other cell type undergoes such extreme morphologic change while carrying out both genetic recombination and a reduction in chromosomal ploidy (Hecht 1998).

The total duration of spermatogenesis is approximately 41 days in boars (estimated to be 30–75 days in mammals) and is under the control of the germ cell gene (França *et al.* 2005, Almeida *et al.* 2006). As in all mammals, the spermatogenic process in pigs is composed of three functionally and morphologically distinct phases: spermatogonial (proliferative or mitotic), spermatocytary (meiotic) and spermiogenic (differentiation), which are under the control of specific regulatory mechanisms. The duration of the spermatogonial phase is approximately 14 days, and the life spans of primary spermatocytes and spermatids are approximately 13 days and 14 days, respectively. The differentiation of spermatogenic cells involves profound changes in the protein composition and ultrastructure of the nucleus, flagellum, mitochondria and Golgi apparatus (reviewed in Kleene 2003). When considering the changes in nuclear shape and chromatin texture, three categories of spermatids can be distinguished: the early spermatids with round nuclei, the intermediate spermatids with elongating nuclei, and the mature spermatids with condensed nuclei.

2.1.1 Proliferative Phase

At the start of spermatogenesis, diploid undifferentiated spermatogonia (stem cells, located in the seminiferous epithelium) proliferate producing three populations of cells with markedly different destinies (Hecht 1998). A part transform into differentiated spermatogonia, type A1. Through a well-controlled series of mitotic cell divisions, spermatogonia A1 give rise to a high number of more advanced spermatogonia (França *et al.* 2005). Four classes of spermatogonia are present in pigs: undifferentiated type A, differentiated type A, intermediate and type B spermatogonia.

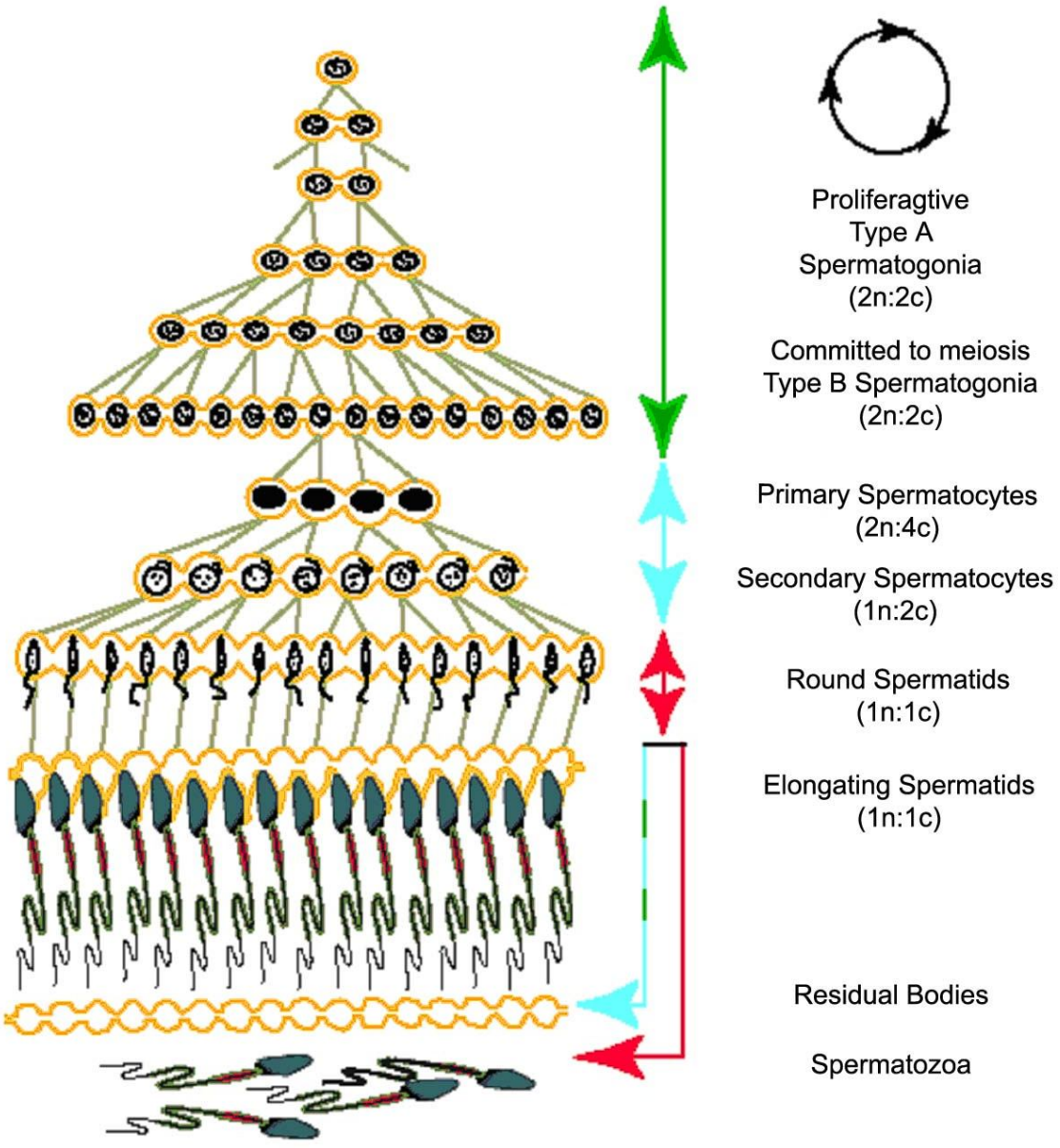

Figure 1: The spermatogenesis (taken from Miller and Ostermeier 2006). On the right side, (n:c) represents the ploidy and the number of chromatids, respectively.

For instance, in several mammalian species such as rats, mice, hamsters, rams and boars, one spermatogonial stem cell goes through approximately 10 mitotic divisions before

differentiation into a spermatocyte occurs (Figure 1). That results e.g. in 4 to 40 million spermatozoa daily per gram of testis tissue (reviewed in França *et al.* 2005).

2.1.2 Meiotic Phase

Type B spermatogonia divide forming the primary spermatocytes, which are the cells that proceed through meiosis. Through the first meiotic division, one primary spermatocyte yields two secondary spermatocytes, which almost immediately enter the second meiotic division resulting in four round spermatids. The haploid round spermatids do not divide further but differentiate into elongating spermatids and ultimately into spermatozoa (Hendriksen 1999).

Meiosis starts by DNA replication that copies each chromosome into two sister chromatids, after which the maternal and paternal copies of each chromosome pair undergo recombination (Hendriksen 1999). The chromosome pairing is preceded by chromosomal condensation in the leptotene and zygotene stages of primary spermatocytes. Thus, a functionally tetraploid primary spermatocyte, which undergoes two divisions without DNA replication, is produced (Hecht 1998). The secondary spermatocytes, formed after the first meiotic division, have a DNA content of 2n. These cells are actually haploid, containing either the X or Y chromosome (Hendriksen 1999). One reason for the meiosis is the reduction of chromosomal ploidy and maybe an opportunity for DNA repair during the lengthy interval of meiotic prophase (Hecht 1998).

2.1.3 Differentiation

After the second meiotic division, the haploid round spermatid differentiates into a spermatozoon (Hecht 1998). Specific characteristics found in different species are the shape and size of the acrosome, the size of the flagellum, and the shape of the mature spermatid nucleus, which is spatulated in boars. The mammalian spermatozoa undergo a variety of morphological, metabolic and biochemical modifications, such as nuclear condensation, development and formation of the acrosome, fibrous sheath development and loss of organelles and a portion of the cytoplasm as a cytoplasmic droplet (França *et al.* 2005).

These modifications start during the development of germ cells up to their maturation in the extratesticular spermducts (Hecht 1998). Sertoli cells divide the seminiferous epithelium into two compartments: a basal compartment where cells are exposed to the surrounding milieu

and a luminal compartment where cells are sequestered behind the blood-testis barrier formed by junctional complexes between Sertoli cells, thereby removing these germ cells from direct effects of circulatory factors. The spermatocytes move from the basal membrane to the adluminal compartment of the seminiferous tubules. The meiotic phase occurs in the basal compartment, while the meiotic and postmeiotic phases occur in the luminal compartment (Eddy 2002). A close interaction between the developing germ and Sertoli cells is established, allowing intercellular transport of nutrients and bidirectional exchanges of regulatory molecules (Hecht 1998).

2.1.4 Spermiogenesis

Spermiogenesis can be divided into two phases. In the first phase, the nucleus is round, contains histones as the major basic nuclear proteins and is transcriptionally active. The second phase involves dramatic changes in chromatin structure, nuclear shaping, and condensation (Meistrich *et al.* 2003). After nuclear elongation starts, the histones are replaced by transition proteins (TNP1 and TNP2) which are replaced by protamines (PRM1 and PRM2) constituting the nuclear proteins of elongated spermatids and mature sperms (reviewed in Steger 2001 and Meistrich *et al.* 2003). This chromatin condensation inhibits the transcription of new RNA. A rigid sperm nucleus may be required 1) for successful transport across the female reproductive tract; 2) for penetration of the oocyte vestments, or 3) for stabilization of the chromatin to compensate the lack of DNA repair enzymes (Dadoune 2003).

2.2 Transcripts, Proteins and their Regulation in Sperms

2.2.1 Nucleus Proteins

Studies performed by França *et al.* (2005) showed that a striking increase in cell size is observed in the primary spermatocytes of pigs from preleptotene to diplotene. This increase is followed by a dramatic decrease in cell size during spermiogenesis due to changes in chromatin and nuclear condensation. The nuclear volume before spermiation attains only 2% of its initial volume. In the course of vertebrate spermatogenesis, the somatic histones undergo several transitions. These involve the replacement by different testis-specific

histones before the final replacement by protamines at the end of spermatogenesis (Wouters-Tyrou *et al.* 1998). During elongation and condensation of the spermatid nucleus, transition proteins are synthesized prior to the deposition of protamines (Dadoune 2003).

It is known that protamines are phosphorylated before binding to DNA and a substantial dephosphorylation takes place concomitant with nucleoprotamine maturation. After binding to the DNA, the formation of disulphide bonds between protamines also stabilizes the nucleoprotamine complex. Protamines are proteins that contain increased numbers of positively charged amino acids, particularly arginine. This positive charge allows the formation of a highly condensed complex with the paternal genomic DNA, which has a strong negative charge (Wouters-Tyrou *et al.* 1998). The central polyarginine-rich segment of protamines lies within the minor groove of DNA, cross-linking and neutralizing the phosphodiester backbone of the macromolecule, while the carboxy- and amino-terminal residues of the protein participate in the formation of inter- and intraprotamine bridges and hydrophobic and disulfide bonds (Oliva 2006).

Sperm DNA is organized in the most tightly compact chromatide, at least six times more condensed than DNA in mitotic chromosomes (Ward 1993). The chromatin condensation results in the cessation of transcription at a time when many proteins need to be synthesized and assembled for the complete condensation of the chromatin, acrosome development, and flagellum formation (Steger 2001). Specific mRNA mechanisms are implicated by the delays in the translation, such as the mouse protamine 1 and 2 mRNA. These are translationally inert in round spermatids and active in elongated spermatids (Kleene 2003). Premature translation of *Prm1* mRNA, in mice, causes early condensation of spermatid nuclear DNA, abnormal head morphogenesis, and incomplete processing of Prm2, resulting in male infertility (Lee *et al.* 1995).

2.2.2 Transcription

A high rate of RNA synthesis was observed in type A spermatogonia while it gradually decreased in intermediate and type B spermatogonia and in preleptotene spermatocytes. No RNA synthesis was found in leptotene, zygotene and early pachytene spermatocytes (Söderström and Parvinen 1976). A sudden increase of transcriptional activity occurrs in mid-pachytene primary spermatocytes. The rate of RNA synthesis gradually decreases towards the end of the pachytene stage, and cells in diplotene and at the meiotic divisions do not

synthesize RNA at all. A low rate of RNA synthesis occurs in secondary spermatocytes and in the early round-nucleated spermatids up to the acrosome phase of spermiogenesis. No RNA synthesis was detected during late spermiogenesis (Kierszenbaum and Tres 1975, Söderström and Parvinen 1976, Kleene 2003), and ejaculated spermatozoa did not show synthesis of new RNA (Grunewald *et al.* 2005 – Figure 2).

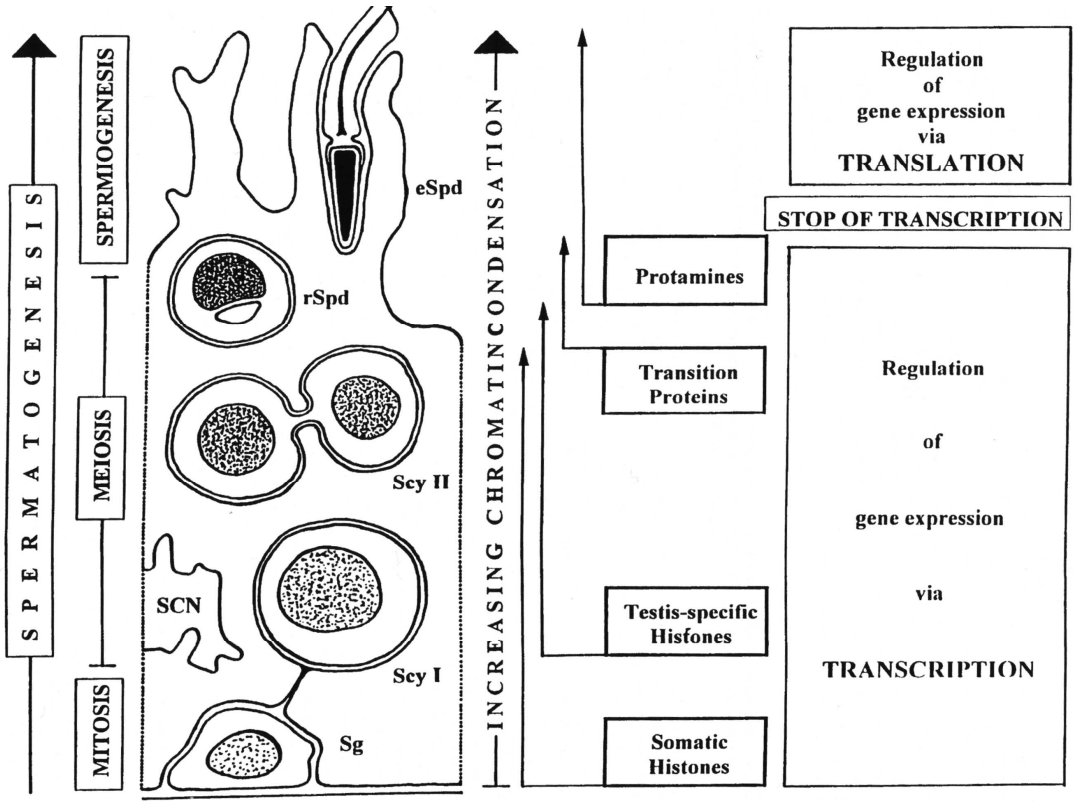

Figure 2: Control of gene expression during spermatogenesis. Sg: Spermatogonia; ScyI: Primary Spermatocyte; ScyII: Secondary Spermatocyte; rSpd: Round Spermatid and eSpd: Elongating Spermatid (taken from Steger 2001).

The time of transcription and degradation of some mRNAs is regulated individually during spermatogenesis (Shih and Kleene 1992). An intrinsic germ cell genetic program probably controls the transcription and translation steps through mitotic, meiotic and post-meiotic phases of development (Eddy 1999, Miller *et al.* 1999). In mammals, round spermatids

contain a number of transcripts that are produced either throughout early spermatogenesis (Eddy 2002) or during spermiogenesis from the haploid genes coding for sperm-specific proteins such as transition proteins and protamines (Steger 2001) or sperm tail cytoskeletal proteins implied in the molecular make-up of the outer dense fibres (Petersen *et al.* 1999) and fibrous sheath (Eddy *et al.* 2003).

2.2.3 Transcripts

Testis-specific mRNAs are synthesized and placed under stringent translational control to ensure appropriate temporal and spatial expression (Penttiliä *et al.* 1995). The virtuall absence of cytoplasm and the insufficiency of ribosomes inhibit the spermatozoa to support normal translation (Miller and Ostermeier 2006). Using mass spectrometry on whole sperm of *Drosophila melanogaster*, Dorus *et al.* (2006) identified 342 unique proteins, many involved in energy production and cytoskeletal functions. Peddinti *et al.* (2008) showed that expression of 2,051 and 2,281 proteins was specific to high and low fertility bull spermatozoa, respectively and 1,518 proteins were common to both. However, there is no evidence for the presence of all proteins corresponding to the majority of mRNAs found in mature spermatozoa (Dadoune *et al.* 2005).

Ejaculated spermatozoa contain high levels of repeated transcripts that are more accessible to cDNA conversion than ‘‘functional’’ mRNA (Miller 2000). Gilbert *et al.* (2007) suggested that the spermatic RNA population is mainly composed of naturally truncated mRNAs, which are at least partly adenylated. Zhao *et al.* (2006) characterized the transcripts from human ejaculated spermatozoa in nucleic acid binding, catalytic activity, signal transducer and structural molecule. Three main categories of transcripts were identified by Lalancette *et al.* (2008) in bovine spermatozoa: mitochondrial, ribosomal and unknown.

Many genes that are expressed in spermatogenic and somatic cells, like α-tubulin or heat shock protein 70 (*HSP70*), produce transcripts that differ in structure due to alternative promoters, alternative splicing, and upstream polyadenylation sites (Willison and Ashworth 1987, Kleene 2005). A variety of mRNA from factors involved in mRNA biogenesis and translational initiation are expressed at 50–1,000 fold higher levels in spermatogenic cells than in somatic cells, e.g. TATA-binding protein or the elongation initiation factor eIF-4E (Kleene *et al.* 2003). Some sperm cell specific transcripts encode truncated proteins, like vasopressin or α-tubulin, that lack domains essential for the functions of the proteins encoded by the same genes in somatic cells (reviewed in Kleene 2001).

2.2.4 Posttranscriptional Regulation

Numerous cellular events, including pre mRNA splicing, mRNA editing, export of the mRNA from the nucleus to the cytoplasm, the stability and translational control of mRNAs, provide opportunities to regulate gene expression at the RNA level (Zielinski *et al.* 2006). Posttranscriptional processing of the mRNA plays an important role in the regulation of spermatogenesis because transcription terminates in the early stages of haploid germ cell differentiation. Furthermore, there is a selective storage and activation of mRNA groups that encode proteins needed during the terminal development of spermatozoa (Hecht 1998, Kleene 2003, Yang *et al.* 2006). RNA interactions with proteins control translation at many levels, including mRNA processing, movement, stability, localization and translation (Gu *et al.* 1998). The chromatoid body has also been suggested to have a functional relationship with the posttranscriptional regulation through the translational apparatus (Kotaja and Sassone-Corsi 2007).

2.2.5 Alternative Transcripts

The real structure of all mRNAs and their isoforms is difficult to determine, since the traditional Affymetrix expression arrays do not identify 11% of the known exons because they are too short for the criteria used to identify exons (Carninci 2006) and the presence of polymorphisms in spermatozoa could impair qRT-PCR assays in some cases (Lalancette *et al.* 2008). Defining the actual structure of mRNA in the presence of multiple, alternatively spliced forms requires either intensive full-length cDNA sequencing or the cloning of products obtained after rapid amplification of cDNA ends (Carninci 2006).

Splicing variations significantly enhance the proteome diversity (Zavolan *et al.* 2003). According to Eddy (2002), alternate transcripts may be beneficial to male germ cells since the addition of new domains can expand the function of the protein. These variations offer modified cell-specific functions and also the replacement of one domain by another can alter the distribution of the protein within the cell.

2.2.6 Non-coding RNA

The role of the unknown RNAs, found in the spermatozoa, remains to be defined. It is presumed that these RNAs have specific roles on spermatogenesis, fertilization or early embryo development (Ostermeier *et al.* 2004, Miller *et al.* 2005). Until recently most of the known non-coding RNAs (ncRNAs) fulfilled generic cell functions, such as the rRNAs and tRNAs involved in mRNA translation, small nuclear RNA (snRNA) involved in splicing and small nucleolar RNA (snoRNA) involved in the modification of rRNA (Mattick and Makunin, 2006). The number of known functional ncRNA genes has risen dramatically in recent years. Mattick and Makunin (2006) suggest that the developmental programming and the phenotypic difference between species and individuals are heavily influenced, if not fundamentally controlled, by the repertoire of regulatory ncRNAs.

Some ncRNAs are derived from the further processing of exons and introns of protein-coding genes as well as from ncRNA genes. Around 97–98% of the transcriptional output of the human genome is ncRNA. This estimate is based upon the estimative that intronic RNA constitutes 95% of primary protein-coding transcripts (pre-mRNA) and that there are a large numbers of ncRNA transcripts that do not contain substantial open reading frames and which may represent at least half of all transcripts (reviewed in Mattick 2003).

The term transcriptional unit (TU) describes a segment of the genome from which transcripts are generated; a single cDNA sequence may define a TU (Okazaki *et al.* 2002). However, a single TU could potentially produce more than one mRNA species through alternative splicing, and alternative transcriptional initiation and termination. The identification of many splicing isoforms indicates that more proteins can be generated from a specific area of genomic DNA than was previously expected. The RIKEN (japanese abreviation of Rikagaku Kenkyūsho, which means Institute of Physical and Chemical Research) project has obtained 60,770 full-length cDNAs from more than 250 normalized and subtracted cDNA libraries from mice. All these cDNAs were clustered into 33,409 representative TUs (Okazaki *et al.* 2002). From them, 17,594 TUs have a coding potential and the remaining 15,815 TUs were defined as non-coding. The extensive but regulated generation of these RNAs might therefore expand the functional diversity of the genome (Suzuki and Hayashizaki 2004). Zavolan *et al.* (2003) estimate that the frequency of mouse genes that have multiple splicing forms is as high as 60%, since many of the variants generated *in vivo* are rare, specific to tissues and developmental stages.

2.2.7 Micro RNA and RNA Interference

The microRNA (miRNA) and RNA interference (RNAi) control mechanisms use small RNA molecules to inhibit gene expression at the level of mRNA degradation, translational repression, or chromatin modification and silencing (Kotaja and Sassone-Corsi 2007). The small interference (si)RNA mediates cleavage of mRNAs whose sequences are fully complementary to their sequence, whereas miRNA is proposed to regulate gene expression by inhibiting protein synthesis or affecting the stability of mRNA through imperfect base-pairing to the 3' untranslated regions of target mRNA (reviewed in Kotaja *et al.* 2006, Kotaja and Sassone-Corsi 2007). Rassoulzadegan *et al.* (2006) showed recently that the exposure of the early embryonic genome from mice to miRNA induced permanent and heritable epigenetic change in gene expression. It has been suggested that the total mouse sperm RNA contains at least 20% of miRNA, which could be released into the oocyte during fertilization (Amanai *et al.* 2006).

A hallmark of RNAi is the production of short double-stranded (ds)RNA molecules, with 21-28 nucleotides in length, by the specialized RNase III protein Dicer (Nicholson and Nicholson 2002, MacRae *et al.* 2007). Members of the ribonuclease III superfamily of (ds)RNA-specific endoribonucleases participate in diverse cellular RNA maturation and degradation pathways (Nicholson and Nicholson 2002). SiRNA and miRNA precursors are processed to mature small RNAs in the cytoplasm of cells (Kotaja and Sassone-Corsi 2007). They differ in the origin; siRNAs arise from long dsRNAs, whereas miRNAs are processed from hairpin precursors (Grivina *et al.* 2006a). Mouse dicer contains a DExH/DEAH helicase motif, a PAZ (Pinwheel-Argonaut-Zwille) domain, a tandem repeat of RNase III catalytic domain sequences and a (ds)RNA-binding motif (Nicholson and Nicholson 2002).

Members of the PAZ/PIWI domain (PPD) and Dicer are central to RNAi (Tahbaz *et al.* 2004). MacRae *et al.* (2007) showed that Dicer binds (ds)RNA 3' ends using the PAZ domain and positions the (ds)RNA substrate along the flat face of the enzyme. A stable binding between PPD proteins and Dicer is dependent on the activity of HSP90 (Tahbaz *et al.* 2004). RNAi is enforced by RNA-induced silencing complex (RISC), a protein-RNA effector nuclease complex. The products cleaved by Dicer are incorporated into PPD-containing RISCs (Hannon 2002). The RISC is composed of various members of the Argonaute subfamily protein (AGO1 to 4) that are characterized by the presence of two homology regions, the PAZ domain and the PIWI domain (Hannon 2002, Kotaja and Sassone-Corsi 2007). RISC mediates cleavage or translational repression of target mRNA through binding to small RNA

and other components of small RNA pathways (Tahbaz *et al.* 2004, Kotaja and Sassone-Corsi 2007).

2.2.8 PIWI-Interacting RNA

The PIWI (P-element induced wimpy testis) protein is a highly basic protein member of a large gene family with an overall homology of ~20% amino acid identity and the conserved PIWI domain (Cox *et al.* 1998) that is essential for its function (Kuramochi-Miyagawa *et al.* 2001, Deng and Lin 2002). Miwi, murine member of PIWI protein, may be involved in spermatocyte and spermatid development since in *Miwi* null mice the spermatogenesis is uniformly arrested at the round spermatid stage (Deng and Lin 2002, Kuramochi-Miyagawa *et al.* 2004).

Lau *et al.* (2006) designated RNAs found binding rat homologue of PIWI protein as PIWI-interacting RNA (piRNA). PIWI may form complexes with RNAs to ensure their stability (Deng and Lin 2002). PiRNAs map to intergenic, RNA-coding, and repetitive sites in the mouse genome that accumulate during spermiogenesis in association with PIWI family protein (Grivina *et al.* 2006a). Their features are clearly distinct from those of miRNAs because they have poor primary sequence conservation, immense complexity and cluster to discrete originating loci (Watanabe *et al.* 2006, Seto *et al.* 2007). The length of piRNA ranges between 19 and 33 nt, most sequences are beginning with U and have a non-uniform distribution in the genome (Lau *et al.* 2006, Girard *et al.* 2006, Watanabe *et al.* 2006). They are evolutionary conserved, predominantly sense sequences and their precursor is a long, single-stranded transcript that is cleaved, preferentially at U residues (Lau *et al.* 2006, Brennecke *et al.* 2007).

The model of Brennecke *et al.* (2007) showed that the piRNA loci encode a diversity of small RNA fragments that have the potential to recognize invading genetic elements. These loci comprise mainly defective transposon sequences, and some have been identified as master regulators of transposon activity. The mouse piRNAs are predominantly produced post-meiotically in early round spermatids in a Miwi-dependent fashion. Grivina *et al.* (2006b) hypothesed that it remains possible that Miwi interaction with Dicer or a Dicer-like enzyme is required for the production of piRNA. Miwi protein was detectable up at least maturation phase of mice elongating spermatids (Deng and Lin 2002). However, piRNAs were only

detected in testes and their absence in the caudal epididymus suggests that they are not loaded into mature sperm (Girard *et al.* 2006, Grivina *et al.* 2006a).

2.2.9 Translational Regulation

All mRNAs in meiotic and haploid mammalian spermatogenic cells are partially translationally repressed, but individual mRNAs exhibit differences in the extent of repression (reviewed in Kleene 2001, Eddy 2002). Novel protein translation was demonstrated in human, mouse, bovine, and rat spermatozoa, indicating that sperm cells are able to translate (Gur and Breitbart 2006). Amino acid incorporation was completely blocked by the mitochondrial translation inhibitors, but not by the cytoplasmic 80S ribosomal inhibitor cycloheximide, indicating that protein translation in sperm involves mitochondrial but not cytoplasmic ribosomes. Moreover, inhibition of protein translation significantly reduced sperm motility, capacitation and *in vitro* fertilization rate (Gur and Breitbart 2006).

In fact, only a few examples of positive and negative controls over translation, like mRNA deadenylation and RNA binding proteins, have been rigorously demonstrated by *in vivo* genetic approaches (Kleene 2003). Some mRNAs were found to have different lengths of poly(A) tails during differentiation (Penttiliä *et al.* 1995). The efficiency of translation of the mRNA encoding transition proteins and protamines depends on the rate of initiation, which is reduced when they are not associated with ribosomes or if ribosomes are spaced more than 100 bases apart on the polysomes. Kleene (1989) compared the quantity and the length of the poly(A) tracts of five haploid-expressed mRNAs in the polysomal and nonpolysomal fractions of round and elongating spermatids in mice. The translationally repressed forms were long and homogenous in size, whereas the polysomal forms were shorter and more heterogenous due to shortening of their poly(A) tracts.

Cis-acting RNA regulatory sequences and a number of nonspecific or sequence-specific RNA binding proteins are also involved in translational control of the transition proteins and protamines. Translational control elements have been identified in both the 5’ and 3’ untranslated regions (UTR) of mRNAs (Dadoune 2003).

2.2.9.1 TB-RBP

Gu *et al.* (1998) identified a testicular and brain RNA-binding protein (TB-RBP) that binds to highly conserved sequence elements present in the 3' untranslated regions of numerous testicular and brain mRNAs. TB-RBP is involved in translational control of testicular mRNAs and cellular transportation of brain mRNAs. *In situ* hybridization, immunohistochemistry and electron microscopy have demonstrated that in mice specific postmeiotic mRNAs (e.g. protamines and clusterin) from male germ cells are transported intercellularly in association with TB-RBP (Morales *et al.* 1998, Chennathukuzhi *et al.* 2003). The mRNAs are translated following the release of the mRNA from the RNA/protein complex, suggesting an involvement of TB-RBP in the translational suppression of linked mRNA in haploid male germ cells (Iguchi *et al.* 2008).

2.2.9.2 Y-Box Proteins

The Y-box proteins can inhibit or stimulate mRNA translation (e.g. clusterin, actin, *Tnp2* and *Prm2*), depending on their concentration (Yang *et al.* 2005b). Furthermore sucrose gradient fractionation demonstrated the presence of the mouse Y-box protein Msy2 in ribonucleoproteins (Yang *et al.* 2006). Binding assays suggested that these proteins bound RNA in a sequence independent form.

Many of the mRNAs bound to Msy2 were meiotic and post-meiotic gamete-specific transcripts that are known to be stored as ribonucleoprotein (RNP). These are critical for germ cell development, like protamine, transition proteins and A kinase anchor protein 4 (AKAP4). Y-box proteins are major constituents of RNP, which contain translationally silent mRNA in gametic cells (Giorgini *et al.* 2001). The absence of Msy2 in mice reduced the postmeiotic germ-cell-specific mRNAs, like protamine, transition proteins or actin, compared to their wild-type littermates (Yang *et al.* 2005b). Yang et al. (2005a) suggested that this contributes to stabilization and/or regulation of the translation of paternal mRNAs.

2.2.9.3 Protamine RNA binding protein

Protamine RNA binding protein (PRBP) is restricted to the testis germ cells (Lee *et al.* 1996). Zhong *et al.* (1999) have suggested that the mice Prbp functions as a translational regulator during mouse spermatogenesis and that it is required for proper translational activation of the mRNAs encoding the protamines. Lee *et al.* (1996) hypothesized that in mice, Prbp firstly

binds to a specific site in the 3' UTR of *Prm-1* RNA and then oligomerizes to form a filamentous ribonucleoprotein particle that is translational inactive. Mice with a targeted disruption of the *Prbp* gene are sterile or severely oligospermic (Zhong *et al.* 1999).

2.3 Chromatoid Body

The chromatoid body (CB) was first described by Benda in 1891 (in Parvinen 2005). Its origin from nucleus or nucleolus was proposed but also a derivation from an intermitochondrial dense material during late pachytene stage of the prophase of first meiotic division was indicated (reviewed in Parvinen 2005). CB material is first apparent in late pachytene spermatocytes prior to the first meiotic division as intermitochondrial fibrous structures that disperse during meiotic division (Kotaja and Sassone-Corsi 2007). Large dense bodies (0.5 µm) appear in the cytoplasm of newly formed secondary spermatocytes. Immediately after the second meiotic division, the CB and the Golgi complex are dispersed in the cytoplasm of the young spermatid as small dark bodies. First, these organelles assume a compact form then it moves rapidly on the nuclear envelope and makes transient contacts with the Golgi complex. At spermiation, the CB is transferred to the residual body and disappears by the detachment (Parvinen 2005).

The finding that poly(A)-containing RNAs accumulate in the CB of spermatids indicates that the process of translational repression, when the spermatid nucleus becomes gradually inactive, is organized in this cytoplasmic structure (Parvinen 2005, Kotaja and Sassone-Corsi 2007). Specifically, the presence of Dicer, Argonaute proteins and miRNAs reveals that the chromatoid body occupies a privileged position in posttranscriptional control of gene expression through the small RNA pathway (Kotaja *et al.* 2006).

2.4 Residual Bodies

Breucker *et al.* (1985) studied in human testis the formation of the residual body of the spermatid and its morphological changes during and after spermiation. The caudal cytoplasmic mass of the late spermatid contains a Golgi complex, mitochondrion, a chromatoid body, flower-like structures, ribosomes, a few large vacuoles, myelin-like membrane profiles and sporadic lipid droplets (Breucker *et al.* 1985). When, by detachment of the caudal cytoplasm from the free spermatozoon, the residual body is formed, the chromatoid body disappears; the mitochondrion is clustered peripherally; the ribosomes

appear as a single complex in contact with a large vacuole containing granular material; in place of the Golgi complex aggregations of vesicles are present. The lipid droplets remain unchanged. The residual bodies or their fragments are either extruded via the seminiferous tubular lumen into the excurrent ducts or they are engulfed by Sertoli cells where in the supranuclear region the successive steps of decomposition can be observed (Breucker *et al.* 1985).

Electron microscopy confirmed that the ultrastructural features of these bodies are compatible with those of apoptotic bodies (Blanco-Rodríguez and Martínez-García 1999). Specialized regions of cytoplasm, the caudal tags of the maturing spermatids, can undergo apoptosis and form apoptotic bodies, the so-called residual bodies, while the cell remains healthy and releases as a free spermatozoon. The main function of residual bodies is a selective removal of some components that are not required by the mature spermatozoon, e.g. the chromatoid body (Blanco-Rodríguez and Martínez-García 1999).

2.5 Ribosomal Proteins and Translation Factors

Translation initiation in eukaryotes depends on many eukaryotic initiation factors (EIFs) that stimulate both recruitment of the initiator tRNA and mRNA to the 40S ribosomal subunit and subsequent scanning of the mRNA for the AUG start codon (Hinnebusch 2006). The mRNAs encoding many constituents of the translational apparatus, including ribosomal proteins, the elongation initiation factors EIF4E and EIF4G, and poly(A) binding protein, are overexpressed in malignant cells (Miyagi *et al.* 1995, Kleene 2005) and, interestingly, are overexpressed in spermatogenic cells, too (Kleene 2003, Miyagi *et al.* 1995).

Various ribosomal proteins are assembled on the RNA core, each protein recognizing its specific binding site. The ribosome performs translation, i.e., it sequentially reads out a genetic message from mRNA chain and concomitantly synthesizes a cognate polypeptide chain (Spirin 2004). Additionally, EIF4E may play a role in mRNA sequestration and stability in cytoplasmic processing bodies. Culjkovic *et al.* (2007) suggested in their recent studies that EIF4E governs the cell cycle. In the testis it may play an important role in spermatogenesis through translational regulation of stage-specific mRNAs during germ cell development (Miyagi *et al.* 1995). The high expression of ribosomal proteins, initiation and elongation factors may follow the fast replication and changes past by the cells during the spermatogenesis (Goossens *et al.* 2007).

2.6 Apoptosis

A high level of germ cell apoptosis occurs during the spermatogonial phase through a density-dependent regulation and during meiosis due to chromosomal damage (Almeida *et al.* 2006). The localization and identification of physiological apoptotic cells by Blanco-Rodríguez *et al.* (2003) and the subsequent observation of the same tubule in the same section provided evidence of a correlation between the cellular death and DNA replication in the spermatogonial generations (Blanco-Rodríguez *et al.* 2003, Heninger *et al.* 2004). In addition, apoptosis might be the mechanism responsible for the formation of residual bodies from the highly specialized region constituting the caudal cytoplasm of maturing spermatids.

In domestic pigs up to 15% of germ cells are lost during spermiogenesis (França *et al.* 2005, Almeida *et al.* 2006). Each Sertoli cell is in contact simultaneously with three or four layers of germ cells that are in all three phases of spermatogenesis (Eddy 2002) and this loss ensures an optimal number of germ cells that can be supported by the Sertoli cells. This mechanism it also plays an important role in removing abnormal sperm (França *et al.* 2005, Cagan 2003). A close interaction between the developing germ cells and Sertoli cells is established, allowing intercellular transport of nutrients and bidirectional exchanges of regulatory molecules (Hecht 1998). It is assumed that the number of germ cells should fit the capability of the Sertoli cells to support them, so this number is adjusted by cell death when proliferation of differentiating spermatogonia begins (Blanco-Rodríguez *et al.* 2003). It can be said that the Sertoli cell “dictates” the magnitude of the sperm production (França *et al.* 2005).

During the release of spermatozoa from the seminiferous epithelium, the residual body of maturing spermatids is collected and removed (Blanco-Rodríguez and Martínez-García 1999, Cagan 2003). Some ultrastructural aspects of residual bodies have been related to a degeneration process, whereas phagocytosis of these bodies by the Sertoli cells is also well established. Blanco-Rodríguez and Martínez-García (1999) realized that the high basophilia shown by residual bodies is similar to that of apoptotic germ cells and that apoptotic factors are somehow segregated to the cytoplasm and this segregation permits the emerging sperm to utilize the apoptotic machinery without dying (Cagan 2003).

2.7 Cytoplasmic Bridges

During spermatogenesis, male germ cells are connected transiently by intercellular bridges, which are transformed into stable intercellular bridges, allowing the midbody and central spindle to break down without abscission. Intercellular bridges are evolutionarily conserved from invertebrates to humans (reviewed in Greenbaum *et al.* 2006 and 2007). They are produced by mitotic and meiotic divisions and create large connections linking the cytoplasm of generations of daughter cells. Cells originating from the same type A1 spermatogonia form a syncytium that connects them by intercellular bridges. More than 50 spermatids remain connected until the end of spermatogenesis when the elongated spermatids are released into the lumen of the seminiferous tubules (Hendriksen 1999). It was demonstrated that mature intercellular bridges grow in diameter as germ cells develop. Type A spermatogonial bridges have a diameter of 1 µm in mice. The bridges expand to 1.5 µm in spermatocytes and 2–3 µm in spermatids (Greenbaum *et al.* 2007).

The male germ cells undergo massive differentiation and rapid polarization. This requires transport and specific localization of vesicles, proteins, mRNAs, and organelles. Chromatoid bodies, mRNA and organelles have been observed to move between haploid spermatids, sharing of essential signals for the synchronous cell divisions or for meiosis (Parvinen 2005, Greenbaum *et al.* 2006). In mice, that were hemizygous for a human growth hormone transgene, only half of the spermatids contained the transgene, but the hormone protein was present in nearly all spermatozoa (Braun *et al.* 1989). Other examples of proteins found to be shared are the t-complex protein 1, protamine and testis-brain RNA-binding protein that may function in the storage and transportation of mRNA to specific intracellular sites where they are translated (Morales *et al.* 1998, Hendriksen 1999, Steger 2001).

After meiosis, cytoplasmic sharing may be necessary for haploid germ cells to remain phenotypically diploid. Those cytoplasmic bridges connecting the clone of spermatids are assumed to facilitate the sharing of haploid gene products and synchronous development of the cells (Greenbaum *et al.* 2006). Post-meiotic gene expression could lead to gametic differences, some of which might lead to preferential transmission of certain alleles over others. Loss of mammalian germ cell bridges prevents syncytium formation, disrupts spermatogenesis and results in sterility (Greenbaum *et al.* 2006).

2.8 Mitochondrion

The flagella of mammalian sperm are modified by three sperm-specific accessory structures, the mitochondrial sheath, the outer dense fibers, and the fibrous sheath, which provide mechanical factors that enhance motility and a scaffold for localization of proteins that produce ATP and regulate motility (Hawthorne *et al.* 2006). Mitochondria generate ATP through oxidative phosphorylation, which takes place within the electron transfer chain and is more productive than anaerobic generation of ATP. In mature mammalian sperm, there are 22 to 75 mitochondria localized in the midpiece (Amaral *et al.* 2007). Mitochondrial membrane potential is a parameter that reflects mitochondrial functionality and provides information about the fertility potential in men (Marchetti *et al.* 2002, Gallon *et al.* 2006).

With ^{32}P label, Premkumar and Bhargava (1972) wrongly concluded that the RNA population of spermatozoa was exclusively of mitochondrial origin. Villegas *et al.* (2000) found six different clones from mitochondrial RNA in the mouse spermatozoa nucleus instead mitochondrion, they believe that these transcripts were results of post-transcription reaction. Mitochondrial transcripts do not change qualitatively or quantitatively in enriched populations of type A and type B spermatogonia obtained from 8-day-old mice but analysis from pachytene spermatocytes, round spermatids, and residual bodies reveals a gradual decrease in the level of mtRNA transcripts, with mRNA levels lower than those detected for equal aliquots from total testes (Alcivar *et al.* 1989).

The mammalian mitochondrial DNA, double-stranded and circular, is transcribed as a single unit from two separate and distinct origins of replication (Alcivar *et al.* 1989). Transcription has been shown to be dependent upon a mitochondrial polymerase and a mitochondrial transcription factor; the RNA is then cleaved into the different mRNAs and tRNAs by RNases at specific recognition sequences (reviewed in Saunders *et al.* 1993). Mouse sperm mitochondria possess several unique proteins or protein isoforms that are not found in somatic mitochondria, these include sperm-specific isoforms of lactate dehydrogenase and hexokinase (Turner 2006).

The proteins encoded by the mitochondrial genome have been identified as components of the respiratory chain: subunits I, II, and III of cytochrome c oxidase (COI, COII, COIII), subunits 6 and 8 of adenosine triphosphatase (ATPases 6 and 8), apocytochrome b (CYTB), and part of Complex I, the seven subunits of the respiratory chain NADH dehydrogenase (ND1, ND2, ND3, ND4, ND4L, ND5, and ND6). Cytochrome c has two apparently separable roles in a cell; it acts in the respiratory chain to regulate energy metabolism within the

mitochondria (Saunders *et al.* 1993) and has been shown to be an important regulator of the "intrinsic" apoptotic pathway (Blanco-Rodríguez and Martínez-García 1999).

2.9 Polyspermy

Extensive attempts have been made to reduce the penetration of the porcine oocytes by more than a single spermatozoon (polyspermy) in *in vitro* fertilization (IVF). The high incidence of polyspermy remains a major impediment of porcine IVF systems (Almiñana *et al.* 2005), since it results in anomaly during the development of the zygote (Funahashi 2003). *In vitro* matured oocytes are exposed to an excessive number of spermatozoa for a longer period of time. Theoretically, the problem of polyspermy could be overcome by reducing the number of spermatozoa within the IVF drops. However, in most cases, such adjustments are associated with low oocyte penetration rates (Abeydeera 2002).

Several natural mechanisms work to avoid the polyspermy, e.g. oviduct environment and the reaction from the zona pellucida further sperm penetration. Just after fusion of spermatozoa with the plasma membrane of the oocyte, exocytosis of the cortical granule contents to the perivitelline space leads to zona modifications and inactivation of sperm receptors (Funahashi 2003). At least in the pig, development of the full machinery to prevent polyspermy, by the oocyte, seems to be complete only after exposure to the oviduct (Abeydeera 2002). However, the reasons for the polyspermy are not totally elucidated.

2.10 Classical Techniques for Semen Evaluation

Semen evaluation is important to obtain success in reproduction. Prediction of the fertilizing ability of sperm is of great economic importance for breeding herds, since it leads to the selection of males with semen with e.g. better freeze/thaw rates needed for artificial insemination (Gadea 2005, Turba *et al.* 2007). Factors, like breed, age and housing also affect the sperm production (Kunavongkrit *et al.* 2005).

Male fertility depends upon a heterogeneous population of spermatozoa interacting at various levels of the female genital tract and oocyte. Semen samples can be subjected to functional *in vitro* tests that are able to discern the ability of spermatozoa to undergo specific steps of fertilization, but these tests usually assess only one or a few of these events at the same time (Rodrígues-Martínez 2003, Gadea 2005). At the moment there is no test reliable

enough to predict the potential fertility of given semen sample or sire as indicated by poor correlations obtained between *in vitro* results and fertility (Rodrígues-Martínez 2003).

The classical *in vitro* evaluation of semen (spermiogram) is based on the examination of the cell structure and is poor in predicting subsequent sperm fertility because only the samples with markedly low quality can be detected and excluded (Gadea 2005). The spermiogram is based on the application of a battery of tests, complementary to the clinical examination, that are simple to perform and with relatively low cost. It includes an immediate assessment of volume, aspect, concentration and motility, as well as a later determination of sperm morphology and the presence of foreign cells (Rodrígues-Martínez 2003). Sperm morphology provides information about the status of spermatogenesis and the evaluation of motility is a good indicator of the intactness of the membranes and functionality; but it is a subjective measure that depends on the individual observer. The most valid assessment of fertility is the pregnancy rate. However, it is also imprecise concerning the high variability associated with the female and with the conditions of insemination (Gadea 2005).

The generation of new knowledge in molecular and genomic technology will help us to evaluate the potential fertility of male ejaculates more accurately (Lalacette *et al.* 2007). Molecular aspects of the semen are an important issue to be implemented. At the same time, it is necessary to understand where these molecules exert their function and which morpho-functional aspects they are able to influence (Turba *et al.* 2007).

2.11 Molecular Biology

2.11.1 Overview

Several methods, like reverse transcriptase-polymerase chain reaction (RT-PCR), *in situ* hybridization and arrays, can be used to investigate spermatozoa RNA (Ostermeier *et al.* 2002, Wang *et al.* 2004, Zhao *et al.* 2006, Carreau *et al.* 2007, Gilbert *et al.* 2007, Jones *et al.* 2007). Some transcripts were already described as related to male fertility (Table 1).

Table 1: Different regulated transcripts according to the literature.

Author	Identity	Specie	Regulation	Case
Lambard et al. 2004	PRM-1, eNOS, nNOS	Human	Up	Low motile spermatozoa
Wang et al. 2004	TPX-1, LDHC	Human	Down	Low motile spermatozoa
Jedrzejczak et al. 2006	CCR5, P450arom	Human	Down	Infertile men
Amaral et al. 2007	POLG, TFAM, COI	Human	Down	Oligoasthenoteratozoospermy
Jedrzejczak et al. 2007	HILS1, TPN-1 and 2	Human	Down	Low motile spermatozoa
Platts et al. 2007	Ubiquitin-proteasomal pathway e.g. Protein-group specific E2, E1 and E3A	Human	Down	Teratozoospermy
Platts et al. 2007	Apoptotic pathway e.g. INSR, CASP1 and 8, BIAP1	Human	Up	Teratozoospermy
Gau et al. 2008	HSP 105, HSPa4l, THAP4	Swine	Up	Heat treated germ cells (39°C)
Lalancette et al. 2008	12S rRNA, 18S rRNA and Ribosomal proteins	Bovine	Up	Low fertile bulls

PRM-1 – protamine 1; eNOS – endothelial nitric oxide synthase; nNOS – neuronal nitric oxide synthase ;TPX-1 – testis specific protein 1; LDHC – lactate dehydrogenase C, variant 1; CCR5 – C-C chemokine receptor 5; P450arom – cytochrome P450 aromatase; POLG – polymerase gamma; TFAM – mitochondrial transcription factor A; COI – cytocrome c oxidase; HILS1 – spermatid-specific linker histone H1; TPN – transition protein; INSR – insulin receptor; CASP – caspace; HSP – heat shock protein; THAP 4 – THAP domain containing 4.

The transcriptome covers all RNAs synthesized in an organism or tissue, including protein-coding, non-protein-coding, alternatively spliced, polyadenylated or initiated, sense, antisense and RNA-edited transcripts. The mammalian transcriptome has been based upon sequences from large-scale expressed-sequence-tag (EST) sequencing projects (Okazaki *et al.* 2002). Conservative estimates suggest that there are approximately 20,000 protein-coding genes in the mammalian genome. According to Ensemble annotations 23,000 protein-coding genes. Tiling array (microarray designed to cover at regular interval whole chromosomes or genome regardless the genome annotation) and tag sequencing data agree that much more of the genome is transcribed than was previously thought (Carninci 2006). Comparative transcriptomics is useful for the identification of alternative splices, alternative promoters, ncRNA, metabolic pathways, imprinting genes and small proteins.

The use of microarrays for the study of sperm mRNA began 2002 with studies by Ostermeier and collaborators. They developed a bundle of microarrays containing 27,016 unique ESTs

to investigate cDNA from human ejaculate spermatozoa. This approach has been applied for studies of fertility and semen quality in men, bovine and more recently in boar (Wang *et al.* 2004, Zhao *et al.* 2006, Gau *et al.* 2008, Lalancette *et al.* 2008).

A study using microarray-based data coupled with real-time PCR detected, in men, a significant decrease in *TPX-1* (testis-specific protein 1) and *LDHC* (lactate dehydrogenase C, transcript variant 1) mRNAs from spermatozoa with low motility (Wang *et al.,* 2004). Jedrzejczak *et al.* (2007) evaluated spermatid-specific linker histone H1-like protein (HILS1) and transition protein (TNP1 and TNP2) transcript levels by reverse-transcriptase quantitative PCR (RT-qPCR) in spermatozoa isolated from normozoospermic and asthenozoospermic (spermatozoa with normal and low motility, respectively) men. They found significantly lower levels of these transcripts in asthenozoospermic compared to normozoospermic spermatozoa. The same group found, in human spermatozoa selected by the swim-up technique, that the decrease in CCR5 (C-C chemokine receptor 5) and cytochrome P450 aromatase (P450arom) transcript levels may be associated with infertility (Jedrzejczak *et al.* 2006).

Gau *et al.* (2008) used a porcine testis cDNA microarray to identify early responding genes induced by a short and mild heat stimulus and tried to elucidate their possible relationship with semen quality in mature boars. Two of the upregulated genes were for heat shock proteins (HSP). Proteins, such as HSP, have been detected in sperm cells and, particularly the 70 kDa HSP (HSP70) appears to be correlated with semen quality, playing crucial roles in fertilization and early embryonic development (Turba *et al.* 2007). Jones *et al.* (2007) discovered, recently, a new morphological region in bovine, ovine and boar spermatozoa, which is assembled during epididymal maturation and has an unusual concentration of HSP70 and phosphorylated proteins. The partial blockage of fertilization of both zona-intact and zona-free oocytes by anti-HSP70 antibody suggested that this molecule plays an important role in porcine gamete interaction during the sperm–egg fusion (Spinaci *et al.* 2005).

Also using microarray technology, Platts *et al.* (2007) showed that human spermatozoa with morphological defects have transcriptional perturbation on remodeling pathways, such as the ubiquitin-proteasome pathway. Another applied study, using microarray, showed substantial differences in the transcript representation of spermatozoa from bulls with different levels of fertility (Lalancette *et al.* 2008). For example, bulls with high non-return rates (>71%) exhibited a greater portion of transcripts associated with translation, glyicosylation and protein degradation, when compared with bulls with low non-return rates (<65%).

2.11.2 Microarray techniques

Microarray technology has the potential to measure simultaneously tens of thousands of genes expressed in a biological sample. At present, serial analysis of gene expression (SAGE), oligo microarrays, cDNA microarrays and Affymetrix GeneChips are the most widely used techniques for determining gene expression levels and gene expression ratios (van Ruissen *et al.* 2005). Comparative genomics, transcriptomics and proteomics help us to understand the mechanism of diseases, in order to identify and develop new effective diagnostics and therapeutics (Suzuki and Hayashizaki 2004).

Techniques, like cDNA libraries and microarrays, can be helpful as complementary diagnosis and for the prognosis of the sperm fertilization capacity. The most common application of cDNA libraries is the identification of genes of interest (Ying 2004). The spermatozoa have mRNAs that differ from somatic cells and are found only in male germ cells (Kleene 2001). Coupled with microarray technologies tissue-specific cDNA libraries can be used to monitor a large number of genes and provide a powerful tool for assessing differential mRNA expression levels (Ying 2004).

Both techniques are complementary and the application of a cDNA library associated with the microarray technology in sperm expands the knowledge of the normal values of fertile male (Ostermeier *et al.* 2002). The Porcine GeneChip from Affymetrix contains 23,937 probe sets representing 20,201 transcripts based on known EST sequences. However, the advantage of a microarray based on a spermatozoa cDNA library over the commercial chip is that the cDNA library contains transcripts specific for sperm cells.

2.12 Damaged DNA and paternal contribution

The nucleus of a normal spermatozoon has a highly condensed chromatin built by association of double stranded DNA with protamines but contains sites at which the chromosomal DNA is sensitive to nuclease activity (Giordano *et al.* 2000). The integrity of the chromatin is directly related to the fertility (Gadea 2005) but the origin of sperm DNA breaks is not defined (Baldi *et al.* 2000). Fatehi *et al.* (2006) induced with gamma or x-rays irradiation DNA damag in bovine sperm. This damag did not impair sperm functioning and fertilization but affect embryo development after the first cleavages. The study also shows that repair of paternal DNA damage alone or by the fertilized oocyte is relatively limited and

that the first cleavages of the embryo are not dependent on the embryonic genome but on a large mRNA pool, which is already present in the cytosol of the unfertilized oocyte.

In the pig, the components of the sperm tail and mid-piece are not essential for fertilization and embryo development in the case of intracytoplasmic sperm injection (Lee and Yang, 2004). Amanai *et al.* (2006) suggest that the sperm miRNAs play a limited role during fertilization or early preimplantation development, since they do not alter significantly the maternal population of miRNA. However, RNAs of paternal origin can play a role in modulating gene expression in the embryo, like epigenetic inheritance and RNA interference (Rassoulzadegan *et al.* 2006). It is known that mature spermatozoa of all species spontaneously take up foreign DNA molecules and can deliver them to oocytes at fertilization (Giordano *et al.* 2000, Smith and Spadafora 2005, Pittoggi *et al.* 2006).

Giordano *et al.* (2000) showed evidence for the presence of a reverse transcriptase (RT) enzyme, localized on the sperm nuclear scaffold, that is activated when spermatozoa are incubated with foreign RNA. The reverse transcribed sequences are not integrated in the host genome but are probably replicated independent of the host genome (Smith and Spadafora 2005). The RT-mediated process is not only triggered when spermatozoa are exposed to exogenous RNA molecules, but is also activated when they interact with DNA molecules (Pittoggi *et al.* 2006). Chang *et al.* (2002) demonstrated that exogenous DNA is specifically bound to the sperm cell surface *via* the linker protein (mAb C) through ionic interaction. Exogenous DNA templates incubated with mature sperm cells spontaneously undergo a stepwise process of transcription, RNA splicing, and retrotransposition (Pittoggi *et al.* 2006).

Sperm-mediated gene transfer (SMGT) may be able to provide efficient, rapid and low-cost protocols for animal transgenesis (Smith and Spadafora 2005). Transgenic animals such as pigs and mice can be generated with a germ-line transmission to the F1 generation of 37.5% in pigs (Chang *et al.* 2002). However, SMGT has not yet become established as a reliable form of genetic manipulation because stable transgenic integration has rarely been detected with this protocol (Smith and Spadafora 2005).

3. Material and Methods

3.1 Material

3.2 Equipment

2100-Bioanalyzer	Agilent Technologies, St. Clara, USA
3100-Avant Genetic Analyzer	Applied Biosystems, Darmstadt, Germany
Analyse scales MJ 3000	Chyo, Chyo, Japan
Illumninator Eagleeye II	Stratagene, Heidelberg, Germany
Magnetic stir	Ika Labortechnik, Staufen, Germany
Multi-channel pipettes	Eppendorf, Hamburg, Germany
Microwave	Siemens, Munich, Germany
NanoDrop 1000A Spectrophotometer	Ambion, Austin, USA
Pipettes	Gilson, Bad Camberg, Germany
Robot liquid handling system Multiprobe II Plus ex	Perkin Elmer, Boston, USA
Shaker KS 250 basic	Ika Labortechnik, Staufen, Germany
Thermomixer 5436	Eppendorf, Hamburg, Germany
Vortex MS1	Ika Labortechnik, Staufen, Germany

3.2.1 Equipment for Gel Electrophoresis

Power PAC 300	BIORAD, Oberkochen, Germany
Power Supply PPS200-1D	MWG-BIOTECH, Heidelberg, Germany

Chamber Owl Scientific	Owl Scientific, Woburm, Germany

3.2.2 PCR Equipment

Thermocycler UNOII	Biometra, Göttingen, Germany
Thermocycler T1	Biometra, Göttingen, Germany
Mastercycler Gradient	Eppendorf, Hamburg, Germany
96-well low profile PCR plate	Abgene, Hamburg, Germany
384-well storage plate	Abgene, Hamburg, Germany
Adhesive PCR Foil Seals	Abgene, Hamburg, Germany
PCR Tube 0.2 ml	Eppendorf, Hamburg, Germany
PCR Micro Amp Full Plate Cover	Perkin Elmer, Connecticut, USA

3.2.3 Centrifuges

Centrifuge 5417 R	Eppendorf Hamburg, Germany
Centrifuge GS-15R	Beckman, Munich, Germany
Centrifuge EBA8S	Hettich, Tuttlingen, Germany

3.2.4 Chemicals

2-Mercaptoethanol	Sigma, Deisenhofen, Germany
Agarose	Invitrogen, Karlsruhe, Germany
Ampicillin	Boehringer, Mannheim, Germany
Boric acid	Merck, Darmstadt, Germany
Bromophenol blue	Sigma, Deisenhofen, Germany

Chloroform	Merck, Darmstadt, Germany
Disodium hydrogen phosphate	Merck, Darmstadt, Germany
Ethanol	Roth, Karlsruhe, Germany
Ethidium bromide (EtBr)	Merck, Darmstadt, Germany
Ethylene diaminetetraacetate (EDTA)	Merck, Darmstadt, Germany
Guanidinium Thiocyanate	Roth, Karlsruhe, Germany
Isoamyl alcohol	Merck, Darmstadt, Germany
Isopropanol (2-propanol)	Roth, Karlsruhe, Germany
Magnesium chloride	Merck, Darmstadt, Germany
Monopotassium phosphate	Merck, Darmstadt, Germany
N-Louroylsarcosine	Sigma, Deisenhofen, Germany
Phenol	Roth, Karlsruhe, Germany
Potassium chloride	Merck, Darmstadt, Germany
Sodium acetate (NaOAc)	Merck, Darmstadt, Germany
Sodium chloride	Roth, Karlsruhe, Germany
Sodium citrate	Roth, Karlsruhe, Germany
Sodium dodecyl sulfate (SDS)	Roth, Karlsruhe, Germany
Sodium hydroxide	Roth, Karlsruhe, Germany
RNAlater	Ambion, Austin, USA
Tris-(hydroxymethyl)-aminomethane (Tris)	Roth, Karlsruhe, Germany
Triton X-100	Sigma, Deisenhofen, Germany
Trizol Reagent	Invitrogen, Karlsruhe, Germany

Trypton	Roth, Karlsruhe, Germany

3.2.5 Solutions and Reagents

3.2.5.1 Solutions and Media

Denaturing Solution

0.1 M	2-Mercaptoethanol
0.5 % (w/v)	N-Lauroylsarcosinate Sodium Salt
4 M	Guanidinium Thiocyanate
25 mM	Sodium Citrate (pH 7.0)

Hypotonic Solution

10 mM	Tris HCl (pH 8.4)
50 mM	KCl
2.5 mM	$MgCl_2$
4 mM	DTT
0.05 %	SDS

Luria-Bertani broth (LB medium)

1 %	Trypton
0.5 %	Yeast extract
1 %	NaCl

The pH was adjusted at 7-7.5 with 0.1% 2 M NaOH.

3.2.5.2 Buffers

2x spotting buffer

40 mM	Tris-HCl (pH 8.0)
2 M	NaCl
2 mM	EDTA
0.01 %	Bromphenol blue

10x TBE electrophoresis buffer

108 g	Tris-base
55 g	Boric acid
40 ml	0.5 M EDTA (pH 8.0)
	add water to 1000 ml

25% SDS

1 kg	SDS
	add water to 4000 ml

Running buffer for agarose gels

0.5 µg/ml	EtBr on TBE (10x)

TE buffer

10 mM	Tris-HCl (pH 8.0)

	1 mM	EDTA

10x Phosphate buffered saline (PBS)

Solution A	250 mM	Na_2HPO_4
	1.25 mM	NaCl
Solution B	250 mM	KH_2PO_4
	1.25 mM	KCl

The solution B was added to the solution A until it reaches the pH 7.5

3.2.6 Water for PCR and RNA

For PCR and RNA reactions ultra-pure water after ultrafiltration (Milli-RO Plus 60, Millipore, Schwalbach, Germany) was filtered through a Millex GP filter with pore diameter of 0.22 µm.

3.2.7 Marker

pUC Mix from 19 (MspI) and	
57 (DraI+HindIII)	MBI Fermentas, St.Leon-Rot, Germany
Lambda DNA (*Eco*RI+*Hind*III)	MBI Fermentas, St.Leon-Rot, Germany
100 bp Ladder Plus	MBI Fermentas, St.Leon-Rot, Germany
1 kp Ladder Plus	MBI Fermentas, St.Leon-Rot, Germany

3.2.8 Reagents and Enzymes

10x Reaction Buffer for cDNA amplification	Amersham, München, Germany
5x First Strand Buffer	Invitrogen, Karlsruhe, Germany
5x Second Strand Buffer	Invitrogen, Karlsruhe, Germany

E. coli DNA Ligase (10 U/µl)	Invitrogen, Karlsruhe, Germany
E. coli DNA polymerase I (10 U/µl)	Invitrogen, Karlsruhe, Germany
DNase I, RNase free	Roche, Mannheim, Germany
dNTP Mix 100 mM	Invitrogen, Karlsruhe, Germany
DTT 100 mM	Invitrogen, Karlsruhe, Germany
RNase, DNase free	Roche, Indianapolis, USA
RNase Inhibitor (40 U/µl)	Roche, Indianapolis, USA
RNase H, *E. coli* (2 U/µl)	Invitrogen, Karlsruhe, Germany
FirePol Taq polymerase	ATG-Biosynthetics, Freiburg, Germany

3.2.9 Other Materials

Cell counting chamber Neubauer improved	Marienfeld, Lauda-Königshofen, Germany
Dialysis Membrane Millipore 0.25 µm	Millipore, Schwalbach, Germany
Filter Millex GP pore diameter 0.22 µm	Millipore, Schwalbach, Germany
Needle 30 G	BD, Heidelberg, Germany
Safe-lock tubes (2, 1.5, 0.5, 0.2 ml)	Eppendorf, Hamburg, Germany
Scalpel blade 18	Aesculap, Tuttlingen, Germany
Syringes 1-20 ml	Codan Medical ApS, Lensahn, Germany
Multiguard Barrier tips 1-1,000 µl	BioScience, Heidelberg, Germany
Cell culture test plate 96-well (ELISA)	TPP, Basel, Switzerland

3.2.10 Oligonucleotides

cDNA primer

3'-GATAT$_{20}$VN (V=A, C or G)-5'

NLT7

5’-GCG TAA TAC GAC TCA CTA TAG G-3’

LT3

5’-CTC ACT AAA GGG AAC AAA AGC TG-3’

3.3 Methods

3.3.1 Agarose Gel Electrophoresis

For separation of RNA or (ds)cDNA 1 % TBE agarose gels were routinely used and electrophoresis was carried out at an electric field strength of 4,5 V/cm for 45-60 min in a submarine slab gel apparatus. Prior use for analysis of RNA, the electrophoresis apparatus, trays and combs were washed with 10 % SDS to avoid nuclease contaminations e.g. RNases. The amplified cDNA fragments of the normalized library were run on a 0.8 % agarose gels. The PCR products were separated at 3 V/cm for 10 min and additional 20 min at 5 V/cm. For visualization of the bands, gels were photographed under UV light (312 nm) using the INTAS UV system (INTAS Science Imaging Instruments; Germany).

3.3.2 Sperm Recovery

The ejaculates of five boars were recovered at the Institute of Molecular Animal Breeding and Biotechnology – Moorversuchsgut (Oberschleißheim). The spermatozoa were investigated for motility and morphology. Only ejaculates from normospermic boars were used for further experiments.

The ejaculate, without the urethral bulb secretion, was transferred to 50 ml tubes and centrifuged for 10 min at 1,200 x *g* at RT. The supernatant was discarded and the pellet washed three times with 1x PBS. The concentration of sperms was determined using a Neubauer improved counting cell chamber, under a microscope. Sperm cells ($1x10^8$) were pellet and added 5 volumes from RNAlater®. The samples were transferred to 2.0 ml tubes and incubated in RNAlater overnight at 4°C to allow thorough penetration, then frozen in liquid nitrogen and transfer to -80°C until use.

3.3.3 Handling RNA

Before starting work with RNA, the lab bench, equipment and pipettes were cleaned with water. Also material free of nucleases was preferred. For every work with RNA gloves were weared and frequently changed to protect the resulting RNA from nucleases (present on human skin).

3.3.4 RNA Extraction

3.3.4.1 TRIzol

Total RNA of spermatozoa was extracted with TRIzol® RNA isolation reagent (Invitrogen, Carlsbad, USA), according to the manufacturer's protocol with minor modifications.

Homogenization

Samples were thawed on ice and then centrifuged for 5 min at 5,000 x *g* at 4°C. The supernatant was discarded and the pellet resuspended with 1 ml hypotonic solution with 0.5 % of Triton-X. Samples were incubated for 10 min on ice for lysis of somatic cells. After a centrifugation step (5 min at 5,000 x *g* at 4°C) the hypotonic/triton-X solution was discarded and the pellet washed with PBS. Sample was resuspended in TRIzol (1 ml for 10^8 cells in suspension). Subsequently, the solution was passed three times through a 30 G needle and vortexed vigorously for 30 sec. After a 5 min incubation step at RT the homogenized samples were centrifuged at 12,000 x *g* for 10 min at 4°C. This resulting pellet consists of insoluble material, such as membranes, polysaccharides and high molecular weight DNA. The supernatant, containing the RNA, was transferred to a fresh tube refilled on 1 ml TRIzol.

Afterwards 200 µl of chloroform was added. The samples were homogenized by vortexing for 30 sec and incubated for 5 min at RT. This was follwed by a centrifugation at 12,000 x *g* for 15 min at 4°C. During the centrifugation, the mixture separates into an organic phase, an interphase and an upper aqueous phase that contains the RNA.

3.3.4.2 Guanidium Thiocyanate/Phenol–Chloroform Extraction

The extraction with acid guanidinium thiocyanate/phenol-chloroform (AGPC) was carried out as described by Chomczynski & Sacchi (1987) with few modifications. The spermatozoa

(10^8) were disrupted in 1 ml of denaturing solution by vortexing for 30 min. To shear the DNA, the lysate was passed several times through a 30 G needle.

Sequentially, 100 µl of 2 M sodium acetate (pH 4), 1 ml of phenol and 200 µl of chloroform–isoamyl alcohol mixture (49:1) were added to the homogenates. The samples were mixed by vortexing and incubated for 5 min at RT, followed by a centrifugation at 12,000 x *g* for 15 min at 4°C. The upper aqueous phase was transferred to a new tube and the RNA precipitated.

RNA Precipitation

The procedure for the RNA precipitation was identical to the TRIzol protocol. The aqueous phase was transferred to a fresh tube and mixed with 1 volume isopropanol (100%), 0.1 volumes 3M NaOAc and 1 µl glycogen (15 µg/µl). The samples were incubated for 30 min at RT and then centrifuged at 20,000 x *g* for 30 min at 4°C.

RNA Wash and Dissolving

The supernatant was removed and the pellets washed with 75 % EtOH. Samples were centrifuged at 12,000 x *g* for 5 min at 4°C. The EtOH was discarded and the pellets dried at RT. Samples were dissolved in 20 µl water and a heating step for 15 min at 58°C was performed.

3.3.4.3 Heated TRIzol

For the heated TRIzol protocol, prior extraction, the sperm cells were heated in TRIzol at 65°C for 30 min (Gilbert *et al.* 2007).

Quantification and Assessment of Purity

The RNA concentrations were determined by measuring absorbance at 260 nm (NanoDrop 1000A Spectrophotometer). The RNA samples (1 µl) were separated on a micro-electrophoresis gel for the assessment of quality and quantity. As size standards Lambda EcoRI+HindIII and pUC Mix (19/MspI and 57/DraI+HindIII), were used.

3.3.5 Generation of a Normalized cDNA Library

The preparation of normalized cDNA libraries was done by the company Vertis Biotechnologie AG (Freising, Germany), from 1.8 µg pooled total RNA. First strand cDNA synthesis was carried out with an oligo(dT) primer. The obtained cDNA population was normalized, in order to adjust the levels of cDNA species without the loss of diversity, so meaningful biological comparisons can be made. Normalization was achieved by one cycle of denaturation and reassociation of the no normalized cDNA (N0), resulting in normalized cDNA (N1). Reassociated (ds)cDNA was separated from the remaining (ss)cDNA (normalized cDNA) by passing the mixture over a hydroxylapatite column. After hydroxylapatite chromatography, the N1 (ss)cDNAs were amplified with 10 PCR cycles (Figure 3).

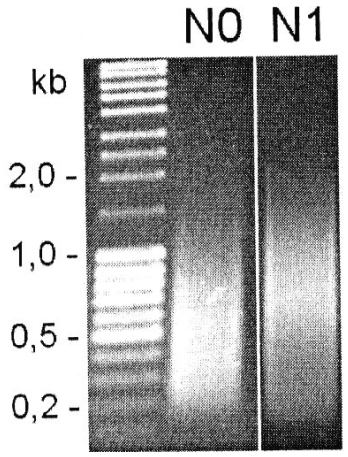

Figure 3: N0 cDNA and N1 normalized cDNA

For cloning the N1-cDNA was subjected to a limited exonuclease treatment to generate 5' overhangs at both ends of the cDNA. The clones were then ligated into the *Eco* RI and *Bam* HI sites of the plasmid vector pBS II sk+. The ligation reactions were electroporated into Transfor Max™ EC100™-T1^{R} electro-competent cells. After transformation, glycerol was added to a final concentration of 15 % (v/v) and for each sublibrary (8 x 550 µl aliquots) the cells were frozen at -80°C. After a freeze-thaw cycle, the titer of the sublibraries was determined to be about 1,000 colony-formation unit (cfu) per microlitre bacterial suspension resulting in a total number of about 4,400,000 recombinant clones.

3.3.5.1 Picking of bacterial clones

One tube containing bacterial suspension of the normalized library was put on dry ice. With a sterile scalpel frozen bacterial suspension was scraped off and transferred to a 1.5 ml reaction tube. One microliter of the suspension was diluted with 200 µl LB medium and distributed on an LB agar plate (12 x 12 cm) with ampicillin (100 µg/ml). The plates were incubated over night at 37°C. 4,224 bacterial colonies were randomly picked using a picking wheel (wheel with eight pins, manufactured in the workshop of the Gene Center) and transferred to 96-well cell culture plates with 100 µl LB/ampicillin (100 µg/ml) medium per well. Each well contained one clone and each clone corresponded to one cDNA fragment.

The 96-well plates were incubated over night (maximal 16 hours) at 37°C on a shaker table (400 rpm). The plates were numbered and each clone was named (addressed) with the name of the library, the plate number and the coordinate on this plate (e.g. PSN1-01-A01).

3.3.5.2 Dilution and lysis of bacteria

96-well PCR plates were filled with 100 µl Tris/EDTA buffer (TE-pH 8.0) per well and 5 µl bacterial suspension was added with a 12-channel-pipette. The plate was covered with a clean silicone mat and incubated in a thermocycler for 15 min at 96°C. This step served to lyse bacteria and to release the plasmid DNA. Afterwards the plates were sealed with an aluminum foil, frozen on dry ice and stored at -20°C.

3.3.5.3 Amplification

The cDNA fragments were amplified, sequenced and analyzed with the BLAST-algorithm (National Centers for Biotechnology Information) for comparison with the sequences of public databases. For the PCR amplification 96-well PCR plates were filled with 18 µl of master mix 1 per well. Subsequently, 2 µl of the diluted cells (bacteria/TE after lysis) was added.

Preparation of the Reaction Solutions

Master Mix 1

NLT7 15 pml/µl	0.5 µl

LT3 15 pmol/µl	0.5 µl
dNTP's 10 mM	0.5 µl
10X Reaction Buffer	1.875 µl
Taq Polymerase (PeqLab) 5 U/µl	0.125µl
ddH_2O	14.5 µl
Final volume	18 µl

Master Mix 2

10 x buffer	0.625 µl
ddH_2O	4.25 µl
Taq polymerase	0.125 µl
Final volume	5 µl

Prior start of the PCR the thermocycler lid was preheated to 99°C. The PCR program included the following steps:

Step 1:	Denaturation	96°C	pause
Step 2:	Denaturation	96°C	120 sec
Step 3:	Denaturation	94°C	25 sec
Step 4:	Annealing	64°C	25 sec
Step 5:	Elongation	72°C	90 sec (24 x from step 3)
Step 6:	Cooling	8°C	pause

Afterwards 5 µl from the Master Mix 2 was applied and a second PCR round was run. The second PCR program covered the following steps:

Step 1:	Denaturation	94°C	pause
Step 2:	Denaturation	94°C	25 sec
Step 3:	Annealing	62°C	25 sec
Step 4:	Elongation	72°C	90 sec (14 x from step 2)
Step 5:	Cooling	20°C	pause

3.3.5.4 Sequencing

After PCR amplification steps, 3 µl from the selected PCR products were transferred onto a 47 mm diameter filter with pores of 0.25 µm for dialysis of remaining primers, buffer salts, and dNTPs, using a pipetting robot (Multiprobe II Automated Liquid Handling System – Figure 4; Perkin Elmer, Rodgau, Germany). Dialysis was done for 50-60 min against 40 ml 25 % TE (10 ml TE and 30 ml water) in a Petri dish at RT. A small magnetic stirer (at approx. 200 rpm) was used to accelerate dialysis. Afterwards, the samples (approx. 2 µl) were transferred to a new 96-well PCR plate containing 1 µl of T7 primer (3.5 pmol) and 4 µl of water using the pipetting robot. DNA sequencing was performed by the company GATC Biotech (Konstanz, Germany).

Figure 4: Multiprobe II Automated Liquid Handling System.

3.3.5.5 Sequence Analysis

The sequences obtained from GATC Biotech were checked and trimmed (removal of vector and adaptor sequences) using Chromas (version 2.3; Technelysium, Australia). Batch FASTA files were generated and compared with public sequence databases using the basic local alignment search tool ('discontinuous Mega BLAST') at the NCBI (www.ncbi.nlm.nih.gov/blast/blast.cgi). cDNAs without results in the 'nr' (non-redundant) database were in addition compared with the 'est' (expressed sequence tag) database. The online DAVID software (http://david.abcc.ncifcrf.gov/) was utilized for functional gene classification.

3.3.6 cDNA array Hybridization

3.3.6.1 Production of cDNA arrays

Fifteen microlitres of PCR reactions of the cDNA fragments that were successfully sequenced were transferred to 384-well microtitre plates (Nunc) containing 15 µl 2-fold spotting buffer (40 mM Tris-HCl pH 8, 2 M NaCl, 2 mM EDTA, bromophenol blue). PCR products were spotted onto nylon membranes (Nytran Supercharge; Schleicher & Schuell, Dassel, Germany) on an area of 20 x 50 mm using an Omnigrid Accent microarrayer (Figure 5 – GeneMachines, San Carlos, CA, USA) and solid pins (Figure 5-a; SSP015, diameter 0.015 inches; Telechem International, Sunnyvale, CA, USA). Spotting was done six times for each PCR product on the same position for sufficient and equal application (Figure 5-b). The spotted cDNAs were denatured on the arrays by incubation on filter paper (Schleicher & Schuell) soaked with 0.5 M NaOH for 20 min at RT. DNA was fixed by baking at 80°C for 30 min and UV cross-linking (120 mJ/cm2) (XL-1,500 UV Crosslinker; Spectronics Corp., New York, USA).

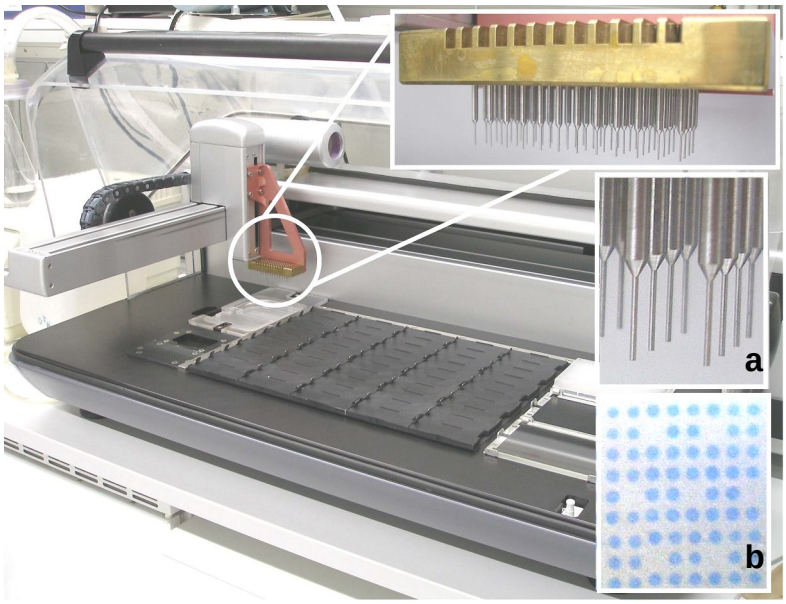

Figure 5: Omnigrid Accent microarrayer. **a**- Solid pins and **b**- spotted nylon membrane

3.3.6.2 Hybridization

For radioactive labeling, cDNA was heat-denatured for 10 min at 96°C and then chilled on ice. High Prime reaction mixture (Roche), dNTP mixture (dCTP final concentration 10 pmol/µl, dATP, dGTP, and dTTP final concentration each 100 pmol/µl) and 90 µCi [α-^{33}P]dCTP were added to a final volume of 20 µl. Reactions were incubated for 1 h at 37°C and subsequently purified with ProbeQuant G50 (Amersham Biosciences, Germany) to remove unincorporated nucleotides and to estimate labeling efficiency.

Hybridization was done as follows: pre-hybridization was done for up to six arrays together in one 15 cm glass hybridization bottle: 3 x 10 min with 10 ml 1 x PBS/10 % SDS at 65°C; 2 x 10 min with 10 ml 0.1 x PBS/1 % SDS at 85°C; 3 x 10 min with 10 ml 1 x PBS/10 % SDS at 65°C. Hybridization probes were denatured for 15 min at 96°C immediately before the hybridization solution was added. Hybridization was done in plastic vials (Poly-Q vials, 18 ml; Beckman Coulter, Munich, Germany) in 2 ml 1 x PBS, pH 7.5/10% SDS for 45 h at 65°C. After hybridization, arrays were put together in one 15 cm glass hybridization bottle and washed as follows: 3 x 5 min with 10 ml 1 x PBS/10% SDS at 65°C; 3 x10 min with 10 ml 1 x PBS/10 % SDS at 65°C; 3 x 10 min with 10 ml 0.1xPBS/1% SDS at 65°C and finally 2 x 5 min with 10 ml 1 x PBS/1 % SDS/2 mM EDTA at RT. Filters were dried by baking at 80°C for 20 min and exposed for 8, 18 and 100 h to an imaging plate BAS-SR 2025 (Fuji Photo Film Co.).

Imaging plates were scanned with a phosphor imager (Typhoon; GE Healthcare Biosciences), which has a linear measuring range of 5 orders of magnitude. This high dynamic range allowed the simultaneous analysis of mRNAs with low and high expression levels.

3.3.6.3 Array analysis

Array evaluation was done using AIDA Image Analyzer software (Version 3.52; Raytest, Straubenhardt, Germany). At first the software “Image Analyzer“ identified the positions of the spots on the array and secondly extracted the signal intensities. Background was subtracted with the ‘Lowest grid dot’ function. Raw data (integral minus background) obtained by AIDA Array software were exported and normalized using the BioConductor package vsn. Normalized data were analyzed with the significance analysis of microarrays method (SAM).

4. Results

4.1 RNA Extraction

Prior the RNA extraction sperm cells were treated with Trition X-100 to lyse contaminating somatic cells (Figure 6a). After treatment with Triton X-100 spermatozoa were still intact but somatic cells were seen no longer (Figure 6b). After homogenization in TRIzol, examination of the homogenate under a light microscope showed that only the heads and tails of spermatozoa were separated but not completely dissolved (Figure 1c). After treatment with chaotropic solution (guanidinium thiocyanate) spermatozoa were completely dissolved (Figure 1d). Homogenates were centrifuged to eliminate cellular debris forming a pellet at the bottom of the tube (Figure 7a). It was noted that the pellet was larger than those formed after homogenization of other tissues.

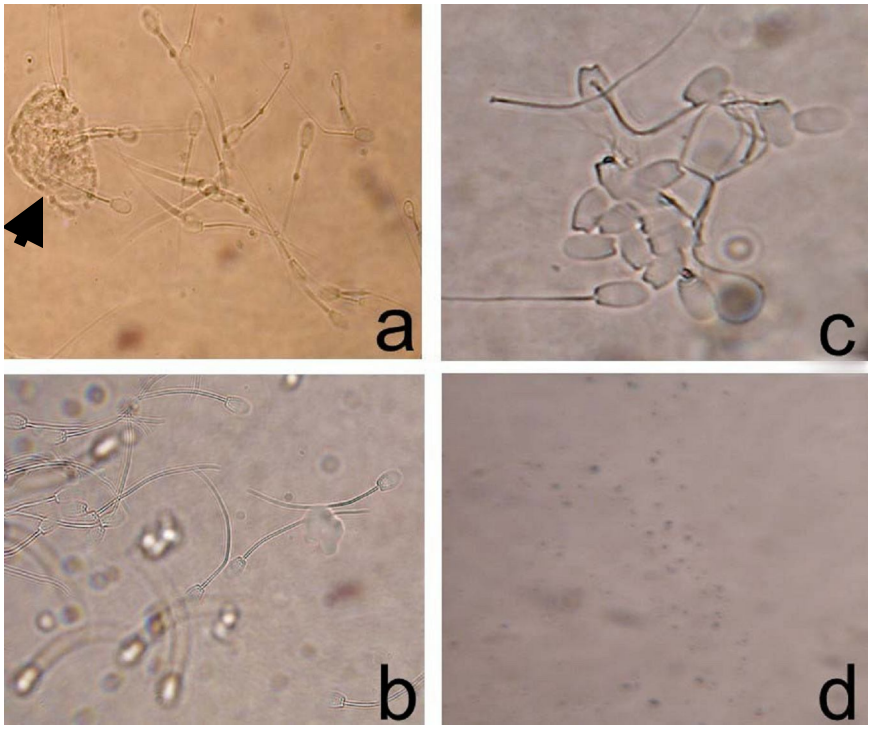

Figure 6**:** Spermatozoa after different treatments (with different digital magnifications) before and after treatment with Triton X-100, TRIzol and guanidinium thiocyanate. **a** – Ejaculate washed with Phosphate-buffered saline (PBS); the arrow indicates a somatic cell. **b** – Ejaculate after treatment for 10 min on ice

with hypotonic solution + 0.5% of Triton X-100. **c** – Spermatozoa after 30 min homogenization in TRIzol at RT. **d** – Spermatozoa after 30 min treatment with acid guanidinium thiocyanate at RT.

To test if there was any RNA remaining in the incompletely homogenized sperm cells, the pellet formed after TRIzol treatment and centrifugation, was dissolved in chaotropic solution (Figure 7b). RNA was extracted with phenol/chloroform (AGPC protocol). The total amount of RNA extracted with the first TRIzol extraction was compared with the RNA amount extracted from the cell pellet, after the AGPC extraction. The TRIzol method extracted an average of 612 ng of nucleic acid from $1x10^8$ cells while the AGPC recovered 297 ng from the pellet according to the UV measurement.

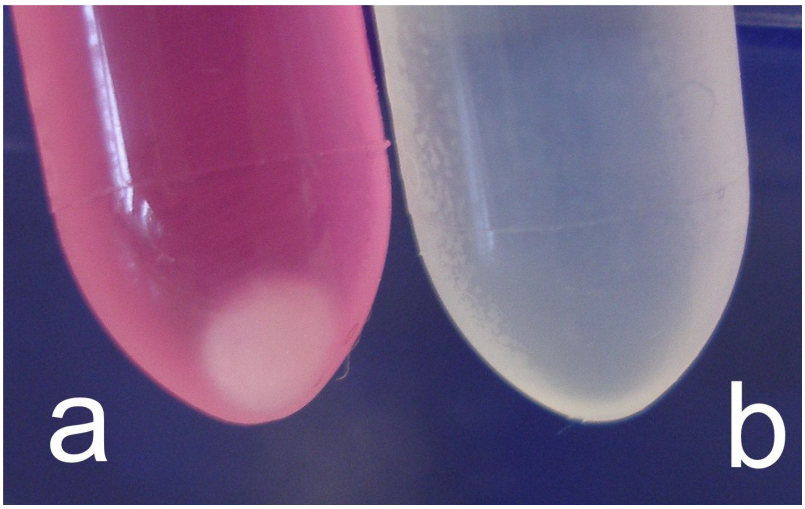

Figure 7**:** RNA extraction from the pellet after TRIzol treatment. **a** – Spermatozoa after 30 min in TRIzol at RT and 5 min centrifugation at 5,000 x *g*; **b** – The pellet formed after treatment with TRIzol was dissolved in chaotropic solution (guanidinium thiocyanate) for 30 min at RT and centrifuged for 5 min at 5,000 x *g*.

The results showed that almost all of the RNA was extracted with the TRIzol protocol. After AGPC extraction of the TRIzol insoluble material only a DNA contamination was observed (Figure 8, lane 2). After digestion with DNase I this DNA contamination disappeared and no

RNA was detectable in the agarose gel (Figure 8, lane 4). In contrast, after DNase I treatment of the RNA isolated with TRIzol only a faint high molecular weight band disappeared.

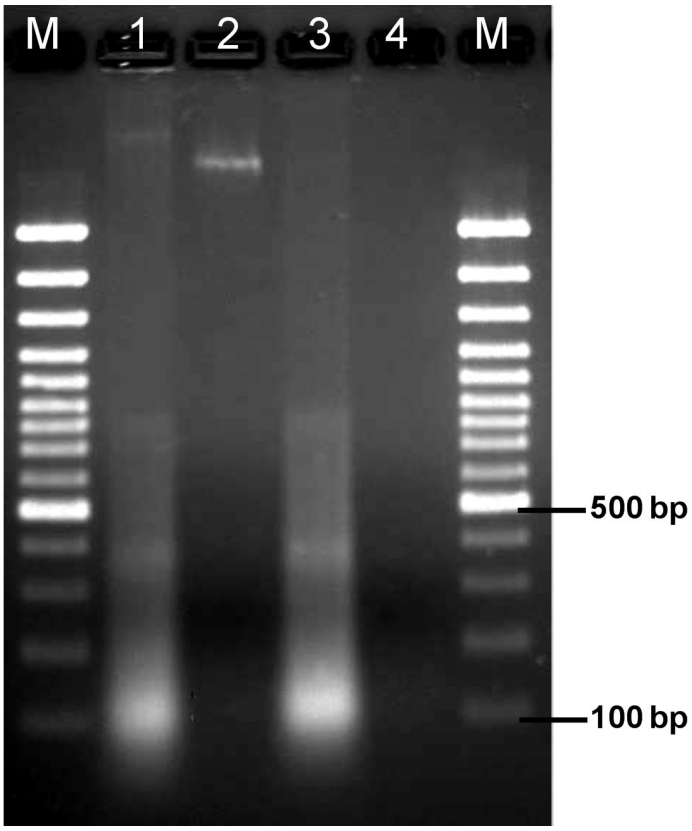

Figure 8**:** Agarose gel analysis of sperm RNA after TRIzol extraction and AGPC extraction of the pellet after TRIzol homogenization. 1% agarose/1xTBE, 4.5 V/cm 45 min. **1** – RNA extracted with TRIzol prior DNase I digestion; **2** – RNA extracted, from the pellet formed after TRIzol treatment and centrifugation, with AGTC prior DNase I digestion. **3** – RNA extracted with TRIzol after DNase I digestion; **4** – RNA extracted from the pellet with AGTC after DNase I digestion. **M** – 100 bp Ladder plus.

Furthermore, the TRIzol standard protocol was compared with the heated TRIzol protocol from Gilbert *et al.* (2007). No significant difference was found in the total RNA extraction

between the two treatments, approx. 472 ng and 454 ng (for $1x10^8$ cells) for heated TRIzol and TRIzol, respectively (Figure **9**).

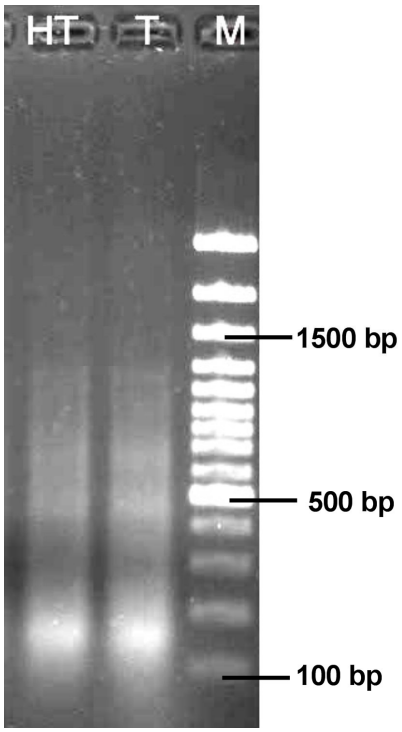

Figure 9: Comparison between standard TRIzol (T) protocol and heated TRIzol (HT) extraction. 1% agarose/1x TBE, 4.5 V/cm 45 min. **HT** – Sperm cells were treated with TRIzol at 65°C for 30 min before extraction; **T** –RNA extracted with TRIzol according to manufacturer's protocol. **M** – 100bp Ladder plus.

Overall, RNA was extracted from semen samples of 8 different boars, 5 of them were preserved in RNAlater and the remaining 3 were frozen for artificial insemination, with high reproducibility. The extracted total RNA from sperm always showed no defined bands for ribosomal RNA (18S and/or 28S) and appeared like having a lower integrity. The total RNA from spermatozoa was compared with total RNA from HeLa cells (300 ng from each cell type) on a 1% agarose gel (Figure 10).

The sperm RNA appeared like a smear without defined bands ranging from approx. 20 bp to 700 bp in Figure 5 and from 50 bp to 1500 bp in Figure 4 compared to the DNA molecular weight marker. A strong signal was always detected at approximately 100 bp relative to the DNA molecular weight marker.

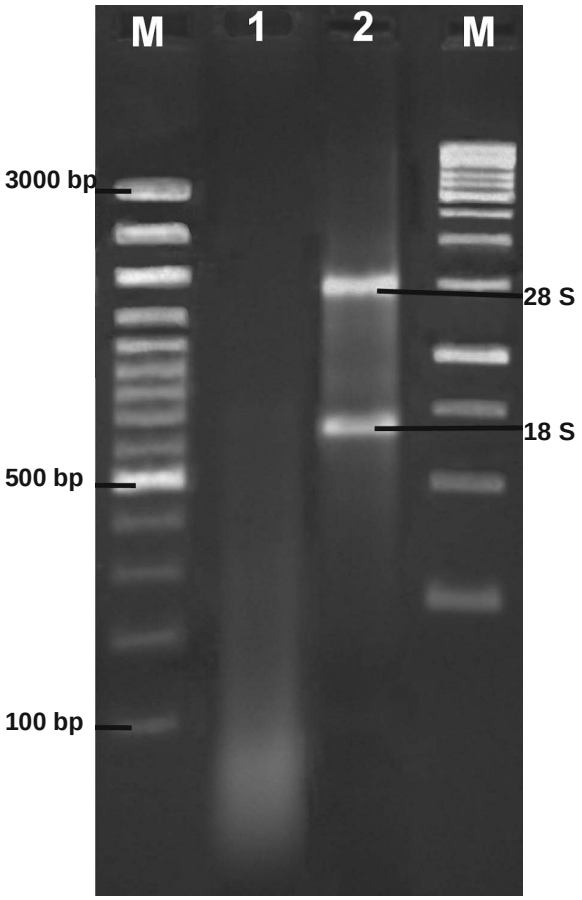

Figure 10: Comparison of 300 ng total RNA extracted from spermatozoa and RNA from HeLa cells. 1% agarose/1x TBE, 4.5 V/cm 60 min. **m** – 100 bp Ladder plus; **1** – RNA from spermatozoa; **2** – RNA from somatic cells and **M** – 1 kb Ladder.

To obtain a better electrophoretic separation the total RNA from boar spermatozoa was analyzed with a Bioanalyzer (Agilent Technologies, Santa Clara - USA), a microfluidics-based platform for the qualitative analysis, which allows the visualization of very low amounts of RNA. The results from the Bioanalyzer Pico chip confirmed that the RNA in the spermatozoa is composed basically of relatively short RNA molecules, most of them between 500 and 25 bp (Figure 11). Peaks for the large ribosomal RNAs were not detectable.

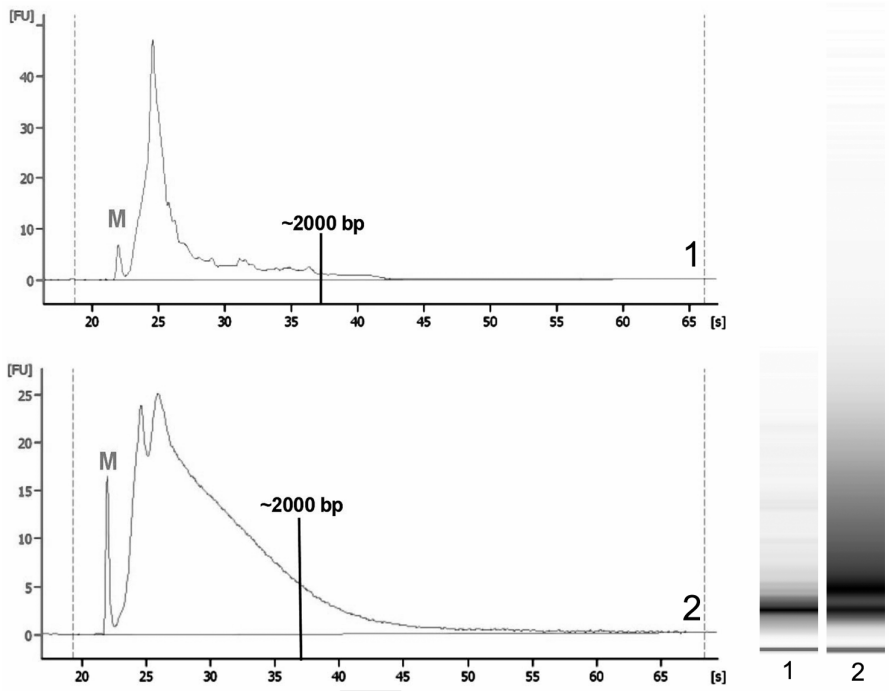

Figure 11: Bioanalyzer-electrophoresis profile of spermatozoa RNA. FU – fluorescence units; s – seconds; M – 25 bp marker. At the right side, an agarose gel electrophoresis-like picture of the analyzed RNA is shown; the line corresponds to the 25 bp marker. 1 – RNA from porcine spermatozoa, 2 – Amplified RNA from porcine spermatozoa.

Obtaining RNA free of DNA, protein and salt contaminations is crucial for quantitative studies. For this purpose the ratio of absorbance at 260 and 280 nm is used to assess the purity of RNA. The aromatic ring structure of the purine and pyrimidine (nucleoside) bases of DNA and RNA are responsible for absorbance of UV light very near of 260 nm (Prathapkumar and Sureshkumar 2002). A ratio of 1.7-2.1, depending on whether the RNA is measured (water-low ratio, TE-2.0-2.1), is generally accepted as “pure” for RNA. A lower ratio may indicate the presence of protein or other contaminant that may absorb strongly at or near 280 nm. The ratio of the absorbances at 260 and 230 nm is a secondary measurement of nucleic acid purity. Organic contaminants like phenol and other aromatic compounds used in RNA extraction, absorb light at a wavelength of about 230 nm. The 260/230 ratio for pure nucleic acid is commonly in the range of 1.8-2.2 and a lower ratio indicates the presence of co-purified contaminants.

The NanoDrop spectrophotometer results showed an unusual absorbance curve for spermatozoa RNA. The curve was similar to results observed in samples, which were contaminated with solvents, salts or proteins. The RNA of spermatozoa shows a strong absorbance between 220 and 230 nm (Figure 12, arrow no. 1) and a second absorbance peak at 270 nm (Figure 12, arrow no. 2). In contrast, RNA isolated from placenta showed an absorbance peak at 260 nm and normal 260/230 and 260/280 ratios. Very low values were observed for the ratios of 260/280 nm and 260/230 nm, lower than 1.8 and 1.0 respectively. These resulted in imprecise measurements leading to artificially elevated RNA concentrations. This finding was further supported by the discrepancies between the measured RNA quantities and the estimation of the quantities seen on agarose gels. The RNA smears on the gels were always suggesting lower RNA amounts compared to the results of the UV measurements. For example, in Figure 10 300 ng total RNA from both samples were loaded.

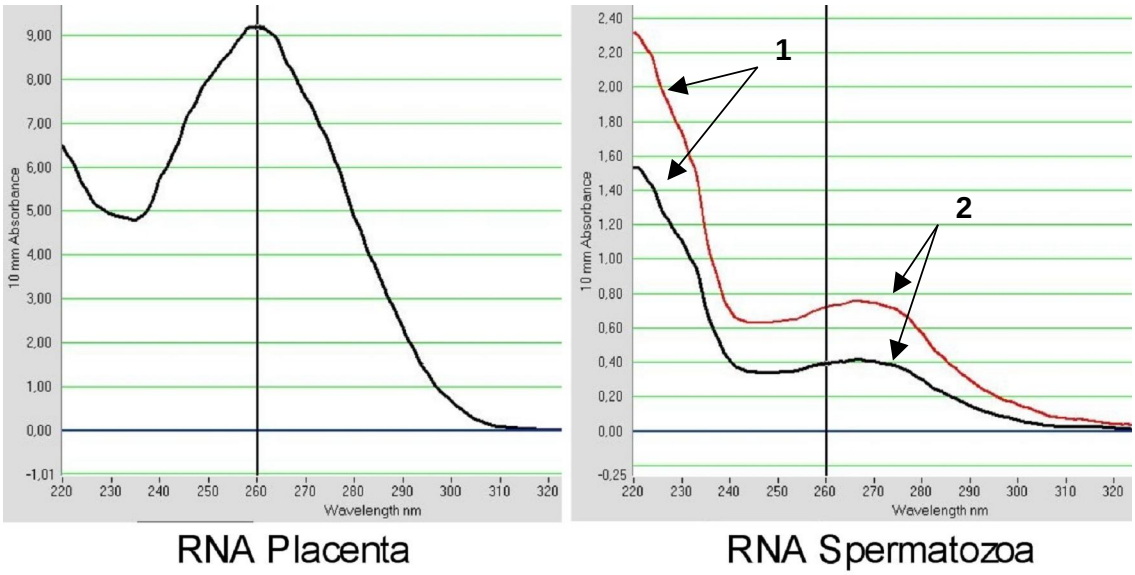

Figure 12: Assessment of RNA quality using spectrophotometric absorbance. Left, typical absorbance spectrum of high quality RNA from placental tissue. Right, typical absorbance spectrum of RNA from spermatozoa (two samples).

Polysaccharides, like glycogen, are known to decrease the 260/230 and 260/280 ratios (Mach *et al.* 1992). As glycogen was used as co-precipitant for the small amounts of RNA

from spermatozoa, the effects of glycogen on the RNA precipitation and the concentration measurement were analyzed. RNA was extracted from somatic cells with or without the addition of glycogen. The pellet from the RNA precipitation with glycogen appeared larger and more compact. A decrease of the 260/230 ratio was seen in the RNA precipitated with glycogen when compared with the RNA precipitated without glycogen, 1.12 and 2.03 respectively. The glycogen did not affect the 260/280 ratio (Figure 13). However, both ratios, obtained from sperm RNA, even when precipitated without glycogen, were much lower than 2.0. Investigation with mass spectrometry detected no protein in the RNA extract from porcine spermatozoa (data not shown).

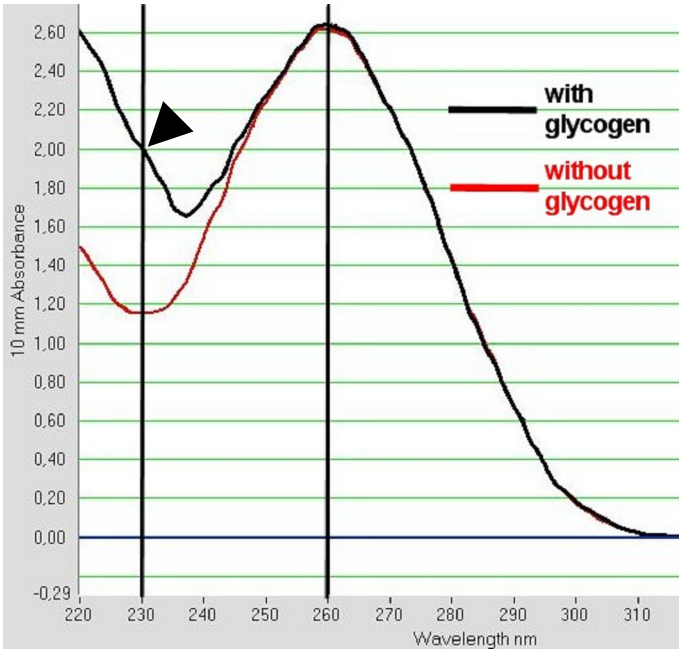

Figure 13: Absorbance curve of RNA extracted from cells COLO 357 precipitated with or without glycogen as co-precipitant. The black curve represents the RNA precipitated with 15 µg glycogen after extraction with TRIzol. The red curve represents the RNA precipitated without glycogen.

Other co-precipitants, like lithium chloride, ammonium acetate and sodium citrate did not result in a visible pellet when small amounts of spermatozoal RNA were precipitated. Furthermore, the precipitation of sperm RNA with 5 M of sodium citrate and increasing concentrations of isopropanol (10, 15, 20%) was tested to eliminate contaminants. The RNA

concentration and the amount of small RNA fragments increased with rising isopropanol concentrations (Figure 14). The fraction of larger RNAs was precipitated with low isopropanol concentration (10%) and small RNAs stayed in suspension being recovered with 20% isopropanol. Due to the low RNA concentrations, the contrast in the area of the agarose gel containing the RNA smears was increased using a image editing software (Adobe Photoshop) for better visualization. Precipitation with 5 M sodium citrate failed to avoid contaminations as shown by the 260/280 and the 260/230 absorbance ratios (Table 2).

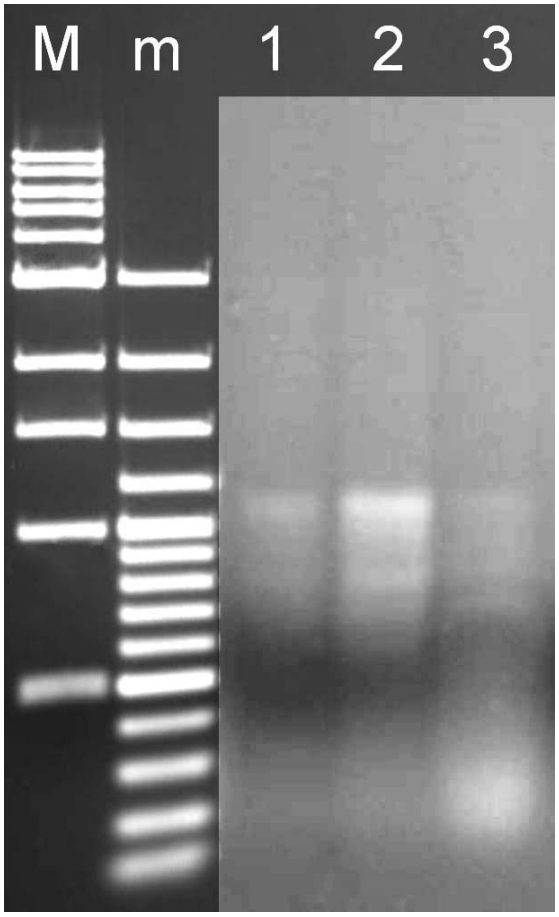

Figure 14: Sperm RNA extracted with TRIzol and precipitated with: **1**- 5 M Na-citrate with 10% isopropanol, **2**- 5 M Na-citrate with 15% isopropanol, **3**- 5 M Na-citrate with 20% isopropanol, **M**- 1 Kb Ladder, **m**- 100 bp Ladder. 1% agarose/1x TBE, 3.5 V/cm for 10 min and 4.5 V/cm for 90 min.

Table 2: 260/280 and 260/230 ratios from RNA precipitated with 5 M Na-citrate and different concentrations of isopropanol.

Isopropanol concentration	260/280	260/230
10%	0.98	0.20
15%	0.72	0.10
20%	0.70	0.09

The problem of the UV measurement of RNA samples derived from spermatozoa was also evident when RNA from spermatozoa and HeLa cells was reverse transcribed with an oligo(dT) primer containing a T7 promoter sequence and amplified with a T7 RNA polymerase (MessageAmp II aRNA Amplification Kit – Ambion). The linear amplification started from 100 ng (according to UV measurement) of RNA from spermatozoa and 100 ng of total RNA from HeLa cells. The obtained amount of RNA from HeLa total RNA was twice as high as that obtained from spermatozoa RNA (9.5 and 4.5 µg, respectively) (Figure 15). This difference was probably due to the lower amount of mRNA templates in the spermatozoa RNA. Based on a mRNA proportion of 3%, 100 ng HeLa total RNA contain only approx. 3 ng mRNA. Given a similar amplification efficiency, 100 ng sperm RNA (according to UV measurement) would contain only 1.5 ng mRNA. It was further observed that only the larger mRNA molecules were amplified and not the RNA fraction that is always seen at 100 bp compared to the DNA molecular weight standard.

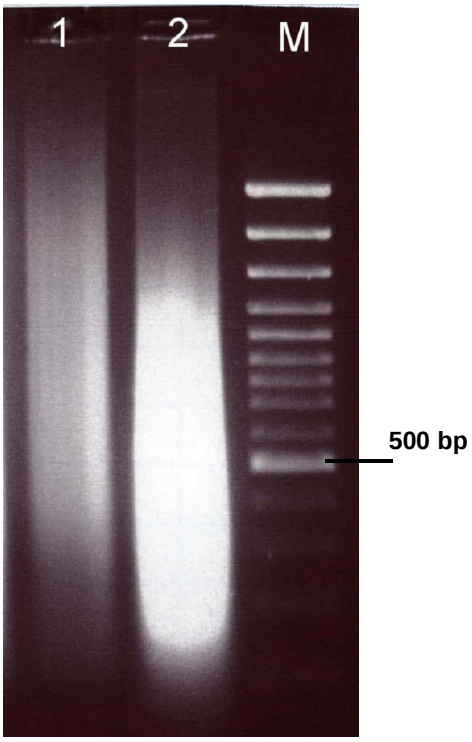

Figure 15: Distribution of amplified RNA derived from spermatozoa and HeLa cells after one-step amplification (for 12 h) with the MessageAmp II aRNA Amplification Kit. 1% agarose/1xTBE, 4.5 V/cm for 45 min. **1** – amplified RNA from sperm cells; **2** – amplified RNA from HeLa cells; **M** – 100 bp Ladder.

To investigate if the unknown contaminations inhibit cDNA synthesis or amplification that is needed for the production of a cDNA library and for the synthesis of hybridization probes, the RNA from sperm cells was mixed with RNA from HeLa cells (50:50) and compared with the amplification starting from 100 ng total RNA of each cell type. The amplification results indicate that the addition of RNA from spermatozoa did not inhibit the RNA amplification. The resulting RNA amount increased proportionally from 100 ng sperm RNA to half sperm/half HeLa RNA to 100 ng HeLa RNA (Table 3).

Table 3: Amplification of total RNA from spermatozoa and HeLa cells pure or mixed.

	ng/µl	260/280	260/230
100 ng spermatozoa RNA	59	2.45	0.5
50 ng spermatozoa, 50 ng HeLa RNA	245	2.33	1.6
100 ng HeLa RNA	405	2.22	1.7

4.2 Normalized cDNA Library

A cDNA library represents the mRNAs that are expressed in a given cell at a given time point. Spotted on cDNA arrays, cDNA libraries permit the study of gene expression at the mRNA level. Due to the very different frequencies of mRNAs in a cell (from one copy to tens of thousand copies per cell), so-called normalized cDNA libraries are usually used as a source for the generation of microarrays to avoid highly redundant cDNAs on the array. For porcine sperm cells normalized cDNA library was constructed by Vertis Biotechnologie AG (Freising-Weihenstephan, Germany).

In total a number of 4,224 cDNA clones was picked from this library. The cDNA fragments of the cDNA clones were amplified via PCR. The quality of all PCR products was analyzed on agarose gels. Complementary DNA clones that yielded no PCR products, double bands (Figure 16, example 1) or products smaller than 500 bp (example 2) were discarded, since first sequencing tests showed that fragments smaller than 500 bp were always derived from clones containing only the empty cloning vector. 3,222 cDNA clones were successfully sequenced. Analysis of these sequences resulted in 1,892 different transcripts (Figure 17). All obtained sequences were compared to GenBank (databases nr/nt, est, est_others) using the NCBI BLAST tool (discontiguous BLAST for more dissimilar sequences). For 15 % of the sequences a known porcine gene was identified. For 1,409 cDNAs porcine EST sequences were found. Except for 404 sequences all cDNAs were assigned to a known gene, in the majority of the cases to human or mouse orthologous genes.

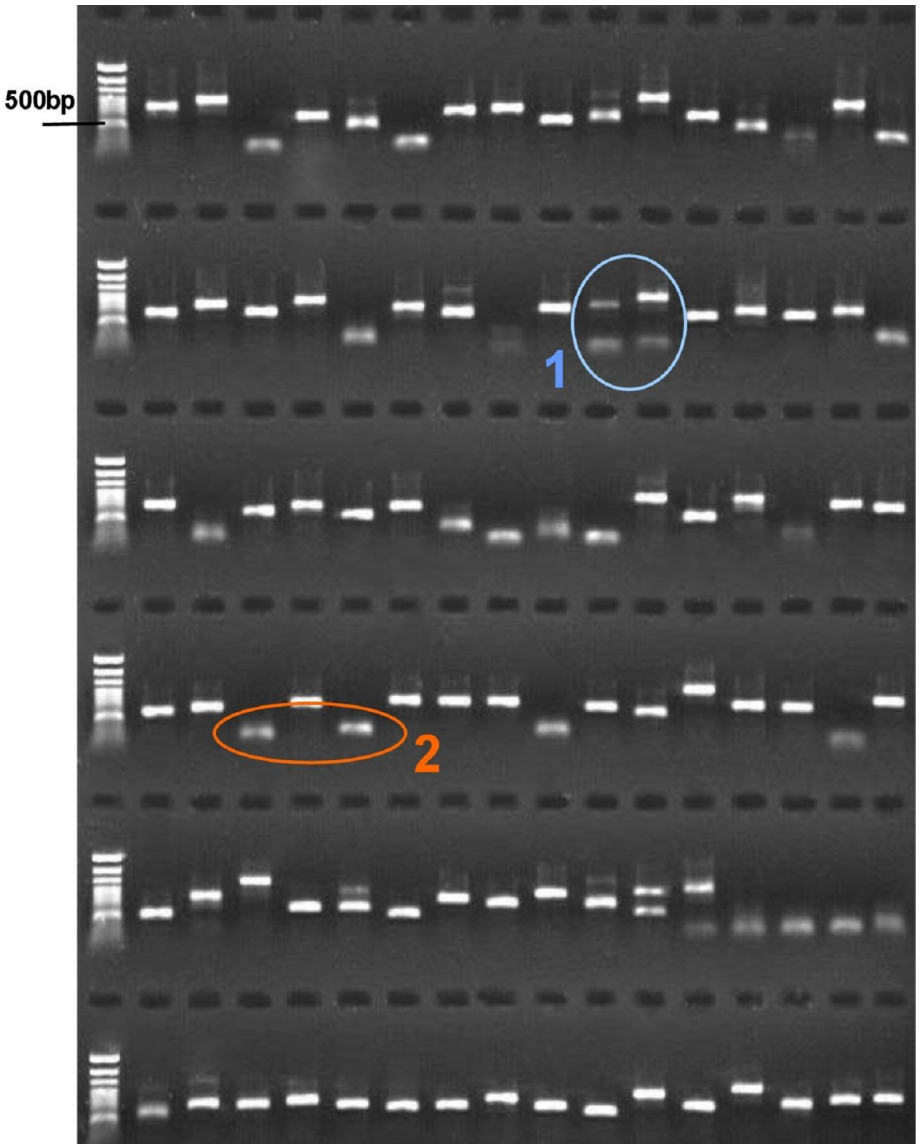

Figure 16: Agarose gel electrophoresis of amplified cDNA clones. 0.8% agarose/1xTBE, 3 V/cm for 10 min and 20 min at 5 V/cm. 0.5 µl of each PCR product were analyzed. The 100 bp Ladder plus was used as molecular weight standard. 1: cDNA clones yielding double bands; 2: cDNA clones yielding PCR products smaller than 500 bp.

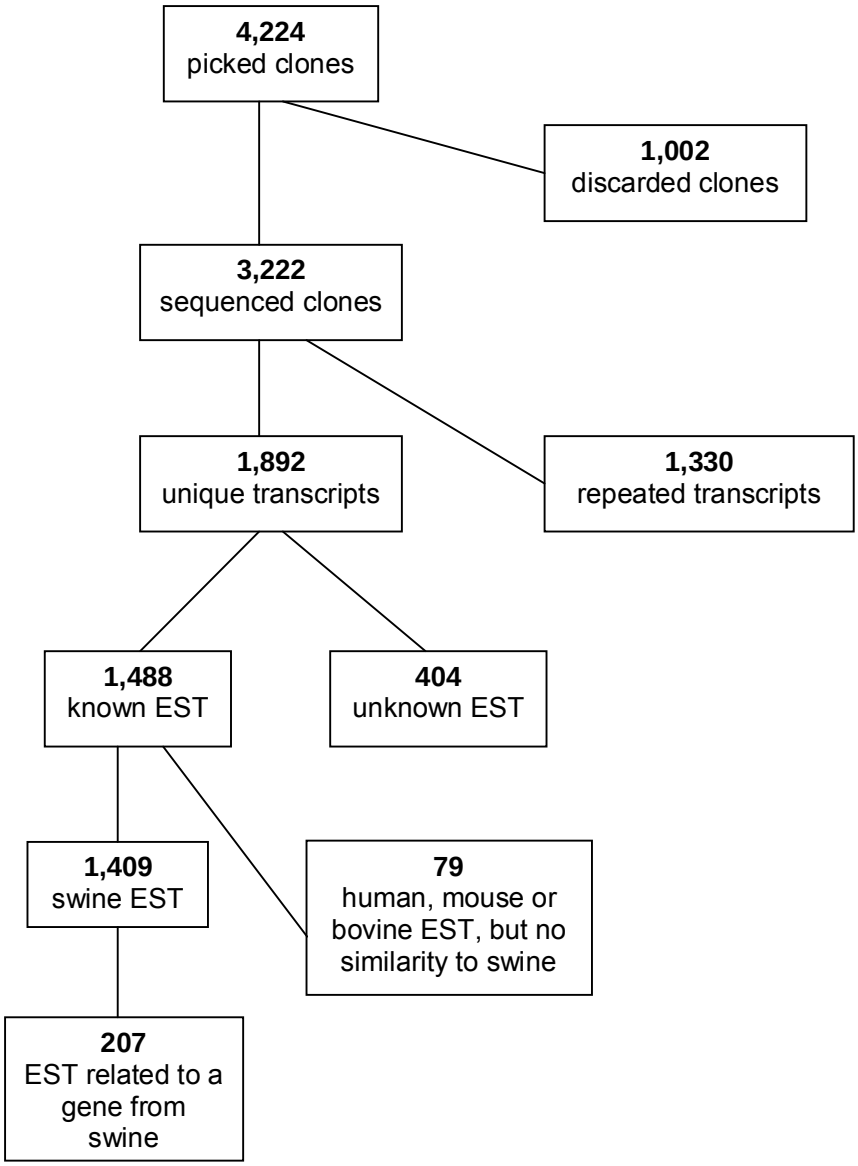

Figure 17: Overview of DNA sequencing results.

A functional gene classification was done based on the human orthologous genes using the DAVID software (http://david.abcc.ncifcrf.gov/). For this purpose, the gene symbols from the transcripts were imputed and the software gives the major biological functions associated with the genes. One hundred and thirty-one (131) functional groups were obtained and the twenty most prominent ones are shown in Table 4. Some transcripts are included in more than one functional group. Approximately 8% of the cDNA sequences corresponded to genes with unknown function. In Table 5 the transcripts with higher occurrence are listed.

Table 4. The twenty most prominent functional groups of genes in the normalized cDNA library for boar spermatozoa using the DAVID software suite.

Function Description	no. of genes
Cellular process	561
Metabolism	409
Organelle	387
Cytoplasm	256
Protein binding	246
Catalytic activity	245
Protein metabolism	211
Nucleus	197
Establishment of localization	138
Biosynthesis	115
Nucleotide binding	107
Ribosome, ribonucleoprotein complex	65
Mitochondrion	46
Golgi apparatus	28
mRNA processing	18
Regulation of translation	15
Mitotic cell cycle	14
Male gamete generation	11
Sperm motility	3

Table 5. List of the 30 most frequent transcripts found in the normalized cDNA library for ejaculated spermatozoa from boar.

Porcine cDNA and homologue (similarity)	Gene symbol	GenBank accession no.	Protein function	Biological process (frequency)
Sus scrofa mitochondrion, 16S rRNA	16S rRNA	DQ518915	Mitochondrial regulation of translation	Translation (30)
Sus scrofa mitochondrion, cytochrome c oxidase subunit I	COI	AF034253	Cytochrome c is the last enzyme in the respiratory electron transport chain of mitochondria.	ATP synthesis (23)
Sus scrofa 18S ribosomal RNA	LOC448984	NR_002170	Regulation of translation	Translation (22)
Sus scrofa mitochondrion, ATPase subunit 6	ATPase 6	AF034253	Enzymes that catalyze the decomposition of adenosine triphosphate (ATP) into adenosine diphosphate (ADP) and a free phosphate ion.	Respiratory chain (18)
Sus scrofa mitochondrion, NADH dehydrogenase subunit 2	NADH2	AF034253	NADH dehydrogenase (quinone) activity	Oxyredutase (18)
Sus scrofa mRNA for spermadhesin PSP-I	PSP-I	AJ853852	Family of secretory proteins of the male genital tract, peripherally associated to the sperm surface	Sperm associated protein (18)
Sus scrofa mitochondrion, cytochrome c oxidase subunit III	COIII	AF034253	Cytochrome c is the last enzyme in the respiratory electron transport chain of mitochondria.	ATP synthesis (17)
Sus scrofa isolate PN149 cytochrome b gene	CYTB	AY830187	Component of the ubiquinol-cytochrome c reductase complex, which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis	ATP synthesis (16)
Sus scrofa ribosomal protein L10	RPL10	NM_001044543	Form the stalk in eukaryotic ribosomes	Translation (16)
Sus scrofa mRNA, clone:OVR010084G10 ; Homo sapiens ribosomal protein, large, P0 (99%)	RPLP0	AK239428	Form the stalk in eukaryotic ribosomes	Translation (15)
Sus scrofa mitochondrion 12S rRNA	12S rRNA	AF034253	Mitochondrial regulation of translation	Translation (14)

Table 5. continued

Porcine cDNA and homologue (similarity)	Gene symbol	GenBank accession no.	Protein function	Biological process (frequency)
Sus scrofa ferritin L subunit	FTL	AF288821	Cellular iron ion homeostasis	Metal ion binding (14)
Sus scrofa mRNA for spermadhesin PSP-II	PSP-II	AJ853853	Family of secretory proteins of the male genital tract, associated to the sperm surface	Sperm associated protein (14)
Sus scrofa mRNA for spermadhesin sperm associated AWN protein	AWN	AJ853850	Family of secretory proteins of the male genital tract, associated to the sperm surface	Sperm associated protein (13)
Sus scrofa eukaryotic translation elongation factor 1 alpha	EEF1A	DQ673096	Stimulates the recruitment of the initiator tRNA and the mRNA to the 40S ribosomal subunit.	Translation (13)
Sus scrofa mRNA for ferritin heavy-chain	FTH1	397030	Cellular iron ion homeostasis	Metal ion binding (13)
Sus scrofa acrosin inhibitor	LOC396905	NM_213877	Important capacitation factors, protecting zona binding sites during sperm uterine passage, and then dissociating to allow sperm binding to the zona pellucida of the oocyte.	Sperm capacitation (12)
Sus scrofa translationally controlled tumor protein	TCTP	AY072784	Involved in calcium binding and microtubules stabilization	Ion binding (11)
Sus scrofa cDNA clone scan0007.g.14 5prim; Homo sapiens keratin 13 (99%)	KRT13	BX919803	Structural integrity	Cytoskeletal keratin (10)
Sus scrofa mRNA, clone:LVR010019A07; Homo sapiens tetraspanin 6 (99%)	TSPAN6	AK232321	Plays a role in the regulation of cell development, activation, growth and motility.	Signaling pathway (10)

Table 5. continued

Porcine cDNA and homologue (similarity)	Gene symbol	GenBank accession no.	Protein function	Biological process (frequency)
Sus scrofa cDNA clone PDUts2091D06 5', H.s. cysteine-rich secretory protein 2 (98%)	CRISP2	CX064592	Component of the sperm acrosome that remains associated with sperm after capacitation and acrosome reaction. Is relevant for sperm-oocyte interaction.	Spermatogenesis (10)
Sus scrofa cDNA clone scan0031.e.20 5prim, Bos taurus similar to probable protease inhibitor splice variant WAP8a (98%)	LOC521427	BX915529		Protease inhibitor (10)
Sus scrofa secreted phosphoprotein 1 (osteopontin)	SPP1	NM_214023	Significantly decreases the incidence of polyspermy during IVF in pig (Hao *et al.* 2006).	Physiological process (9)
jns18bA04.f jns Sus scrofa cDNA 5' ; Homo sapiens poly(A) binding protein, cytoplasmic 1 (99%)	PABPC1	NM_002568	Poly(A) shortening and translation initiation. Is found complexed to the 3-prime poly(A) tail of eukaryotic mRNA	Translation (9)
Sus scrofa mitochondrion NADH dehydrogenase subunit 6	NADH6	AF034253	NADH dehydrogenase (quinone) activity	Oxiyredutase (9)
Sus scrofa beta-2-microglobulin	B2M	DQ845172	Beta-chain of major histocompatibility complex class I molecules.	Antigen presentation (8)
Sus scrofa peroxiredoxin 5	PRDX5	NM_214144	Antioxidant protective role in different tissues under normal conditions and during inflammatory processes	Antioxidant (7)
Sus scrofa cytoskeletal beta actin; Homo sapiens actin, gamma 1 (98%)	ACTB	AK240355	Involved in various types of cell motility, and maintenance of the cytoskeleton	Motility (7)
Sus scrofa mRNA, clone:UTR010025G02; Homo sapiens myosin, light chain 6 (99%)	MYL6	AK240054	Is a hexameric ATPase cellular motor protein.	Motility (7)

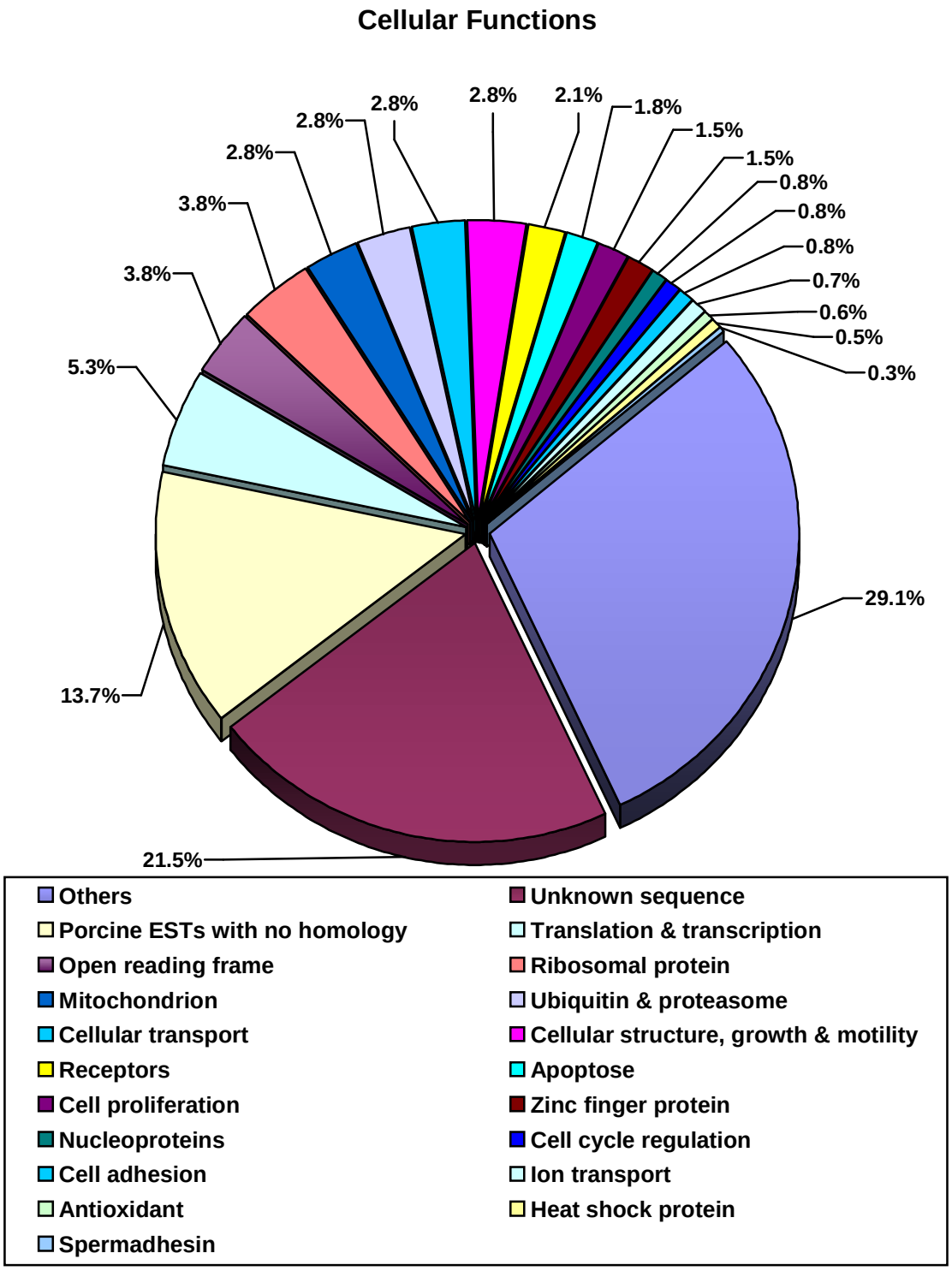

Figure 18: Illustration of the cellular functions of the genes found in the cDNA library.

As shown in Table 4 and Figure 18, the transcripts cover a wide variety of functions, such as spermatogenesis (*DDX* family, *CPEB2*, *TXNDC8*, *TEKT1*, *TRIM69*), mitosis and meiosis (*SEH1L, CCNB1, CETN1, LMNB2, LZTFL1, PTMA, PLK2, MEIG1*), cell cycle regulation (*S100, CALM1, CCPG1*), ubiquitin and proteasome (*UBB, UBA52, UBE2D3, PSM* family), apoptosis regulation (*PDCD2L, ITM2B, SERINC3, TXN2, DAD1, GLUL*), cell motility (*TSPAN, AKAP4, S100A11, ACTG1, CAPG*), acrosome reaction and cell adhesion (*CRISP, SPACA, SPA17, CDH1, SDCBP, TACSTD1, CD9*).

Transcripts for mitochondrial proteins, heat shock proteins and ribosomal proteins had a strong presence in the cDNA library. There were 53 different ESTs, including two mitochondrial rRNAs (16S mitochondrial rRNA was the most frequent RNA in the library) and cDNAs for eight mitochondrial ribosomal proteins. Ten different transcripts for members of the HSP90 and HSP70 family and approximately 100 transcripts related to transcription and translation were found in the cDNA library. Nineteen transcripts are coding for initiation factors, including subunits, and 6 for elongation factors. More than 70 transcripts for ribosomal proteins were identified, most frequent the ribosomal proteins P0 and L10 (RPLP0 and RPL10).

4.3 Comparative cDNA Array Hybridization of Sperm RNA from a Polyspermy Boar and a Normal Control Boar

A total of 3,009 cDNA clones were spotted onto two membranes to generate a cDNA microarray. To investigate if there are differences in gene expression in spermatozoa derived from a boar that showed polyspermy *in vitro* and a normal control boar, RNA was extracted from the sperm of the two boars and hybridized with the arrays. For every boar three replicate hybridizations were performed. For better comparison between the different array hybridizations the membranes were exposed for different times (8, 18 and 100 hours). The spot positions on the array were identified and the signal intensities extracted using AIDA Array Evaluation software (Figure 19). The pink circles in Figure 14 represent a signal derived from a cDNA with differential expression between the polyspermic boar and the boar with normal fertilization rate.

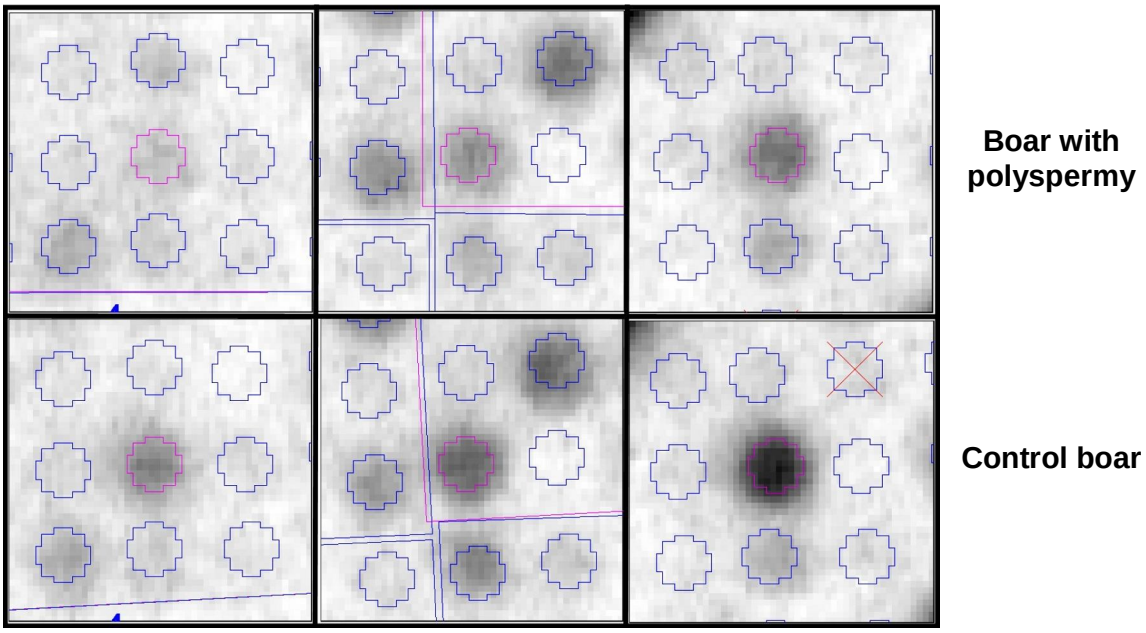

Figure 19: Array evaluation using AIDA Image Analyzer software (Version 3.52; Raytest, Straubenhardt, Germany).

The hybridization with the sample derived from the boar without polyspermy revealed 60 spots with higher and 5 with lower hybridization signals when compared with the sample from the boar with polyspermy as obtained by statistical analysis using the Significance Analysis of Microarray (SAM) method, with a false discovery rate of approx. 5 % for both array membranes (Figure 20). The transcripts with higher abundance in the samples from the boar without polyspermy are shown in Table 6. All transcripts, except the last one, code for mitochondrial proteins and five ESTs that had no similarity with GenBank sequences (they were not included in Table 6). From the 5 down regulated genes, 3 code for transcripts similar to bacterial 16S rRNA, one is highly similar to open reading frame 24 from the chromosome 16 (from Homo sapiens with 99 % similarity) and the last was a short sequence with 31 bp for that no similarity to known sequences was found.

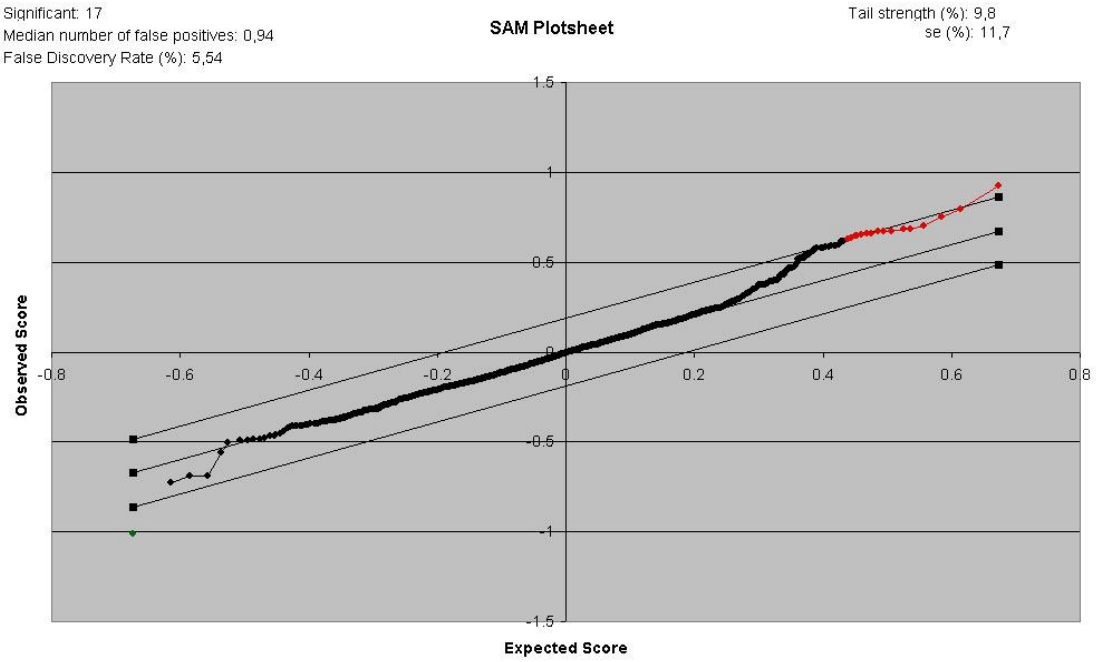

Figure 20: Significance analysis of microarray (SAM) plot for the data derived from one microarray membrane. Red dots: data points of cDNA clones with higher expression levels in spermatozoa of the control boar; green dots: data points of cDNA clones with lower levels in spermatozoa of the control boar.

More than 170 cDNA clones showed very strong signals, whereas the other signals were dramatically lower (Figure 21). The fifty cDNAs of each membrane with the strongest hybridization signals were further analyzed. Of these cDNAs, 25 sequences matched with porcine ESTs but not with known genes. Five cDNAs showed a very high similarity (>98%) with human genes. The identified transcripts code for glycoprotein IIIa (GPIIIa), catenin alpha 1 (CTNNA1), both related to cell adhesion; S100 calcium binding protein A12 (S100A12), zinc finger and BTB domain containing 20 (ZBTB20), related to cell differentiation, and for trace amine associated receptor 6 (TAAR6). For the remaining seventy percent of the cDNA clones that showed very strong signals no similar sequences were found in GenBank.

Table 6: Genes with higher RNA concentrations in samples derived from the boar without polyspermy.

Best BLAST hit	Gene Symbol	Accession No	Frequency	q-value%	Fold Change
Sus scrofa mitochondrion, complete genome	12S rRNA	AF486866	6	3.1	2.3
Sus scrofa mitochondrion, complete genome	ATPase 6	AF034253	8	1.9	2.5
Sus scrofa mitochondrion, complete genome	ATPase 8	AF034253	2	2.8	2.8
Sus scrofa mitochondrion, complete genome	COI	AF034253	4	0	2.4
Sus scrofa mitochondrion, complete genome	COII	AF034253	3	2.3	2.2
Sus scrofa mitochondrion, complete genome	COIII	AF034253	8	2	2.4
Sus scrofa complete mitochondrial DNA	CYTB	AF034253	6	3.3	2.3
Sus scrofa mitochondrion, complete genome	NADH1	AF034253	4	3.6	2.1
Sus scrofa mitochondrion, complete genome	NADH2	AF034253	7	2.4	2.5
Sus scrofa isolate PN149 cytochrome b	NADH4	AJ002189	3	0.8	2.6
Sus scrofa mitochondrion, complete genome	NADH5	AF034253	3	4.4	2.6
Sus scrofa mitochondrion, complete genome	NADH6	AF034253	5	2.7	2.5
Sus scrofa mRNA, clone: AMP010048E04, expressed in alveolar macrophage		AK230819	1	2.3	2.0

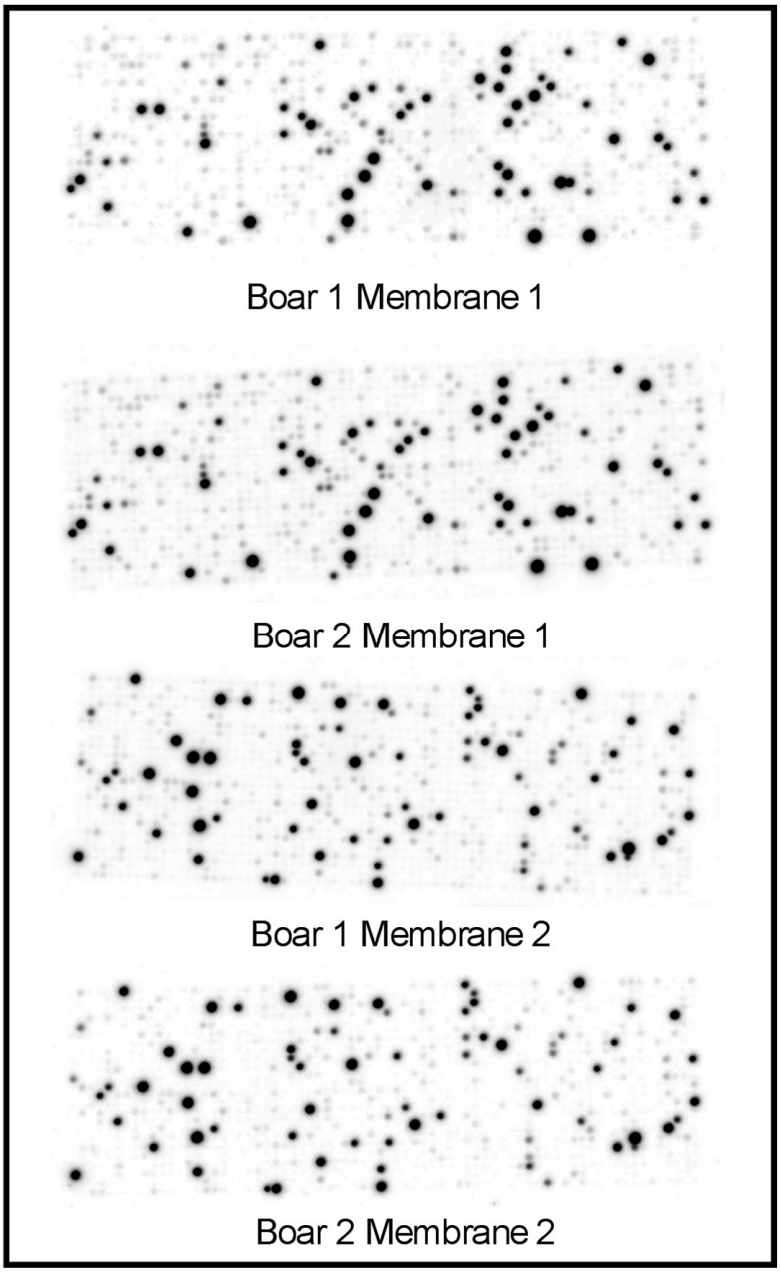

Figure 21: Hybridization of RNA from two boars with the microarrays derived from the normalized cDNA library.

5. Discussion

5.1 Extraction and Characterization of RNA Isolated from Porcine Spermatozoa

Miller and Ostermeier (2006) speculated that the spermatozoon does not only transport the paternal genome to the oocyte, but stored RNA could be also delivered, being essential for the gene expression in the zygote prior to the activation of the embryonic genome. RNA from spermatozoa has been isolated by a number of different groups for characterization of single mRNAs or large sets of mRNAs using techniques such as quantitative real-time RT-PCR or microarrays. The techniques for RNA extraction are standardized for most cells and tissues of different origin and several commercial products are available for this purpose. In the case of spermatozoa the limitations of all RNA isolation techniques are the low amount of cytoplasm in sperm cells, low concentration of total RNA and the resistance of the cellular membrane to lysis.

The most widly used protocol for RNA extraction from ejaculated sperm cells is the method of Chomczynski and Sacchi (1987) using acid guanidinium isothiocyanate phenol-chloroform extraction (AGPC) (Alcivar *et al.* 1989, Miller *et al.* 1994, Ostermeier *et al.* 2002, Lambard *et al.* 2004). In advancement of this method, the extraction with TRIzol RNA isolation reagent (Invitrogen, Carlsbad, CA, USA) has been used with few modifications (Dadoune *et al.* 2005, Zhao *et al.* 2006, Gilbert *et al.* 2007). Another possibility for RNA isolation is based on the disruption of cells with a chaotropic solution and subsequent filtering through a silica-based column that selectively and quantitatively binds mRNA and larger rRNAs (Ostermeier *et al.* 2005, Gur and Breitbart 2006). That excludes on the one hand the largest RNA population in spermatozoa, small RNAs, which do not efficiently bind to the silica matrix and, on the other hand, a large portion of the RNA is irreversibly bound leading to decreased RNA yields.

The isolation of RNA from porcine spermatozoa using Trizol resulted in 4 to 10 ng total RNA obtained from $1x10^6$ spermatozoa, being in accordance with Lalancette *et al.* (2007) who recovered an average of 12 ng total RNA from $1x10^6$ bovine sperm cells. For human sperm cells a RNA yield of approx. 15 ng total RNA per $1x10^6$ cells has been described (Miller *et al.* 2005). According to the TRIzol manual, the expected amount of total RNA obtained from $1x10^6$ cultured cells is between 5 and 15 µg. The amount of total RNA recovered from spermatozoa is similar between species but 1,000 times lower than the amount of total RNA isolated from somatic cells. This implies that there are either only a few copies per transcript

or much lower numbers of different transcripts in a single spermatozoon, as there are only 10-20 femtogram (10^{-15} g) total RNA per spermatozoon (Krawetz, 2005). Using qPCR, Savchuk (2006) found only 10^6 mRNA copies of the *PRM2* gene in 10^8 porcine sperm cells. The concentration of 1 transcript per 100 spermatozoon was attributed to losses during spermatogenesis or during the RNA extraction.

In order to obtain a sufficient amount of RNA for differential gene expression studies, it is necessary to start the RNA isolation from high amounts of cells or to maximize the yield of RNA using optimized protocols. During pre-experiments different protocols cited in the literature were tested (Lambard *et al.* 2004, Dadoune *et al.* 2005, Ostermeier *et al.* 2005, Gur and Breitbart 2006, Zhao *et al.* 2006, Gilbert *et al.* 2007) and the best results, regarding quantity and quality, were obtained with methods based on the protocol of Chomczynski and Sacchi (Savchuk 2006). However, in this work no difference in RNA yield was found compared with the TRIzol method. After cell lysis with TRIzol and subsequent centrifugation the formation of a larger pellet was observed, bigger than for other tissues, on the bottom of the tube. The TRIzol was unable to completely dissolve the membrane of the sperm cells. A complete dissolution of the sperm cells was only achieved with chaotropic salt solution. This method has also been used by Miller and others (1994). However, the dissolution of the TRIzol insoluble pellet with chaotropic salt solution showed that there was no or little RNA left and only genomic DNA was isolated.

The electrophoretic separation of the isolated RNA on agarose gels always showed that the porcine spermatozoa RNA population is comprised of a large portion of relatively short RNAs, ranging from 25 to 200 bases compared to a DNA molecular weight standard. Separation on a microfluidic chip for small amounts of RNA confirmed this result that is in agreement with the work published by Gilbert *et al.* (2007). They showed that RNA isolated from bovine spermatozoa also contains a large amount of small RNA molecules, which they interpreted as a natural segmentation of the mRNA population. The use of lower concentrations of isopropanol in combination with a higher concentration of salts allowed a selective precipitation of larger RNA molecules leaving the small fragments in the supernatant. With this modified precipitation method it is possible to selectively enrich larger RNA molecules that potentially represent messenger RNAs. However, the small RNA fragments should be further characterized in future studies as they could have important regulatory roles like micro RNAs (miRNA). Amanai *et al.* (2006) suggested that the total RNA of mouse sperm contains approx. 20% miRNAs and the exposure of the early embryonic genome of mice to miRNA could induce permanent and heritable epigenetic changes in gene

expression (Rassoulzadegan *et al.* 2006). The small fragments of RNA found in porcine spermatozoa could, in part, be composed of miRNAs or other regulatory RNAs acting in epigenetic regulation during embryonic development. The results of the extraction of porcine sperm poly(A)$^+$ RNA with oligo(dT) beads and the amplification of poly(A)$^+$ RNA with an oligo(dT)/T7 promoter primer-based approach suggested that most of these small RNA fragments have no poly(A) tail and that the total amount of polyadenylated RNA is much lower than indicated by the UV measurement.

A further hurdle for the analysis of spermatozoa RNA was the abnormal absorption spectrum. Very low 260/280 nm and 260/230 nm ratios were measured using the NanoDrop spectrophotometer indicating salt, protein or polysaccharide contamination. Due to an overlaying absorption with a maximum at approx. 270 nm the unknown contamination artificially increased the RNA concentration, which did not correspond to results of electrophoretic analysis on agarose gels. One source for this contamination could be proteins, which were not removed by the phenol/chloroform extraction. Yamaguchi *et al.* (1992) suggested that some protein contamination could remain in the aqueous/organic phase during the RNA extraction resulting in low 260/280 values. But the analysis with highly sensitive mass spectrometry showed that no protein is detectable in the isolated RNA. Another possibility was that the glycogen, which was added to support the precipitation of small amounts of total RNA, resulted in the low 260/230 ratio. It is known that some polysaccharides can decrease this ratio (Mach *et al.* 1992). In fact, the 260/230 ratio of total RNA from somatic cells precipitated with glycogen was affected, whereas the 260/280 ratio was not influenced, showing that the glycogen indeed produced low 260/230 ratios. However, even after precipitating sperm RNA without glycogen the low 260/230 ratio remained. As alternative, the RNA could be precipitated with linear acrylamide (Gaillard and Strauss 1990). Also tests with selective precipitation failed to remove the unknown contamination. Although the contamination interfered with RNA quantification no inhibitory effects on subsequent enzymatic reactions were observed as indicated by cDNA synthesis and RNA amplification starting from a mixture of spermatozoa RNA and RNA derived from HeLa cells. In the literature, there are no commentaries about absorbance ratios of sperm RNA measured by spectrophotometry.

All analyses of sperm RNA, which were performed in this work, revealed that sperm cells contain very small amounts of mRNA. This small amount of RNA is limiting the use for hybridization experiments. To circumvent this limitation a linear amplification of the poly(A)$^+$ RNA was performed based on T7 RNA polymerase. The suitability of different amplification

approaches for sperm RNA has been shown by a number of different groups (Platts *et al.* 2007, Gau *et al.* 2008, Gilbert *et al.* 2007, Lalancette *et al.* 2007). For example, the amplification via T7 RNA polymerase is standard for the preparation of hybridization probes for Affymetrix GeneChips®, the most widespread commercial microarray technology. The amplification from porcine sperm RNA resulted in a much lower yield than expected from the UV measurement of the starting material. Based on the results of the amplification of RNA derived from HeLa cells that was done in parallel, 100 ng sperm RNA (according to UV measurement) contain only 1.5 ng poly(A)$^+$ RNA. These data indicate that 1x10^8 porcine sperm cells contain only 6-15 ng poly(A)$^+$ RNA.

In conclusion, the quantitative and qualitative analyses of porcine sperm RNA showed that sperm RNA contains a large proportion of short RNA molecules and the proportion of poly(A)$^+$ RNA is very low and can be estimated using T7 RNA polymerase amplification technique. The UV measurement results are indicative of an unknown contamination that artificially increases the estimated RNA concentration and that cannot be removed with the techniques used in this work. However, the contamination does not disturb subsequent enzymatic steps.

5.2 Characterization of a Normalized cDNA Library Produced from Porcine Spermatozoa

In order to obtain a representative source for mRNAs of porcine sperm cells a normalized cDNA library was produced. To selectively enrich for poly(A)$^+$ RNA, reverse transcription was done with a poly(A)-specific cDNA synthesis primer. For the characterization of the normalized cDNA library more than four thousand bacterial clones (4,224) were picked to obtain cDNA clones for DNA sequencing and the generation of cDNA microarrays. All cDNA clones resulting in too short cDNA fragments (<500 bp) or more than one cDNA fragment after PCR amplification were sorted out (in total 1,002). The residual cDNA clones (3,222) were sequenced in 5´-3´ direction with long reads (up to 800 bp). For many (2,544) of the cDNA clones the 3´ end containing a poly(A) tail was found in the sequence due to the relatively short length of most cDNA fragments. The analysis of the obtained cDNA sequences revealed 1,892 different transcripts. For 1,409 cDNA sequences a corresponding porcine cDNA, either a known mRNA or EST sequences only, was found and 207 corresponded to a known porcine gene. For 404 sequences no similarity with sequences of

the GenBank database was found. For 79 cDNAs similarity with human, mouse or bovine mRNAs was found, but no relation to porcine sequences.

Unknown Transcripts. The complex variety of transcripts from spermatozoa was already proven and, like in previous publications, there were several transcripts with no matches to GenBank sequences in the porcine sperm cDNA library (Miller *et al.* 1999, Zhao *et al.* 2006, Gilbert *et al.* 2007, Lalancette *et al.* 2007). Also the majority of the cDNA clones showing very strong signals in the microarray hybridization represented transcripts not found in the databases. These sequences could be from unknown genes, alternative transcripts or from non-coding RNA (ncRNA).

The classification of ncRNAs is very difficult, especially when the full genome sequence is not available, because many of them are simply unprocessed primary transcripts or, in other cases, are formed from exons of spliced transcripts. These may also be alternatively spliced or polyadenylated in a developmentally regulated fashion (Mattick 2003). Only for a short time ncRNAs have been studied systematically and their significant role in cell and especially developmental biology has been recognized. There are many presumptions about the specific functions of these RNAs via a variety of RNA-DNA, RNA-RNA and RNA-protein interactions, such as RNA-induced silencing complex (RISC) or regulatory RNA sequences in the untranslated regions (UTRs) of mRNAs (Kotaja and Sassone-Corsi 2007). Both translation and splicing require core infrastructural RNAs not only for sequence-specific recognition of RNA substrates, but also for the catalytic process itself (Mattick 2003, Mattick and Makunin 2006). Many ncRNAs carry a poly(A) signal for post-transcriptional polyadenylation, suggesting that they might be the products of polymerase II-mediated transcription and allowing them to be reverse transcribed (Suzuki and Hayashizaki 2004). According to Mattick (2003) the conclusions about genetic regulatory networks derived from microarray analyses of transcriptome activity in complex organisms may be substantially incomplete if ncRNA sequences are not included.

The understanding of mammalian genes and genomes and their annotations are still simplistic because transcription can start from multiple positions and in both directions (Carninci 2006). Zhao *et al.* (2006) considered that 54 overlapped tags obtained from spermatozoa SAGE from fertile men were novel genes and Peddinti *et al.* (2008) identified 10.6% and 9.8% of hypothetical proteins specific to high and low fertility bovine spermatozoa, respectively. Also the production of alternative transcripts may be beneficial to the spermatozoa. These variations could modify proteins acting in sperm specific functions or

replacing the protein within the spermatozoon (Eddy 2002). Some proteins or functional ncRNAs might have diverse roles in different cells, and their orthologs might have varied functions in different organisms. In such cases, additional functional analyses accompanied by comparative genomics and proteomics are needed to demonstrate their biological significance (Suzuki and Hayashizaki 2004). In the porcine spermatozoa cDNA library more than 259 ESTs were found corresponding to porcine (database) ESTs derived from cDNA libraries of other tissues but with no detectable similarity to human, mouse or bovine sequences (using gapped BLAST algorithm), and more than 400 ESTs were found in the GenBank database not at all. At least a part of these sequences may have some testis or sperm specific function, since approx. 70% of the cDNA clones showing very strong signals in the microarray hybridization corresponded to cDNAs with no homologies to known mRNAs.

Mitochondrial Genes. Mitochondrial integrity is related to fertility and mitochondrial transcripts are also related to spermatozoa fertility (Amaral *et al.* 2007, Lalancette *et al.* 2007). The functionality of the mitochondrial membrane is the most sensitive test to determine sperm quality (Marchetti *et al.* 2002). Human spermatozoa with high mitochondrial membrane potential have higher fertilization capacity because they have a high proportion of motile and morphologically normal sperm able to undergo the acrosome reaction (Gallon *et al.* 2006). As energy source and possible address for translation (Gur and Breitbart 2006), much attention must be paid to this organelle. Mitochondrial transcripts from components from the respiratory chain (cytochrome c oxidase, subunits 6 and 8 of adenosine triphosphatase, apocytochrome b and subunits of the NADH dehydrogenase) appeared in large scale not only in this porcine sperm cDNA library but also in bovine (Lalancette *et al.* 2007) and human libraries (Zhao *et al.* 2006). Some of these transcripts were also down-regulated in the spermatozoa of the boar with polyspermy and will be discussed below.

Transcripts related with apoptosis. Apoptosis occurs normally during spermatogenesis, playing a critical role in determining the size of the sperm output (Heninger *et al.* 2004, Franca *et al.* 2005). It is believed that in pigs up to 15% of germ cells are lost during spermiogenesis (França *et al.* 2005, Almeida *et al.* 2006). Spermatozoa with pathological morphologies have an elevation in transcripts related to the apoptotic pathway (Platts *et al.* 2007) and apoptosis might be the mechanism responsible for the formation of residual bodies of maturing spermatids (Blanco-Rodríguez *et al.* 2003). Some aspects of residual bodies have been related to a degeneration process, the high basophilia shown by residual bodies is similar to that of apoptotic germ cells and apoptotic factors are segregated to the

cytoplasm permitting the spermatozoon to utilize the apoptotic machinery without dying (Blanco-Rodríguez and Martínez-García 1999, Cagan 2003).

Almost 2% of the transcripts found in porcine spermatozoa were related to apoptosis regulation. Transcripts for positive and negative regulation of apoptosis were also found in human and bovine spermatozoa (Ostermeier *et al.* 2002 and Gilbert *et al.* 2007), moreover, transcripts from the apoptotic pathway are up-regulated in human spermatozoa with morphological anomalies (Platts *et al.* 2007).

Ribosomal Protein, Initiation and Elongation Factor Transcripts. Two important factors in the initiation of translation, EIF4E and the poly(A) binding protein, are grossly overexpressed in spermatogenic cells (Kleene 2003). Therefore, it is a striking observation that EIF4E is expressed at much higher levels in spermatogenic cells than in malignant cells (Miyagi *et al.* 1995). A significant fraction of EIF4E resides in the nucleus where it regulates the nuclear export of specific mRNAs (Culjkovic *et al.* 2007) and is a key component in the regulation of the translational efficiency of mRNAs (Miyagi *et al.* 1995).

A large number of cDNA clones of the normalized cDNA library corresponded to ribosomal proteins. More than 70 different cDNAs for ribosomal proteins were identified; the most frequent were the ribosomal proteins L10 (RPL10, 16 times) and P0 (RPLP0, 15 times). The RPLP0 together with proteins RPLP1 and RPLP2 form the stalk in eukaryotic ribosomes. The RPLP0 is important for the interaction of proteins P1/P2 with the ribosome, but it is not essential for protein synthesis (Santos and Ballesta 1995). Furthermore, 19 transcripts coding for initiation factors, including subunits were identified and six elongation factors. The translation elongation factor 1 alfa (EEF1A) was the most frequent in the elongations factor list (13 times). This factor has been described as up regulated during cell death (Lamberti *et al.* 2004). The apoptosis process is an essential part of spermatogenesis during the residual body formation (Blanco-Rodríguez and Martínez-García 1999).

EEF1A1, among others, is highly expressed at the 8-16 cell stage of bovine embryos, when compared to the blastocyst stage. EEF1A1 and RPL10 are between the 15 transcripts most frequent in the boar spermatozoa cDNA library and they are examples for mRNAs that could be transmitted by the male gamete influencing early embryogenesis (Goossens *et al.* 2007). However, the lower expression of genes involved in protein synthesis, like EEF1A1 and RPL10, in *in vitro* produced 8-cell embryos might indicate that the embryonic protein synthesis is affected or delayed in embryos produced *in vitro*, suggesting that these transcripts are produced during the maturation phase and not during fertilization (Goossens *et al.* 2007).

Heat Shock Protein Transcripts. Protective molecules, such as heat shock proteins (HSP), have been detected in porcine sperm cells and are strongly related to sperm fertility (Turba *et al.* 2007). Proteins of the heat shock protein 70 (*HSP70*) family are expressed in spermatogenic and somatic cells but the transcripts can differ in structure due to alternative promoters, alternative splicing, and upstream polyadenylation sites (Kleene 2005). Different experiments in different species indicate that the heat shock proteins and their phosphorylation are important steps in spermatozoa capacitation and sperm/embryo interaction (Ecroyd *et al.* 2003, Spinaci *et al.* 2005, Jones *et al.* 2007, Gau *et al.* 2008).

During capacitation of human, mouse and rat spermatozoa, HSP90 becomes tyrosine phosphorylated (Ecroyd *et al.* 2003). An equatorial subsegment of the sperm head in bovine, ovine and porcine spermatozoa, which is assembled during epididymal maturation, has an unusual concentration of HSP70 and phosphorylated proteins (Jones *et al.* 2007). The partial blockage of fertilization of both zona-intact and zona-free oocytes by anti-HSP70 antibodies suggested that this molecule plays an important role in porcine gamete interaction during the sperm-egg fusion (Spinaci *et al.* 2005). The *HSP* genes are up-regulated in porcine germ cells with heat stress (Gau *et al.* 2008). Thus, HPSs are potential markers for semen quality. Ten different transcripts for heat shock proteins, including the HSP90 and HSP70 family, were present in the normalized porcine cDNA library confirming their importance. The presence of these transcripts on the microarray will permit studies related to heat stress and sperm cell capacitation or adhesion, where HSP seems to be involved.

Spermadhesins. The major proteins of mammalian seminal plasma are members of a group of proteins named spermadhesins (Bergeron *et al.* 2005). Spermadhesins are expressed in the vesicular glands, epididymis and rete testis of the boar, stallion and bull. They have been found to be peripherally associated with the sperm surface (Strzezek *et al.* 2005). The spermadhesins adhere to sperm by binding to phosphatidyl ethanolamine present in the sperm membrane (Bergeron *et al.* 2005). Members of the spermadhesin family share 60%-98% sequence identity but they are not functionally equivalent. Spermadhesins can be divided into two groups, depending on their ability to either bind heparin (AQN-1, AQN-3, AWN-1, AWN-2, DQH) or not (PSPI/PSP-II heterodimer) (Caballero *et al.* 2004).

Generally, seminal plasma contains specific protein factors that influence both the fertilizing ability of spermatozoa and exert important effects on the female reproductive physiology. Seminal plasma proteins are involved e.g. in the control of molecular mechanisms accompanying sperm transport in the female reproductive tract, in the suppression of the immune response against sperm antigens and in the gamete interaction following fertilization (Strzezek *et al.* 2005). The interaction of seminal plasma proteins bound to the sperm

surface with polysaccharides or glycosaminoglycans of oviduct epithelial cells is the most probable event leading to sperm capacitation in the female reproductive tract (Jonáková *et al.* 2000).

Spermadhesin transcripts were also found in the normalized porcine sperm cDNA library. So far the expression of spermadhesins in spermatozoa was not described in the literature. They are normally produced by accessory gland cells and were not detected in spermatozoa. However, this could e an indication that spermadhesin transcripts are residual contaminations from testis somatic cells. Maybe the Triton-X treatment has not eliminated all somatic cells and remaining spermadhesin mRNAs were relatively enriched during normalization resulting in a similar frequency to spermatozoa mRNAs. Garcia *et al.* (2007) found PSP-I and II mRNA in testis and accessory glands. When these RNAs are considered as a contaminant from epithelial cells, they could be used as a control for contamination in spermatozoa RNA isolates, like leukocyte contamination can be detected using primers targeting the CD4 and CD45 antigens sequences (Gilbert *et al.* 2007) or protamine 2 to verify the presence of spermatozoa RNA (Miller *et al.* 1999).

5.3 Comparative Microarray Hybridization with Samples Derived from a Boar with Polyspermy and a Control Boar

The successfully sequenced cDNA clones of the normalized porcine sperm cDNA library were spotted onto two membranes to generate a cDNA microarray. RNA from two boars was extracted, reverse transcribed, labeled and hybridized with the microarray. The sperms of one boar showed a higher polyspermy rate of fertilized oocytes in the *in vitro* production of embryos and the second was used as control (low polyspermy rate).

The microarray hybridization results showed that the sperm cells from the boar with high polyspermy rate had a lower transcript levels of several mitochondrial genes (12S rRNA, ATPase 6, ATPase 8, COI, COII, COIII, CYTB, NADH1, NADH2, NADH4, NADH5 and NADH6). The spermatozoa from bulls with high fertility have significantly increased expression from ATPase und COIII protein genes (Peddinti *et al.* 2008) and human good-quality sperm had fewer copies of mitochondrial DNA were found but significantly higher expression of nuclear-encoded polymerase gamma (POLG), mitochondrial transcription factor A (TFAM) and COI (Amaral *et al.* 2007) has been shown. Lalancette *et al.* (2007) found an elevated number of 12S rRNA in low fertile bulls. The hybridization also showed

that the concentration of the 16S rRNA was higher in sperm cells of the boar with polyspermy in comparison with the control.

The reasons for polyspermy are not known, but can be related to sperm, e.g. partial acrosome reaction or failure in the sperm membrane composition (Abeydeera 2002, Yi *et al.* 2007). Almiñana *et al.* (2005) suggested that the qualification of boar spermatozoa for IVF varies among boars, different males do not respond equally to IVF conditions. Variation in the quality of oocytes (e.g. age) or *in vitro* fertilization conditions (sperm concentration or medium composition) could also result in polyspermy (Funahashi 2003). Unfortunately, important information about the boars for more detailed investigations, like sperm motility and vigor prior the cryconservation and *in vitro* production rates (e.g. polyspermy rates), were not available. Moreover, only one boar with polyspermy was investigated.

5.4 Molecular Biology Tools for the Investigation of Infertility

Bovine spermatozoa with high and low fertility were described to have 125 proteins with differential expression These proteins are potential biomarkers for bull fertility, also at the mRNA levels (Peddinti *et al.* 2008). Gur and Breitbart (2006) showed that the inhibition of protein translation by the mitochondrion in mammalian spermatozoa significantly reduced sperm motility, capacitation and in vitro fertilization rate. Thus, there are proteins important to male fertility and some of them seem to be translated after the loss of the residual body and synthesized within the mitochondrion. This finding is further supporting the importance of transcriptome analyses in context of sperm function. Molecular biology techniques can be helpful as complementary diagnosis and for the prognosis of the sperm fertilization capacity by the investigation from such kind of transcripts. Particularly powerful are techniques for the analysis of the transcriptome like Serial Analysis of Gene Expression (SAGE) or the microarray technology (van Ruissen *et al.* 2005). There is also a large commercially available porcine microarray from Affymetrix that contains 23,937 probe sets. However, in the case of the analysis of specialized cells, like sperm cells, the use of a custom microarray based in a normalized cDNA library from porcine spermatozoa has several advantages over the Affymetrix array. The sequence analysis of the normalized library revealed many cDNAs that were not found in the public databases, which is the basis for the design of the Affymetrix array. Furthermore, transcripts of spermatozoa often differ from transcripts from somatic cells in the use of alternative promoters, alternative splicing, and upstream polyadenylation sites (Willison and Ashworth 1987, Kleene 2005). More than 20% of the transcripts found in the normalized porcine spermatozoa cDNA library were unknown. Moreover, the unknown

sequences showed strong hybridization signals. This indicates that these transcripts are present at high levels in sperm cells but their function is still to be defined.

This is the first systematic investigation of isolation of RNA from porcine spermatozoa. The obtained results are in accordance with studies from human and bovine spermatozoa but showed that the amount of poly(A)$^+$ RNA is very low in porcine sperm cells. The characterization of porcine sperm cell RNA and the generated microarray based on the normalized cDNA library from porcine spermatozoa will be the basis for future studies comparing sperm cells from boars with high versus low fertility.

6. Summary

Qualitative and quantitative analysis of porcine sperm transcripts and characterization of a normalized cDNA library

During the development of the spermatozoon the nuclear protein histone is replaced by protamine, resulting in a higher condensation of the chromatin. This terminates at the late spermatogenesis the production of messenger RNA until the protamines are replaced by histones after fertilization. But (male) transcripts are needed for the fertilization and also in the late spermatogenesis and must be produced and stored in earlier stages. Furthermore, the translation stays under a stringent control to ensure appropriate temporal and spatial expression. However, the origin, the quantity and functions of RNA from sperm cells are not completely understood.

The aim of this work was to analyze the quantity of RNA in ejaculated boar sperms and furthermore the characterization of the RNA population to identify e.g. genes with different transcript abundance associated to different levels of male fertility. To characterize the RNA population in boar spermatozoa the ejaculates of five males were recovered and washed, and the total RNA was isolated using a modified TRIzol protocol. For the study of differentially regulated mRNA in boar ejaculate a cDNA library and a microarray were constructed. For the construction of the normalized cDNA library the total RNA was extracted and first strand cDNA synthesis was carried out with an oligo(dT) primer. The resulting cDNA was normalized and ligated into a plasmid vector, followed by electroporation into electro-competent cells. The bacterial suspension was transferred to agar plates and clones were randomly picked. After growth in the incubator, the bacteria were lysed to release the plasmid DNA. The cDNA fragments were amplified, sequenced and analyzed with the BLAST-Algorithm (National Centers for Biotechnology Information) for comparison with the sequences of public databases.

The total RNA extracted from boar spermatozoa is dominantly composed of small fragments between 20 to 700 bp. The sperm cells carry a high variety of transcripts that can be successfully extracted with standard protocols like TRIzol. These RNAs are represented in low copy numbers and mostly composed of low molecular fragments, whereas the fragments lower than 100 bp have no poly(A) tail. Four to ten ng total RNA was recovered from $1x10^6$ boar spermatozoa, representing only a few copies per transcript per spermatozoon. The cDNA library resulted in 1,892 unique transcripts covering functions from basic cell

metabolism and cell division to transcripts needed in the late spermatogenesis, capacitation and fertilization. However, only 15% were from known porcine genes and 404 sequences were not found in the NCBI GenBank databases.

A total of 3,009 cDNA clones were spotted onto two membranes to generate a cDNA microarray based on the porcine sperm normalized cDNA library. Microarray hybridization was carried out using ^{33}P- labeled cDNA probes obtained from a boar with a high polyspermy rate in the *in vitro* fertilization and a boar with a low rate. For every boar three replicate hybridizations were performed. The hybridization signals were analyzed with the AIDA software and the normalized data were analyzed with the significance analysis of microarrays method (SAM). Seventy percent of the spots with high hybridization signal were from unknown sequences, but with no difference between the two boars tested. According to SAM results there were 60 genes with high and 5 genes with low transcript abundance in the boar with normospermy when compared with the polyspermic boar, with a false discovery rate of approx. 5 %. The hybridization signals showed that the boar with polyspermy had lower hybridization signal from mitochondrial genes 12S rRNA, ATPase 6, ATPase 8, COI, COII, COIII, CYTB, NADH1, NADH2, NADH4, NADH5 and NADH6 and from a unknown EST (AK230819). These results showed that the microarray based in the normalized cDNA library from boar spermatozoa is sensitive enough for studies of infertility and fertility, respectively.

The RNA from boar spermatozoa can be extracted with simple protocols like TRIzol and it acts like a “fingerprint” representing the past, e.g. spermatogenesis, and the future, represented by stored transcripts related to capacitation and fertilization. The application of a cDNA library associated with the microarray technology in sperms expands the knowledge of the normal values of fertile male. As shown here it is helpful as complementary diagnosis of infertility or as tool for the investigation of new target molecules related to the prognosis of the fertilization capacity.

7. Zusammenfassung

Qualitative und quantitative Analyse porciner Spermientranskripte und Charakterisierung der normalisierten cDNA-Bank

Während der Spermienentwicklung werden die nuklearen Histonproteine durch Protamine ersetzt. Dies führt zu einer höheren Verdichtung des Chromatins. Dadurch wird auch die Transkription der Messenger-RNA in den späten Stadien der Spermatogenese solange inhibiert, bis nach der Verschmelzung mit der Eizelle diese Protamine wieder durch Histone ersetzt werden. Neuere Arbeiten zeigten aber, dass für die Befruchtung und in den späten Stadien der Spermatogenese (väterliche) Transkripte notwendig sind. Diese müssen deshalb in früheren Stadien der Spermatogenese produziert und auch gespeichert werden. Dies bedeutet aber auch, dass die Translation unter stringenter Kontrolle stehen muss, um eine geeignete zeitliche und räumliche Expression zu gewährleisten. Bis heute sind die Herkunft, die Anzahl und die Funktionen der Spermien-RNA noch nicht vollständig bekannt.

Das Ziel dieser Arbeit war die Analyse der Transkript-Menge in ejakulierten Eberspermien und darüber hinaus die Charakterisierung dieser RNA-Population, um z.B. unterschiedlich hoch regulierte Gene zu identifizieren, die mit unterschiedlichen Fertilitätsgraden assoziiert werden können. Für die Charakterisierung der Spermien RNA-Population wurden Ejakulate von fünf Eber gesammelt, gewaschen und die Gesamt-RNA mit einem modifierten TRIzol-Protokoll isoliert. Für die Untersuchung der differenziell exprimierten Transkripte wurden eine cDNA-Bank und ein Microarray entwickelt. Für die Herstellung der normalisierten cDNA-Bank wurde die Gesamt-RNA gewonnen und die Erststrang-Synthese mit einem Oligo (dT) Primer durchgeführt. Diese cDNA wurde normalisiert, in Plasmid-Vektoren ligiert und diese Plasmide wurden in kompetente Bakterien elektroporiert. Diese Bakteriensuspension wurde auf Agar-Platten transferiert und die Kolonien (Klone) wurden nach Inkubation zufällig gepickt und in einem Flüssigmedium weiterkultiviert. Nach Kultivierung und Isolierung der Plasmide wurden die cDNA-Fragmente amplifiziert, sequenziert und mit Sequenzen aus öffentlich zugänglichen Datenbanken verglichen.

Die isolierte Gesamt-RNA aus Eberspermien wird von kleinen Fragmenten zwischen 20 bis 700 bp Länge dominiert. Insgesamt existiert mRNA in Spermien in nur geringer Menge und besteht vor allem aus niedermolekularen Fragmenten, wobei die Fragmente kürzer als 100 bp keinen Poly(A) Schwanz haben. Vier bis zehn ng Gesamt-RNA wurden aus $1x10^6$ Eberspermien extrahiert, so dass sich rechnerisch nur wenige Transkripte eines Genes pro

Spermium ergeben. Die Untersuchung der cDNA-Bank führte zu 1.892 verschiedenen Transkripten, die u.a. grundlegende Funktionen des Zellstoffwechsels, der Zellteilung, der Kapazitation und auch der Befruchtung umfassen. Allerdings wurden nur 15% der Transkripte in den öffentlich zugänglichen Schweine-Datenbanken gefunden und 404 Sequenzen konnten auch in anderen Spezies-Datenbanken nicht identifiziert werden.

Insgesamt wurden zur Generierung eines cDNA-Microarray 3.009 cDNA-Klone auf zwei Membranen gespottet, der auf den Transkripten der normalisierten cDNA-Bank basierte. Die Microarray-Hybridisierung wurde mit ^{33}P-markierten cDNA-Proben von einem Eber mit hoher Polyspermierate in der *in vitro* Befruchtung und einem Eber mit geringer Polyspermierate durchgeführt. Für jeden Eber wurden drei Hybridisierungsreplikate durchgeführt. Die Hybridisierungssignale wurden mit der AIDA-Software und die normalisierten Daten mit der *significance analysis of microarrays method* (SAM) analysiert. Siebzig Prozent der Spots mit hohem Hybridisierungssignal stammten von unbekannten Sequenzen, die aber zwischen den beiden Ebern nicht unterschiedlich waren. Nach der SAM Analyse ergaben sich beim Kontroll-Eber im Vergleich mit dem Eber mit hoher Polyspermierate 60 Transkripte mit höherer und 5 Transkripte mit niedrigerer Abundanz (*false discovery rate* von ca. 5%). Die Hybridisierungsignale zeigten zudem, dass der Eber mit hoher Polyspermierate niedrigere Expressionen der mitochondrialen Gene 12S rRNA, ATPase 6, ATPase 8, COI, COII, COIII, CYTB, NADH1, NADH2, NADH4, NADH5, NADH6 und von einem unbekannten EST (AK230819) aufwies. Diese Ergebnisse zeigen, dass der Microarray, basierend auf einer normalisierten cDNA-Bank, sensitiv genug ist, um Studien zur Frucht- bzw. Unfruchtbarkeit anhand von Eberspermien durchzuführen.

Die Gesamt-RNA kann mit einfachen RNA-Isolierungsprotokollen, wie dem TRIzol-Protokoll, isoliert werden. Diese RNA stellt einen „Fingerprint“ der Vergangenheit (Spermatogenese) und in Form der gespeicherten Transkripte einen Abdruck der Zukunft (Kapazitation, Fertilisierungt) dar. Die Anwendung einer cDNA-Bank/Microarray zur Untersuchung von Spermien kann somit eine Erweiterung der klassischen Fruchtbarkeitsparameter bedeuten und es ist zudem denkbar, dass diese Technik als ergänzende Diagnostik der Fruchtbarkeit von Prüftieren zukünftig angewendet werden kann.

8. References

ABEYDEERA L.R. (2002). In vitro production of embryos in swine. Theriogenology, 57:257-273.

ALCIVAR A.A., HAKE L.E., MILLETTE C.F., TRASLER J.M. AND HECHT N.B. (1989). Mitochondrial Gene Expression in Male Germ Cells of the Mouse. Developmental Biology, 135 263-271.

ALMEIDA F.F.L., LEAL M.C., AND FRANÇA L.R. (2006). Testis Morphometry, Duration of Spermatogenesis, and Spermatogenic Efficiency in the Wild Boar (Sus scrofa scrofa). Biology of Reproduction, 75 792–799.

ALMIÑANA C., GIL M.A., CUELLO C., ROCA J., VAZQUEZ J.M., RODRIGUEZ-MARTINEZ H. AND MARTINEZ E.A. (2005). Adjustments in IVF system for individual boars: Value of additives and time of sperm–oocyte co-incubation. Theriogenology, 64 1783–1796

AMANAI M., BRAHMAJOSYULA M. AND PERRY A.C.F. (2006). A Restricted Role for Sperm-Borne MicroRNAs in Mammalian Fertilization. Biology of Reproduction, 75 877–884.

AMARAL A., RAMALHO-SANTOS J. AND JOHN J.C.ST. (2007). The expression of polymerase gamma and mitochondrial transcription factor A and the regulation of mitochondrial DNA content in mature human sperm. Human Reproduction, 22(6) 1585–1596.

BALDI E., LUCONI M., BONACCORSI L., MURATORI M. AND FORTI G. (2000). Intracellular events and signaling pathways involved in sperm acquisition of fertilizing capacity and acrosome reaction. Frontiers in Bioscience, 5 110-123.

BERGERON A., VILLEMURE M., LAZURE C. AND MANJUNATH P. (2005). Isolation and Characterization of the Major Proteins of Ram Seminal Plasma. Molecular Reproduction and Development 71 461–470.

BLANCO-RODRIGUEZ J. AND MARTINEZ-GARCIA C. (1999). Apoptosis Is Physiologically Restricted to a Specialized Cytoplasmic Compartment in Rat Spermatids. Biology of Reproduction, 61 1541–1547.

BLANCO-RODRÍGUEZ J., MARTÍNEZ-GARCÍA C. AND PORRAS A. (2003). Correlation between DNA synthesis in the second, third and fourth generations of spermatogonia and the occurrence of apoptosis in both spermatogonia and spermatocytes. Reproduction, 126, 661–668.

BRAUN R.E., BEHRINGER R.R., PESCHON J.J., BRINSTER R.L. AND PALMITER R.D. (1989). Genetically haploid spermatids are phenotypically diploid. Nature, 337 373-376.

BRENNECKE J., ARAVIN A.A., STARK A., DUS M., KELLIS M., SACHIDANANDAM R. AND HANNON G.J. (2007). Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila. Cell, 128 1089–1103.

BREUCKER H., SCHÄFER E. AND HOLSTEIN A.F. (1985). Morphogenesis and fate of the residual body in human spermiogenesis. Cell and Tissue Research, 240 303-309.

CABALLERO I., VAZQUEZ J.M., GIL M.A., CALVETE J.J., ROCA J., SANZ L., PARRILLA I., GARCIA E.M., RODRIGUEZ-MARTINEZ H. AND MARTINEZ E.A. (2004). Does Seminal Plasma PSP-I/PSP-II Spermadhesin Modulate the Ability of Boar Spermatozoa to Penetrate Homologous Oocytes In Vitro? Journal of Andrology, 25(6) 1004-12.

CAGAN R.L. (2003). Spermatogenesis: Borrowing the Apoptotic Machinery. Current Biology, 13 600–602.

CARNINCI P. (2006). Tagging mammalian transcription Complexity. TRENDS in Genetics, 22(9) 501-510.

CARREAU S., LAMBARD S., SAID L., SAAD A. AND GALERAUD-DENIS I. (2007). RNA dynamics of fertile and infertile spermatozoa. Biochemical Society Transactions, 35(3) 634-636.

CHANG K., QIAN J., JIANG M., LIU Y.-H., WU M.-C., CHEN C.-D., LAI C.-K., LO H.-L., HSIAO C.-T., BROWN L., BOLEN JR. J., HUANG H.-I, HO P.-Y., SHIH P.Y., YAO C.-W., LIN W.-J., CHEN C.-H., WU F.-Y., LIN Y.-J., XU J. AND WANG K. (2002). Effective generation of transgenic pigs and mice by linker based sperm-mediated gene transfer. Biomed central Biotechnology, 2(5) 1-13.

CHENNATHUKUZHI V., MORALES C.R., EL-ALFY M. AND HECHT N.B. (2003). The kinesin KIF17b and RNA-binding protein TB-RBP transport specific cAMP-responsive element modulator-regulated mRNAs in male germ cells. PNAS, 100(26) 15566-15571.

Chomczynski p. And sacchi n. (1987). Single-step of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Analytical Biochemistry, 162 156-159.

COX D.N., CHAO A., BAKER J., CHANG L., QIAO D. AND LIN H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes & Development, 12 3715-3727.

CULJKOVIC B., TOPISIROVIC I. AND BORDEN K.L. (2007). Controlling gene expression through RNA regulons: the role of the eukaryotic translation initiation factor eIF4E. Cell Cycle, 6(1) 65-69.

DADOUNE J.P. (2003). Expression of mammalian spermatozoal nucleoproteins. Microscopy Research and Technique, 61 56–75.

DADOUNE J.P., PAWLAK A., ALFONSI M.F. AND SIFFROI J.P. (2005). Identification of transcripts by macroarrays, RT–PCR and in situ hybridization in human ejaculate spermatozoa. Molecular Human Reproduction, 11 133–140.

DENG W. AND LIN H. (2002). miwi, a Murine Homolog of piwi, Encodes a Cytoplasmic Protein Essential for Spermatogenesis. Developmental Cell, 2 819–830.

DORUS S., BUSBY S.A., GERIKE U., SHABANOWITZ J., HUNT D.F. AND KARR T.L. (2006). Genomic and functional evolution of the Drosophila melanogaster sperm proteome. Nature Genetics, 23 1440-1445.

ECROYD H., JONES R.C. AND AITKEN R.J. (2003). Tyrosine Phosphorylation of HSP90 During Mammalian Sperm Capacitation. Biology of Reproduction, 69 1801–1807.

EDDY E.M. (1999). Role of heat shock protein HSP70-2 in spermatogenesis. Reviews of Reproduction, 4 23–30.

EDDY E.M. (2002). Male Germ Cell Gene Expression. Recent Progress in Hormone Research, 57 103-28.

EDDY E.M., TOSHIMORI K. AND O'BRIEN D.A. (2003). Fibrous Sheath of Mammalian Spermatozoa. Microscopy Research and Technique, 61 103–115.

FATEHI A.N., BEVERS M.M., SCHOEVERS E., ROELEN B.A.J., COLENBRANDER B. AND GADELLA B.M. (2006). DNA Damage in Bovine Sperm Does Not Block Fertilization and Early Embryonic Development But Induces Apoptosis After the First Cleavages Journal of Andrology, 27 176-188.

FRANÇA L. R., AVELAR G. F., ALMEIDA F. F. L. (2005). Spermatogenesis and sperm transit through the epididymis in mammals with emphasis on pigs. Theriogenology, 63 300–318.

FUNAHASHI H. (2003). Polyspermic penetration in porcine IMV-IMF systems. Reproduction, Fertility and Development, 15 167-177

GADEA J. (2005). Sperm factors related to in vitro and in vivo porcine fertility. Theriogenology, 63 431–444.

GAILLARD C. AND STRAUSS F. (1990). Ethanol precipitation of DNA with linear polyacrylamide as carrier. *Nucleic Acids Research,* 18 378.

GALLON F., MARCHETTI C., JOUY N. AND MARCHETTI P. (2006). The functionality of mitochondria differentiates human spermatozoa with high and low fertilizing capability. Fertility and Sterility, 86 1526-1530.

GARCÍA E.M., VÁZQUEZ J.M., PARRILLA I., ORTEGA M.D., CALVETE J.J., SANZ L., MARTÍNEZ E.A., ROCA J. AND RODRÍGUEZ-MARTÍNEZ H. (2007). Localization and expression of spermadhesin PSP-I/PSP-II subunits in the reproductive organs of the boar. International Journal of Andrology, 1-10.

GAU B.-H., CHU I.-M., HUANG M.-C., YANG K.-T., CHIOU S.-H., FAN Y.-H., CHEN M.-Y., LIN J.-H., CHUANG C.-K., HUANG S.-Y. AND LEE W.-C. (2008). Transcripts of enriched germ cells responding to heat shock as potential markers for porcine semen quality. Theriogenology, in press.

GILBERT I., BISSONNETTE N., BOISSONNEAULT G., VALLEE M. AND ROBERT C. (2007). A molecular analysis of the population of mRNA in bovine spermatozoa. Reproduction, 133 1073–1086.

GIORDANO R., MAGNANO A.R., ZACCAGNINI G., PITTOGGI C., MOSCUFO N., LORENZINI R. AND SPADAFORA C. (2000). Reverse Transcriptase Activity in Mature Spermatozoa of Mouse. The Journal of Cell Biology, 148(6) 1107–1113.

GIORGINI F., DAVIES H.G. AND BRAUN R.E. (2001). MSY2 and MSY4 bind a conserved sequence in the 3' untranslated region of Protamine 1 mRNA in vitro and in vivo. Molecular and Cellular Biology, 21 7010–7019.

GIRARD A., SACHIDANANDAM R., HANNON G.J. AND CARMELL M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature, 442(13) 199-202.

GOOSSENS K., VAN SOOM A., VAN POUCKE M., VANDAELE L., VANDESOMPELE J., VAN ZEVEREN A. AND PEELMAN L.J. (2007). Identification and expression analysis of genes associated with bovine blastocyst formation. BioMed Central Developmental Biology, 7:64 1-12.

GREENBAUM M.P., YAN W., WU M.-H., LIN Y.-N., AGNO J.E., SHARMA M., BRAUN R.E., RAJKOVIC A. AND MATZUK M.M. (2006). TEX14 is essential for intercellular bridges and fertility in male mice. PNAS, 103 4982–4987.

GREENBAUM M.P., MA L. AND MATZUK M.M. (2007). Conversion of midbodies into germ cell intercellular bridges. Developmental Biology, 305 389–396.

GRIVNA S.T., BEYRET E., WANG Z. AND LIN H. (2006a). A novel class of small RNAs in mouse spermatogenic cells. Genes and Development, 20 1709-1714.

GRIVNA S.T., PYHTILA B. AND LIN H. (2006b). MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. PNAS, 103(36) 13415-13420.

GRUNEWALD S., PAASCH U., GLANDER H.-J. AND ANDEREGG U. (2005). Mature human spermatozoa do not transcribe novel RNA. Blackwell Publishing Ltd Andrologia, 37 69–71.

GU W., WU X.-Q., MENG X.-H., MORALES C., EL-ALFY M. AND HECHT N.B. (1998). The RNA- and DNA-Binding Protein TB-RBP Is Spatially and Developmentally Regulated During Spermatogenesis. Molecular Reproduction and Development 49 219–228.

GUR Y. AND BREITBART H. (2006). Mammalian sperm translate nuclear-encoded proteins by mitochondrial-type ribosomes. Genes & Development 20 411–416.

HANNON G.J. (2002). RNA interference. Nature, 418 244-251.

HAWTHORNE S.K., GOODARZI G., BAGAROVA J., GALLANT K.E., BUSANELLI R.R., OLEND W. J. AND KLEENE K.C. (2006). Comparative genomics of the sperm mitochondria-associated cysteine-rich protein gene. Genomics, 87 382–391.

HAYASHI S., YANG J., CHRISTENSON L., YANAGIMACHI R. AND HECHT N.B. (2003). Mouse preimplantation embryos developed from oocytes injected with round spermatids or

spermatozoa have similar but distinct patterns of early messenger RNA expression. Biology of Reproduction, 69 1170–1176.

HECHT N.B. (1998). Molecular mechanisms of male germ cell differentiation. BioEssays, 20 555–561.

HENDRIKSEN P.J.M. (1999). Do X and Y spermatozoa differ in proteins? Theriogenology, 52 1295-1307.

HENINGER N. L., STAUB C., BLANCHARD T. L., JOHNSON L., VARNER D. D. AND FORREST D. W. (2004). Germ cell apoptosis in the testes of normal stallions. Theriogenology, 62 283-297.

HINNEBUSCH A.G. (2006). eIF3: a versatile scaffold for translation initiation complexes. Trends in Biochemical Science, 31(10) 553-62.

IGUCHI N., XU M., HORI T., AND HECHT N.B. (2008). Non-coding RNAs of the mammalian testis: the meiotic transcripts Nct1 and 2 encode piRNAs. Annals of the New York Academy of Sciences, 1120 84-94.

JEDRZEJCZAK P., JANUCHOWSKI R., TASZAREK-HAUKE G., LADDACH R., PAWELCZYK L., JAGODZINSKI P.P. (2006). Quantitative analysis of CCR5 chemokine receptor and cytochrome P450 aromatase transcripts in swim-up spermatozoa isolated from fertile and infertile men. Archives of Andrology, 52 335-41.

JEDRZEJCZAK P., KEMPISTY B., BRYJA A., MOSTOWSKA M., DEPA-MARTYNOW M., PAWELCZYK L. AND JAGODZINSKI P.P. (2007). Quantitative Assessment of Transition Proteins 1, 2 Spermatid-Specific Linker Histone H1-Like Protein Transcripts in Spermatozoa from Normozoospermic and Asthenozoospermic Men. Archives of Andrology, 53 199-205.

JONÁKOVÁ V., MANÁSKOVÁ P., KRAUS M., LIBERDA J. AND TICHÁ M. (2000). Sperm Surface Proteins in Mammalian Fertilization Molecular Reproduction and Development, 56 275–277.

JONES R., JAMES P.S., HOWES L., BRUCKBAUER A. AND KLENERMAN D. (2007). Supramolecular organization of the sperm plasma membrane during maturation and capacitation. Asian Journal of Andrology, 9(4) 438–444.

KIERSZENBAUM A.L. AND TRES L.L. (1975). Structural and transcriptional features of the mouse spermatid genome. Journal of Cell Biology, 65(2) 258-70.

KLEENE K.C. (1989). Poly(A) shortening accompanies the activation of translation of five mRNAs during spermiogenesis in the mouse. Development, 106 367-373.

KLEENE K.C. (2001). A possible meiotic function of the peculiar patterns of gene expression in mammalian spermatogenic cells. Mechanisms of Development, 106 3-23.

KLEENE K.C. (2003). Patterns, mechanisms and functions of translation regulation in mammalian spermatogenic cells. Cytogenetic and Genome Research, 103 217–224.

KLEENE K.C. (2005). Sexual selection, genetic conflict, selfish genes, and the atypical patterns of gene expression in spermatogenic cells. Developmental Biology, 277 16–26.

KOTAJA N., BHATTACHARYYA S.N., JASKIEWICZ L., KIMMINS S., PARVINEN M., FILIPOWICZ W. AND SASSONE-CORSI P. (2006). The chromatoid body of male germ cells: Similarity with processing bodies and presence of Dicer and microRNA pathway components. PNAS, 103 2647–2652.

KOTAJA N. AND SASSONE-CORSI P. (2007). The chromatoid body: a germ-cellspecific RNA-processing centre. Nature Reviews, Molecular Cell Biology, 8 85-90.

KRAWETZ S.A. (2005). Paternal contribution: new insights and future challenges. Nature reviews Genetics, 8 633-642.

KUNAVONGKRIT A., SURIYASOMBOON A., LUNDEHEIM N., HEARD T. W. AND EINARSSON S. (2005). Management and sperm production of boars under differing environmental conditions. Theriogenology, 63 657–667.

KURAMOCHI-MIYAGAWA S., KIMURA T., IJIRI T.W., ISOBE T., ASADA N., FUJITA Y. IKAWA M., IWAI N., OKABE M., DENG W., LIN H., MATSUDA Y. AND NAKANO T. (2004). Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development, 131(4) 839-849.

LALANCETTE C., THIBAULT C., BACHAND I., CARON N. AND BISSONNETTE N. (2008). Transcriptome Analysis of Bull Semen with Extreme Nonreturn Rate: Use of Suppression-Subtractive Hybridization to Identify Functional Markers for Fertility. Biology of Reproduction, 78 618-635.

LAMBARD S., GALERAUD-DENIS I., MARTIN G., LEVY R., CHOCAT A. AND CARREAU S. (2004). Analysis and significance of mRNA in human ejaculated sperm from

normozoospermic donors: relationship to sperm motility and capacitation. Molecular Human Reproduction, 10 535–541.

LAMBERTI A., CARAGLIA M., LONGO O., MARRA M., ABBRUZZESE A. AND ARCARI P. (2004). The translation elongation factor 1A in tumorigenesis, signal transduction and apoptosis: review article. Amino Acids, 26(4) 443-448.

LAU N.C., SETO A.G., KIM J., KURAMOCHI-MIYAGAWA S., NAKANO T., BARTEL D.P. AND KINGSTON R.E. (2006). Characterization of the piRNA Complex from Rat Testes. Science, 313 363-367.

LEE J.-W. AND YANG X. (2004). Factors Affecting Fertilization of Porcine Oocytes Following Intracytoplasmic Injection of Sperm. Molecular Reproduction and Development, 68 96–102.

LEE K., HAUGEN H.S., CLEGG C.H. AND BRAUN R.E. (1995). Premature translation of protamine 1 mRNA causes precocious nuclear condensation and arrests spermatid differentiation in mice. Proceedings of National Academy of Sciences of the USA, 92(26) 12451-12455.

LEE K., FAJARDO M.A., AND BRAUN R.E. (1996). A Testis Cytoplasmic RNA-Binding Protein That Has the Properties of a Translational Repressor. Molecular and Cellular Biology, 16(6) 3023–3034.

MACH H., MIDDAUGH C.R., AND LEWIS R.V. (1992). Detection of proteins and phenol in DNA samples with second-derivative absorption spectroscopy. Analytical Biochemistry, 200 20-26.

MACRAE I.J., ZHOU K. AND DOUDNA J.A. (2007). Structural determinants of RNA recognition and cleavage by Dicer. Nature Structural and Molecular Biology, 14(10) 934-940.

MARCHETTI C., OBERT G., DEFFOSEZ A., FORMSTECHER P. AND MARCHETTI P. (2002). Study of mitochondrial membrane potential, reactive oxygen species, DNA fragmentation and cell viability by flow cytometry in human sperm. Human reproduction, 17 1257-1265.

MATTICK J.S. (2003). Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. BioEssays, 25 930–939.

MATTICK J.S. AND MAKUNIN I.V. (2006). Non-coding RNA. Human Molecular Genetics, 15 17–29.

MILLER D., TANG P.-Z., SKINNER C. AND LILFORD R. (1994). Differential RNA fingerprinting as a tool in the analysis of spermatozoal gene expression. Human Reproduction, 9(5) 864-869.

MILLER D., BRIGGS D., SNOWDEN H., HAMLINGTON J., ROLLINSON S., LILFORD R. AND KRAWETZ S. A. (1999). A complex population of RNAs exists in human ejaculate spermatozoa: implications for understanding molecular aspects of spermiogenesis. Gene, 237 385–392.

MILLER D. (2000). Analysis and Significance of Messenger RNA in Human Ejaculated Spermatozoa. Molecular Reproduction and Development, 56 259–264.

MILLER D., OSTERMEIER C.G. AND KRAWETZ S.A. (2005). The controversy, potential and roles of spermatozoal RNA. Trends in Molecular Medicine, 11 156-163.

MILLER D. and ostermeier C.G. (2006). Towards a better understanding of RNA carriage by ejaculate spermatozoa. Human Reproduction Update, 12(6) 757-767.

MIYAGI Y., KERR S., SUGIYAMA A., ASAI A., SHIBUYA M., FUJIMOTO H. AND KUCHINO Y. (1995). Abundant expression of translation initiation factor EIF-4E in post-meiotic germ cells of the rat testis. Laboratory Investigation, 73(6) 890-898.

MEISTRICH M.L., MOHAPATRA B., SHIRLEY C.R., ZHAO M. (2003). Roles of transition nuclear proteins in spermiogenesis. Chromosoma, 111 483–488.

MORALES C.R., WU X.Q., AND HECHT N.B. (1998). The DNA/RNA-Binding Protein, TB-RBP, Moves from the Nucleus to the Cytoplasm and through Intercellular Bridges in Male Germ Cells. Developmental Biology, 201 113–123.

NICHOLSON R.H., NICHOLSON A.W. (2002). Molecular characterization of a mouse cDNA encoding Dicer, a ribonuclease III ortholog involved in RNA interference. Mammalian Genome, 13(2) 67-73.

OKAZAKI Y., FURUNO M., KASUKAWA T., ET AL. (2002). Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature, 420 563-573.

OLIVA R. (2006). Protamines and male infertility. Human Reproduction Update, 12 417–435.

OSTERMEIER C.G., DIX D.J., MILLER D., KHATRI P., KRAWETZ S. A. (2002). Spermatozoal RNA profiles of normal fertile men. Mechanisms of disease, 360 772-777.

OSTERMEIER G.C., MILLER D., HUNTRISS J.D., DIAMOND M.P., KRAWETZ S.A. (2004). Delivering spermatozoan RNA to the oocyte Nature, v.429.

OSTERMEIER G. C., GOODRICH R. J., MOLDENHAUER J. S., DIAMOND M. P. AND KRAWETZ S. A. (2005). A Suite of Novel Human Spermatozoal RNAs. Journal of Andrology, 26 70-74.

PARVINEN M. (2005). The chromatoid body in spermatogenesis. International journal of andrology, 28 189–201.

PEDDINTI D., NANDURI B., KAYA A., FEUGANG J.M., BURGESS S.C. AND MEMILI E. (2008). Comprehensive proteomic analysis of bovine spermatozoa of varying fertility rates and identification of biomarkers associated with fertility. *BMC Systems Biology,* 2 19 in press.

PENTTILIÄ T.-L., YUAN L., MALI P., HÖÖG C., AND PARVINEN M. (1995). Haploid Gene Expression: Temporal Onset and Storage Patterns of 13 Novel Transcripts during Rat and Mouse Spermiogenesis. Biology of Reproduction, 53 499-510.

PETERSEN C., FÜZESI L. and hoyer-fender s. (1999). Outer dense fibre proteins from human sperm tail: molecular cloning and expression analyses of two cDNA transcripts encoding proteins of approximately 70 kDa. Molecular Human Reproduction, 5 (7) 627-635.

PITTOGGI C., BERALDI R., SCIAMANNA I., BARBERI L., GIORDANO R., MAGNANO A.R., TOROSANTUCCI L., PESCARMONA E. AND SPADAFORA C. (2006). Generation of Biologically Active Retro-genes Upon Interaction of Mouse Spermatozoa With Exogenous DNA. Molecular Reproduction and Development, 73 1239–1246.

PLATTS A.E., DIX D.J., CHEMES H.E., THOMPSON K.E., GOODRICH R., ROCKETT J.C., RAWE V.Y., QUINTANA S., DIAMOND M.P., STRADER L.F. AND KRAWETZ S.A. (2007). Success and failure in human spermatogenesis as revealed by teratozoospermic RNAs. Human Molecular Genetics, 16 763–773.

PRATHAPKUMAR S. AND SURESHKUMAR B.V. (2002). Rapid and Simple Approach to Determine Nucleic Acid Purity by Ratio method (A260/280) using Bio-Spectrophotometer. A.P.I.E., Sanathnagar,Hyderabad, India.

PREMKUMAR E., BHARGAVA P.M. (1972). Transcription and translation in bovine spermatozoa. Nature: New Biology, 240 139-43.

RASSOULZADEGAN M., GRANDJEAN V., GOUNON P., VICENT S., GILLOT I. AND CUZIN F. (2006). Nature, 431, 469-474.

RODRIGUES-MARTINEZ H. (2003). Laboratory Semen Assessment and Prediction of Fertility: still Utopia? Reproduction in Domestic Animals, 38 312–318.

SANTOS C. AND BALLESTA J.P. (1995). The highly conserved protein P0 carboxyl end is essential for ribosome activity only in the absence of proteins P1 and P2. Journal of Biological Chemistry, 1;270(35):20 608-614.

SAUNDERS P.T.K., MILLAR M.R., WEST AP. AND SHARPE R.M. (1993). Mitochondrial Cytochrome C Oxidase II Messenger Ribonucleic Acid Is Expressed in Pachytene Spermatocytes at High Levels and in a Stage-Dependent Manner during Spermatogenesis in the Rat. Biology of Reproduction, 48 57-67.

SAVCHUK L. (2006). Versuche zur Etablierung eines Verfahrens zur Isolierung von RNA aus Eberspermien. *Doktorarbeit Dissertation,* aus dem Institut für Tierzucht der Tierärztlichen Fakultät der Ludwig-Maximilians-Universität München.

SETO A.G., KINGSTON R.E. AND LAU N.C. (2007). The Coming of Age for Piwi Proteins. Molecular Cell, 26 603-609.

SHIH D.M. AND KLEENE K.C. (1992). A Study by In Situ Hybridization of the Stage of Appearance and Disappearance of the Transition Protein 2 and the Mitochondrial Capsule Seleno-Protein mRNAs During Spermatogenesis in the Mouse. Molecular Reproduction and Development, 33 222-227.

SMITH K. AND SPADAFORA C. (2005). Sperm-mediated gene transfer: applications and implications. BioEssays, 27 551–562.

SÖDERSTRÖM, K.O. AND PARVINEN, M. (1976). RNA synthesis in different stages of rat seminiferous epithelial cycle. Molecular and Cellular Endocrinology, 5 181–199.

SPINACI M., VOLPE S., BERNARDINI C., DE AMBROGI M., TAMANINI C., SEREN E. AND GALEATI G. (2005). Immunolocalization of Heat Shock Protein 70 (Hsp 70) in Boar Spermatozoa and Its Role During Fertilization. Molecular Reproduction and Development, 72 534–541.

SPIRIN A.S. (2004). The ribosome as an RNA-based molecular machine. RNA Biology, 1(1) 3-9.

STEGER K. (2001). Haploid spermatids exhibit translationally repressed mRNAs. Anatomy and Embryology 203 323–334.

STRZEZEK J, WYSOCKI P, KORDAN W, KUKLIŃSKA M, MOGIELNICKA M., SOLIWODA D. AND FRASER L. (2005). Proteomics of boar seminal plasma – current studies and possibility of their application in biotechnology of animal reproduction. Reproductive Biology, 5(3):279-290.

SUZUKI M. AND HAYASHIZAKI Y. (2004). Mouse-centric comparative transcriptomics of protein coding and non-coding RNAs. BioEssays, 26 833–843.

TAHBAZ N., KOLB F.A., ZHANG H., JARONCZYK K., FILIPOWICZ W. AND HOBMAN T.C. (2004). Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer. European Molecular Biology Organization-reports, 5(2) 189–194.

TURBA M.E., FANTINATI P., BERNARDINI C., GENTILINI F., BACCI M.L. AND FORNI M. (2007). Relationships between innovative and traditional parameters to investigate semen quality in pigs. Animal Reproduction Science, 99 72–81.

TURNER R.M. (2006). Moving to the beat: a review of mammalian sperm motility regulation. Reproduction, Fertility and Development, 18 25-38.

VAN RUISSEN F., RUIJTER J.M., SCHAAF G.J., ASGHARNEGAD L., ZWIJNENBURG D.A., KOOL M. AND BAAS F. (2005). Evaluation of the similarity of gene expression data estimated with SAGE and Affymetrix GeneChips. BMC Genomics, 6(91).

VILLEGAS J., ZÁRRAGA A.M., MULLER I., MONTECINOS L., WERNER E., BRITO M., MENESES A.M. and BURZIO L.O. (2000). A Novel Chimeric Mitochondrial RNA Localized in the Nucleus of Mouse Sperm. DNA and Cell Biology, 19(9), 579–588.

WANG H., ZHOU Z., XU M., LI J., XIAO J., XU Z-Y. AND SHA J. (2004). A spermatogenesis-related gene expression profile in human spermatozoa and its potential clinical applications. *Journal of Molecular Medicine,* 82 317-324.

WARD W.S. (1993). Deoxyribonucleic Acid Loop Domain Tertiary Structure in Mammalian Spermatozoa. Biology of Reproduction, 48 1193-1201.

WATANABE T., TAKEDA A., TSUKIYAMA T., MISE K., OKUNO T., SASAKI H., MINAMI N. AND IMAI H. (2006). Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes & Development, 20 1732-1743.

WILLISON K. AND ASHWORTH A. (1987). Mammalian spermatogenic gene expression. Trends in Genetics, 3 351-355.

WOUTERS-TYROU D., MARTINAGE A., CHEVAILLIER P. AND SAUTIÈRE P. (1998). Nuclear basic proteins in spermiogenesis. Biochimie, 80 117-128.

YAMAGUCHI M., DIEFFENBACH C.W., CONNOLLY R., CRUESS D.F., BAUR W., AND SHAREFKIN J.B. (1992). Effect of different laboratory techniques for guanidinium- phenol-chloroform RNA extraction on A260/A280 and on accuracy of mRNA quantitation by reverse transcriptase-PCR. PCR Methods and Applications, 1 286-290.

YANG J., MEDVEDEV S., REDDI P.P., SCHULTZ R.M. AND HECHT N.B. (2005a). The DNA/RNA-binding protein MSY2 marks specific transcripts for cytoplasmic storage in mouse male germ cells. PNAS, 102 1513–1518.

YANG J., MEDVEDEV S., YU J., TANG L.C., AGNO J.E., MATZUK M.M., SCHULTZ R.M. AND HECHT N.B. (2005b). Absence of the DNA/RNA-binding protein MSY2 results in male and female infertility. PNAS, 102 5755–5760.

YANG J., MEDVEDEV S., YU J., SCHULTZ R.M. AND HECHT N.B. (2006). Deletion of the DNA/RNA-binding protein MSY2 leads to post-meiotic arrest. Molecular and Cellular Endocrinology, 250 20–24.

YI Y.-J., MANANDHAR G., SUTOVSKY M., LI R., JONÁKOVÁ V., OKO R., PARK C.-S., PRATHER R.S. AND SUTOVSKY P. (2007). Ubiquitin C-Terminal Hydrolase-Activity Is Involved in Sperm Acrosomal Function and Anti-Polyspermy Defense During Porcine Fertilization. Biology of Reproduction, 77 780–793.

YING S.-Y. (2004). Complementary DNA Libraries: An Overview. Molecular Biotechnology, 27 245-252.

ZAVOLAN M., KONDO S., SCHÖNBACH C., ADACHI J., HUME D.A., RIKEN GER GROUP AND GSL MEMBERS, HAYASHIZAKI Y. AND GAASTERLAND T. (2003). Impact of alternative initiation, splicing, and termination on the diversity of the mRNA transcripts encoded by the mouse transcriptome. Genome Research, 13 1290–1300.

ZHAO Y., LI Q., YAO C., WANG Z., ZHOU Y., WANG Y., LIU L., WANG Y., WANG L. AND QIAO Z. (2006). Characterization and quantification of mRNA transcripts in ejaculated spermatozoa of fertile men by serial analysis of gene expression. Human Reproduction 21 1583–1590.

ZIELINSKI J., KILK K., PERITZ T., KANNANAYAKAL T., MIYASHIRO K. Y., EIRÍKSDÓTTIR E., JOCHEMS J., LANGEL U. AND EBERWINE J. (2006). In vivo identification of ribonucleoprotein–RNA interactions. PNAS 103 1557–1562.

ZHONG J., PETERS A.H.F.M., LEE K. AND BRAUN R. E. (1999). A double-stranded RNA binding protein required for activation of repressed messages in mammalian germ cells. Nature Genetics, 22 171-174.

9. Acknowledgements

I would like to thank Prof. Dr. Eckhard Wolf for receiving me in Germany and allowing me to carry out this research work at the Gene Center of the Ludwig-Maximillians-University of Munich. Thanks for his supervision, encouragement and for critically reviewing this manuscript.

I am particularly thankful to Dr. Helmut Blum for giving me the opportunity to develop this research at the Institute of Molecular Animal Breeding and Functional Genome Analysis (LAFUGA). His constant support and advice were decisive in guiding this work.

I am very grateful to Dr. Stefan Bauersachs and Dr. Marc Boelhauve who mentored the execution of this project, including reviewing this manuscript. I am also grateful to Dr. Stefan Krebs for their competent scientific advice and Mrs. Andrea Klanner for her help in the experiments and sequencing. Thanks to Dr. Christian Simmet for the opportunity to work in this project.

Special thanks to all past and present members of the first floor form the Gene Center for their help, advice and good company. I am extremely grateful to my dear friends for their help during decisive stages of this work.

I would like to thank the financial support from the “Deutscher Akademischer Austausch Dienst” (DAAD) and my contact person, Mrs. Maria Salgado Martinez, who spared no effort to help me in these 3,5 years.

Lastly, I am grateful for my wife Eloah who grants me hope, love and continuous support during all stages of this work. My heartfelt thanks to my parents Lídio and Marial and my sisters Paula and Roberta for their love, support and encouragement throughout the years.