RNA in human sperm

Rui Pires Martins¹, Stephen A. Krawetz¹,²,³

¹Center for Molecular Medicine and Genetics, ²Department of Obstetrics and Gynecology, ³Institute for Scientific Computing, Wayne State University School of Medicine, 275 E. Hancock, Detroit, MI 48201, USA

Abstract

We have yet to develop a fundamental understanding of the molecular complexities of human spermatozoa. This encompasses the unique packaging and structure of the sperm genome along with their paternally derived RNAs in preparation for their delivery to the egg. The diversity of these transcripts is vast, including several anti-sense molecules resembling known regulatory micro-RNAs. The field is still grasping with its delivery to the oocyte at fertilization and possible significance. It remains tempting to analogize them to maternally-derived transcripts active in early embryo patterning. Irrespective of their role in the embryo, their use as a means to assess male factor infertility is promising. (Asian J Androl 2005 Jun; 7: 115-120)

Keywords: RNA; siRNA; miRNA; sperm; microarray; fertilization; diagnostics

1 Introduction

Mature spermatozoa serve as a repository for information regarding both genetic and environmental influences. They seed some of the many forms of male factor infertility. Within the past few decades, there has been a decline in human male fertility [1–4]. The direct causes of this reduction remain enigmatic and controversial, but some work suggests that increased environmental and systemic exposure to pesticides, herbicides, estrogenic compounds, heavy metals and reactive oxygen species [5–7] may play a causative role. Concurrent with and perhaps contributory to the decrease in male fertility, there has been a corresponding increase in the incidence of testicular cancer and cryptorchidism [8, 9]. The medical costs associated with treating the rising number of infertile men and men with a fertility-affecting pathology is mounting. These theories remain speculative because little is known about the genetic and molecular basis of infertility.

Of the genes reported to correlate with male factor infertility, most have been identified via non-systematic, or limited surveys; the focus has been on specific genes that were of interest to an individual investigator rather than a genome-wide approach [10]. For example, single nucleotide polymorphisms in one of the genes involved...
in repackaging the male haploid genome, \textit{prm2} [11] were found in infertile males; one of which leads to a truncated protein. Decreased levels of the estrogen receptor [12] in spermatozoan nuclear matrices was also found to occur in ideopathically infertile men. Studies using animal models have pointed to a number of factors that when deficient or inappropriately expressed, lead to infertility [13–18]. Using a more comprehensive approach, recent systematic analyses of the human Y chromosome have identified a number of candidate genes for male factor infertility. For example, deletions identified in the non-recombining region of the Y chromosome lead to spermatogenic failure [19, 20]. Other testis-specific defects have been demonstrated in men with deletions of the Y chromosome encompassing one or more genes [21]. Some of the genes in this region have since been determined to not have any clinical relevance with regard to infertility [22], while other factors demonstrate a clear relationship to an infertile phenotype. It is reasonable to assume that the majority of idiopathic male factor infertilities are likely multifactorial disorders [23–26]. As shown by the recent transcriptional profiling of the CREM knock out mouse model [27] this is reflected by substantial changes in the presence or absence of various multiple members of the affected pathways. Determining exactly how all of these genes are involved, however, remains necessary.

2 RNA in sperm

It has long been held that the tightly packaged chromatin within mature spermatozoa is transcriptionally inert [28]. Despite this, RNA was observed in the mature sperm nucleus of \textit{Scolopendrium} [29] and in rodents and other species [28–33]. \textit{In situ} hybridization from normal fertile human ejaculate localized both beta-actin and \textit{prm2} mRNAs to the head region [34], suggesting that this RNA was a unique sperm head component. Spermatozoa transcript complexity was first addressed by analyzing a series of randomly selected cDNA clones [35]. Sequence comparison showed that 11 of the hybridizing cDNAs were unique within the expressed sequence tag (ESTs) and non-redundant databases whereas five showed no similarity to any of the sequences in the database. Additionally, one was a member of the SINE, i.e., short interspersed repetitive element family, e.g. Alu and another contained a CA\textsubscript{N} repetitive element comprised of a repeating string of CA nucleotides. Together these studies independently confirmed that spermatozoa contain a wealth of both known and unknown protein-encoding and non-coding RNAs. The presence of this suite of RNAs is intriguing when we consider that mature sperm have little if any cytoplasm [36]. These methods of characterization proved a rather laborious means to profile the set of transcripts present in the mature spermatozoa. Since then, other large-scale strategies have been employed.

3 Molecular characterization of sperm transcripts

Characterizing the transcript complexity of the spermatozoa has been refined by the use of microarrays. For example, using mRNA from both testes and ejaculate spermatozoa, a suite of cloned EST microarrays were probed [37, 38]. Stringent precautions were employed to avoid somatic cell contamination. This included two sequential centrifugations through a 40:80 discontinuous gradient of percoll followed by treatment with Triton-X 100 plus sodium dodecyl sulphate (SDS) to remove even a hint of residual somatic cells. cDNA probes from the pooled poly(A\textsuperscript{+}) RNA fraction and the single total RNA fraction were then hybridized to a series of six microarrays containing 27 016 unique ESTs. All transcripts isolated from human sperm were present in testes, but all testes transcripts were not present in sperm. The pooled testes probe hybridized to 26 % of the ESTs whereas the pooled poly(A\textsuperscript{+}) RNA fertile spermatozoa probe identified 12 % of the ESTs. The total RNA spermatozoal probe from the single fertile male ejaculate hybridized to 10 % of the ESTs. The spermatozoal sequences comprised a discrete subset of those identified with the testes probe. As expected the population of RNAs represented by the poly(A\textsuperscript{+}) RNA fraction was similar to that of the total RNA fraction.

To assess variance, different preparations of sperm RNA from different individuals have been compared. Representative results are summarized in Figure 1. Spermatozoal RNAs were isolated and array specific labeled probes were constructed. Each probe was then individually hybridized to an array spotted with 1.176 unique ESTs, shown in panel A. Panels B and C demonstrate a greater than 90 % concordance, between individuals, of both positively hybridized ESTs (\textit{+ve}: 132 of \textasciitilde 170 ESTs), as well as with those ESTs that failed to hybridize (\textit{−ve}: 938 or about \textasciitilde 1015 ESTs). This supports the notion that a core set of invariant fertile transcripts will be identified.

Subsequent data analysis from these and other stud-
ies has revealed that hydrolases, many of which are found in the acrosome and DNA binding proteins, associated with the extensive restructuring of the nucleus during spermiogenesis [39] are among the largest protein-encoding groups identified. In comparison, when queried by cellular component, the largest protein-encoding groups were the plasma membrane, nucleus and cytoplasm. A series of spermatozoal transcripts have been identified that are concordant with mRNAs known to participate in fertilization and embryonic development. A comparison of this suite of sperm RNAs with those present in human and mouse oocyte cDNA libraries revealed that several sperm derived transcripts essential for early development were not present in the oocyte. These encode a series of proteins associated with fertilization (i.e., SP-40, sulfated glycoprotein 2, calmegin and several heat shock response products) that are important for embryo development [40]. Their presence in human testes and sperm and their absence in the unfertilized egg as assessed by RT-PCR and the delivery of some of these RNAs from the sperm to the egg upon fertilization has been confirmed [41]. This has led to the obvious yet unanswered question, do spermatozoal RNAs, encode unique function(s) in the developing zygote and/or embryo (Figure 2A)? Perhaps they help sustain viability of the embryonic genome?

Of the ribonucleic species delivered to the oocyte at fertilization, a large proportion of them are of low molecular weight and are shown to be in the anti-sense orientation (Figure 2A) [42]. A total of 68 different siRNAs have been identified in human spermatozoa (http://compbio.med.wayne.edu/Sperm_RNAi.htm) of which 13 show significant similarity with those previously implicated in RNA-mediated regulation [42]. Could these function in a manner similar to small interference RNAs (siRNAs) like lin-4 and let-7? These are well known as regulators of development and differentiation in Caenorhabditis elegans regulating the timing of larval development [43, 44] and the transition from late larval to adult stages [45, 46], respectively. The small sperm RNA transcripts show a wide tissue distribution including a number of early embryonic tissues. Several of the transcripts are implicated in embryonic development including DKK2, TIA and fat-3. For example, dickkopf2 (DKK2) is known to inhibit the WNT signaling pathway [47–50] regulating cell fate and pattern generation dur-
ing embryogenesis whereas, *fat-3*, is perhaps a key player in the differentiation of mesoderm, somites and neurons in mammals [51], as well as growth and neuronal signaling in *C. elegans* [52].

With such a large number and diversity of transcripts found in mature spermatozoa, it is tempting to draw comparisons with maternal RNAs, stored in the oocyte. The importance of these in the early patterning and development in *Drosophila* [53] and *Xenopus* [54, 55] embryos is well documented. In mammals, their importance was first only inferred [56] and later confirmed through the discovery of specific genes and by experiments examining the onset of zygotic transcription [57, 58]. As sperm have been shown to deliver RNA to the oocyte, the presence of antisense RNAs within this pool could provide a means to regulate early parental-based gene activity, by selectively targeting maternally or paternally derived RNAs for degradation. Furthermore, with siRNAs being known to persist through several cell divisions, it is possible that these transcripts could help regulate early genetic activity through to the multi-cell embryo.

### 4 Sperm RNA as a clinical tool

The application of microarray technology to spermatozoal RNA isolated from ejaculate presents a unique opportunity to globally address the mechanisms that control the differentiation of the male gamete during normal, perturbed and diseased states (Figure 2B). They provide a useful molecular record to assess environmental insult and/or genetic status since spermatogenesis is highly sensitive to environmental exposures including chemical, thermal and biological agents [59]. Recently two independent studies have used RNA profiling techniques to address the relationship of motility and the RNA population between normal and motility impaired sperm [60, 61]. Interestingly several transcripts were identified that varied in a significant manner between the normal and motility impaired samples. These included testis specific protein 1 and lactate dehydrogenase C transcript variant 1 [61]. This clearly points to the potential of this strategy to be used as a clinical assay to provide a panoramic view of testis gene expression [37] that can be difficult to achieve from a testicular biopsy. Defining the “fertile male fingerprint” would have a significant impact on diagnosis, treatment and counseling. Realization would signal a major advance in the field of andrology.

### Acknowledgment

The authors gratefully acknowledge the Michigan Economic Development Corporation and the Michigan Technology Tri-corridor for the support of this research program (Grant 085P4001419). The authors would like to thank the laboratory personnel and many collaborators who have contributed to this work.

### References

5808–20.


RNA in human sperm

403: 901–6.


