Journal of Andrology, Vol. 26, No. 1, January/February 2005

A Suite of Novel Human Spermatozoal RNAs

G. CHARLES OSTERMEIER,*† ROBERT J. GOODRICH,*† JULIE S. MOLDENHAUER,*
MICHAEL P. DIAMOND,* AND STEPHEN A. KRAWETZ*†‡

*From the *Department of Obstetrics and Gynecology, †Center for Molecular Medicine and Genetics, and ‡Center for Scientific Computing, Wayne State University, Detroit, Michigan.*

ABSTRACT: We recently described a complex population of spermatozoal coding RNAs that are delivered to the oocyte on fertilization. These are derived throughout spermatogenesis, representing a record of past events. Recently, evidence has been provided that micro-RNAs are present in testes, suggesting that they might also be carried in ejaculate spermatozoa. To directly test this hypothesis, a unique microarray system capable of directly identifying antisense RNAs and predicted transcripts was utilized. RNA isolated from the ejaculate spermatozoa of 6 normal fertile men was directly hybridized to sense oligonucleotide arrays containing 10 000 elements. This revealed 68 shared RNAs, some of which are similar to those previously defined as micro-RNAs, whereas others were the antisense of previously in silico–predicted transcripts. The results and implications of this study are described in this communication.

Key words: Development, fertility, cloning, microarray, novel transcripts, antisense RNA, spermatozoa.

J Androl 2005;26:70–74

Small noncoding RNAs are receiving increased attention for their role in regulating gene expression in animals, plants, and fungi. The first such regulatory RNA to be identified was *lin-4*. This small interfering RNA (siRNA) controls the timing of *Caenorhabditis elegans* larval development (Lee et al, 1993; Wightman et al, 1993). A second RNA regulator, *let-7*, is expressed later in development and appears to trigger the transition into the late larval and adult stages (Reinhart et al, 2000; Slack et al, 2000). Human spermatozoa have recently been documented to contain a wide spectrum of RNAs (Ostermeier et al, 2002b), some of which are delivered as part of the payload to the oocyte on fertilization (Ostermeier et al, 2004). On the basis of the observation that spermatozoal RNAs are heterogeneously sized (Ostermeier et al, 2002b) and the recent interest in both micro-RNAs (miRNAs) and silencing RNAs (siRNAs), the hypothesis was put forth that spermatozoal RNAs might also include these novel classes of RNAs. To test this tenet, a unique microarray system ideal for directly identifying novel RNAs (eg, antisense), predicted transcripts, or both was employed.

Materials and Methods

RNA Extraction

Six human ejaculates were collected by masturbation and allowed to liquefy for 30 minutes. The semen was washed 3 times in frozen storage buffer (50 mM HEPES buffer, pH 7.5; 10 mM NaCl; 5 mM Mg-acetate; and 25% glycerol), then stored at −80°C. Subsequent to storage, the samples were thawed, washed twice in 12 mL of PBS, then resuspended in a somatic cell lysis buffer (0.1% SDS, 0.5% Triton X-100 in dH_2O). This procedure produces essentially pure spermatozoa (Ostermeier et al, 2002b). Cell concentrations were determined with a hemacytometer. Ribonucleic acid was extracted from the purified spermatozoa with the RNeasy kit (Qiagen Inc, Valencia, Calif) with minor modifications. Lysis buffer was added to the samples at 600 $\mu L/10^7$ cells. The lysates were homogenized with a 26-gauge needle and heated for 30 minutes at 65°C. After incubation, the samples were homogenized a final time to ensure shearing of the DNA. This sample was then loaded onto the RNeasy column and processed essentially as described by the manufacturer. The RNA was eluted from the column with 2 (50-μL) RNase-free H_2O washes. Dithiothreitol at a final concentration of 5 mM and 40 units of RNase Block (Strategene, La Jolla, Calif) were then added to the samples before a 20-minute treatment with RNase-free DNase I (Ambion, Austin, Tex) at 37°C. The amount of RNA was then determined with the RiboGreen RNA quantifying kit (Molecular Probes, Eugene, Ore) as previously described (Goodrich et al, 2003). The samples were stored at −80°C until use.

Quality Control

Sample purity was established by reverse transcription polymerase chain reaction (RT-PCR), employing the intron-spanning protamine 2 (PRM2) primers. Only a 149-bp intronless product should be observed. The presence of contaminating DNA, as evidenced by a 310-bp product, prompted an additional DNase I treatment. RT-PCR was performed with oligo-dT primers and the SuperScript III enzyme (Invitrogen Co, Carlsbad, Calif) and 100–200 ng of total RNA. PCRs were performed for 40 cycles with a cDNA copy of the purified RNA as template, along with the following intron-spanning PRM2 primers (for-

This work was supported by WSU Ob/Gyn grant 95200 to S.A.K.

Correspondence to: Stephen A. Krawetz, Charlotte B. Failing Professor of Fetal Therapy and Diagnosis, 253 C.S. Mott Center, 275 East Hancock, Detroit, MI 48201 (e-mail: steve@compbio.med.wayne.edu).

Received for publication June 20, 2004; accepted for publication August 2, 2004.

ward: -tat agg cgc aga cac tgc; reverse -gcc ttc tgc atg ttc tct) and the Hot-Star Taq Polymerase system (Qiagen).

Labeling and Hybridization

Hybridizations were carried out with the Hybrid Capture HC Express Array Kit (Digene, Gaithersburg, Md). Briefly, 500 ng of isolated RNA was denatured at 95°C for 2 minutes in 25 μL of hybridization buffer. MWG sense oligonucleotide arrays (MWG Biotech, Ebersberg, Germany) under LifterSlips (Electron Microscopy Sciences, Washington, Pa) were then overlayed with the hybridization mix by capillary action. The hybridization chambers were sealed and then incubated overnight at 65°C. Following hybridization, the arrays were washed, allowed to air dry for 1 minute, then returned to the hybridization chambers. RNase (50 μL) containing primary RNA/DNA hybrid antibody was placed under the LifterSlips, and the sample was incubated for 60 minutes at room temperature. The microarrays were then washed and stained with 50 μL of the secondary antibody at room temperature. Samples were then washed a final time before signal enhancement, as described by the manufacturer. The microarrays were dried then scanned with a Typhoon 9210 scanner (Amersham Pharmacia Biotech, Piscataway, NJ).

Analysis

The scanned images were viewed with Quantity One software (BioRad, Hercules, Calif) and saved in the TIFF format. The saved images were analyzed with the Imagene software (BioDiscovery, El Segundo, Calif) with default settings for threshold (ie, 2 SD above background) and spot detection. A list of positive hybridizations was generated for each microarray and hybridizations were compared with Statistical Analysis Software (SAS, Cary, NC) as described by Ostermeier et al (2002a).

Results and Discussion

To examine whether ''noncoding'' strand RNAs were present in human spermatozoa, the MWG 30k (A, B, and C) 50-mer sense oligonucleotide arrays (MWG Biotech) were directly probed with RNA isolated from ejaculate spermatozoa. All RNA-DNA oligonucleotide hybrids were detected with a fluorescently tagged RNA-DNA antibody (Digene). To maximize specificity, stringent hybridization temperatures and washes were employed. The calculated melting temperature (Howley et al, 1979) of each arrayed 50-mer element indicated that a perfectly matched 20-mer could yield a hybridization signal. Only the C-array that predominantly contained in silico–identified genes showed significant hybridization. Specificity was further confirmed because the 104 *Arabidopsis* controls present on each chip did not hybridize to the added RNA probe. It was thus selected for further analysis of spermatozoal RNAs from 6 different men. Each of the ejaculate profiles from the C-chip was then compared. A core set of 68 candidate spermatozoal sense-hybridizing RNAs shared among all 6 individuals was identified, as summarized in Table 1. The 68 core sequences represented by each 50-mer were compared with the mature segment of the miRNAs of the human genome (Lim et al, 2003). Eighteen of the 50-mers corresponding to the candidate spermatozoal RNAs that hybridized to the sense array had an alignment *z*-score of greater than 100 (http://compbio.med.wayne.edu/Sperm_RNAi_seq.htm). The alignment of one of these miRNAs contained in the human miRNA registry of complete miRNA sequences (http://www.sanger.ac.uk/Software/Rfam/mirna/index.shtml) could be extended throughout its corresponding 50-mer. This sequence aligned with mir-182. Given that at least 15% of the miRNAs remain to be identified (Lim et al, 2003), these data support the notion that novel spermatozoal noncoding antisense RNAs are present.

To begin to define their role, the potential targets of these 68 shared sequences were mined with the ENSEMBL (http://www.ensembl.org/) and SOURCE databases (http://genome-www5.stanford.edu/cgi-bin/source/sourceSearch) to identify cytogenetic location and overlap with known transcripts, as well as their tissue and functional distributions in early development. As summarized in Table 2, the corresponding transcripts that have been associated with early development were evenly distributed within the genome and appeared to be expressed in a wide variety of tissues, including the testes and early embryonic tissues. These include DKK2, TIA, and FAT3. For example, dickkopf2 (DKK2) has been shown to inhibit the wingless type (WNT) signaling pathway (Mao et al, 2001; Mao and Niehrs, 2003) that has been implicated in oncogenesis and in several developmental processes, including the regulation of cell fate and pattern formation during embryogenesis. Thus, silencing of DKK2 with antisense RNAs delivered by the spermatozoa could ensure a functional WNT pathway.

The results presented in this study provide additional evidence for the existence of a wide spectrum of spermatozoal RNAs. On one hand, these exciting observations provide clear evidence that previously unknown or in silico–predicted transcripts can be identified by the application of microarrays to spermatozoal RNAs. This is interesting in light of the recent and fruitful efforts to test these various predictions (Rinn et al, 2003). These studies have indeed shown that our information content is larger than expected. On the other hand, as we have shown, this population houses a suite of noncoding strand RNAs that could include miRNAs. Consistent with our previous observations that spermatozoa contain testis transcripts (Ostermeier et al, 2002b), these are likely similar to the micro-RNAs recently identified in human testis (Liu et al, 2004). It is tempting to speculate that the delivery of these spermatozoal ''antisense'' RNAs on fertilization enables their participation in early postfertilization processes. They may provide a new level of control that helps to

Table 1. *Hybridizing spermatozoa RNAs: siRNA candidates**

Accession No.†	Gene Name‡
AB058706 1	ZNF462: zinc finger protein 462
AF009225 1	*CHUK*: conserved helix-loop-helix ubiquitous kinase
AF103305 1	Immunoglobin heavy chain variable region
AF279899 1	*PNRC1*: proline-rich nuclear receptor coactivator 1
AJ247651 1	Immunoglobulin heavy chain variable region; ighlsg6.1
AL139429.8.1.90052.1	Sodium/potassium/calcium exchanger 3 precursor (Na(+)/K(+)/CA(2+)-exchange protein 3): SLC24A3
AC002556.1.1.123189.2	Olfactory receptor 8D1 (olfactory receptor–like protein JCG9)
AP002534.1.1.300000.2	Gamma interferon–inducible protein IFI-16 (interferon-inducible myeloid differentiation transcriptional activator): IFI16
AC004590.1.1.135259.2	Voltage-dependent T-type calcium channel alpha-1G subunit: *CACNA1G*
AB019437.1.1.200000.1	IG heavy chain V-I region V35 precursor
AC005412.6.1.183700.2	Beta crystallin A3 [contains: beta crystallin A1]: CRYBA1
AL157877.11.1.198567.1	ETS-related transcription factor *ELF-1* (E74-like factor 1): *ELF1*
AL031115.2.1.96558.2	No genes within 10 kb
AL391478.13.1.180625.1	Pro-neuregulin-3 precursor (*PRO-NRG3*)
AL121713.24.1.67575.1	No description, hypothetical protein
ENSG00000073578	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial precursor: *SDHA*
L03175 1	Immunoglobulin gamma chain, v region; ighve
M71229 1	Immunoglobulin heavy chain vdjc region
X99367 1	Immunoglobulin heavy chain
Y14938 1	igg4 Rheumatoid factor immunoglobulin heavy chain; ighv
Z30558 1	ig heavy chain (vh4) v region (vdj)
AC011737.8.93599.176848.2	*ALS2CR17* protein (fragment)
AC000382.1.1.620.1	No genes within 10 kb
AL590041.5.1.110189.1	No genes within 10 kb
AC019127.7.67762.75714.1	Latent transforming growth factor beta binding protein, isoform 1S precursor (*LTBP-1*)
AC008936.4.66581.102307.1	No genes within 10 kb
AC003099.1.1.95129.1	Dickkopf-related protein-2 precursor: *DKK2*
AC090886.1.1.168778.4	No description, hypothetical protein
AC019336.5.3553.5907.1	No genes within 10 kb
AC004040.1.1.59411.1	No genes within 10 kb

Table 1. *Continued*

Accession No.†	Gene Name‡
AC000127.1.1.41306.1	No description, hypothetical protein
AL138928.6.47715.179479.2	No genes within 10 kb
AC027288.25.1.167394.2	No genes within 10 kb
AC016725.2.174323.203356.1	No description, hypothetical protein
AC018542.8.51167.92203.1	No genes within 10 kb
AC017061.5.61887.90968.2	Solute carrier family 26, member 7 isoform A; sulfate anion transporter: *SLC26A7*
AL353804.22.1.167668.2	No genes within 10 kb
AL078602.13.1.152130.2	No genes within 10 kb
AC022206.3.1.183224.3	No genes within 10 kb
AL136525.16.1.191682.8	No description, hypothetical protein
AC025953.2.116948.127472.1	No genes within 10 kb
AC002080.1.1.171905.2	Hepatocyte growth factor receptor precursor: *MET*
AC023391.4.35989.52052.1	WD repeat and FYVE domain–containing 3 isoform 1: WDFY3
AL035691.17.1.129968.1	Cation-independent mannose-6-phosphate receptor precursor (CI MAN-6-P receptor) (*CI-MPR*) (insulinlike growth factor II receptor): *IGF2R*
AL031771.1.1.87552.1	No genes within 10 kb
AL049591.12.1.122400.2	No genes within 10 kb
AP001726.1.1.340000.1	No genes within 10 kb
AC007006.3.1.112027.1	No genes within 10 kb
AC021994.3.103121.115550.1	No genes within 10 kb
AL109611.8.1.100204.3	No genes within 10 kb
AC023131.5.113997.162896.2	No description, hypothetical protein
AC027179.3.123461.145612.1	No genes within 10 kb
AC010478.5.1.126852.1	Insulin gene enhancer protein *ISL-1*
AC079338.2.55527.74865.1	Nucleolysin TIA-1
AP003171.1.54891.68656.1	*FAT3* (fragment)
AL392171.3.25707.97190.1	Partitioning-defective 3 homolog: *PARD3*
AP001203.3.1.79432.1	No genes within 10 kb
AL583847.3.160096.168498.1	No genes within 10 kb
AC051643.2.161782.189397.1	No genes within 10 kb
AC025273.4.19196.27310.1	Conserved oligomeric Golgi complex component 7: *COG7*
AP000762.3.1.42839.1	Neurotrimin precursor: *HNT*
AC084367.2.152436.167463.1	No description, hypothetical protein
AC004097.1.1.87077.1	No genes within 10 kb
AC010650.6.89651.91916.1	KIAA1576 protein, hypothetical protein
ENSG00000115257	Polyposis locus protein 1-like 1
AC005683.2.1.74673.2	Neuronal cell adhesion molecule precursor: *NRCAM*
AC009781.8.32872.52388.1	No genes within 10 kb
AC004562.1.1.133925.2	No genes within 10 kb

* For more information, including genomic location and tissue distribution, see http://compbio.med.wayne.edu/Sperm_RNAi.htm.

† ID number provided by MWG.

‡ Information obtained from ENSEMBL.

Table 2. *Examples of candidate early development antisense RNAs identified in sperm**

Gene ID	Chromosomal Location	Tissue Distribution
CACNA1G	17q21.33	Brain, amygdala, subthalamic nuclei, cerebellum, thalamus, heart, *placenta,* kidney, lung, colon, bone marrow, *fetal brain, fetal heart, fetal kidney, fetal lung*
PRO-NRG3	10q23.1	Brain, *testis,* aorta, hypothalamus, *fetal brain,* retina, cochlea, MS, neuroblastoma
ALS2CR17	2q33.2	*Fetal eye,* lung, pancreas, *placenta,* chondrosarcoma
LTBP1	2p22.3	Fibroblasts, breast adenocarcinoma, LCIS, bone, *whole embryo,* amygdala, kidney, ovary, chondrosarcoma, lung tumor
DKK2	4q25	Heart, brain, skeletal muscle, lung, *placenta,* cochlea, Ewings sarcoma, meningioma, *fetal heart,* pregnant uterus, melanocyte, leiomyosarcoma
SCL26A7	8q21.3	Thymus, testis, osteoblastic cells, clear cell tumor, *placenta,* endometrial adenocarcinoma, *fetal eyes,* eyes, pancreatic islet
IGF2R	6q25.3	*Fetal pancreas,* invasive carcinoma, lymphoma, spinal cord, anterior eye, myeloid cells, uterus, lens, optic nerve, peripheral nervous system
KIAA1576	16q23.1	Brain, DRG, astrocytoma, sympathetic trunk, retina, iris, RPE, adrenal adenoma, teratocarcinoma, lens, *embryonal carcinoma, placenta, embryonic stem cells*
FAT3	11q14.3	Wilms, brain, bone, hypothalamus, chondrosarcoma, melanocyte, *fetal heart,* pregnant uterus, lung, thymus, spleen, spinal cord, mammary gland, liver, intestines, skeletal muscle, kidney, heart
COG7	16p12.2	Prostate, brain, adrenal adenoma, renal cell adenoma, *placenta,* optic nerve, NK cells, sciatic nerve, neuroblastoma
HNT	11q25	Brain, oligodendroglioma, *fetal eyes,* eyes, optic nerve, retina, carcinoid, hippocampus, schizophrenic brain, amelanotic melanoma
TIA1	2p13.3	Carcinoma, adenocarcinoma, wilm's tumor, kidney, leukopheresis, thymus, *whole embryo,* malignant melanoma, uterus, hypothalamus
ISL1	5q11.2	Retina, adenocarcinoma, neuroblastoma, *fetal eyes,* optic nerve, *testis, placenta,* insulinoma, adrenal medulla, dorsal root ganglion, pineal, brain
ZNF462	9q32	Kidney, early stage papillary serous carcinoma, posterior rhombomeres, mandible, Ewings sarcomamelanotic melanoma, colonic mucosa, aleolar rhabdomyosarcoma, *fetal eyes,* human retina, *embryonic stem cells, testis*
CHUK	10q24-q25	Normal endometrium, myelogenous leukemia, neuroblastoma, meningiomas, uterine, human retina, sciatic nerve, RPE and Choroid, *placenta,* adrenal cortex carcinoma

* For additional information, see http://compbio.med.wayne.edu/Sperm_RNAi.htm. Tissues representative of early development are italicized.

establish imprints during the transition from maternal to embryonic genome, or both. The latter could be the case given the recent observations of Morris et al (2004) and Fukagawa et al (2004), showing that this class of RNAs in mammals can confer transcriptional silencing by methylation, a known mechanism of imprinting. Their results and the findings reported in this communication should spur our understanding of the functions of spermatozoal RNAs.

Acknowledgment

Normal fertile human sperm samples were a generous gift from the NICDH Reproductive Medicine Network (U10 HD-39005).

References

Fukagawa T, Nogami M, Yoshikawa M, Ikeno M, Okazaki T, Takami Y, Nakayama T, Oshimura M. Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat Cell Biol.* 2004;6:784–791.

Goodrich RJ, Ostermeier GC, Krawetz SA. Multitasking with molecular dynamics Typhoon: quantifying nucleic acids and autoradiographs. *Biotechnol Lett.* 2003;25:1061–1065.

Howley PM, Israel MA, Law MF, Martin MA. A rapid method for detecting and mapping homology between heterologous DNAs. Evaluation of polyomavirus genomes. *J Biol Chem.* 1979;254:4876–4883.

Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to *lin-14. Cell.* 1993;75:843–854.

Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP. Vertebrate microRNA genes. *Science.* 2003;299:1540.

Liu CG, Calin GA, Meloon B, et al. An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. *Proc Natl Acad Sci U S A.* 2004;101:9740–9744.

Mao B, Niehrs C. Kremen2 modulates Dickkopf2 activity during Wnt/LRP6 signaling. *Gene.* 2003;302:179–183.

Mao B, Wu W, Li Y, Hoppe D, Stannek P, Glinka A, Niehrs C. LDL-receptor–related protein 6 is a receptor for Dickkopf proteins. *Nature.* 2001;411:321–325.

Morris KV, Chan SW-L, Jacobsen SE, Looney DJ. Small interfering RNA-induced transcriptional gene silencing in human cells. *Science.* 2004;305:1289–1292.

Ostermeier GC, Dix DJ, Krawetz SA. A bioinformatic strategy to rapidly characterize cDNA libraries. *Bioinformatics.* 2002a;18:949–952.

Ostermeier GC, Dix DJ, Miller D, Khatri P, Krawetz SA. Spermatozoal RNA profiles of normal fertile men. *Lancet.* 2002b;360:772–777.

Ostermeier GC, Miller D, Huntriss JD, Diamond MP, Krawetz SA. Reproductive biology: delivering spermatozoan RNA to the oocyte. *Nature.* 2004;429:154.

Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans. Nature.* 2000;403: 901–906.

Rinn JL, Euskirchen G, Bertone P, et al. The transcriptional activity of human chromosome 22. *Genes Dev.* 2003;17:529–540.

Slack FJ, Basson M, Liu Z, Ambros V, Horvitz HR, Ruvkun G. The lin-41 RBCC gene acts in the *C. elegans* heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Mol Cell.* 2000;5:659–669.

Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans. Cell.* 1993;75:855–862.