

Generation of a bovine oocyte cDNA library and microarray: resources for identification of genes important for follicular development and early embryogenesis

Jianbo Yao, Xiaoning Ren, James J. Ireland, Paul M. Coussens, Timothy P. L. Smith and George W. Smith

Physiol Genomics 19:84-92, 2004. doi:10.1152/physiolgenomics.00123.2004

You might find this additional information useful...

This article cites 42 articles, 21 of which you can access free at:

<http://physiolgenomics.physiology.org/cgi/content/full/19/1/84#BIBL>

This article has been cited by 2 other HighWire hosted articles:

Analysis of the Female Gametophyte Transcriptome of Arabidopsis by Comparative Expression Profiling

H.-J. Yu, P. Hogan and V. Sundaresan

Plant Physiology, December 1, 2005; 139 (4): 1853-1869.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Genes Preferentially Expressed in Bovine Oocytes Revealed by Subtractive and Suppressive Hybridization

S. Pennetier, S. Uzbekova, C. Guyader-Joly, P. Humblot, P. Mermillod and R. Dalbies-Tran

Biol Reprod, October 1, 2005; 73 (4): 713-720.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Medline items on this article's topics can be found at <http://highwire.stanford.edu/lists/artbytopic.dtl> on the following topics:

Biophysics .. Complementary DNA

Cell Biology .. Ovum

Cell Biology .. Microarrays

Developmental Biology .. Embryo Gene Expression

Endocrinology .. Ovaries

Genomics .. Arrays

Updated information and services including high-resolution figures, can be found at:

<http://physiolgenomics.physiology.org/cgi/content/full/19/1/84>

Additional material and information about *Physiological Genomics* can be found at:

<http://www.the-aps.org/publications/pg>

This information is current as of August 7, 2006 .

Generation of a bovine oocyte cDNA library and microarray: resources for identification of genes important for follicular development and early embryogenesis

Jianbo Yao,¹ Xiaoning Ren,^{4,6} James J. Ireland,^{4,5,6} Paul M. Coussens,^{4,6} Timothy P. L. Smith,² and George W. Smith^{3,4,5,6}

¹Division of Animal and Veterinary Sciences, West Virginia University, Morgantown, West Virginia 26506;

²United States Department of Agriculture, Agricultural Research Service, United States Meat Animal Research Center, Clay Center, Nebraska 68933; and ³Laboratory of Mammalian Reproductive Biology and Genomics, Departments of ⁴Animal Science and ⁵Physiology, and ⁶Center for Animal Functional Genomics, Michigan State University, East Lansing, Michigan 48824

Submitted 26 May 2004; accepted in final form 9 July 2004

Yao, Jianbo, Xiaoning Ren, James J. Ireland, Paul M. Coussens, Timothy P. L. Smith, and George W. Smith. Generation of a bovine oocyte cDNA library and microarray: resources for identification of genes important for follicular development and early embryogenesis. *Physiol Genomics* 19: 84–92, 2004; doi:10.1152/physiolgenomics.00123.2004.—The oocyte is a key regulator of ovarian folliculogenesis and early embryonic development. However, the composition of the oocyte transcriptome and identities and functions of key oocyte-specific genes involved in the above processes are relatively unknown. Using a PCR-based cDNA amplification method (SMART technology), we constructed a bovine oocyte cDNA library. Analysis of 230 expressed sequence tags (ESTs) from this library identified 102 unique sequences. Although some correspond to house-keeping genes (e.g., ribosomal protein L15) and some represent genes previously known to be expressed in oocytes and other tissues, most encode for genes whose expression in mammalian oocytes has not been reported previously (e.g., cocaine- and amphetamine-regulated transcript) or genes of unknown function. Sixteen did not show significant sequence similarity to any entries in the GenBank database and were classified as novel. Using over 2,000 unsequenced, randomly selected cDNA clones from the library, we constructed an oocyte microarray and performed experiments to identify genes preferentially expressed in fetal ovary (an enriched source of oocytes) relative to somatic tissues. Eleven clones were identified by microarray analysis with consistently higher expression in fetal ovaries (collected from animals at days 210–260 of gestation) compared with spleen and liver. DNA sequence analysis of these clones revealed that two correspond to JY-1, a novel bovine oocyte-specific gene. The remaining nine clones represent five identified genes and one additional completely novel gene. Increased abundance of mRNA in fetal ovary for five of the six genes identified was confirmed by real-time PCR. Results demonstrate the potential utility of these unique resources for identification of oocyte-expressed genes potentially important for reproductive function.

cattle; egg; expressed sequence tag; ovary

ABUNDANT EVIDENCE HAS BEEN presented over the last decade indicating that the oocyte is a key regulator of ovarian follicular development. A developmental program intrinsic to the oocyte controls the overall rate of follicular development (12),

and oocyte-secreted factors dramatically affect the development and function of follicular granulosa cells (11). Oocytes secrete many proteins as demonstrated by two-dimensional gel electrophoresis (19), but only a small number of these proteins have been identified and characterized so far. Some of the known, well-characterized factors secreted by oocytes include growth differentiation factor-9 (GDF-9) and bone morphogenetic protein-15 (BMP-15). The essential roles of these factors in follicular development have been demonstrated primarily by their targeted deletion in mice and molecular genetics approaches in farm species (7, 9, 13, 40).

Oocyte-expressed genes are not only important for follicular growth and development but also are crucial for early embryogenesis. During the initial cleavage divisions postfertilization, embryonic development is supported by maternal mRNAs and proteins synthesized and stored during oogenesis. These maternal-effect factors are critical for the interval between fertilization and the maternal-embryonic transition when transcriptional activity of the embryonic genome becomes fully functional. More specifically, gene targeting studies have revealed an obligatory role for *MATER* and *Zar1* in promoting early embryogenesis prior to embryonic genome activation (4, 36, 38). However, our understanding of composition of the oocyte transcriptome and the identity of key oocyte-expressed genes with important regulatory roles in folliculogenesis and early embryonic development is far from complete.

A variety of approaches have been used previously to study gene expression in oocytes and embryos. RT-PCR has been the method of choice for examination of expression of known genes of interest on an individual basis. Differential display RT-PCR (21) has been successfully used for identification of novel developmentally regulated genes in oocytes and early embryos (21, 24, 26, 27). Suppressive subtractive hybridization (8) has been utilized for isolation of bovine oocyte transcripts associated with maturation (30), blastocyst-stage-specific transcripts (23), and granulosa cell mRNA associated with oocyte competence (31). The development of DNA microarray technology has permitted the analysis of thousands of genes simultaneously, and the technology has become increasingly popular as a means of identifying differentially expressed genes in complex biological models. However, direct application of this technology to biological studies of oocytes and early embryos has been limited, particularly in nonlaboratory model species.

Article published online before print. See web site for date of publication (<http://physiolgenomics.physiology.org>).

Address for reprint requests and other correspondence: J. Yao, Division of Animal and Veterinary Sciences, West Virginia Univ., Morgantown, WV 26506-6108 (E-mail: jianbo.yao@mail.wvu.edu).

Generation of cDNA libraries from specific tissues and analysis of expressed sequence tags (ESTs) are essential steps toward the discovery of new genes and the development of DNA microarray resources. The limited quantity of RNA in oocytes and early embryos prevents the use of conventional methods to construct cDNA libraries from such materials. The PCR-based cDNA amplification method known as the SMART technology (BD Biosciences, Palo Alto, CA) allows for cDNA library construction from limited quantities of RNA. Using this technology, Yan et al. (39) have constructed a mouse oocyte cDNA library and identified a novel oocyte-secreted factor by differential screening of random clones spotted on membrane arrays. To generate resources for functional genomic studies of oocyte biology and identification of novel oocyte-expressed genes potentially involved in oocyte/follicular development and early embryogenesis in cattle, we constructed a bovine oocyte cDNA library using the SMART technology. Sequencing of a limited number of ESTs led to the discovery of 16 novel sequences. In addition, using over 2,000 randomly selected clones from the library, we developed an oocyte-specific cDNA microarray and utilized this array to identify 6 oocyte-expressed genes preferentially expressed in fetal ovary (an enriched source of oocytes) compared with adult somatic tissues.

MATERIALS AND METHODS

Construction of the bovine oocyte cDNA library. We collected 100 germinal vesicle stage and 100 in vitro matured metaphase II stage bovine oocytes for use in cDNA library construction. Germinal vesicle stage oocytes were obtained from ovaries collected at an abattoir. All oocytes utilized were grade I (surrounded by multiple layers of compact cumulus cells and containing a homogeneously organized ooplasm). Metaphase II stage oocytes were obtained from a commercial source (Cyagra, Elizabethtown PA). Germinal vesicle and metaphase II stage oocytes were denuded of surrounding cumulus cells by incubation for 3–5 min (with gentle pipetting) in microdroplets of HEPES-buffered TCM-199 media (Sigma, St. Louis, MO) containing 1 mg/ml hyaluronidase. Oocytes were then transferred through five microdroplets of HEPES-buffered TCM-199 media to remove dissociated cumulus cells. Oocytes completely denuded of cumulus cells were then individually selected under a stereomicroscope and pooled for subsequent RNA isolation.

Construction of the oocyte cDNA library was performed using the SMART technology (BD Biosciences, Palo Alto CA). Poly(A)⁺ RNA was isolated directly from a pool of 100 germinal vesicle stage and 100 in vitro matured metaphase II stage bovine oocytes using an Oligotex direct mRNA micro kit (Qiagen, Valencia, CA). The eluted mRNA was dried by speed vacuum and resuspended in 8 μ l of water followed by addition of 1 μ l each of a modified oligo-dT primer (1 μ g/ μ l, 5'-GACTAGTTCCTAGATCGCGAGCGGCCGCTTTTTTTTTTTTTTTTTTTTTTTTTTTT-3') containing a *NotI* site and a SMART primer (1 μ g/ μ l, 5'-AAGCAGTGGTAACAACGACAGATCGAATTTCGTCGACGCGGG-3') containing a *SallI* site. The mixture was incubated at 70°C for 5 min and placed immediately on ice for 3 min. The volume was then adjusted to 20 μ l with 4 μ l of 5 \times first-strand cDNA synthesis buffer, 2 μ l of 0.1 M DTT, 1 μ l of 10 mM dNTP, 2 μ l of water, and 1 μ l of SuperScript II reverse transcriptase (200 U/ μ l, Invitrogen Life Technologies, Gaithersburg, MD) followed by incubation at 42°C for 60 min in a thermal cycler. The synthesized cDNA (20 μ l) was amplified by 35 cycles of PCR in a 200- μ l reaction containing 1 \times Advantage 2 PCR buffer (BD Biosciences), 0.4 mM dNTP, 0.3 μ M 5'-SMART PCR primer (5'-AAGCAGTGGTAAACAACGACAGATAC-3'), 0.3 μ M 3'-SMART PCR primer (5'-GACTAGTTCCTAGATCGCGAGCGG-3'), and 4 μ l of 50 \times Advan-

tage 2 polymerase mix (BD Biosciences). The PCR product was extracted with phenol:chloroform:isoamyl alcohol (25:24:1), ethanol-precipitated, and resuspended in 20 μ l of water followed by digestion with *SallI* and *NotI*. The digested cDNAs were purified using a PCR purification kit (Qiagen) and fractionated through a Chroma Spin 400 column (BD Biosciences). The digested/size selected cDNAs were then ligated to *SallI/NotI* predigested pSPORT1 vector (Invitrogen) and electroporated into DH10B competent cells. The library was amplified by overnight culture (the whole library) on Luria broth (LB)/agar plates (~20,000 independent colonies/150-mm plate), and plasmid DNA was isolated using a Qiagen MAXI column. To remove clones with no inserts or very short inserts, a size selection procedure was applied to the isolated library plasmid DNA. Approximately 4 μ g of library plasmid were separated on a 1% agarose gel, the front tip of the plasmid DNA smear (containing mostly supercoiled plasmid with no inserts) was excised, and plasmid from the remaining agarose gel was recovered by phenol extraction followed by ethanol precipitation.

DNA sequencing and EST analysis. Random colonies were picked from the plated library (size-selected) and grown in LB supplemented with 100 μ g/ml ampicillin. Plasmid DNA was isolated from overnight cultures using the QIAprep 96 Turbo miniprep kit (Qiagen). Sequencing reactions were performed using a BigDye terminator sequencing kit (Perkin-Elmer/ABI, Palo Alto, CA) and analyzed on an ABI 3700 DNA analyzer. Sequence data were uploaded to Geospiza software (Geospiza, Seattle, WA) for quality assessment and vector sequence trimming. Clustering analysis was done using stackPACK 2.2 software (Electric Genetics, Observatory, South Africa) to identify unique sequences. Final sequences were subjected to basic local alignment search tool (BLAST) searches against the GenBank nonredundant (nr) and expressed sequence tag (est) databases using the BLASTN and BLASTX programs via netblasting. The known genes were classified functionally based on the Gene Ontology Consortium classifications (<http://www.geneontology.org>).

Construction of the bovine oocyte cDNA microarray. Over 2,000 colonies from the plated library were robotically picked and transferred to LB in 96-well plates. Following plasmid preparation as described above, the cDNA inserts from the plasmids were amplified by PCR using primers flanking the cloning site and purified using the Millipore Multiscreen PCR 96-well purification system (Millipore, Bedford, MA). The purified products were resuspended in 50% DMSO and transferred to 384-well plates as microarray source plates. To confirm that inserts from each clone were amplified and included on the microarray, 2 μ l of each purified amplicon were separated on precast agarose gels using the E-Gel 96 high-throughput agarose electrophoresis system (Invitrogen). All above procedures were performed with a Tecan Genosys 150 liquid handling robot (Tecan US, Research Triangle Park, NC). Microarrays were printed using a GeneTAC G3 arraying robot (Genomic Solutions, Ann Arbor, MI) equipped with a 48-pin head. The pin configuration yields microarrays consisting of 48 patches of spots. A microarray design was developed that allows triple spotting of each clone within a patch and an overall 12 \times 12 spots in each patch. The array also included spots of amplicon from bovine ribosomal protein L15 cDNA as positive control as well as blank spots for background control. Prior to storage, the printed slides were processed by UV cross-linking and blocked in 2% SDS.

Tissue collection and RNA preparation. Bovine adult spleen and liver samples ($n = 3$) and fetal ovaries (from cows at *day 210*, *day 240*, and *day 260* of gestation; $n = 3$) were collected at a local slaughterhouse. Age of fetuses was estimated by measuring crown-rump length (29). All samples were frozen in liquid nitrogen and stored at -80°C until use. Total RNA was isolated using the TRIzol reagent (Invitrogen) according to manufacturer's instruction. Concentrations of isolated RNA were determined by measuring absorbance at 260 nm. Purity of RNA was determined by calculating the ratio of absorbance at 260 nm and 280 nm, and integrity of RNA was determined by agarose gel electrophoresis. For microarray experi-

ments, equal amounts of total RNA from three different samples for each tissue were pooled, and mRNA was isolated using the PolyAT-tract mRNA isolation system (Promega, Madison, WI) according to manufacturer's instruction.

Microarray experiments. Microarray experiments were performed to identify differentially expressed genes in fetal ovary relative to spleen and liver. Separate experiments (fetal ovary vs. spleen and fetal ovary vs. liver) were conducted to compare gene expression differences. Approximately 2 µg of pooled mRNA from fetal ovary and spleen (or liver) were used as templates in reverse transcription reactions incorporating amino-modified dUTPs into the cDNA using the Atlas Glass cDNA labeling system (BD Biosciences). The synthesized cDNAs from the fetal ovary and spleen (or liver) were differentially labeled using *N*-hydroxysuccinate-derived Cy3 and Cy5 dyes, respectively (Amersham Biosciences, Piscataway, NJ). Labeled cDNAs were purified using a PCR purification kit (Qiagen) to remove unincorporated dyes. The Cy3- and Cy5-labeled cDNAs were then combined, and total probe volume was adjusted to 110 µl with SlideHyb 3 solution (Ambion, Austin, TX). Microarray hybridizations were conducted for 18 h using step-down temperatures from 65°C to 42°C in sealed chambers of a GeneTAC hybridization station (Genomic Solutions). Following hybridization, the microarrays were washed twice with medium stringency buffer and once with high-stringency buffer (Genomic Solutions). Washed slides were rinsed in water and dried by centrifugation in a cushioned 50-ml conical centrifuge tube. After washing, microarrays were scanned immediately using a GeneTAC LS IV microarray scanner (Genomic Solutions). GeneTAC Integrator 4.0 software was used to process the array images, find spots, integrate robot-spotting files with the microarray image, and finally to create reports of spot intensity data. The final report was retrieved as raw spot intensities in a comma-separated values file. The microarray data were then normalized using the scatter plot smoother LOWESS (6) from the statistical package SAS ("PROC LOWESS") according to the method of Yang et al. (41). LOWESS-normalized array data were imported into Microsoft Excel for further analysis as described (33). The above experiments were then repeated with reverse-labeled cDNA samples. Data for the microarray experiments can be found in GenBank's Gene Expression Omnibus (GEO) with the series accession number GSE14422.

Quantitative real-time PCR analysis. Quantitative real-time PCR was used to confirm the differentially expressed genes revealed by microarray experiments. Aliquots of individual total RNA samples used in the microarray experiments were treated with RQ1 RNase-free DNase (Promega) for 15 min at 37°C to remove contaminating genomic DNA. Two micrograms of each treated RNA sample was then converted to cDNA using SuperScript II reverse transcriptase (Invitrogen). Real-time PCR primers were designed based on the cDNA sequences of the differentially expressed genes identified from above microarray experiments (Table 1) using Primer Express version 2.0 software (Applied Biosystems, Foster City, CA). Quantitative PCR was performed in duplicate for each cDNA sample on an ABI 7000 sequence detection system (Applied Biosystems) using 2× SYBR Green Master Mix (Applied Biosystems) in 50-µl reaction volumes with 300 nM of each primer and cDNA derived from 0.2 µg of total RNA. Bovine β-actin gene (GenBank accession no.

AY141970) was used as an endogenous control for normalization. Standard curves for each gene and the endogenous control were constructed using 10-fold serial dilutions of the corresponding plasmid. Standard curves were run on the same plate as the samples. Threshold lines were adjusted to intersect amplification lines in the linear portion of the amplification curve, and cycles to threshold (C_t) were recorded. For each sample, the amount of the target gene and endogenous reference were determined from the appropriate standard curve. The amount of the target gene was then divided by the amount of the reference gene to obtain a normalized target value. The mean differences in expression levels between fetal ovary and liver or spleen were determined by two-tailed *t*-test.

Real-time PCR was also utilized to verify that the differentially expressed genes identified by microarray analysis of fetal ovaries were indeed expressed in isolated bovine oocytes from adult ovaries. Collection of germinal vesicle stage bovine oocytes was conducted as described above for library construction. Total RNA from germinal vesicle stage oocytes (in groups of 25) was isolated and simultaneously subjected to DNase treatment using the RNAqueous Micro Scale RNA isolation kit (Ambion). Generation of cDNA and real-time analysis were conducted as described above using a fraction of the cDNA equivalent to that generated from RNA isolated from 0.2 oocyte in each PCR reaction.

RESULTS

Analysis of ESTs generated from the oocyte library. Using Poly(A)⁺ RNA directly isolated from a pool of 200 bovine oocytes, we constructed a directionally cloned cDNA library. The library was estimated to contain a total of 2×10^5 independent clones by titration of the original electroporated bacterial stock. Restriction analysis showed that 13 of 18 randomly selected clones contained inserts with an average size of 450 bp. A total of 230 EST sequences from the library were obtained by 5' end sequencing of randomly selected clones (GenBank accession nos. CN544926–CN545155). The sequence information for these ESTs is also available in our database (<http://nbfgc.msu.edu>) under the BEGG (bovine egg) library. The Turnkey system of the database automatically runs BLASTN and BLASTX searches on each of the ESTs entered in the database against the GenBank database. The top five hits for each BLAST report are retained and can be accessed through a web interface to the database. Clustering analysis revealed that the 230 ESTs represent 102 unique sequences. BLAST analysis revealed that 46 of the unique sequences display significant similarity with known genes in the GenBank database. Of the remaining 56 sequences, 40 matched only to genomic or EST sequences in the database and were classified as unknown genes, and 16 did not show significant similarity to any genes or ESTs in the database and were classified as novel. The gene ontology distribution for the bovine oocyte EST sequences is shown in Fig. 1. Examples of known genes and their functions are listed in Table 2. Among

Table 1. Primers used for quantitative real-time PCR analysis

Clone Name	Forward Primer	Reverse Primer	Product, bp
BEGG11_C8	5'-GCAGCGGGAGAAAATGAAGAC-3'	5'-CAACCACATTCGCCCTCTTCTG-3'	85
BEGG20_H6	5'-GGTCTGCAAGAGCCTCTTTCTG-3'	5'-CCAGCTGCTCCACGTATTCTG-3'	74
BEGG16_B3	5'-GTGGCAGTTCTGTACGGAAAGTC-3'	5'-CACAGCAACGTTCCGAGTACATCTC-3'	72
BEGG08_G9	5'-TCAATCAGTGGACCACAAACG-3'	5'-CTTAAATGGCTTCCCCAGCTT-3'	71
BEGG16-H10	5'-CAAGTCAACTCGGAGGACAAGAG-3'	5'-GACGACGGATCTCATCGTATCTG-3'	91
BEGG19-G10	5'-TCACATCCTAGCCCGATCCA-3'	5'-CTCTGCCTCTGTGAGAGGTGACA-3'	78
Bovine β-actin	5'-CGCCATGGATGATGATATTGC-3'	5'-AAGCCGGCCTTGCACAT-3'	66

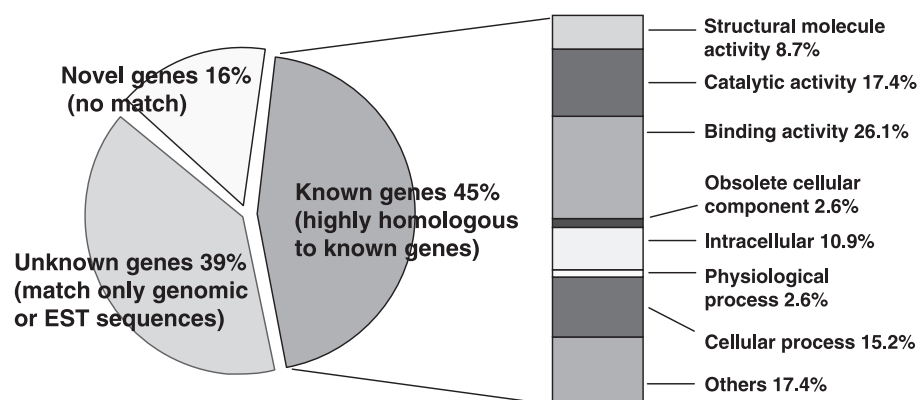


Fig. 1. Ontology classification for the oocyte expressed sequence tag (EST) sequences. The ESTs were identified based on their identity with sequences in the GenBank database. The known genes were classified functionally based on the Gene Ontology Consortium classifications (<http://www.geneontology.org>).

the known genes identified by bovine oocyte EST sequencing, some are housekeeping genes critical to function of all cell types (e.g., ribosomal protein L15) and others were previously known to be expressed in oocytes and other tissues (e.g., dynein). However, for the majority, expression in mammalian oocytes has not previously been reported (e.g., cocaine and amphetamine regulated transcript).

Characterization of the bovine oocyte cDNA microarray. A total of 2,112 anonymous clones (twenty-two 96-well plates) randomly picked from the bovine oocyte library was used to construct a cDNA microarray. The insert amplicons were spotted in triplicate to improve reproducibility and statistical analysis of microarray results. To determine the potential for false-positive expression changes, an initial microarray experiment was performed by hybridizing differentially labeled (Cy3 and Cy5) aliquots of cDNA from the same sample (fetal ovary). Since the Cy3- and Cy5-labeled samples were from identical cDNA sources, there should be no significant differences in gene expression and we would anticipate microarray analysis to yield equal fluorescence intensities for both the Cy3 and Cy5 labels at every spot on the microarray. As expected, there were few differences between the Cy3 and Cy5 signals for most of the microarray spots in this experiment (Fig. 2). Only 10 clones showed significant differential expression of >1.5-fold ($P < 0.05$; false-positive rate: $10/2,112 = 0.47\%$). This result illustrates that the microarray design and subsequent analysis procedures yielded a very low potential false-positive rate.

Identification of oocyte-expressed genes abundant in fetal ovary. The bovine oocyte microarray was used to identify genes highly expressed in fetal ovary (an enriched source of oocytes) relative to adult spleen and liver tissues. The goal of this experiment was to validate utility of the cDNA microarray and to identify additional oocyte-expressed genes abundant in fetal ovary relative to somatic tissues. Microarray experiments with fetal ovary vs. spleen revealed 24 clones with >2-fold higher expression in fetal ovary ($P < 0.05$). Likewise, microarray experiments comparing fetal ovary and liver RNA revealed 98 clones with >2-fold higher expression in fetal ovary ($P < 0.05$). Overall, 11 clones were found to be in common between the lists of genes expressed at higher levels (>2-fold) in fetal ovary than in either spleen or liver. These 11 clones were retrieved from the original source plates and subjected to DNA sequencing. Sequence analysis revealed that these 11 clones represent 7 different genes. Two clones encode for JY-1, a known bovine oocyte-specific gene (43). Eight of the remain-

Table 2. Examples of known genes identified by BLAST analysis of ESTs from the bovine oocyte library

Gene Name	Known (or potential) Functions
Nucleosome assembly protein-1	Signal transduction
RNA binding protein HQK	Cell differentiation, proliferation, apoptotic processes
Bithoraxoid-like protein	Tumor progression
Vitamin A-deficient testicular protein 5	Spermatogenesis
Ed1 gene for ectodysplasin A	Epidermal morphogenesis
Early growth response 4	Regulation of transcription
Splicing factor 30	Regulation of splicing
Dihydropyrimidinase-like 2	Developmental process of the nervous system
Ribosomal protein L15	Structural constituent of ribosome
Mitotic control protein dis3 homolog	Regulation of mitosis
Endophilin B1	Synaptic vesicle endocytosis
CGI-128 protein	Small GTPase-mediated signal transduction.
Rab5-interacting protein	Homotypic endosome fusion
DNA-dependent PK catalytic subunit	DNA repair, signal transduction
Oxysterol binding protein-like 1A	Control of cellular lipid metabolism
Pyruvate carboxylase	Gluconeogenesis
cAMP-dependent PK type I regulatory subunit	Signal transduction
Ring-box 1	Chromosome metabolism, cell cycle control
Junctional adhesion molecule 3	Homotypic and heterotypic intercellular interactions
MUS81 endonuclease	Recombination repair
Interferon-gamma prepeptide	Host defense against infection with <i>Mycobacterium tuberculosis</i>
U6 snRNA-associated Sm-like protein	Nuclear mRNA splicing
Peanut-like 2	Cytokinesis
Dynein, cytoplasmic light polypeptide	Microtubule motor activity
Ubiquitin-protein isopeptide ligase	Motor neuron death
Nerve injury gene 283	Ubiquitin-mediated protein modification
YY1 transcription factor	Regulation of gene expression and initiation of transcription
Bone morphogenetic protein receptor IB	Bone formation during embryogenesis
DEAD/H box polypeptide 10	RNA metabolism, genome replication, repair and expression
Colony stimulating factor 3 receptor	Hematopoiesis, inflammatory reactions
Importin α 1b	Nuclear import of proteins
CDC28 protein kinase 1	Control of cell-cycle progression
Cocaine- and amphetamine-regulated transcript	Neuronal development and survival

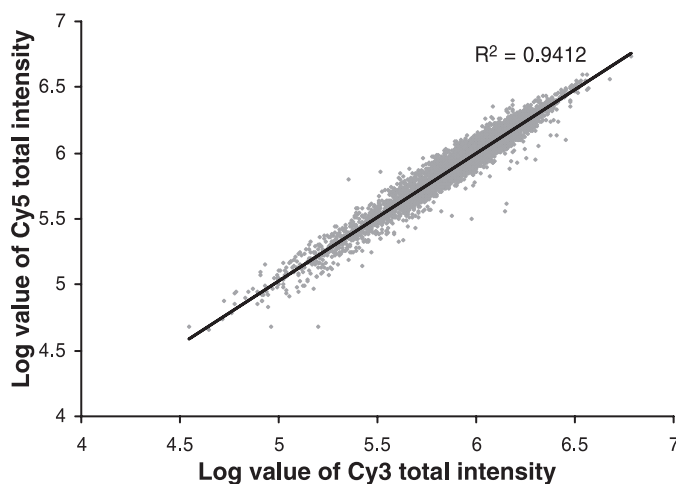


Fig. 2. Scatter plot illustrating false-positive rate via comparison of Cy3 and Cy5 intensities for the same cDNA sample. Differentially labeled (Cy3 and Cy5) cDNA samples generated from an identical source (fetal ovary) were hybridized to the oocyte microarray to determine the false-positive rate. As expected, there were few differences between the Cy3 and Cy5 signal intensities for most of the microarray spots, and the false-positive rate was <1%.

ing nine clones encode for five previously identified genes (ribosomal protein L7a, dynein light chain, Doc2 α , calmodulin, leucine-rich protein). The final clone encodes for a completely novel transcript (Table 3). Since a heterogeneous tissue source including both oocytes and somatic cells (fetal ovary) was used as an enriched source of bovine oocytes for microarray analysis, expression of above-described genes in isolated bovine oocytes (from adult ovaries) was confirmed by real-time RT-PCR analysis (Fig. 3). Quantitative real-time PCR assay also confirmed that expression of the five previously identified genes (described above) was significantly higher in fetal ovary compared with either liver or spleen (Fig. 4, $P < 0.05$). Real-time PCR analysis showed a tendency for higher expression of the novel gene (clone Begg20_H6) in fetal ovary compared with spleen ($P < 0.1$) but failed to detect differential expression of this gene between fetal ovary and liver (Fig. 4). Expression of the novel gene was also detected in five additional bovine tissues (kidney, heart, lung, thymus, and muscle) by real-time PCR analysis (data not shown). However, abundance of this transcript present in all tissues examined (including fetal ovaries) was extremely low relative to the other five identified genes (data not shown).

DISCUSSION

The oocyte sustains a highly orchestrated program of gene expression that is essential for oocyte development, maturation, fertilization, and early embryogenesis. Although several key factors of oocyte origin that are critical for ovarian follicle development and early embryogenesis are known, the composition of the oocyte transcriptome and the identities and functions of the majority of the oocyte-expressed genes remain largely unknown. Results of the present studies illustrate the potential utility of the bovine oocyte cDNA library and microarray for characterization of the oocyte transcriptome and changes in abundance of mRNAs for oocyte-expressed genes potentially involved in the above physiological processes.

Conventional cDNA library construction methods require micrograms of mRNA as starting material. The poly(A)⁺ RNA content of a single bovine oocyte is ~53 pg (20). Thus it would be impractical to collect enough bovine oocytes to isolate sufficient RNA to construct a cDNA library using standard procedures. To overcome this limitation, we used a PCR-based cDNA amplification (SMART) method to construct the oocyte library. Although libraries generated with this method are generally of significantly reduced quality (small insert size) compared with libraries constructed using traditional procedures, it is a viable approach to allow for EST sequencing and generation of novel sequences from individual cell types that are rare or of low abundance in tissues represented in existing EST libraries. In the present studies, 16 of 102 unique sequences identified were novel, despite the fact that 327,664 (TIGR Cattle Gene Index, Release 9.0) bovine EST sequences have been reported to date, with 29,666 sequences derived from mixed tissue libraries in which ovarian tissue (containing oocytes) were included (32).

Our initial analysis of ESTs from the bovine oocyte library revealed novel information about composition of the oocyte transcriptome. Many ESTs represent genes whose expression in oocytes, to our knowledge, has not been reported previously. Interestingly, several ESTs encode for genes previously reported to be primarily of neural origin. For example, one EST of particular interest encodes for cocaine- and amphetamine-regulated transcript (CART; Ref. 10), a potent hypothalamic anorectic peptide (18). Previous reports indicated that CART is expressed solely in the central nervous system or in cells of neuronal origin (35), but the presence of CART mRNA in mammalian oocytes has not been reported previously. Our

Table 3. Clones show consistently higher levels of expression in fetal ovary in comparison to both spleen and liver

Clone Name	Fetal Ovary vs. Spleen		Fetal Ovary vs. Liver		Matching Sequence in GenBank
	Mean expression ratio	P value	Mean expression ratio	P value	
BEGG11_C8	3.00	0.0011	2.16	0.0018	<i>Bos taurus</i> leucine-rich protein
BEGG22_D9	2.85	0.0004	2.83	0.0007	<i>B. taurus</i> leucine-rich protein
BEGG20_H6	2.58	0.0022	2.47	0.0110	Novel
BEGG16_B3	2.84	0.0015	3.03	0.0009	<i>Homo sapiens</i> calmodulin
BEGG08_G9	3.44	0.0013	2.08	0.0114	<i>B. taurus</i> cytoplasmic dynein light chain
BEGG07_B8	3.14	0.0039	2.21	0.0020	<i>B. taurus</i> cytoplasmic dynein light chain
BEGG16_H10	2.48	0.0033	2.30	0.0027	<i>Mus musculus</i> ribosomal protein L7a
BEGG19_G10	2.65	0.0001	2.13	0.0010	<i>M. musculus</i> Doc2 α
BEGG15_G10	2.24	0.0014	3.02	0.0001	<i>M. musculus</i> Doc2 α
BEGG02_E4	2.17	0.0016	3.05	0.0072	<i>B. taurus</i> JY-1
BEGG15_E9	2.13	0.0098	2.81	0.0157	<i>B. taurus</i> JY-1

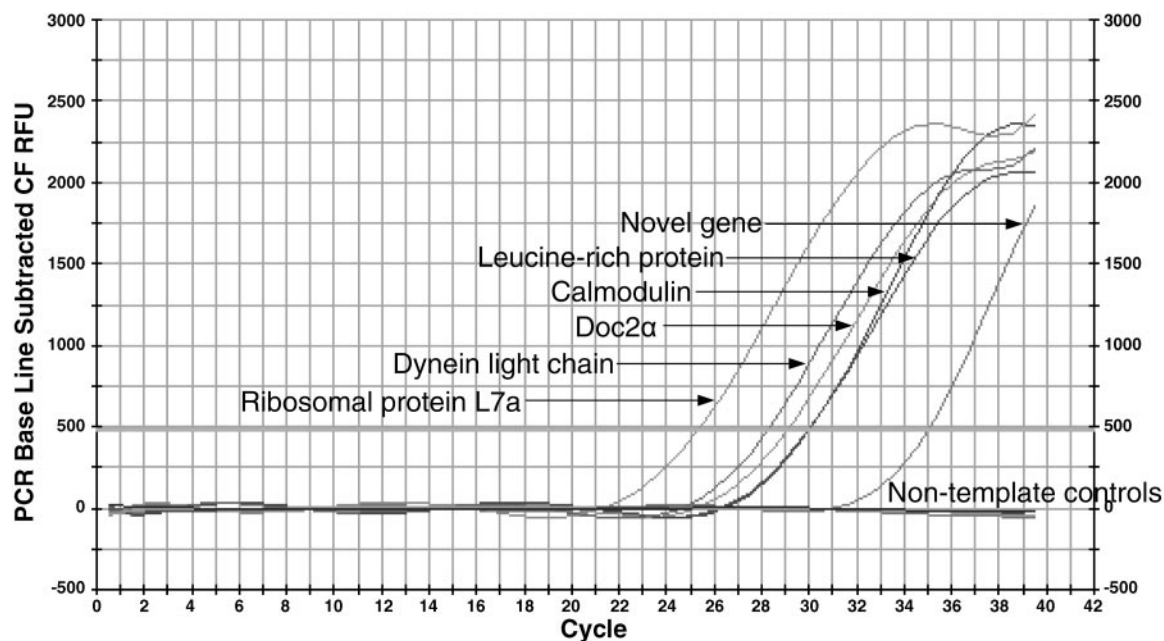


Fig. 3. Confirmation of expression of 6 differentially expressed genes (identified by microarray experiments) in isolated bovine oocytes (from adult ovary) using real-time PCR analysis. CF RFU, curve fit relative fluorescence unit.

recent findings indicate that CART is a novel local regulator of ovarian follicular atresia (17). Other bovine oocyte ESTs of interest that fit this criteria encode for Doc2 α (25), nerve injury gene (1), endophilin B1 (22), and early growth response gene 4 (28).

In addition to the known genes described, nearly 16% of the EST sequences generated in the present studies showed no homology to any sequences, known genes, or ESTs in the GenBank database and therefore were classified as novel. Identification of such a high percentage of novel ESTs is significant, because there are currently over 5 million human, nearly 4 million mouse, and ~322,000 bovine EST sequences deposited in GenBank (TIGR Gene Indices, <http://www.tigr.org>). The unique oocyte origin of these ESTs could account for their novelty, as oocyte ESTs are likely to be grossly underrepresented in mixed tissue or ovarian EST libraries sequenced. Thus the novel transcripts may potentially encode for genes specifically or predominantly expressed in oocytes. Indeed, we have recently confirmed oocyte-specific expression of two of the novel transcripts identified via oocyte EST sequencing (43, 44).

Because of the high cost of sequencing, many gene discovery experiments are performed using anonymous cDNA microarrays (5, 14, 37). Although there are obvious limitations, such arrays generally are easier and cheaper to construct than UniGene or oligonucleotide arrays, and identity of differentially expressed genes of interest revealed using such arrays can be determined retrospectively. Using randomly picked EST clones from the bovine oocyte library, we constructed an oocyte-specific cDNA microarray. The quality of this array was first tested and experiments revealed a very low false-positive rate (<1%). To demonstrate utility of this microarray, we then conducted experiments to identify genes abundantly expressed in fetal ovary (an enriched source of oocytes) compared with other somatic tissues. A similar approach has been

applied successfully by (39) to identify a novel gene expressed in mouse oocytes. As was true for library construction, execution of such experiments using standard amounts of RNA from isolated oocytes would not have been practical or would have required use of linear amplification procedures. Thus fetal ovaries were utilized as an enriched source of oocyte RNA in the microarray experiments. Oocytes present in fetal ovaries are primarily at the earlier stages of folliculogenesis and thus are immature and not fully grown. Therefore, potential genes differentially expressed specifically in full-grown, mature oocytes cannot be identified using this strategy. Despite the limitation, one gene (JY-1) that is known to be oocyte specific in cattle (43) was identified using this approach. The other six genes [ribosomal protein L7a, dynein light chain, Doc2 α , calmodulin, leucine-rich protein, and the novel gene clone (Begg20_H6)] identified as differentially expressed in fetal ovary vs. liver and spleen do not appear to be oocyte specific (2, 15, 16, 25). However, differential (increased) expression of five of the six genes in fetal ovary vs. spleen and liver was confirmed by real-time PCR analysis. Results of real-time PCR experiments also confirmed that mRNA for the above genes identified via microarray experiments (using fetal ovary RNA) is present in isolated bovine oocytes (from adult ovaries). Thus we are confident that identified genes are truly oocyte expressed. Great care was undertaken in the present studies to only select oocytes for RNA isolation that were completely denuded of adjacent cumulus cells, and the likelihood of cumulus cell contribution to RNA obtained is very minimal. However, we cannot draw conclusions from present experiments on the relative contribution of other ovarian cell types, including cumulus cells, to the high levels of expression detected in fetal ovaries via microarray analyses, and we acknowledge that alternative approaches, such as in situ hybridization, would help further verify oocyte expression of identified genes of interest.

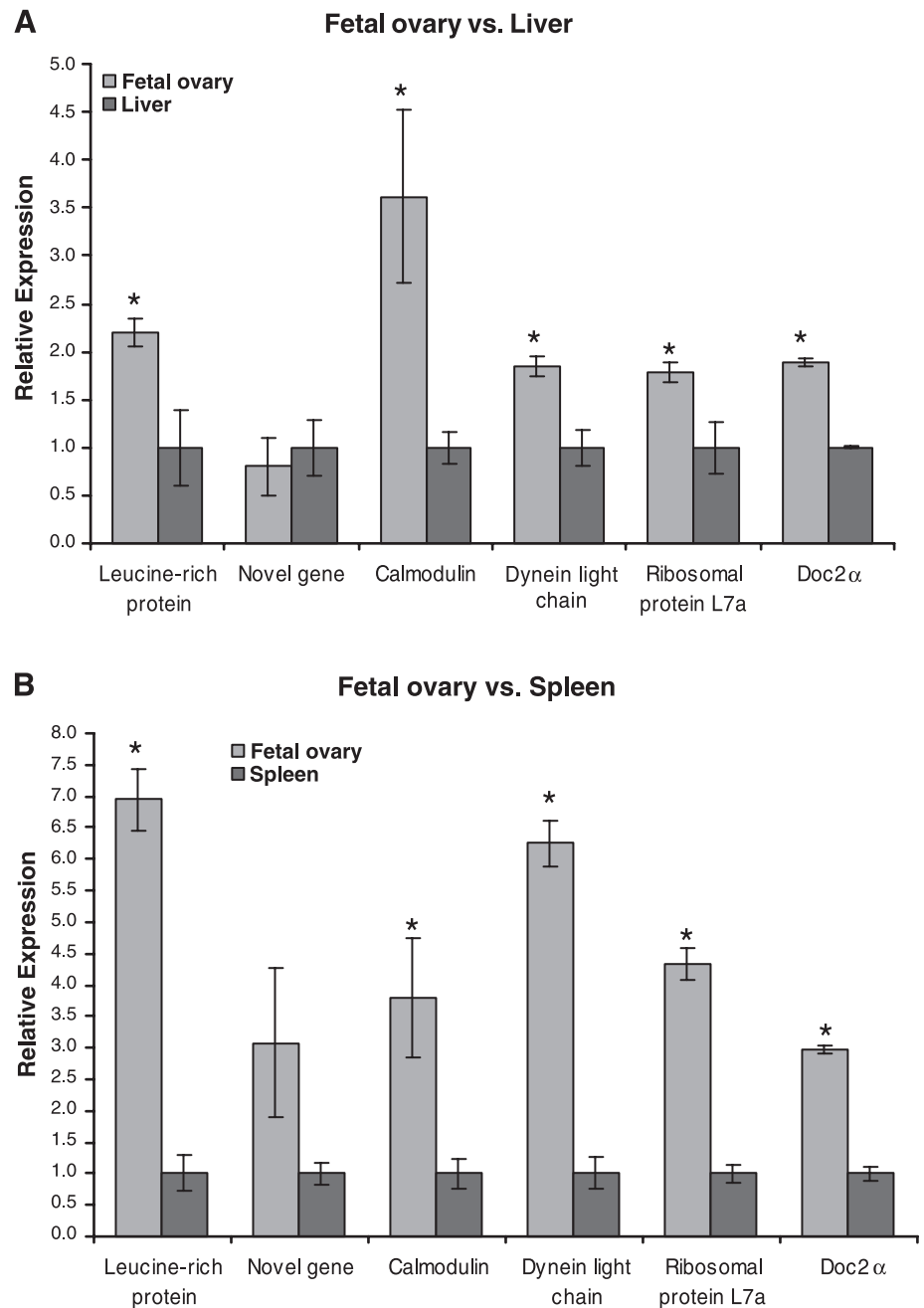


Fig. 4. Quantitative real-time PCR confirmation of differential expression in fetal ovary vs. liver (A) and fetal ovary vs. spleen (B) for the 6 differentially expressed genes identified by microarray analysis. * $P < 0.05$

Microarray technologies have proven to be valuable for understanding biological pathways important in various physiological processes. Microarray resources are now available for many species, including agriculturally important species. In cattle, a number of relatively small cDNA microarrays using ESTs derived from specific tissue types have been reported previously (3, 33, 42). Recently, a large bovine microarray containing over 18,000 EST clones was developed (34). Although this array covers a significant portion of the bovine genome, transcripts of oocyte origin may be underrepresented, since the ESTs used for construction of this large array were derived from libraries of mixed tissue origin. Indeed, BLAST analysis of the oocyte EST sequences generated in the present studies vs. the 18,000 EST clone set (<http://nbfgc.msu.edu>)

revealed that 53 of the 102 oocyte ESTs generated in the present studies are not represented on the large array. Therefore, the bovine oocyte-specific array generated in this study should supplement existing bovine microarray resources and prove of utility to researchers interested in functional genomics studies of oocyte biology and early embryogenesis in cattle.

In summary, we have described the generation of a bovine oocyte cDNA library and an oocyte-specific microarray. Initial characterization of a limited number of EST clones from the library has revealed novel information about composition of the bovine oocyte transcriptome and a significant portion of unique sequences. Experiments using the oocyte microarray have led to the identification of several oocyte-expressed genes abundant in fetal ovary and illustrated the utility of this novel

resource for future functional genomic studies of oocyte biology and early embryogenesis in cattle.

ACKNOWLEDGMENTS

We thank Dr. Osman Patel, Anilkumar Bettegowda, and Dr. Jose Cibelli for help with oocyte collection and processing, and we thank Dr. Jonathan Green at the University of Missouri-Columbia for helpful insight on library construction.

GRANTS

This work was supported through funds provided by Rackham Foundation, the Michigan Agriculture Experiment Station, the Michigan State University Office of the Vice President for Research and Graduate Studies, and the West Virginia University Agriculture and Forestry Experiment Station. This paper is published with the approval of the Director of the West Virginia University Agriculture and Forestry Experiment Station as scientific paper no. 2873.

REFERENCES

- Araki T, Nagarajan R, and Milbrandt J. Identification of genes induced in peripheral nerve after injury. Expression profiling and novel gene discovery. *J Biol Chem* 276: 34131–34141, 2001.
- Armes N and Fried M. The genomic organization of the region containing the *Drosophila melanogaster* rpl7a (Surf-3) gene differs from those of the mammalian and avian Surf1 loci. *Mol Cell Biol* 15: 2367–2373, 1995.
- Band MR, Olmstead C, Everts RE, Liu ZL, and Lewin HA. A 3800 gene microarray for cattle functional genomics: comparison of gene expression in spleen, placenta, and brain. *Anim Biotechnol* 13: 163–172, 2002.
- Burns KH, Viveiros MM, Ren Y, Wang P, DeMayo FJ, Frail DE, Eppig JJ, and Matzuk MM. Roles of NPM2 in chromatin and nucleolar organization in oocytes and embryos. *Science* 300: 633–636, 2003.
- Clark J, Edwards S, John M, Flohr P, Gordon T, Maillard K, Giddings I, Brown C, Bagherzadeh A, Campbell C, Shipley J, Wooster R, and Cooper CS. Identification of amplified and expressed genes in breast cancer by comparative hybridization onto microarrays of randomly selected cDNA clones. *Genes Chromosomes Cancer* 34: 104–114, 2002.
- Cleveland WS and Grosse E. Computational methods for local regression. *Statistics Computing* 1: 47–62, 1991.
- Davis GH, McEwan JC, Fennessy PF, Dodds KG, McNatty KP, and O WS. Infertility due to bilateral ovarian hypoplasia in sheep homozygous (FecXI FecXI) for the Inverdale prolificacy gene located on the X chromosome. *Biol Reprod* 46: 636–640, 1992.
- Diatchenko L, Lau YF, Campbell AP, Chenchik A, Moqadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov ED, and Siebert PD. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci USA* 93: 6025–6030, 1996.
- Dong J, Albertini DF, Nishimori K, Kumar TR, Lu N, and Matzuk MM. Growth differentiation factor-9 is required during early ovarian folliculogenesis. *Nature* 383: 531–535, 1996.
- Douglass J and Daoud S. Characterization of the human cDNA and genomic DNA encoding CART: a cocaine- and amphetamine-regulated transcript. *Gene* 169: 241–245, 1996.
- Eppig JJ. Oocyte control of ovarian follicular development and function in mammals. *Reproduction* 122: 829–838, 2001.
- Eppig JJ, Wigglesworth K, and Pendola FL. The mammalian oocyte orchestrates the rate of ovarian follicular development. *Proc Natl Acad Sci USA* 99: 2890–2894, 2002.
- Galloway SM, McNatty KP, Cambridge LM, Laitinen MP, Juengel JL, Jokiranta TS, McLaren RJ, Luiro K, Dodds KG, Montgomery GW, Beattie AE, Davis GH, and Ritvos O. Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner. *Nat Genet* 25: 279–283, 2000.
- Hayward RE, Derisi JL, Alfadhli S, Kaslow DC, Brown PO, and Rathod PK. Shotgun DNA microarrays and stage-specific gene expression in *Plasmodium falciparum* malaria. *Mol Microbiol* 35: 6–14, 2000.
- Jiang J, Yu L, Huang X, Chen X, Li D, Zhang Y, Tang L, and Zhao S. Identification of two novel human dynein light chain genes, DNLC2A and DNLC2B, and their expression changes in hepatocellular carcinoma tissues from 68 Chinese patients. *Gene* 281: 103–113, 2001.
- Klee CB and Vanaman TC. Calmodulin. *Adv Protein Chem* 35: 213–321, 1982.
- Kobayashi Y, Jimenez-Krassel F, Ireland JJ, and Smith GW. Evidence that cocaine- and amphetamine-regulated transcript (CART) is a novel intraovarian regulator of follicular atresia. *Biology of Reproduction 37th SSR Meeting, Vancouver, August 1–4, 2004*.
- Kristensen P, Judge ME, Thim L, Ribel U, Christjansen KN, Wulff BS, Clausen JT, Jensen PB, Madsen OD, Vrang N, Larsen PJ, and Hastrup S. Hypothalamic CART is a new anorectic peptide regulated by leptin. *Nature* 393: 72–76, 1998.
- Latham KE, Bautista FD, Hirao Y, O'Brien MJ, and Eppig JJ. Comparison of protein synthesis patterns in mouse cumulus cells and mural granulosa cells: effects of follicle-stimulating hormone and insulin on granulosa cell differentiation in vitro. *Biol Reprod* 61: 482–492, 1999.
- Lequarre AS, Traverso JM, Marchandise J, and Donnay I. Poly(A) RNA is reduced by half during bovine oocyte maturation but increases when meiotic arrest is maintained with cdk inhibitors. *Biol Reprod* 71: 425–431, 2004.
- Liang P and Pardee AB. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257: 967–971, 1992.
- Modregger J, Schmidt AA, Ritter B, Huttner WB, and Plomann M. Characterization of endophilin B1b, a brain-specific membrane-associated lysophosphatidic acid transferase with properties distinct from endophilin A1. *J Biol Chem* 278: 4160–4167, 2003.
- Mohan M, Ryder S, Claypool PL, Geisert RD, and Malayer JR. Analysis of gene expression in the bovine blastocyst produced in vitro using suppression-subtractive hybridization. *Biol Reprod* 67: 447–453, 2002.
- Monk M, Holding C, and Goto T. Isolation of novel developmental genes from human germ cell, oocyte and embryo cDNA by differential display. *Reprod Fertil Dev* 13: 51–57, 2001.
- Naito A, Orita S, Wanaka A, Sasaki T, Sakaguchi G, Maeda M, Igarashi H, Tohyama M, and Takai Y. Molecular cloning of mouse Doc2a and distribution of its mRNA in adult mouse brain. *Brain Res Mol Brain Res* 44: 198–204, 1997.
- Natale DR, De Sousa PA, Westhusin ME, and Watson AJ. Sensitivity of bovine blastocyst gene expression patterns to culture environments assessed by differential display RT-PCR. *Reproduction* 122: 687–693, 2001.
- Natale DR, Kidder GM, Westhusin ME, and Watson AJ. Assessment by differential display-RT-PCR of mRNA transcript transitions and alpha-amanitin sensitivity during bovine preattachment development. *Mol Reprod Dev* 55: 152–163, 2000.
- Nissen PH, Thomsen B, Offenberger H, Thomsen PD, and Bendixen C. Cloning and characterization of the bovine EGR4 gene and evaluation as candidate gene for bovine spinal dysmyelination. *Anim Genet* 34: 124–131, 2003.
- Richardson C, Jones PC, Barnard V, Hebert CN, Terlecki S, and Wijeratne WV. Estimation of the developmental age of the bovine fetus and newborn calf. *Vet Rec* 126: 279–284, 1990.
- Robert C, Barnes FL, Hue I, and Sirard MA. Subtractive hybridization used to identify mRNA associated with the maturation of bovine oocytes. *Mol Reprod Dev* 57: 167–175, 2000.
- Robert C, Gagne D, Bousquet D, Barnes FL, and Sirard MA. Differential display and suppressive subtractive hybridization used to identify granulosa cell messenger RNA associated with bovine oocyte developmental competence. *Biol Reprod* 64: 1812–1820, 2001.
- Smith TP, Grosse WM, Freking BA, Roberts AJ, Stone RT, Casas E, Wray JE, White J, Cho J, Fahrenkrug SC, Bennett GL, Heaton MP, Laegreid WW, Rohrer GA, Chitko-McKown CG, Pertege G, Holt I, Karamycheva S, Liang F, Quackenbush J, and Keele JW. Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle. *Genome Res* 11: 626–630, 2001.
- Suchyta SP, Sipkovsky S, Halgren RG, Kruska R, Eftman M, Weber-Nielsen M, Vandehaar MJ, Xiao L, Tempelman RJ, and Coussens PM. Bovine mammary gene expression profiling using a cDNA microarray enhanced for mammary-specific transcripts. *Physiol Genomics* 16: 8–18, 2003. First published October 14, 2003; doi:10.1152/physiolgenomics.00028.2003.
- Suchyta SP, Sipkovsky S, Kruska R, Jeffers A, McNulty A, Coussens MJ, Tempelman RJ, Halgren RG, Saama PM, Bauman DE, Boisclair YR, Burton JL, Collier RJ, DePeters EJ, Ferris TA, Lucy MC, McGuire MA, Medrano JF, Overton TR, Smith TP, Smith GW,

- Sonstegard TS, Spain JN, Spiers DE, Yao J, and Coussens PM. Development and testing of a high-density cDNA microarray resource for cattle. *Physiol Genomics* 15: 158–164, 2003. First published September 16, 2003; doi:10.1152/physiolgenomics.00094.2003.
35. Thim L, Kristensen P, Nielsen PF, Wulff BS, and Clausen JT. Tissue-specific processing of cocaine- and amphetamine-regulated transcript peptides in the rat. *Proc Natl Acad Sci USA* 96: 2722–2727, 1999.
36. Tong ZB, Gold L, Pfeifer KE, Dorward H, Lee E, Bondy CA, Dean J, and Nelson LM. Mater, a maternal effect gene required for early embryonic development in mice. *Nat Genet* 26: 267–268, 2000.
37. van Hemert S, Ebbelaar BH, Smits MA, and Rebel JM. Generation of EST and microarray resources for functional genomic studies on chicken intestinal health. *Anim Biotechnol* 14: 133–143, 2003.
38. Wu X, Viveiros MM, Eppig JJ, Bai Y, Fitzpatrick SL, and Matzuk MM. Zygote arrest 1 (Zar1) is a novel maternal-effect gene critical for the oocyte-to-embryo transition. *Nat Genet* 33: 187–191, 2003.
39. Yan C, Pendola FL, Jacob R, Lau AL, Eppig JJ, and Matzuk MM. Oosp1 encodes a novel mouse oocyte-secreted protein. *Genesis* 31: 105–110, 2001.
40. Yan C, Wang P, DeMayo J, DeMayo FJ, Elvin JA, Carino C, Prasad SV, Skinner SS, Dunbar BS, Dube JL, Celeste AJ, and Matzuk MM. Synergistic roles of bone morphogenetic protein 15 and growth differentiation factor 9 in ovarian function. *Mol Endocrinol* 15: 854–866, 2001.
41. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, and Speed TP. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 30: e15, 2002.
42. Yao J, Burton JL, Saama P, Sipkovsky S, and Coussens PM. Generation of EST and cDNA microarray resources for the study of bovine immunobiology. *Acta Vet Scand* 42: 391–405, 2001.
43. Yao J, Huang R, Kobayashi Y, Bettgowda A, Coussens PM, Ireland JJ, and Smith GW. Identification and characterization of expression of JY-1: a novel oocyte specific gene in cattle. *Biol Reprod* 68: 112, 2003.
44. Yao J and Smith GW. Identification of a novel oocyte-expressed transcript in cattle. *Biology of Reproduction 37th SSR meeting, Vancouver, August 1–4, 2004.*

