

Research article

Open Access

## Cross-species hybridizations on a multi-species cDNA microarray to identify evolutionarily conserved genes expressed in oocytes

Maud Vallée<sup>1</sup>, Claude Robert<sup>1</sup>, Steve Méthot<sup>2</sup>, Marie-France Palin<sup>2</sup> and Marc-André Sirard\*<sup>1</sup>

Address: <sup>1</sup>Centre de Recherche en Biologie de la Reproduction, Département des Sciences Animales, Université Laval, Québec, G1K 7P4, Canada and <sup>2</sup>Dairy and Swine Research and Development Center, Agriculture and Agri-Food Canada, Lennoxville, Québec, J1M 1Z3, Canada

Email: Maud Vallée - [maud.vallee.1@ulaval.ca](mailto:maud.vallee.1@ulaval.ca); Claude Robert - [clauderobert@san.ulaval.ca](mailto:clauderobert@san.ulaval.ca); Steve Méthot - [methots@agr.gc.ca](mailto:methots@agr.gc.ca); Marie-France Palin - [palinmf@agr.gc.ca](mailto:palinmf@agr.gc.ca); Marc-André Sirard\* - [Marc-Andre.Sirard@CRBR.ulaval.ca](mailto:Marc-Andre.Sirard@CRBR.ulaval.ca)

\* Corresponding author

Published: 10 May 2006

Received: 13 October 2005

*BMC Genomics* 2006, **7**:113 doi:10.1186/1471-2164-7-113

Accepted: 10 May 2006

This article is available from: <http://www.biomedcentral.com/1471-2164/7/113>

© 2006 Vallée et al; licensee BioMed Central Ltd.

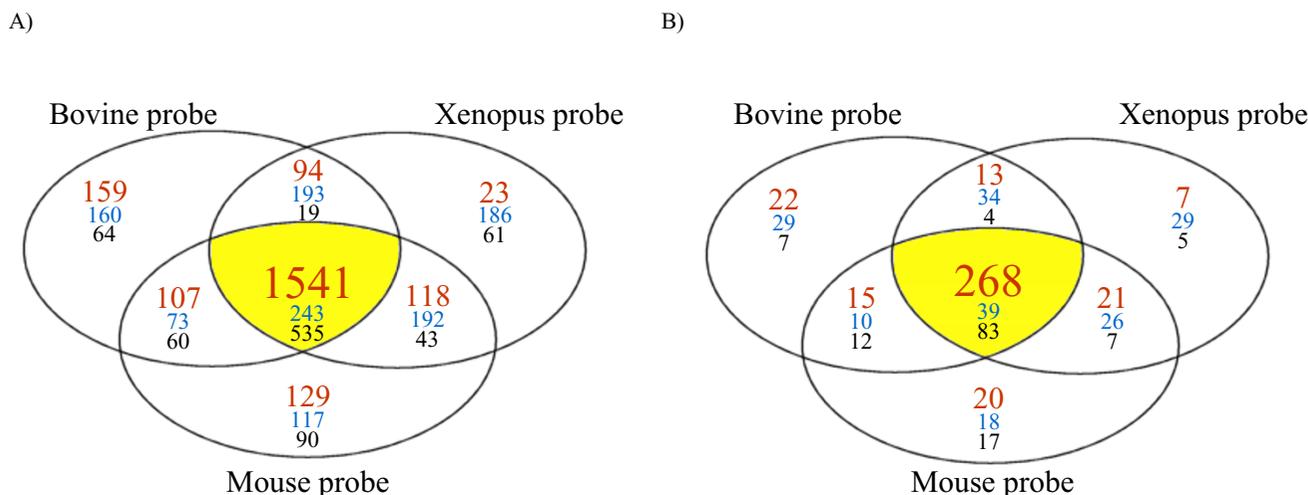
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Comparative genomic analysis using cDNA microarray is a new approach and a useful tool to identify important genetic sequences or genes that are conserved throughout evolution. Identification of these conserved sequences will help elucidate important molecular mechanisms or pathways common to many species. For example, the stockpiled transcripts in the oocyte necessary for successful fertilization and early embryonic development still remain relatively unknown. The objective of this study was to identify genes expressed in oocytes and conserved in three evolutionarily distant species.

**Results:** In this study we report the construction of a multi-species cDNA microarray containing 3,456 transcripts from three distinct oocyte-libraries from bovine, mouse and *Xenopus laevis*. Following the cross-species hybridizations, data analysis revealed that 1,541 positive hybridization signals were generated by oocytes of all three species, and 268 of these are preferentially expressed in the oocyte. Data reproducibility analyses comparing same-species to cross-species hybridization indicates that cross-species hybridizations are highly reproducible, thus increasing the confidence level in their specificity. A validation by RT-PCR using gene- and species-specific primers confirmed that cross-species hybridization allows the production of specific and reliable data. Finally, a second validation step through gene-specific microarray hybridizations further supported the validity of our cross-species microarray results. Results from these cross-species hybridizations on our multi-species cDNA microarray revealed that *SMFN* (Small fragment nuclease), *Spin* (Spindlin), and *PRMT1* (Protein arginine methyltransferase 1) are transcripts present in oocytes and conserved in three evolutionarily distant species.

**Conclusion:** Cross-species hybridization using a multi-species cDNA microarray is a powerful tool for the discovery of genes involved in evolutionarily conserved molecular mechanisms. The present study identified conserved genes in the oocytes of three distant species that will help understand the unique role of maternal transcripts in early embryonic development.



Present; 16/16 replicates

Ambiguous; 8-15/16 replicates

Absent; 0-7/16 replicates

**Figure 1**

**Venn diagram representing clones present in oocytes of one, two or all three species.** Clones are considered as present when their signal intensities are above threshold in every replicates (16/16), they are considered as ambiguous when their signal intensities are above threshold in 8 to 15 replicates, and finally are considered absent when their signal intensities are above threshold in less than 7 replicates. A) All the clones on the microarray slides are considered, B) only the clones preferentially expressed in the oocyte are considered. For the clones present in all three species, the number of clones derived from each oocyte library is given in table 1.

**Background**

Evolutionarily distant animals exhibit common mechanisms and pathways involved in early development. One of the characteristics conserved across species is that the oocyte arrests during the first meiotic division, where a stockpile of transcripts and proteins that are synthesized and stored will subsequently support early development [1,2]. The maternal transcripts that are stored in the oocyte will drive meiotic resumption of the oocyte and early cleavage divisions of the embryo up to zygotic genome activation [3]. In Xenopus, major zygotic genomic activation takes place after 12 rapid synchronous cleavage divisions generating > 4,000 cells, while in the bovine and mouse it occurs at the eight- to sixteen-cell stages and two-cell stage, respectively [4-6]. It is speculated that several hundred maternal transcripts play an active role in early development, although only a few have been identified to date [7]. Information for only a limited number of these genes is currently known, meaning that our basic understanding of gene expression patterns driving pre-implantation development is still very restricted. A few maternally expressed genes with important functions

related either to oogenesis, folliculogenesis, fertilization, and or early embryonic development have been discovered in the mouse oocyte, such as *Mos* (Moloney sarcoma oncogene) [8], *Zp3* (Zona pellucida glycoprotein 3) [9], *Zp2* (Zona pellucida glycoprotein 2) [10], *Zp1* (Zona pellucida glycoprotein 1) [11], *Gdf9* (Growth differentiation factor 9) [12], *Fig 1α* (Factor in the germline alpha) [13], *Bmp15* (Bone morphogenetic protein 15) [14], *H1foo* (H1 histone family, member O, oocyte-specific) [15], *Zar1* (Zygote arrest 1) [16], *Mater* (Maternal antigen that embryos require) [17], *Npm2* (Nucleophosmin/nucleoplasmin 2) [18], and *Msy2* (Y box protein 2) [19]. Some of these oocyte-specific genes have been identified via model organisms; for instance the mammalian oocyte-specific cleavage stage linker histone *H1foo* and the *Msy2* gene were both first identified in the *Xenopus laevis* oocyte [15,19]. Therefore, our ability to compare the conserved maternal genes across evolutionarily distant species that share common mechanisms, such as the *Xenopus laevis*, mouse and the bovine, will contribute by identifying functionally important genes involved in early development.

**Table 1: Origin of clones conserved in all three species.**

	Bovine oocyte library	Mouse oocyte library	Xenopus oocyte library	Total
Conserved in all three species	718	476	347	1541
Conserved in all three species and preferentially expressed in oocyte	120	96	52	268

In the past, embryonic development has been studied through time-consuming gene-by-gene analyses that characterized only very specific molecular mechanisms. The need for large-scale genomic approaches is required to analyze a large cohort of genes simultaneously. Suppressive subtractive hybridization (SSH) and differential display (DDRT) have been successfully applied to early developmental studies [20-25]. Analysis of expressed sequence tags (EST) has also been used to study the gene expression that occurs during early development [7,26]. Furthermore, the large amount of sequence information that has been placed into public databases over the last decade has allowed for the use of In Silico approaches to identify oocyte-specific transcripts [27,28]. In Silico approaches are ideal for mouse studies due to the large quantity of genomic information available for this species. Unfortunately, genomic information is quite limited for bovine and Xenopus, which renders this approach less feasible in this case. Thus, when the objective is to compare those three species, the In Silico analysis cannot be used as the main approach, but the information that it can provide can help support as in this case, the microarray results. Recently, microarrays have been widely used for large-scale transcriptome analyses and have proven to be a powerful approach to study molecular mechanisms underlying early development [24,29-33]. However, DNA arrays are currently available for only a limited number of species and to overcome this limitation, cross-species hybridization has been utilized as one potential solution. For example, human arrays have been used to study gene expression patterns in both the swine and bovine species [34,35]. Gene expression profiles have been compared between the human and canine, pig, bovine, and chimpanzee using human arrays [29,36-39]. While studies using cross-species hybridization have used different platforms (oligo or cDNA arrays), they have all used arrays that were designed based on the sequence of only one species. In an attempt to address this cross-species issue, another study constructed a multiprimate cDNA array to study the effect of sequence divergence on gene expression analysis [40]. It is now possible to apply microarrays beyond the conventional usage, proving that this technique can be flexible, however one has to be careful and respect the limits of cross-species hybridization.

Here we report the construction and hybridization of a multi-species cDNA array representing 3,456 oocyte tran-

scripts from bovine, mouse and Xenopus laevis. A total of 1,152 clones from each species were randomly selected from three distinct oocyte-subtracted libraries generated through SSH for the construction of this multi-species cDNA microarray. Hybridizations allowed the identification of candidate genes conserved in those three species and those candidates that are preferentially expressed in the oocyte. We show herein that cross-species comparison using these arrays is a powerful tool for the discovery of evolutionarily conserved molecular mechanisms related to the unique genes and functions found in the oocyte.

## Results

### Microarray data analysis

#### Transcripts present in oocytes

The transcripts considered as present in the bovine, mouse or Xenopus laevis oocytes, were selected based on the following criteria. Normalized and log transformed data points above a calculated threshold were considered as "present". The threshold was calculated with the intensity values of the 424 negative control spots present on the array. 75% of the clones showed signal intensity above the threshold, thus considered as expressed in oocytes of at least one species. The distribution of the expressed clones was as follows; 35% bovine clones, 31% mouse clones and 34% Xenopus laevis clones. The percentage of expressed clones is in close agreement with the percentage of insert-containing clones found in the libraries spotted on our array [24].

#### Transcripts common in oocytes of all three species

The clone distribution for all the clones on the array is presented in a Venn diagram according to their detection in oocytes of one, two, or three species based on a three level score (present, ambiguous, absent) (Fig. 1A). Transcripts that were above the calculated threshold for all data points (48 data points were generated for each clone; 4 spotted clones × 4 hybridizations replicates × 3 species) were selected as transcripts present in oocytes of all three species (Table 1, Fig. 1A). These analyses revealed that 45% of the transcripts (1,541) gave hybridization signals above the calculated threshold, thus were considered as present in oocytes from all three species. More specifically 718 transcripts are from the bovine oocyte library, 476 and 347 transcripts are from the mouse and Xenopus oocyte libraries, respectively (Table 1).

**Table 2: Clones preferentially expressed in oocyte and conserved across species.**

Species <sup>a</sup>	Gene name (symbol) <sup>bc</sup>	GenBank Accession no <sup>b</sup>	% identity <sup>b</sup>	e value <sup>b</sup>
Bos taurus	Bt 675_bovine oocyte cDNA subtracted library	<a href="#">CX123770</a>	100	0.0
Xenopus laevis	XI AGENCOURT_10748348 Wellcome CRC pSK egg clone	<a href="#">CA974130</a>	92	e-153
Bos taurus	Hs B-cell translocation gene 4 ( <i>BTG4</i> )	<a href="#">NM_017589</a>	89	0.0
Mus musculus	Mm Bone morphogenetic protein 15 ( <i>Bmp15</i> )	<a href="#">BC055363</a>	97	0.0
Mus musculus	Mm Expressed sequence C87414	<a href="#">BC052888</a>	100	0.0
Bos taurus	Bt Growth differentiation factor 9 ( <i>GDF9</i> )	<a href="#">AF307092</a>	99	0.0
Bos taurus	Bt Hypothetical protein LOC616433	<a href="#">XM_881654</a>	99	0.0
Mus musculus	Mm NACHT, leucine rich repeat and PYD containing 14 ( <i>Nalp14</i> )	<a href="#">NM_001002894</a>	98	0.0
Xenopus laevis	XI NICHD_XGC_OOI clone IMAGp998K1111210	<a href="#">BX843422</a>	99	0.0
Mus musculus	Mm Nucleophosmin/nucleoplasmin 2 ( <i>Npm2</i> )	<a href="#">NM_181345</a>	98	0.0
Mus musculus	Mm Oogenesis 1 ( <i>Oog1</i> )	<a href="#">AB050008</a>	99	0.0
Bos taurus	Hs Polyadenylate binding protein-interacting protein 1 ( <i>PAIP1</i> )	<a href="#">NM_183323</a>	93	0.0
Xenopus laevis	XI Protein arginine methyltransferase 1 ( <i>PRMT1</i> ) <sup>d</sup>	<a href="#">AB085173</a>	99	0.0
Xenopus laevis	XI Protein tyrosine phosphatase, receptor type Q ( <i>PTPRQ</i> )	<a href="#">XM_370699</a>	95	0.0
Mus musculus	Mm Similar to Nur77 downstream protein 1 ( <i>Ndg1</i> )	<a href="#">XM_355193</a>	99	0.0
Bos taurus	Bt Similar to zinc finger protein 678 ( <i>ZF678</i> )	<a href="#">XM_601114</a>	88	6e-68
Bos taurus	Bt Small fragment nuclease ( <i>SMFN</i> ) <sup>d</sup>	<a href="#">XM_591769</a>	99	0.0
Mus musculus	Mm Spindlin ( <i>Spin</i> ) <sup>d</sup>	<a href="#">BC016517</a>	100	0.0
Mus musculus	Mm Zona pellucida glycoprotein 2 ( <i>Zp2</i> )	<a href="#">NM_011775</a>	99	0.0

<sup>a</sup> Species of the clone expressed in oocyte and conserved across species.

<sup>b</sup> Based on the BLAST results of the clone sequences compared against the GenBank database.

<sup>c</sup> Bt; Bos taurus, Hs; Homo sapiens, Mm; Mus musculus, XI; Xenopus laevis.

<sup>d</sup> Candidates validated by RT-PCR and gene-specific microarray.

#### Oocyte-specific transcripts common in all three species

To further characterize this subpopulation of genes conserved in all three species, another classification was performed in order to identify the ones that are preferentially expressed in oocytes. This was done by comparing results from a previous study [24] where transcripts preferentially expressed in oocytes compared to somatic tissues were identified. In this previous study, a list of genes preferentially expressed in the oocyte was obtained through subtractive hybridization and microarray experiments in the bovine, mouse and *Xenopus laevis*. For the present study, the genes found to be conserved in all three species (1,541) were compared against the list of genes preferentially expressed in the oocyte previously obtained. The comparison of these two lists revealed that 268 clones are preferentially expressed in oocytes and also conserved across all three species (Table 1, Fig. 1B). That list can be further subdivided as follows; 120 transcripts originate from the bovine oocyte library, 96 from the mouse oocyte library, and 52 from the *Xenopus laevis* oocyte library (Table 1). A representative list of transcripts preferentially expressed in the oocyte and conserved across species is presented in table 2.

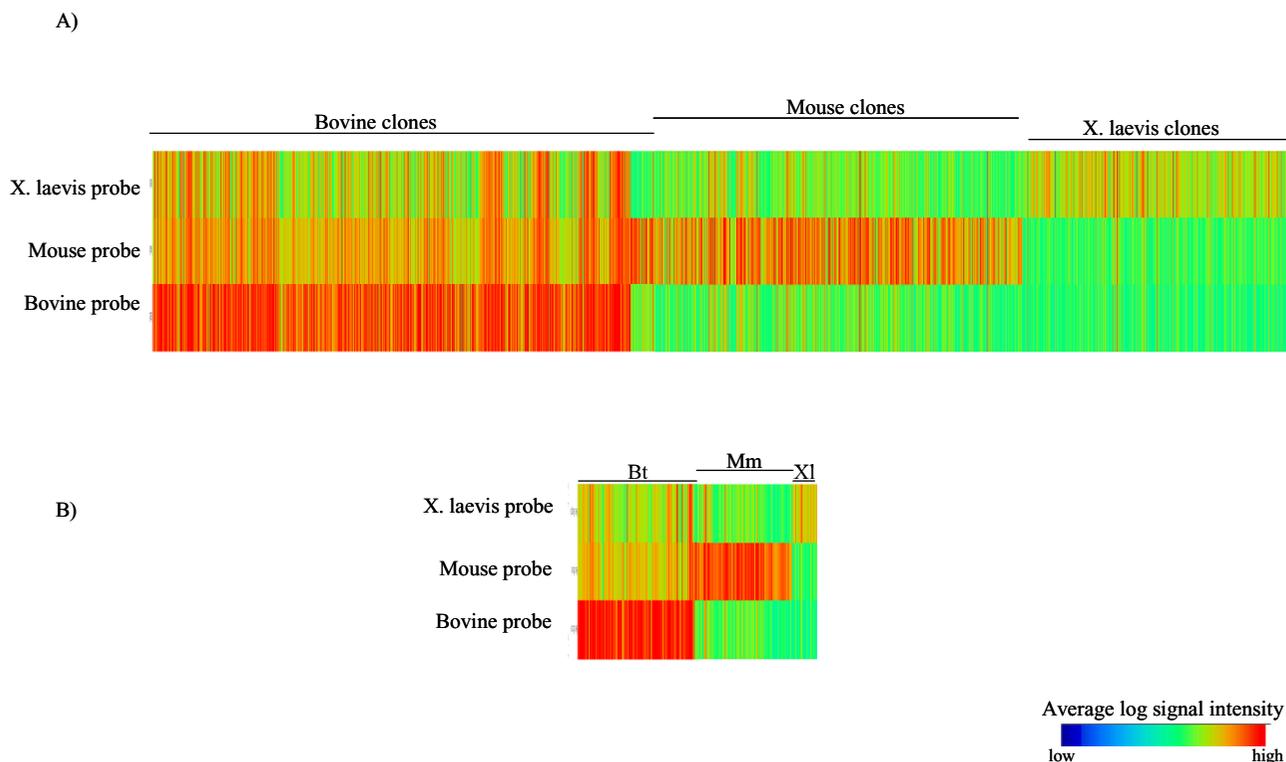
#### TMeV visualization of microarray data

To better visualize the microarray results, a versatile microarray data analysis tool, TIGR Multiexperiment Viewer (TMeV), was used. The average normalized log intensities for the 1,541 transcripts listed above are repre-

sented in figure 2A. Results are shown for each probe corresponding to one of the three species, and transcripts were ordered manually by clone number. Clearly evident with this representation is the fact that same-species hybridizations produced globally higher signal intensities than cross-species hybridizations. However, the cross-species hybridizations also generated signal intensities that were visibly and significantly above background. Also, bovine clones generate higher signal intensities than clones from the other two species, independent of the probe used for the hybridization. This is mainly a methodological artifact as the spotted bovine clones were more concentrated than the others (average concentration of spotted clones; bovine: 157 ng/ul, Mouse: 55 ng/ul, and *Xenopus laevis*: 59 ng/ul). Also, it was not due to a probe effect since this was controlled with our positive control, a cDNA fragment of the Green Fluorescent Protein (*GFP*). Figure 2B shows the 268 transcripts considered as genes preferentially expressed in oocytes and conserved in all three species.

#### Analysis of cross-species versus same-species hybridization Reproducibility

An essential criterion for the application of cross-species experiments is data reproducibility. To test this, we calculated the correlation coefficients of signal intensities between replicated experiments in a pair-wise manner (Table 3). For each experiment, the signal intensities generated from a probe corresponding to one species was cal-



**Figure 2**  
**TIGR Multiexperiment Viewer representation of microarray data.** Transcripts were ordered horizontally by clone number and grouped by species. Average log signal intensities are presented for clones hybridized with probes corresponding to one of the three species. The expressed clones that are conserved across all three species are represented in (A) and those conserved across species and preferentially expressed in the oocyte are represented in (B).

culated by comparing the signal intensities obtained from a replicate experiment. The correlation coefficients were first calculated with the signal intensities of all clones on the microarray slides. Next, the correlation coefficients were calculated with the signal intensities of clones corresponding to one species at a time, thus generating CC for same-species and cross-species hybridizations (Table 3). Results show that when all the spots on the multi-species microarray are considered, the CCs are relatively high and very similar between the three species (0.934 – 0.957). The analysis also revealed that *Xenopus laevis* clones generate the lowest CCs even with the *Xenopus laevis* probe (0.852). Taking this into consideration, the CCs calculated with data from cross-species hybridizations are in an acceptable range (0.817 – 0.965).

*Same species versus cross-species hybridizations*

After assessing the level of data reproducibility, variations between same-species and cross-species hybridizations were analyzed. A candidate gene, member of the *Zp* family (Bt; *ZP4*, Mm; *Zp1*, Xl; *ZPB*), was selected according to its presence in all three libraries, its previously reported

expression in oocytes of the three species being studied, and its evolutionary conserved sequence [Homologene: 33483]. By comparing between hybridizations the average log signal intensities of two clones from each species corresponding to the same *Zp* family member (n = 6), it is possible to observe that all the clones corresponding to this gene are consistent, not only between same-species, but also among the three different species (Fig. 3A). In cross-species hybridizations, all cross-species hybridizations signals are weaker than same-species hybridization signals, but are still relatively high, and always significantly above background (Fig. 3B). A similar analysis was performed with all 3,456 transcripts found on our multi-species microarray (Fig. 4). Once again, as anticipated, same-species hybridizations result in higher signal intensities than cross-species. Nevertheless, cross-species hybridizations average signal intensities are above background intensity suggesting that cross-species hybridizations are specific.

**Table 3: Correlation coefficients of the signal intensities between replicated experiments<sup>a</sup>**

Species <sup>b</sup>	All clones <sup>c</sup>	Same-species <sup>d</sup>	Cross-species <sup>e</sup>		
			Bovine	Mouse	Xenopus
Bovine	0.957	0.973		0.965	0.923
Mouse	0.938	0.969	0.961		0.817
Xenopus	0.934	0.852	0.960	0.945	

<sup>a</sup>Results are presented as the average correlation coefficient.  
<sup>b</sup>Corresponds to the species of the probe used for the hybridization.  
<sup>c</sup>All spots on the microarray were considered irrespective of the species.  
<sup>d</sup>Spots corresponding to the same species as the probe were considered.  
<sup>e</sup>Spots corresponding to the other two species as the probe were considered.

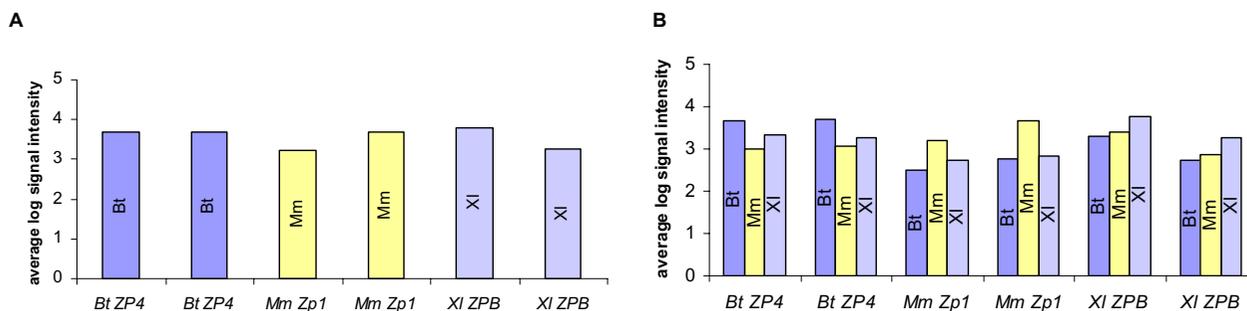
*In Silico sequence similarity analysis*

Another essential criterion for successful cross-species hybridization is adequate homology. To assess the extent of this potential problem, all transcripts identified as being preferentially expressed in the oocyte in all three species were compared to identify possible hits against the other two species using the GenBank database (Table 4). On average, bovine sequences show an 86% and 80% identity rate with the mouse and Xenopus laevis sequences, respectively and between the mouse and Xenopus, an 80% identity rate is observed. The BLAST results in average were given for a target region of an acceptable length (average 243 nt). However, not all transcripts compared against the GenBank database gave positive hits (Table 5). For the mouse, 64% of the transcripts gave no significant positive hit when compared to the other two species, whereas 41% of the Xenopus laevis and 29% and bovine transcripts also resulted in no significant positive hit. Amongst the Xenopus transcripts analyzed, 48% gave

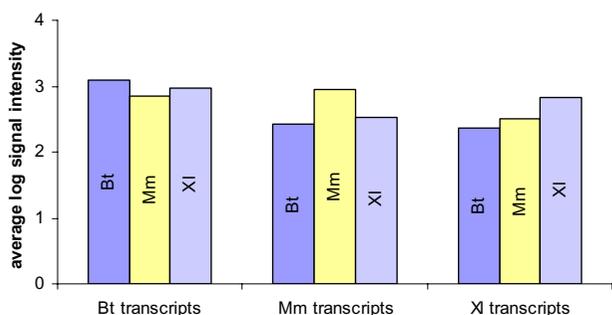
positive hits for the other two species. As for bovine and mouse it was even lower, where 37% and 12% of the transcripts gave positive hits for the other two species.

**Validation of cross-species hybridization specificity**

In order to support the results obtained with our microarray analysis, we performed a two-step validation process. First, this validation consisted of a standard detection test using RT-PCR and secondly a microarray hybridization analysis using specific PCR products as probes. Three candidates were selected for this validation process, one for each species, based on criteria originating from the microarray results. First the transcript had to generate a positive hybridization signal in all three species (in all 16 repetitions) and had to be preferentially expressed in the oocyte. Next, the candidate had to have a known ortholog gene in the two other species, an essential criterion if we wanted to design species-specific primers. Finally, we selected candidates that were not yet reported to be



**Figure 3**  
**Same-species versus cross-species hybridization.** The hybridization pattern of one member of the Zp gene family is represented. Average log signal intensities for two different clones corresponding to the same Zp genes from each species for A) same-species hybridization and B) cross-species hybridization. Probe species are indicated inside each bar. Bt; Bos taurus, Mm; Mus musculus, XI; Xenopus laevis.



**Figure 4**  
**Average log signal intensity for all 3,456 clones.** Shown here are average global log signal intensities values for same-species and cross-species hybridization. Results are separated according to the respective species of the clone in order to compare same-species versus cross-species hybridization. Probe species are indicated inside each bar. Bt; Bos taurus, Mm; Mus musculus, XI; Xenopus laevis.

expressed in the oocyte for the two other species since we wanted to know if this cross-species hybridization approach allowed the production of precise and reliable data across species that could lead to the discovery of novel transcripts present in the oocyte.

*RT-PCR amplification*

The RT-PCR amplifications were performed on cDNA generated from oocytes total RNA using gene- and species-specific primers (Table 6). For each candidate gene, the amplifications were performed on the other two species, not the species in which the candidate was first identified since the expression in the oocyte was already known for this species. The three candidate genes selected for validation process were SMFN (Small fragment nuclease) from the bovine oocyte-subtracted library, Spin (Spindlin) from the mouse oocyte-subtracted library and PRMT1 (Protein arginine methyltransferase 1) from the Xenopus laevis oocyte-subtracted library. This RT-PCR validation process revealed that all three candidates showed amplification products in oocytes of the other two species (Fig. 5). Resulting amplicons were sequenced to check for spe-

cificity. Identity was compared between the sequences of the PCR amplicons and the clone found on the array, and an acceptable identity rate was obtained (average 85%, min 76 %, and max 94%). Thus further supporting our multi-species microarray results.

*Gene-specific microarray hybridization*

The second step in this validation process was to perform gene-specific microarray hybridization on the multi-species cDNA microarray. Following the RT-PCR amplification reaction, the amplicons were labeled and hybridized to our multi-species microarray slide in order to assess the efficiency of gene-specific but cross-species hybridizations. This experiment was performed in three replicates and hybridizations were performed with either probes of both species simultaneously or only one species at a time. We reasoned that the validation could be considered successful only if the signals detected significantly above background corresponded to the selected candidate spotted on the array. This validation process revealed that all three candidates showed specific cross-species hybridization on our multi-species microarray slide supporting once again the idea that our cross-species hybridizations are specific even between not so closely related species (Fig. 6).

**Discussion**

The microarray gene chip platform is a powerful tool allowing for the analysis of thousands of genes simultaneously. In this study, we explored the technical feasibility of utilizing cross-species hybridizations to identify genes expressed in the oocyte and that are conserved across three species. Our experimental strategy was twofold. The first goal was to test the possibility of cross-species hybridization of three distantly related vertebrates on cDNA arrays, and most importantly, our second goal was to identify genes expressed in oocytes of all three species. The rationale was to use a subtractive strategy to produce libraries enriched in transcripts preferentially expressed in oocytes, and to use the homology requirement for high specificity microarray hybridization to identify those transcripts that are conserved through evolution. Moreover, we believed that transcripts fulfilling both criteria, specificity to the oocyte and evolutionarily conserved, are potentially

**Table 4: BLAST results for transcripts commonly expressed in oocytes of all three species<sup>a</sup>**

Species <sup>b</sup>	Bovine		Mouse		X. laevis	
	Identity rate (%)	Length (nt)	Identity rate (%)	Length (nt)	Identity rate (%)	Length (nt)
Bovine			86 <sub>[73-96]</sub>	259 <sub>[32-721]</sub>	80 <sub>[73-90]</sub>	304 <sub>[41-1143]</sub>
Mouse	86 <sub>[81-94]</sub>	188 <sub>[48-458]</sub>			81 <sub>[74-90]</sub>	149 <sub>[42-323]</sub>
X. laevis	80 <sub>[74-90]</sub>	217 <sub>[45-869]</sub>	79 <sub>[72-92]</sub>	338 <sub>[27-933]</sub>		

<sup>a</sup> The average identity rates is given in percentage (%) and the minimum and maximum in brackets. The average length is given in nucleotide (nt) and the minimum and maximum are in brackets.

<sup>b</sup> Species of the transcripts compared against GenBank database.

**Table 5: Transcripts that generated a significant positive hit when compared to GenBank database<sup>a</sup>**

Species <sup>b</sup>	no hit <sup>c</sup>	1 hit <sup>d</sup>	2 hits <sup>e</sup>
Bovine	29%	71%	37%
Mouse	64%	36%	12%
X. laevis	41%	59%	48%

<sup>a</sup> Significant positive hit when  $e$  value  $< 10^{-5}$ .

<sup>b</sup> Species of the transcripts.

<sup>c</sup> The percentage of transcripts that generated no significant hits when compared to the GenBank database for the other two species.

<sup>d</sup> The percentage of transcripts that generated a significant hit for at least one species.

<sup>e</sup> The percentage of transcripts that generated a significant hit for both species.

important maternal genes involved in key functions of oocyte maturation and early development. By working simultaneously with three evolutionarily distant species, it increased our efficiency at identifying novel oocyte-specific genes and elucidating the important evolutionarily conserved mechanisms in different species. Also, this approach facilitates the identification of new genes not previously identified due to their low expression level in a particular species.

One of the unique features of preimplantation embryo development is that it occurs in the presence of maternally stored RNAs in oocytes as the embryonic genome has yet to be activated. These transcripts have specific functions either in oogenesis, oocyte maturation, fertilization and/or the early phase of preimplantation development. Only a few of these genes are well known and have been characterized, as Sharov et al., have demonstrated that, in the mouse, 119 out of the 196 oocyte-specific ESTs were unknown genes in 2003 [7]. Identification and characterization of these genes will enable us to better understand the unique molecular mechanisms present in the oocyte.

In the present study we designed a cDNA multi-species microarray containing 3,456 transcripts from bovine, mouse, and *Xenopus laevis* oocytes. Transcripts found on the array were randomly selected from oocyte-subtracted libraries constructed in a previous study [24]. The use of clones coming from three different subtracted libraries created a variation between the three species in the average concentration of the clones spotted on the slides. Based on our multi-species cDNA microarray results, 1,541 transcripts in total gave positive hybridization signals in oocytes across all three species and of these, 268 transcripts in total were found to be preferentially expressed in oocytes for all three species. However, the higher concentration of the spotted bovine clones resulted in more bovine clones being identified as conserved in oocytes of all three species when compared to the number of identified mouse and *Xenopus*

*laevis* clones. The difference in the concentration of the spotted clones also explains the general higher signal intensities seen in bovine clones. However, although this created a distortion in species representation in the cross-species hybridization results, the average concentration of spotted clones is within acceptable range for each species according to our quality control experiment and the slides manufacturer. The distribution of clones considered as present is practically equal between the three species (35%, 31% and 34% for bovine, mouse, and *Xenopus laevis* clones respectively). This methodological artifact might have resulted in a failure of identifying all transcripts conserved across species but nevertheless did not affect the validity of the results obtained.

In this study, results obtained from the reproducibility analysis increased the confidence in the data generated from our hybridizations. The CCs calculated with all the clones on the array showed that signal intensities from all three probes used were highly reproducible (0.934 – 0.957). For the bovine and mouse probes, cross-species hybridizations showed a slightly lower correlation coefficient compared with the same-species experiments. On the other hand, *Xenopus laevis* clones always generated the lowest CCs even with the *Xenopus laevis* probe, which might be in part related to a lower amount of cDNA spotted onto the array. Nevertheless, the correlation coefficients for the cross-species hybridization were sufficiently high to assure that reproducibility between replicated experiments is acceptable, increasing our confidence in the validity of cross-species hybridizations results. Also, the sequence mismatches present between the three species should be taken into account since they probably contribute to the lower correlations observed in cross-species hybridizations.

The degree of homology between probes and targets when performing cross-species hybridizations is extremely variable. In the presence of sequence mismatches, relative hybridization intensities will reflect both differences in transcript abundance levels, as well as differences in hybridization kinetics. In addition, it can even be variable between two different cross-species hybridizations, especially when the studied species are not equally divergent. Due to these limitations, the goal of this study was not to assess gene expression levels but instead, to survey the products of three subtracted libraries in order to identify transcripts present in oocytes of all three species. Nevertheless, sequence homology had to be sufficiently high in the target region to result in a proper hybridization, since mismatches will inevitably occur in evolutionarily distant animals. To assess this issue, all transcripts identified as being preferentially expressed in the oocyte and present in all three species were compared to identify possible hits with the other two species using the GenBank database.

On average, our bovine sequences showed an 86% and 80% identity rate with the mouse and *Xenopus laevis* sequences, respectively and between the mouse and *Xenopus*, an 80% identity rate was observed. In the last few years, a number of studies have successfully used cross-species hybridizations [29,34-40]. Ji et al., created a simple mathematic model for cross-species hybridization and concluded that a contiguous matched oligo of 16 bp long was sufficient to generate a specific hybridization signal [41]. Kane et al., have also reached similar conclusions where their results showed that specificity of the probe requires target-genes to be at least 75% similar over the target region [42]. In addition, if the target region is marginally similar (50–75%), a stretch of complementary sequence of more than 15 contiguous bases will allow hybridization [42]. These studies further support the results obtained from cross-species hybridization. However, it has to be considered that this can also be regarded has a limitation inherent to cDNA arrays, since it may also allow some cross-hybridization with other isoforms and/or non-target transcripts and therefore allow non-specific hybridization signals to contribute to the overall signal. Validation by PCR with gene-specific primers can verify this limitation.

Furthermore, the present study has demonstrated that cross-hybridization results can be confirmed by both RT-PCR reactions and gene-specific hybridizations. In order to ascertain that the observed signals were not originating from the annealing of random non-specific sequences, specificity validation was conducted through a simple detection test using species- and gene-specific primers and by gene-specific microarray hybridization. The labeled amplification products corresponding to three independ-

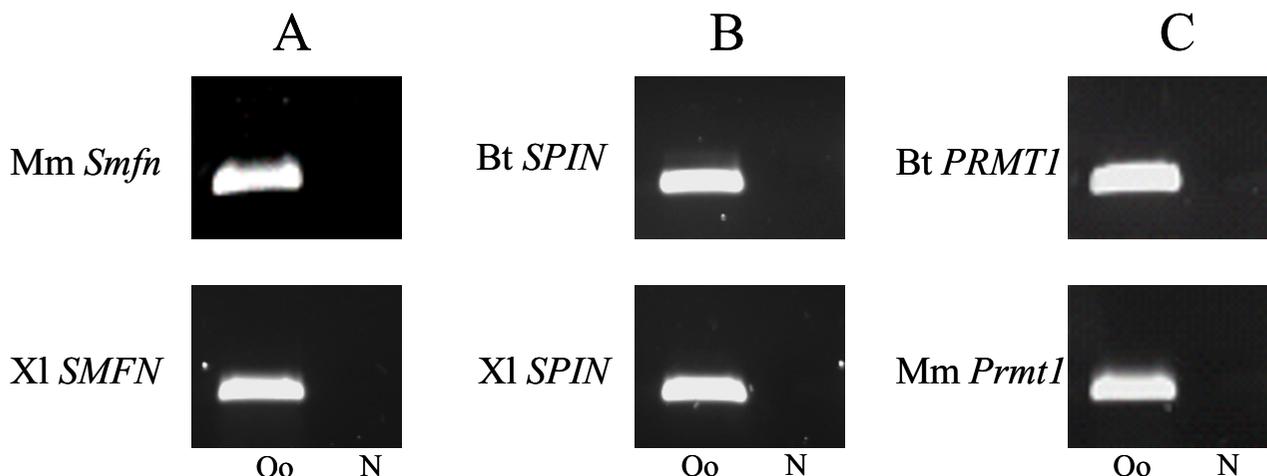
ent single genes showed positive signals only for their corresponding target even across species, lending support to the validity of our cross-species hybridizations.

Three candidate genes, one from each species, were selected for these validation processes. Our first candidate gene used for validation was bovine *SMFN*, also known as *CGI-114*, which is homologue of *Orn*, a 3-prime-to-5-prime exoribonuclease of *E. coli*. The ORN protein is known to attack the free 3-prime hydroxyl group on single-stranded RNA, releasing 5-prime mononucleotides in a sequential manner [GeneID: 25996] [43]. In human, a study by Nguyen et al., suggests a role for *SMFN* in cellular nucleotide recycling [44]. In the mouse, *Smfn* gene has also been characterized and is reported to be expressed in a variety of tissues including testis, uterus and embryo, but to date, no report indicates its expression in mouse oocytes [Unigene: Mm.21911]. The similarity between the mouse *Smfn* gene sequence and our bovine clone sequence is relatively high; 91% on 304 bp. A Blast search of our bovine clone sequence against *Xenopus laevis* sequence revealed that a cDNA clone IMAGE 7205916 [GenBank: [BC087528](#)] has an acceptable identity rate with our bovine clone; 79% on 206 bp. This clone is reported to be testis specific [Unigene: Xl.9259]. Like in the mouse, no expression was reported in the *Xenopus laevis* oocyte for this transcript. With our cross-species microarray hybridization we were able to detect the presence of *SMFN* transcripts in bovine, mouse and *Xenopus laevis* oocytes, and this was also confirmed by RT-PCR and gene-specific cross-species microarray hybridizations.

Our second candidate gene used for validation was mouse *Spin* (Spindlin), an abundant maternal transcript present

**Table 6: Sequence of gene-and species-specific primers.**

Gene Name	Species	Primer Sequence	amplicon size
<i>Smfn</i>	Mouse	up 5'- TTT GTA CGA CAG CAG ACT CCT C -3' low 5'- TGA AGC TCT TTG ATG CTT TCA C -3'	245
<i>SMFN</i>	<i>Xenopus</i>	up 5'- GAG CAC TGT GAA AGA ATT GTG C -3' low 5'- CTT CGA GTC AGA TGT TGG TCA C -3'	225
<i>SPIN</i>	Bovine	up 5'- TGG CAT GGT CAT TCA TCA AG -3' low 5'- TTC TGC TGG CAG AGA TTG TG -3'	233
<i>SPIN</i>	<i>Xenopus</i>	up 5'- CAG CCT TGT AGG GAA GCA AG -3' low 5'- TGT TTG GCA ACA AGT CAA GG -3'	184
<i>PRMT1</i>	Bovine	up 5'- ACC CTC ACA TAC CGC AAC TC -3' low 5'- TTC CCC TTG ATG ATG GTC AC -3'	218
<i>Prmt1</i>	Mouse	up 5'- ACC CTC ACA TAC CGC AAC TC -3' low 5'- TGA TGG TCA CCA CAT GGT CT -3'	208
<i>GAPD</i>	Bovine	up 5'- CCA ACG TGT CTG TTG TGG ATC TGA -3' low 5'- GAG CTT GAC AAA GTG GTC GTT GAG -3'	226
<i>Gapdh</i>	Mouse	up 5'- ATG TCG TGG AGT CTA CTG GTG TC -3' low 5'- CAR ACT TGG CAG GTT TCT CCA G -3'	486
<i>GAPD</i>	<i>Xenopus</i>	up 5'- TGT AGT TGG CGT GAA CCA TGA G -3' low 5'- CAG CAT CAA AGA TGG AGG AGT G -3'	500

**Figure 5****Validation of microarray result by simple detection analysis using RT-PCR for three candidate transcripts.**

Amplifications are performed in oocytes of two species with gene- and species-specific primers designed according to the known sequences in respective species. A) Amplification in mouse and *Xenopus laevis* oocytes of *SMFN* transcript originally found in the bovine oocyte-library. B) Amplification in bovine and *Xenopus laevis* oocytes of *Spin* transcript originally found in the mouse oocyte-library. C) Amplification in bovine and mouse oocytes of *Prmt1* transcript originally found in the *Xenopus laevis* oocyte-library. Bt; *Bos Taurus*, Mm; *Mus musculus*, Xl; *Xenopus laevis*, *Smf*n; Small fragment nuclease, *Spin*; Spindlin, *Prmt1*; Protein arginine methyltransferase I, Oo; oocyte, N; negative control.

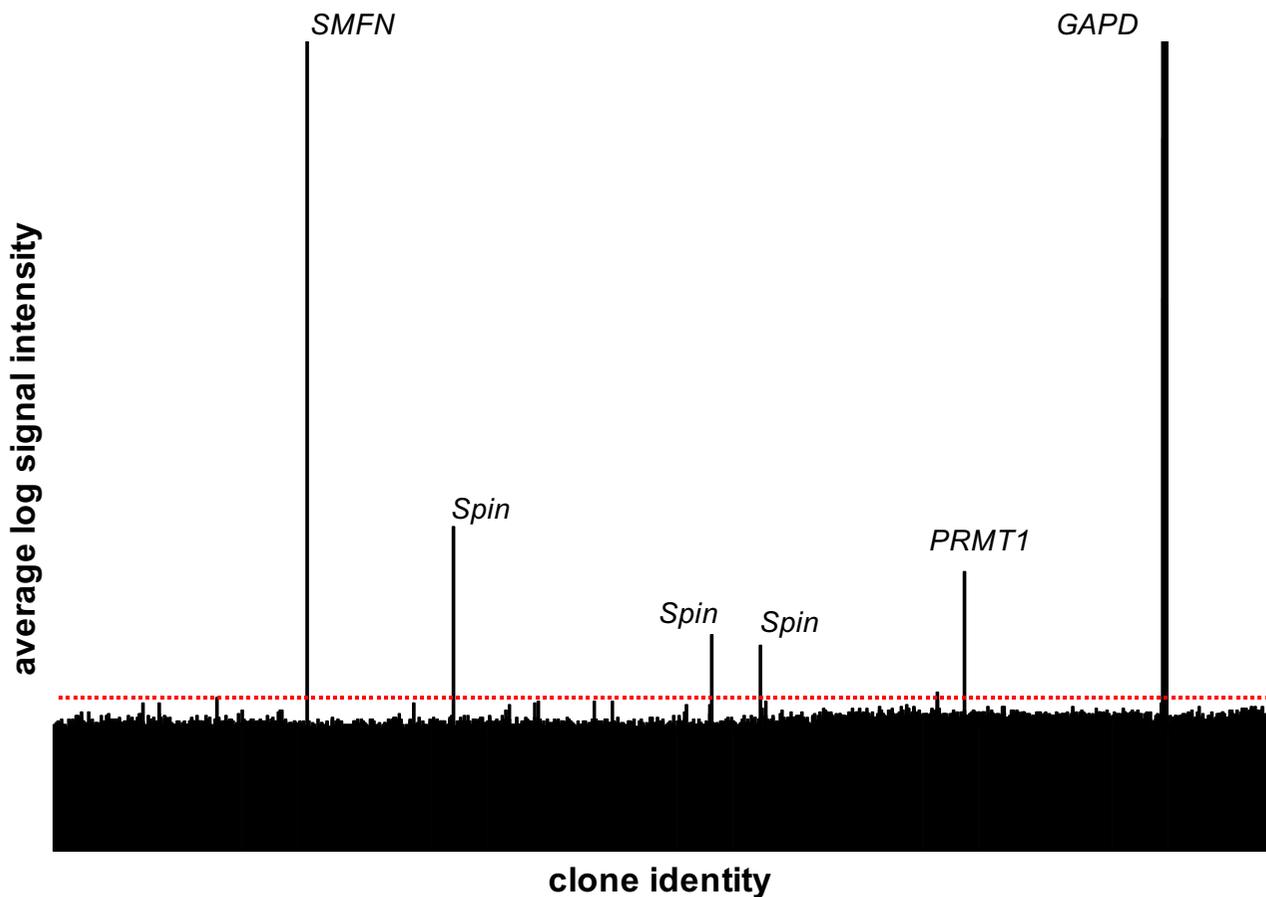
in the unfertilized egg and 2-cell, but not 8-cell mouse embryo [45]. SPIN protein associates with the meiotic spindle and is modified by phosphorylation in a cell-cycle-dependent fashion, and is suggested to play a role in cell-cycle regulation during the transition from gamete to embryo [45]. Also, further studies imply that SPIN is a substrate in the MOS/MAP kinase pathway and that this phosphorylation of Spin may be essential for its interaction with the spindle [46]. The *SPIN* gene has not been identified and characterized thus far in the bovine, however there is a predicted sequence from an automated computational analysis in the NCBI database [GenBank: [XM\\_614403](#)]. This predicted sequence has 94% similarity (234 bp) with our mouse *Spin* clone. Since it is a predicted sequence, no report of its expression pattern is available. A Blast search of our mouse clone sequence against *Xenopus laevis* sequences revealed that a cDNA clone IMAGE 6324148 [GenBank: [BC097748](#)] has an acceptable identity rate with our clone; 84% on 146 bp. No report of the expression pattern for the bovine and *Xenopus laevis* are currently available. Once again, with our cross-species microarray hybridization we were able to detect the presence of *SPIN* transcripts in bovine, mouse and *Xenopus laevis* oocytes, and this was also confirmed by RT-PCR and gene-specific cross-species microarray hybridizations.

The third and last candidate gene used for validation is the *Xenopus laevis* Protein arginine methyltransferase 1 (*PRMT1*), an xCirp2-binding protein. The methylation of

xCIRP2 (cold-inducible RNA binding protein 2) by PRMT1 results in the accumulation of xCIRP2 in the cytoplasm [47]. It is also known that xCIRP2, which is highly expressed in *Xenopus laevis* oocytes, is associated with ribosomes, suggesting that it participates in translational regulation in oocytes [48]. Bovine *HRMT12* (Hmt1 hnRNP methyltransferase-like 2) gene possesses a high similarity with *Xenopus PRMT1*; 78% on 546 bp. This bovine gene is reported to be expressed in different tissues such as fetus and adult brain but there is no report available indicating its expression in the oocyte [Unigene: Bt.4871]. Mouse Heterogeneous nuclear ribonucleoproteins methyltransferase-like 2 (*Hrmt112*) also possess a high identity rate with *Xenopus PRMT1*; 78% on 413 bp. Mouse *Hrmt112* gene is reported to be expressed in a variety of tissues including testis, ovary and embryo, but there are no reports indicating its specific expression in the oocyte [Unigene: Mm.21911]. Once more, we were able to detect the presence of *PRMT1* transcripts in bovine, mouse and *Xenopus laevis* oocytes, and this was also confirmed by RT-PCR and gene-specific cross-species microarray hybridizations.

**Conclusion**

In summary, these results prove the feasibility of cross-species hybridization and the utility of a multi-species microarray. Our results demonstrate that cross-species hybridization is not only useful for studying species for which microarrays are not yet available, but are also very



**Figure 6**  
**Gene-specific cross-species hybridization on our multi-species array.** Example of a representative result for average log signal intensities is shown for all the 3,456 transcripts on the array. The red dotted line represents background intensity. The three candidate genes and our positive control *GAPD* are identified. Note that *Spin* transcript is present 3 times on the multi-species microarray. *Smfjn*; Small fragment nuclease, *Spin*; Spindlin, *Prmt1*; Protein arginine methyltransferase 1.

powerful in elucidating the important evolutionarily conserved mechanisms in different species. The identification of all genes expressed in oocytes will allow a better understanding of the mechanisms and pathways regulating gametogenesis and embryogenesis.

## Methods

### cDNA Multi-species microarray preparation

The multi-species cDNA microarray used in this study contained transcripts from three oocyte-specific libraries constructed previously using Suppressive Subtractive Hybridization (SSH) [24]. Briefly, total RNA from a pool of somatic tissues was subtracted to total RNA from oocytes to generate three libraries enriched in oocyte-specific transcripts (mouse, bovine, and *Xenopus laevis*). The complete procedure used for slide preparation has been described previously [24]. The array was strictly divided

into three equal sections, each corresponding to one of the three species. There were a total of 3,456 oocyte transcripts represented on the array, thus 1,152 clones per species. Each transcript was spotted four times for a total of 13,824 spots. It is important to note that the libraries did not only consist of unique transcripts, it was possible that more than one sequence and/or several copies of a sequence were present on the slide. In addition, negative and positive controls were randomly distributed on the cDNA multi-species array for diverse quality controls; three different SpotReport Alien cDNA Array Validation System (Stratagene, La Jolla, CA) were used as negative controls ( $n = 424$ ) and a cDNA fragment of the Green Fluorescent Protein (*GFP*) was used as an exogenous positive control ( $n = 260$ ).

### **DNA sequencing and analysis**

DNA sequencing was performed as previously described [24]. The resulting sequence traces were visualized with the online freeware Chromas 1.45 [49] and uploaded into a cDNA Library Manager program (Genome Canada Bioinformatics) that automates and facilitates sequence analysis and clone identification. Briefly, sequence traces were uploaded into the cDNA Library Manager, trimmed (Phred software) and compared against a locally installed GenBank database [50].

### **Labeling probes for the cDNA Multi-species microarray**

Forward-subtracted PCR products from the subtracted libraries (oocyte minus somatic tissues) corresponding to oocyte-libraries were used as probes to hybridize the cDNA multi-species microarray as previously described [24]. Briefly, probes were labeled with Alexa Fluor 555 and 647 reactive dye packs (Molecular Probe, Burlington, ON, Canada) using Amino Allyle dUTP (Ambion, Austin, TX) according to the manufacturer's instructions.

### **Array hybridization**

Slides were prehybridized with DIG buffer (Roche Diagnostics, Laval, QC, Canada) supplemented with yeast tRNA (4 mg/ml, Invitrogen) and Cot-1 DNA (1 mg/ml, Invitrogen) for 1 hour at 37°C. Slides were then hybridized overnight at 37°C with labeled purified probes added to fresh prehybridization solution. Hybridizations were performed in the ArrayBooster using the Advacard AC3C (The Gel Company, San Francisco, CA). Slides were then washed once with  $1 \times$  SSC-0.2% SDS for 10 min at RT,  $1 \times$  SCC-0.2% SDS for 10 min at 55°C, and for 5 min at RT with  $0.1 \times$  SCC-0.2% SDS. Hybridizations were always performed with probes corresponding to two different species, in an all pair design (bovine-mouse; mouse-Xenopus laevis; Xenopus laevis-bovine). A dye swap experiments was included for each sample to take into consideration the variation in dye incorporation efficiency. Moreover, two biological replicates for each sample were used. Thus, for each species, the hybridizations were carried out twice with one dye and twice with the other dye, giving four technical replicate hybridizations per species. Since each clone was replicated four times on the microarray, a total of 16 data points were generated for all the candidates per species-specific hybridization.

Considering that the same spots were also hybridized during cross-species hybridizations, 48 data points were generated for each spotted clone (4 spotted clones  $\times$  4 hybridizations  $\times$  3 species). Quality control was performed through the addition of 424 negative and 260 positive controls included on the array. GFP cDNA fragments were added to the probes in equal amounts, before labeling, to use as positive controls.

### **Microarray image processing**

Slides were scanned using the VersArray ChipReader System (Bio-Rad, Mississauga, ON, Canada) and visualized with the ChipReader software (Media Cybernetics, San Diego, CA). Microarray image processing was performed with the ArrayPro Analyzer software (Media Cybernetics, San Diego, CA). Local background was subtracted and data were normalized (LOWESS). Microarray experiments presented in this study adhere to the standards proposed by the Microarray Gene Expression Data Society [51]. Raw and normalized data for the microarray experiments reported herein are stored in the public repositories ArrayExpress (accession no E-MEXP-488) [52].

### **Microarray data analysis**

#### *Transcripts present in oocytes*

First, data were log transformed before proceeding with a simple analysis where thoughtful criteria were applied to minimize, to the extent possible, the false positive rate. The analysis consisted of a pretreatment to eliminate uninformative data according to a calculated threshold;  $t = m + 2 \times sd$  (where 't' is the calculated threshold, 'm' is the mean and 'sd' is the standard deviation of the negative control data,  $n = 424$ ). Transcripts above the calculated threshold were considered as present in the bovine, mouse, or Xenopus laevis oocytes. A second independent analysis with the NIA Array Analysis tool was also conducted [53,54]. Briefly, raw data from the multi-species microarray hybridizations were uploaded and background threshold was determined according to the plot of error function (standard deviation, SD (= square root of the error variance), versus expression level (Log intensity)). Clones with a mean log signal intensity above the calculated background threshold (log 2.5) were considered as expressed in bovine, mouse, or Xenopus oocytes. Since both independent methods of analysis generated similar results, the list of genes expressed in the oocyte from the first method was used for the following steps.

#### *Transcripts common in oocytes of all three species*

Subsequently, clones for which all three species and all replicates were above the calculated threshold were selected to generate a list of transcripts common in oocytes of all three species (48 data points were generated for each clone; 4 spotted clones  $\times$  4 replicate hybridizations  $\times$  3 species).

#### *Oocyte-specific transcripts common in all three species*

To further characterize this subpopulation of transcripts common in oocytes of all three species, another classification was performed in order to identify the ones that are preferentially expressed in oocytes. This was done by comparing results from a previous study [24] where transcripts preferentially expressed in oocytes compared to somatic tissues were identified. In this previous study, a list of

genes preferentially expressed in the oocyte was obtained through subtractive hybridization and microarray experiments in the bovine, mouse and *Xenopus laevis*. For the present study, the genes found to be conserved in all three species were compared against the list of genes preferentially expressed in the oocyte previously obtained. The combination of these two sets of results generated a list of oocyte-specific transcripts common in all three species.

#### Reproducibility

Data reproducibility was assessed by calculating the correlation coefficients between signal intensities across replicated experiments in a pair-wise manner with the NIA Array Analysis tool [53]. For each experiment, the signal intensities generated from a probe corresponding to one species was calculated by comparing the signal intensities obtained from a replicate experiment. The correlation coefficients were first calculated with the signal intensities of all clones on the microarray slides. Next, the correlation coefficients were calculated with the signal intensities of clones corresponding to one species at a time.

#### Validation

To validate the presence of oocyte-expressed genes detected in all three species, RT-PCR analysis, using gene- and species-specific primers, were performed on three different selected candidate genes (Table 6). Briefly, equal amounts of total RNA isolated from bovine, mouse and *Xenopus laevis* oocytes was used to generate cDNA with an oligo (dT) primer and the Omniscript reverse transcriptase (Invitrogen) according to the manufacturer's instructions. PCR amplifications were performed as mentioned with gene- and species-specific primers in two species, not the species in which the candidate was first identified. Resulting amplicons were sequenced in order to check for specificity. Amplicons were then labeled as described above and used as probes to hybridize the multi-species microarray slide. Also, *GAPD* (Glyceraldehyde-3-phosphate dehydrogenase) amplifications were carried out as a positive control amplification to assure the quality of the cDNAs used for this experiment. *GAPD* PCR products were also labeled to use as a positive control of hybridization to assure the quality of the hybridization.

#### Authors' contributions

MV designed, optimized and carried out the protocols, analyzed the results and wrote the manuscript. CR was involved in the conceptualization of this project and critically reviewed the manuscript for important intellectual content. SM was involved in the microarray data analysis and critically reviewed the manuscript. MFP provided support in the preparation of this manuscript and critically reviewed the manuscript for important intellectual content. MAS was involved in the conceptualization of this project, the preparation of the manuscript and provided

mentorship. All authors read and approved the final manuscript.

#### Acknowledgements

The authors acknowledge Dr Susan Novak for critical reading of the manuscript and providing valuable comments. MV is supported by NSERC fellowship. This project was supported by Canada Research Chair and Natural Science and Engineering Research Council of Canada. Lennoxville Dairy and Swine R & D Center Contribution No. 888.

#### References

- Bachvarova R, Baran MM, Tejblum A: **Development of naked growing mouse oocytes in vitro.** *J Exp Zool* 1980, **211**:159-169.
- LaMarca MJ, Smith LD, Strobel MC: **Quantitative and qualitative analysis of RNA synthesis in stage 6 and stage 4 oocytes of *Xenopus laevis*.** *Dev Biol* 1973, **34**:106-118.
- Telford NA, Watson AJ, Schultz GA: **Transition from maternal to embryonic control in early mammalian development: a comparison of several species.** *Mol Reprod Dev* 1990, **26**:90-100.
- Etkin LD, Balcells S: **Transformed *Xenopus* embryos as a transient expression system to analyze gene expression at the midblastula transition.** *Dev Biol* 1985, **108**:173-178.
- Memili E, First NL: **Control of gene expression at the onset of bovine embryonic development.** *Biol Reprod* 1999, **61**:1198-1207.
- Schultz RM: **Regulation of zygotic gene activation in the mouse.** *Bioessays* 1993, **15**:531-538.
- Sharov AA, Piao Y, Matoba R, Dudekula DB, Qian Y, VanBuren V, Falco G, Martin PR, Stagg CA, Bassey UC, Wang Y, Carter MG, Hamatani T, Aiba K, Akutsu H, Sharova L, Tanaka TS, Kimber WL, Yoshikawa T, Jaradat SA, Pantano S, Nagaraja R, Boheler KR, Taub D, Hodes RJ, Longo DL, Schlessinger D, Keller J, Klotz E, Kelsoe G, Umezawa A, Vescovi AL, Rossant J, Kunath T, Hogan BL, Curci A, D'Urso M, Kelso J, Hide W, Ko MS: **Transcriptome analysis of mouse stem cells and early embryos.** *PLoS Biol* 2003, **1**:E74.
- Mutter GL, Wolgemuth DJ: **Distinct developmental patterns of c-mos protooncogene expression in female and male mouse germ cells.** *Proc Natl Acad Sci USA* 1987, **84**:5301-5305.
- Philpott CC, Ringuelette MJ, Dean J: **Oocyte-specific expression and developmental regulation of ZP3, the sperm receptor of the mouse zona pellucida.** *Dev Biol* 1987, **121**:568-575.
- Liang LF, Chamow SM, Dean J: **Oocyte-specific expression of mouse Zp-2: developmental regulation of the zona pellucida genes.** *Mol Cell Biol* 1990, **10**:1507-1515.
- Epifano O, Liang LF, Dean J: **Mouse Zp1 encodes a zona pellucida protein homologous to egg envelope proteins in mammals and fish.** *J Biol Chem* 1995, **270**:27254-27258.
- McGrath SA, Esqueda AF, Lee SJ: **Oocyte-specific expression of growth/differentiation factor-9.** *Mol Endocrinol* 1995, **9**:131-136.
- Liang L, Soyal SM, Dean J: **FIGalpha, a germ cell specific transcription factor involved in the coordinate expression of the zona pellucida genes.** *Development* 1997, **124**:4939-4947.
- Dube JL, Wang P, Elvin J, Lyons KM, Celeste AJ, Matzuk MM: **The bone morphogenetic protein 15 gene is X-linked and expressed in oocytes.** *Mol Endocrinol* 1998, **12**:1809-1817.
- Tanaka M, Hennebold JD, Macfarlane J, Adashi EY: **A mammalian oocyte-specific linker histone gene H1oo: homology with the genes for the oocyte-specific cleavage stage histone (cs-H1) of sea urchin and the B4/H1M histone of the frog.** *Development* 2001, **128**:655-664.
- Wu X, Viveiros MM, Eppig JJ, Bai Y, Fitzpatrick SL, Matzuk MM: **Zygote arrest I (Zar1) is a novel maternal-effect gene critical for the oocyte-to-embryo transition.** *Nat Genet* 2003, **33**:187-191.
- Tong ZB, Gold L, Pfeifer KE, Dorward H, Lee E, Bondy CA, Dean J, Nelson LM: **Mater, a maternal effect gene required for early embryonic development in mice.** *Nat Genet* 2000, **26**:267-268.
- Burns KH, Viveiros MM, Ren Y, Wang P, DeMayo FJ, Frail DE, Eppig JJ, Matzuk MM: **Roles of NPM2 in chromatin and nucleolar organization in oocytes and embryos.** *Science* 2003, **300**:633-636.

19. Murray MT, Krohne G, Franke WW: **Different forms of soluble cytoplasmic mRNA binding proteins and particles in *Xenopus laevis* oocytes and embryos.** *J Cell Biol* 1991, **112**:1-11.
20. Hwang KC, Park SY, Park SP, Lim JH, Cui XS, Kim NH: **Specific maternal transcripts in bovine oocytes and cleaved embryos: identification with novel DDRT-PCR methods.** *Mol Reprod Dev* 2005, **71**:275-283.
21. Ponsuksili S, Tesfaye D, El-Halawany N, Schellander K, Wimmers K: **Stage-specific expressed sequence tags obtained during preimplantation bovine development by differential display RT-PCR and suppression subtractive hybridization.** *Prenat Diagn* 2002, **22**:1135-1142.
22. Robert C, Barnes FL, Hue I, Sirard MA: **Subtractive hybridization used to identify mRNA associated with the maturation of bovine oocytes.** *Mol Reprod Dev* 2000, **57**:167-175.
23. Robert C, Gagne D, Bousquet D, Barnes FL, Sirard MA: **Differential display and suppressive subtractive hybridization used to identify granulosa cell messenger rna associated with bovine oocyte developmental competence.** *Biol Reprod* 2001, **64**:1812-1820.
24. Vallee M, Gravel C, Palin MF, Reghenas H, Stothard P, Wishart DS, Sirard MA: **Identification of novel and known oocyte-specific genes using complementary DNA subtraction and microarray analysis in three different species.** *Biol Reprod* 2005, **73**:63-71.
25. Zeng F, Schultz RM: **Gene expression in mouse oocytes and preimplantation embryos: use of suppression subtractive hybridization to identify oocyte- and embryo-specific genes.** *Biol Reprod* 2003, **68**:31-39.
26. Ko MS, Kitchen JR, Wang X, Threat TA, Hasegawa A, Sun T, Grahovac MJ, Kargul GJ, Lim MK, Cui Y, Sano Y, Tanaka T, Liang Y, Mason S, Paonessa PD, Sauls AD, DePalma GE, Sharara R, Rowe LB, Eppig J, Morrell C, Doi H: **Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development.** *Development* 2000, **127**:1737-1749.
27. Rajkovic A, Yan MSC, Klysik M, Matzuk M: **Discovery of germ cell-specific transcripts by expressed sequence tag database analysis.** *Fertil Steril* 2001, **76**:550-554.
28. Stanton JL, Green DP: **A set of 840 mouse oocyte genes with well-matched human homologues.** *Mol Hum Reprod* 2001, **7**:521-543.
29. Adjaye J, Herwig R, Herrmann D, Wruck W, Benkahl A, Brink TC, Nowak M, Carnwath JW, Hultschig C, Niemann H, Lehrach H: **Cross-species hybridisation of human and bovine orthologous genes on high density cDNA microarrays.** *BMC Genomics* 2004, **5**:83.
30. Altmann CR, Bell E, Sczyrba A, Pun J, Bekiranov S, Gaasterland T, Brivanlou AH: **Microarray-based analysis of early development in *Xenopus laevis*.** *Dev Biol* 2001, **236**:64-75.
31. Carter MG, Hamatani T, Sharov AA, Carmack CE, Qian Y, Aiba K, Ko NT, Dudekula DB, Brzoska PM, Hwang SS, Ko MS: **In situ-synthesized novel microarray optimized for mouse stem cell and early developmental expression profiling.** *Genome Res* 2003, **13**:1011-1021.
32. Hamatani T, Carter MG, Sharov AA, Ko MS: **Dynamics of global gene expression changes during mouse preimplantation development.** *Dev Cell* 2004, **6**:117-131.
33. Zeng F, Baldwin DA, Schultz RM: **Transcript profiling during preimplantation mouse development.** *Dev Biol* 2004, **272**:483-496.
34. Medhora M, Bousamra M 2nd, Zhu D, Somberg L, Jacobs ER: **Upregulation of collagens detected by gene array in a model of flow-induced pulmonary vascular remodeling.** *Am J Physiol Heart Circ Physiol* 2002, **282**:H414-422.
35. Robert C, Hue I, McGraw S, Gagne D, Sirard MA: **Quantification of cyclin B1 and p34(cdc2) in bovine cumulus-oocyte complexes and expression mapping of genes involved in the cell cycle by complementary DNA macroarrays.** *Biol Reprod* 2002, **67**:1456-1464.
36. Grigoryev DN, Ma SF, Simon BA, Irizarry RA, Ye SQ, Garcia JG: **In vitro identification and in silico utilization of interspecies sequence similarities using GeneChip technology.** *BMC Genomics* 2005, **6**:62.
37. Moody DE, Zou Z, McIntyre L: **Cross-species hybridisation of pig RNA to human nylon microarrays.** *BMC Genomics* 2002, **3**:27.
38. Shah G, Azizian M, Bruch D, Mehta R, Kittur D: **Cross-species comparison of gene expression between human and porcine tissue, using single microarray platform – preliminary results.** *Clin Transplant* 2004, **18**(Suppl 12):76-80.
39. Caceres M, Lachuer J, Zapala MA, Redmond JC, Kudo L, Geschwind DH, Lockhart DJ, Preuss TM, Barlow C: **Elevated gene expression levels distinguish human from non-human primate brains.** *Proc Natl Acad Sci USA* 2003, **100**:13030-13035.
40. Gilad Y, Rifkin SA, Bertone P, Gerstein M, White KP: **Multi-species microarrays reveal the effect of sequence divergence on gene expression profiles.** *Genome Res* 2005, **15**:674-680.
41. Ji W, Zhou W, Gregg K, Yu N, Davis S: **A method for cross-species gene expression analysis with high-density oligonucleotide arrays.** *Nucleic Acids Res* 2004, **32**:e93.
42. Kane MD, Jatkoe TA, Stumpf CR, Lu J, Thomas JD, Madore SJ: **Assessment of the sensitivity and specificity of oligonucleotide (50 mer) microarrays.** *Nucleic Acids Res* 2000, **28**:4552-4557.
43. Zhang X, Zhu L, Deutscher MP: **Oligoribonuclease is encoded by a highly conserved gene in the 3'-5' exonuclease superfamily.** *J Bacteriol* 1998, **180**:2779-2781.
44. Nguyen LH, Erzberger JP, Root J, Wilson DM 3rd: **The human homolog of *Escherichia coli* Orn degrades small single-stranded RNA and DNA oligomers.** *J Biol Chem* 2000, **275**:25900-25906.
45. Oh B, Hwang SY, Solter D, Knowles BB: **Spindlin, a major maternal transcript expressed in the mouse during the transition from oocyte to embryo.** *Development* 1997, **124**:493-503.
46. Oh B, Hampl A, Eppig JJ, Solter D, Knowles BB: **SPIN, a substrate in the MAP kinase pathway in mouse oocytes.** *Mol Reprod Dev* 1998, **50**:240-249.
47. Aoki K, Ishii Y, Matsumoto K, Tsujimoto M: **Methylation of *Xenopus* CIRP2 regulates its arginine- and glycine-rich region-mediated nucleocytoplasmic distribution.** *Nucleic Acids Res* 2002, **30**:5182-5192.
48. Matsumoto K, Aoki K, Dohmae N, Takio K, Tsujimoto M: **CIRP2, a major cytoplasmic RNA-binding protein in *Xenopus* oocytes.** *Nucleic Acids Res* 2000, **28**:4689-4697.
49. Chromas [<http://www.technelysium.com.au>]
50. Basic Local Alignment Search Tool (BLAST) [<http://www.ncbi.nlm.nih.gov/blast/>]
51. Microarray Gene Expression Data Society [<http://www.mged.org>]
52. ArrayExpress [<http://www.ebi.ac.uk/arrayexpress>]
53. NIA Array Analysis [<http://lgsun.grc.nia.nih.gov/ANOVA>]
54. Sharov AA, Dudekula DB, Ko MS: **A web-based tool for principal component and significance analysis of microarray data.** *Bioinformatics* 2005, **21**:2548-2549.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

