What can we learn from gene expression profiling of mouse oocytes?

Toshio Hamatani¹,², Mitsutoshi Yamada¹,², Hidenori Akutsu², Naoaki Kuji¹, Yoshiyuki Mochimaru¹, Mitsuko Takano¹, Masashi Toyoda², Kenji Miyado², Akihiro Umezawa² and Yasunori Yoshimura¹

¹Department of Obstetrics and Gynecology, Keio University School of Medicine, 35 Shinanomachi Shijuku-ku, Tokyo 160-8582, Japan and ²Department of Reproductive Biology and Pathology, National Institute for Child Health and Development, 2-10-1 Okura Setagaya-ku, Tokyo 157-8535, Japan

Correspondence should be addressed to T Hamatani; Email: t-hama@sc.itc.keio.ac.jp

Abstract

Mammalian ooplasm supports the preimplantation development and reprograms the introduced nucleus transferred from a somatic cell to confer pluripotency in a cloning experiment. However, the underlying molecular mechanisms of oocyte competence remain unknown. Recent advances in microarray technologies have allowed gene expression profiling of such tiny specimens as oocytes and preimplantation embryos, generating a flood of information about gene expressions. So, what can we learn from it? Here, we review the initiative global gene expression studies of mouse and/or human oocytes, focusing on the lists of maternal transcripts and their expression patterns during oogenesis and preimplantation development. Especially, the genes expressed exclusively in oocytes should contribute to the uniqueness of oocyte competence, driving mammalian development systems of oocytes and preimplantation embryos. Furthermore, we discuss future directions for oocyte gene expression profiling, including discovering biomarkers of oocyte quality and exploiting the microarray data for ‘making oocytes’.


Introduction

The mammalian oocyte is marked by extraordinary biological competence; it can haploidize its DNA, be fertilized and reprogram the sperm chromatin into a functional pronucleus, induce zygotic genome activation (ZGA), give rise to totipotency, and drive early embryonic development. Using its ability to reprogram a somatic nucleus transferred into an enucleated oocyte, derivation of embryonic stem (ES) cells from cloned blastocysts is pursued for therapeutic cloning. The molecular mechanisms underlining such oocyte competence are largely unknown.

On the other hand, the reproductive capacity of women declines dramatically beyond the mid-30s (van Kooij et al. 1996, ASRM/SART 2000, Armstrong 2001, Klein & Sauer 2001), which is mainly caused by age-related decline in oocyte quality. For example, young women undergoing standard in vitro fertilization (IVF) with their own eggs show a success rate comparable with older women (>40 years) undergoing IVF with eggs donated by this younger subset of women (Navot et al. 1991). To overcome age-related decline in oocyte quality, ooplasmic donation has been performed by injecting ooplasm from a young, healthy donor oocyte into a patient oocyte to improve the outcome of assisted reproduction techniques (Cohen et al. 1997, 1998, Takeuchi et al. 1999). There is, however, little molecular evidence supporting the efficacy and the safety of ooplasmic donation. Furthermore, no molecular biomarker for oocyte quality has been established. Oocyte quality based on a morphological grading system is the only reliable prognostic factor in human IVF programs. Studies of molecular mechanisms involved in oocyte quality could have important implications for the efficacy and safety of clinical ooplasmic donation.

Thus, understanding the molecular mechanisms in oocytes is quite important for both reproductive biology and regenerative medicine. The scarcity of the materials, however, both in size (diameter <100 um) and in quantity (only a few to tens of oocytes from each ovulation in mice), has hampered the molecular analysis of oocytes. Earlier attempts to analyze oocytes employed RT-PCR and differential display using only a few candidate genes. In addition, serial analysis of gene expression (SAGE) and cDNA libraries were generated from mouse and human oocytes, and SAGE tags and expressed sequence tags (ESTs) were sequenced for rapid gene discovery and expression profiling in oocytes.
Global gene expression profiling of mouse preimplantation embryos to dissect maternal transcripts

Two groups simultaneously published the first reports on global gene expression profiling of all stages of pre-implantation embryos (Fig. 1; Hamatani et al. 2004a, Wang et al. 2004). While Wang et al. used the Affymetrix 25-mer oligo DNA microarray system, we used the NIA 60-mer oligo microarray (Agilent Mouse Development 25-mer oligo DNA microarray system, we used the NIA 60-mer oligo microarray (Agilent Mouse Development Array), which is enriched for genes expressed in stem cells and preimplantation embryos (Carter et al. 2003). Taking advantage of 60-mer oligo DNA hybridization kinetics (Hughes et al. 2001), it was also optimized for use with tiny amounts of RNA (Carter et al. 2003). During preimplantation development, 12 179 out of 21 939 gene features on the NIA 60-mer oligo microarray showed statistically significant changes with false discovery rate (FDR) <10% by ANOVA-FDR test. Pair-wise comparison, hierarchical clustering analysis, and principal component analysis (PCA) revealed two major transient waves of de novo transcription (Fig. 1A–C). The first wave corresponds to ZGA. The second wave, mid-preimplantation gene activation (MGA), contributes to dramatic morphological changes during late preimplantation development.

To trace the expression changes of individual genes, 12 179 statistically significant genes were analyzed by k-means clustering method and 9 clusters were identified (Fig. 2). Gene expression patterns of these clusters can be assigned to three main groups. The first group appears to represent ZGA genes that are first activated from the zygotic genome (Clusters 1, 4, 5, and 8). The list of the ZGA genes suggests that ZGA is not promiscuous as previously proposed and contributes mainly to the preparation of basic cellular machinery during the two-cell and the four-cell stages.

The second group represents maternal transcripts with distinctive patterns of degradation during preimplantation development (Clusters 7 and 9). Although the massive maternal RNA degradation pattern by the two-cell stage is confirmed (Cluster 9) as previous studies suggested (Nothias et al. 1995, Schultz 2002), 70.5 and 32.5% of the transcripts in Clusters 7 and 9 respectively further show significant reduction from the four-cell to eight-cell stages. Selective degradation of maternal transcripts during oocyte maturation is, as also shown by the latest study (Su et al. 2007), a developmentally regulated event preceding the transition of gene expression from maternal to zygotic control. Since most genes in Clusters 7 and 9 are not reactivated during preimplantation development, the genes in these clusters are suggested to have specific functions either in oogenesis, oocyte maturation, fertilization, and/or early phases of preimplantation development.

The third group appears to represent genes that follow a combination of these two patterns (Clusters 2 and 3); 3329 genes whose expression first significantly increase from the four-cell to eight-cell stages are identified as the MGA genes, and 82.7 and 12.3% of them fall into Clusters 2 and 3 respectively. Further expression profiling of embryos treated with inhibitors of transcription and translation reveals that the translation of maternal RNAs is required for the initiation of ZGA, suggesting a cascade of gene activation from maternal RNA/protein sets to ZGA gene sets and thence to MGA gene sets (Hamatani et al. 2004a).

By MAPPFinder (Dahlquist et al. 2002, Doniger et al. 2003), which is a tool to identify global biological trends in gene expression data by interacting the annotations of Gene Ontology (GO) terms (Ashburner et al. 2000), the genes in the clusters of maternal transcripts are associated to such GO terms as ‘circadian rhythm,’ ‘M-phase of mitotic cell cycle,’ ‘DNA replication,’ ‘Golgi apparatus/intracellular protein transport,’ ‘adherent junction,’ ‘small GTPase regulatory/interacting protein,’ and ‘intracellular signaling cascade’. The ‘circadian rhythm’ category includes seven mammal known circadian genes: Per1–3, Cry1–2, Bmal1/Arntl, and Clock. The transcripts of Bmal1/Arntl, Clock, Timeless, Cry1, and Csnk1e decrease during the one-cell to two-cell stages as previous reports showed (Johnson et al. 2002).

The egg–sperm fusion at fertilization in mammals releases an oocyte from metaphase II arrest by increasing Ca²⁺ levels, activating Ca²⁺-calmodulin kinase II, and targeting cyclin B and c-mos for degradation via the ubiquitin–proteasome pathway. Ripl4, an E3 ubiquitin protein ligase, regulates the degradation of cyclin B1 (Ccnb1) protein (Cluster 6b) (Suzumori et al. 2003), which is a well-known example of a transcript with a short poly(A) tail that is regulated at the post-transcriptional level in oocytes. Furthermore, Cpeb, Eif4e, Cpsl2, and Stk13/Aurkc, which are involved in the masking and/or translational regulation of transcripts with short

(Ko et al. 2000, Ko 2004, Adjaye 2005, Evsikov et al. 2006). Furthermore, the recent progress in RNA amplification methods and microarray platforms including genes unique to oocytes and preimplantation embryos allows us to apply global gene expression profiling to the studies of the oocytes and preimplantation embryos (Carter et al. 2003). To date, several reports of the oocyte transcriptome using unique biological models have been published (Dobson et al. 2004, Hamatani et al. 2004a, 2004b, Wang et al. 2004, Zeng et al. 2004, Fan et al. 2005, Assou et al. 2006, Kocabas et al. 2006, Yoon et al. 2006). The identification of a large number of genes expressed in oocytes, especially oocyte-specific genes, and multiple signaling pathways in the models by such global gene expression profiling is the first step toward understanding oocyte quality and the molecular mechanisms underlying oogenesis, developmental programs, and totipotency in preimplantation embryos.
poly(A) tails in oocytes (Hodgman et al. 2001, Mendez & Richter 2001), also decrease their transcripts by the two-cell stage. The presence of the ‘DNA replication’ category in oocytes indicates that oocytes are already well equipped with DNA replication machinery, as exemplified by the fact that neither the lack of Zar1 (Wu et al. 2003) nor the presence of jasplakinolide, which is the most powerful known microfilament inhibitor (Terada et al. 2000), can prevent the initiation of DNA replication. In another global gene expression study of preimplantation embryos, DNA repair genes are also over-represented at the oocyte stage when compared with the one-cell through the blastocyst stages in their transcript profiling during preimplantation development (Zeng et al. 2004). Genes that are down-regulated from oocytes to two-cell embryos include many genes involved in DNA repairs, including Orc1l, Orc4l, Orc5l, Orc6l, Mcm4, Pcna, Pola2, Polm, Blm, Top1, and Msh6 (Cluster 9); Msh3 and Mcm7 (Cluster 7); and Cdc711/Cdc7, Cdc45l, Ccnl2, and Dbf4/Ask (Cluster 6). Furthermore, another group searched for maternal transcripts of polarity-regulating genes in mouse oocytes by global gene expression profiling of preimplantation embryos, which may subsequently control polarity in preimplantation embryos (Wang et al. 2004). They focused on three genes whose homologs have been shown to regulate cellular polarity in Drosophila: Flamingo, dystroglycan1 (Dag1), and cornichon (Cnih2) both of which are included by Cluster 3.

Global gene expression changes during oogenesis

Although several groups have studied global gene expression in human and mouse oocytes at the later stages of folliculogenesis (germinal vesicle stage and metaphase II stage; Wang et al. 2004, Cui et al. 2007,
Gasca et al. 2007, Zhang et al. 2007), only two groups successfully performed global gene expression studies using mouse oocytes at the very early stages of folliculogenesis (Pan et al. 2005, Yoon et al. 2006). Pan et al. (2005) compared the transcriptomes of mouse oocytes obtained from day 2 primordial follicles to day 22 equine CG primed large antral follicles. From the primordial to large antral stages, 18 529 probe sets corresponding to 11 766 unigenes detected significant gene expression in oocytes that developed in vivo. The hierarchical clustering dendrogram and PCA analysis showed that the primordial oocyte is separated from oocytes obtained from the other stages. Many important genes encoding ‘secreted proteins’, which are defined on their own terms in that manuscript, display marked upregulation between the primordial and primary follicle stages (e.g., Gdf9, Bmp15, Bmp5, Tgfβ2, Tgfβ3, and several genes related to Notch, Shh, and Egf signaling pathways). Thus, the primordial to primary follicle transition is a major transition and likely reflects the dramatic reorganization in follicle structure and initiation of growth and development. Of the 16 883 probe sets differentially detected between these stages, 5020 display a twofold change in relative abundance. Another apparent transition occurs between oocytes obtained from secondary follicles and those from small antral follicles, which corresponds to the acquisition of meiotic competence. The 736 probe sets of which ~65% are downregulated display a significant twofold change at this transition.

The principal component-based clustering shows three distinct patterns of gene expression. The first pattern shows consistent increase or decrease throughout the oocyte development and the most dramatic changes from the primordial to primary follicle stages, which the bulk of genes (10 117 probe sets) display (Fig. 3A). The second pattern peaks or hits the bottom at the primary follicle stage (Fig. 3B) and the third one shows the dynamic expression changes from the primary to the secondary follicle stages (Fig. 3C). The Expression Analysis Systematic Explorer software (http://david.abcc.ncifcrf.gov/ease/ease.jsp) for discovery of biological themes within the list of genes also shows the over-representation of genes involved in DNA repair and response to DNA damage throughout oocyte development, suggesting a protective mechanism to insure genomic integrity of the female germ line.

In addition, by analysis of global gene expression profiling of oocytes during the germinal vesicle stage to the metaphase II stage, new potential regulators and marker genes for oocyte maturation have been identified: Pacsin2, Map2k (Cui et al. 2007), and the genes related to BRCA1 regulation pathway, including Bard1, Rbpb4, Brap, Rbpb7, Rbl2, Bub3, and Bub1b (Gasca et al. 2007).

Global gene expression changes during loss of oocyte quality
To elucidate factors determining oocyte quality, a mouse model highlighting the age-related decline in fertility and oocyte quality was used (Hamatani et al. 2004b, Steuerwald et al. 2007). The expression profiles of metaphase II oocytes collected from 5- to 6-week-old mice were compared with those collected from 42- to 45-week-old mice using the NIA 60-mer oligo microarray (Hamatani et al. 2004b). Among ~11 000 genes whose transcripts were detected in oocytes, about 5% (530) showed statistically significant expression changes,
excluding the possibility of global decline in transcript abundance. Consistent with the generally accepted view of aging, the differentially expressed genes include ones involved in mitochondrial function and oxidative stress. Interestingly, a new non-invasive and highly sensitive method for measuring cellular respiration with scanning electrochemical microscopy shows that decreased cellular respiration in oocytes from aged mice is associated with impaired preimplantation development (Abe 2007). However, the expression of other genes involved in chromatin structure, DNA methylation, genome stability, and RNA helicases are also altered, suggesting the existence of additional mechanisms for aging in oocytes. For example, the decreased Dnmt1 (Dnmt1o and Dnmt1s) expression and the increased Dnmt3b during aging are observed in oocytes. Because the same pattern of expression change in Dnmt genes has already been reported in aging WI-38 fibroblast cells (Lopatina et al. 2002), the genomic methylation patterns are suggested to be altered in aging cells. Telomerase reverse transcriptase and yeast mutant H/LS mismatch repair gene homologs are also downregulated during aging. Interestingly, more than 30 zinc finger proteins are shown as the downregulated genes during aging. Furthermore, we identified and characterized a group of new oocyte-specific mouse genes, members of the human NACHT, leucine rich repeat and pyrin domain containing (NALP/NLRP) gene family among the transcripts decreased with aging. The Nalp gene family includes Mater/Nalp5/Nltp5 whose null mutant embryos arrest cleavage at the two-cell stage (Tong et al. 2000), suggesting an important role of this gene family in oogenesis, fertilization, and/or preimplantation development. These results have implications for aging research as well as for clinical ooplasmic donation to rejuvenate aging oocytes.

Polycystic ovary syndrome (PCOS) is another good model for studying loss of oocyte quality. The reproductive performance of women undergoing IVF treatment with PCOS is characterized by their good response to ovarian stimulation that yields higher number of oocytes; however, with lower implantation and higher miscarriage rates (Engmann et al. 1999, Ludwig et al. 1999, Mulders et al. 2003). Individual oocytes retrieved from nine women with PCOS and that from ten non-hirsute ovulatory women are used for microarray hybridization (Wood et al. 2007). Of the 8123 transcripts expressed in metaphase II oocytes, 374 show significant differences in mRNA abundance in the PCOS oocyte. The genes associated with chromosome alignment and centrosome, and the genes containing putative androgen receptors and/or PPAR-binding sites are upregulated. The expression of these genes, which is generally not a part of the human oocyte transcriptome, is suggested to contribute to abnormalities in early embryonic development. Furthermore, upregulation of maternal-effect genes are notable. Although only seven mammalian maternal-effect genes (Mater/Nltp5, Hsf1, Dnmt1, Zar1, Npm2, Stella, Fmn2, and Bnc1) have been identified to date, three (Mater/Nltp5, Fmn2, and Bnc1) are upregulated. Increased expression of maternal-effect genes may negatively impact embryonic development.

Dielectrophoresis is a potential non-invasive method to select oocytes of good quality. In fact, dielectrophoretically separated in vitro-derived bovine metaphase II oocytes show a difference in the rate of blastocyst development and significant difference in transcriptional abundance of 36 genes as a result from global gene expression profiling. This suggests that dielectrophoretic behavior and the 36 genes including Anxa2, Ptg2s2, and Dnmt1 are potential biomarkers for oocyte quality (Dessie et al. 2007).

Recently, microarray technology was also applied to screening for chromosomal anomalies: comparative genomic hybridization (CGH) is used to assess the copy number of chromosomes in polar bodies and oocytes (Wells et al. 2002, Fraguoli et al. 2006). CGH has the major advantage that every chromosome is tested, rather than the limited subset assessed using fluorescence in situ hybridization (FISH). The CGH protocols, which allow efficient DNA amplification from single cells and

---

**Figure 3** Principal component-based clustering to analyze gene expression profiles of oocytes during the primordial follicle stage to the large antral follicle stage (Pan et al. 2005) by NIA array analysis tool (http://lgsun.grc.nia.nih.gov/ANOVA/). The NIA Array Analysis tool identifies two clusters associated with a given pattern: genes positively and negatively correlated with the pattern. P, 1, 2, SA, and LA represent the primordial follicle stage, the primary follicle stage, the secondary follicle stage, the small antral follicle stage, and the large antral follicle stage respectively.
reduce the amount of time required for the analysis, are currently undergoing preclinical testing in a number of preimplantation genetic diagnosis studies (Patrizio et al. 2007).

**Identification of oocyte-specific transcripts and their clustering in the mouse genome**

A mammalian oocyte is the only known cell that can activate a zygotic genome after fertilization and reprogram a somatic nucleus transferred from a differentiated cell in cloning experiments. Therefore, several genes specifically expressed in oocytes are likely responsible for the ability to reprogram genomes as well as for oogenesis. It is the case for the so-called maternal genes such as *Mater*, *Zar1*, and *Npm2* that are all required for normal embryonic development beyond the one-cell or two-cell stage (Fig. 4A; Tong et al. 2000, Dean 2002, Burns et al. 2003, Wu et al. 2003). *Gdi9* and *Bmp15* are also known to play important roles in female germ cells during folliculogenesis (Dong et al. 1996, Galloway et al. 2000). Accordingly, genes specifically expressed in the oocyte seem to control oogenesis, ovarian folliculogenesis, and preimplantation development.

In attempts to identify novel oocyte-specific genes, several groups have used mRNA differential display (Zeng & Schultz 2003), suppression subtractive hybridization (Hennebold et al. 2000), and *in silico* subtraction approaches (Rajkovic et al. 2001, Dade et al. 2003). It is, however, essential to analyze all transcripts/genomes in a wide selection of organs and cell types including totipotent fertilized eggs, pluripotent embryonic cells, a variety of adult stem cells, and terminally differentiated cells. Sharov et al. (2003) obtained 249 200 high-quality EST sequences from the NCBI Unigene database that included a broad collection of NIA mouse cDNA libraries and clustered them into ~30 000 gene indexes including 977 previously unidentified genes. By analyzing the expression levels of the gene indexes based on the frequency of the corresponding ESTs in Unigene cDNA libraries, genes that characterize oocytes and preimplantation embryos are identified (Sharov et al. 2003). Furthermore, the gene expression specificity to oocytes or/and preimplantation embryos is validated using gene expression profiling data of female germ cells during oogenesis and preimplantation embryos (Hamatani et al. 2004a, Wang et al. 2004, Pan et al. 2005). Several example of genes preferentially expressed in oocytes are selected and their gene expression levels are demonstrated to increase in oocytes during oogenesis (from the primordial follicle stage to the large antral follicle stage) and decrease during preimplantation development (from unfertilized egg to blastocyst) by the microarray experiments (Fig. 4B). Mager et al. (2006) also identified 51 genes as candidate maternal-effect genes *in silico* (always not present during the two-cell through the eight-cell or at the blastocyst stage), by comparing published results of three independent studies of mouse preimplantation embryo transcriptomes (Hamatani et al. 2004a, Wang et al. 2004, Zeng et al. 2004).

The group that found six genes of the mouse oogenesin family reported that not a few loci near the telomere in the mouse genome contain several genes specifically expressed in oocytes (Paillisson et al. 2005). Mouse oogenesin family genes are expressed exclusively in oocytes and present on chromosomes 4 in a cluster of almost 1 Mb composed of 12 ooggenesin paralogous genes. On the other hand, we also identified nine novel genes presenting similarities with Mater/Nalp5 (Tong et al. 2000) and expression specific to oocytes, seven of which are clustered on a certain locus of chromosome 7 (Fig. 5). The gene expression specificity of the novel Nalp-family genes to oocytes has been experimentally validated using Northern blot and *in situ* hybridization (Hamatani et al. 2004b; Fig. 6). Recently, we further identified a group of oocyte-specific genes encoding zinc finger proteins that clusterize in a near-telomere locus of chromosomes 6 and 11 (unpublished data). Telomeric regions of chromosomes are mainly composed of heterochromatin in most eukaryotic genomes. Because gene silencing near the telomere has been known and called ‘telomere position effect’ in *Drosophila* and yeast, the specific near-telomere position of the clusters of oocyte-specific genes in mice may contribute toward their gene silencing in non-ovarian tissues.

There is another noted attempt to identify important genes that are preferentially expressed in oocytes and conserved in chordates. Eviskov et al. (2006) compared the collection of ESTs from their mouse oocyte libraries to those from the eggs of *Xenopus* and ascidians to extract conserved genes that are expressed in chordate oocytes. More than 50% of the genes expressed in mouse oocyte libraries are also expressed in the eggs of *Xenopus* and ascidians. To investigate the evolutionary hardwired molecular pathways shared among chordates, GO term frequencies in 2090 genes that are commonly expressed in eggs of all three species are compared with those in the entire set of genes expressed in their mouse oocyte library. Although this analysis shows a substantial overlap with GO terms (biological process) associated with housekeeping genes, several GO terms (molecular function) such as ‘motor activity,’ ‘small protein activating enzyme activity,’ ‘transferase activity,’ ‘helicase activity,’ and two specific signal transducer activities (serine/threonine kinase activity and ligand-dependent nuclear receptor activity) are over-represented and provide a snapshot of gene functions shared particularly by the chordate oocytes.

Another study group used a multi-species cDNA microarray containing 3456 transcripts from three distinct cDNA libraries from bovine, mouse, and *Xenopus* oocytes (Vallee et al. 2006). The cross-species hybridizations reveal that 1541 positive hybridization signals are generated by oocytes of all three species, and 268 of these, including SMFN (small fragment nuclease), *Spin* (spindlin), and...
Figure 4  (A) Knockout mouse phenotypes of genes preferentially expressed in oocytes. *Development of embryos from Stella−/− intercrosses starts to be affected from 1.5 dpc onward (the two-cell stage) and only a low percentage reach the blastocyst stage by 3.5 dpc. (B) The gene expression changes of several genes known as oocyte specific. The oocyte-specific genes, including Nalps, showed increased expression during oogenesis and decreased expression during preimplantation development in the global gene expression studies. P, 1, 2, SA, and LA represent the primordial follicle stage, the primary follicle stage, the secondary follicle stage, the small antral follicle stage, and the large antral follicle stage respectively. U, F, 2, 4, 8, M, and B denote unfertilized egg, fertilized egg, two-cell embryo, four-cell embryo, eight-cell embryo, morula, and blastocyst respectively.
PRMT1 (protein arginine methyltransferase 1) transcripts, are preferentially expressed in oocytes (Vallee et al. 2006). Furthermore, an important molecular characteristic of germ cells was also reported; germ cell-specific regulation of core promoter-associated transcription factors is conserved between Xenopus and mice (Xiao et al. 2006). Tbp12/Trf3 and Gtf2a1lf/Alf are demonstrated to be expressed preferentially in oocytes and can form in vitro core promoter complexes with TBP and TFIIA. Therefore, identifying other germ cell-specific transcription factors is necessary to understand the genetic cascades that drive oocyte development and folliculogenesis.

Comparison of oocytes with ES cells in terms of their gene expression profiles

Recent studies on cell fusion between a somatic cell and an ES cell suggest that cytoplasm of ES cells can reprogram an introduced somatic nucleus to confer pluripotency. In this aspect, the cytoplasmic environments of ES cells and oocytes share the capacity to reprogram a somatic nucleus (Tada et al. 2001, Cowan et al. 2005). Accordingly, a set of genes commonly expressed in oocytes and ES cells are likely responsible for reprogramming somatic cells. To identify these genes, gene expression profiling data of human oocytes and human ES cells were explored (Kocabas et al. 2006, Zhang et al. 2007). Compared with reference samples, 5331 and 1626 transcripts are significantly upregulated in human oocytes and ES cells respectively (Kocabas et al. 2006). When the genes differentially upregulated in human ES cells are intersected with those differentially upregulated in human oocytes, 388 transcripts are overlapped. This list of genes, including POU5F1/OCT4, DNMT3b, DAZL, and high-mobility group proteins (HMGB2, HMGB3, and HMGN4) (Kocabas et al. 2006), may provide good candidate genes for the future studies on molecular mechanisms of nuclear reprogramming.

Figure 5 Clusters of oocyte-specific genes on mouse chromosomes 4, 6, 7, 9, 12, and 19.

Figure 6 Oocyte-specific expression of the novel Nalp-family mouse genes (reprinted from 'Age-associated alteration of gene expression patterns in mouse oocytes', Hamatani T et al. 2004 Human Molecular Genetics 13 2263–2278, with permission from Oxford University Press). (A) Northern blot analysis shows their ovary-specific expression and (B) in situ hybridization shows their oocyte-specific expression on ovary sections.
On the other hand, ‘induced pluripotent stem cells (iPS cells)’ were recently generated by forced expression of defined factors: Pou5f1/Oct4, Sox2, Klf4, and Myc (Takahashi & Yamanaka 2006). Surprisingly, iPS cells selected by Nanog expression are capable of germ cell transmission (Okita et al. 2007). These iPS factors, however, show little maternal expression in oocytes (except in the case of Oct4) and increased zygotic expression during preimplantation stages (except in the case of Myc), based on EST frequencies in Unigene cDNA libraries and microarray data during oogenesis to preimplantation development (Fig. 7). Therefore, the mechanism of oocytes to induce pluripotency is likely different from that of ES cells. Although the genes commonly expressed in oocytes and ES cells are not necessarily important to induce pluripotency, maternal factors that can induce zygotic expression of the ‘iPS factors’ (Oct4, Sox2, and Klf4) are rather more substantial in oocytes.

**Perspective**

Oocytes offer a relatively homogeneous biological system that is well adapted to gene expression profiling studies: arrest of cell cycle at the metaphase II stage, quiescence in transcription after germinal vesicle breakdown, and little contamination in oocyte samples with any other types of cells after thorough removal of cumulus cells. There are, however, several limitations in applying microarray technologies to study the molecular mechanisms in oocytes and preimplantation embryos. Although the recent advent of linear RNA amplification (in vitro transcription-based protocols) and exponential amplification (PCR-based strategies) techniques allowed several groups to study oocyte transcriptomes using a tiny amount of RNA even in an individual oocyte (Bermudez et al. 2004, Dobson et al. 2004, Li et al. 2006, Jones et al. 2007), the efficacy of RNA amplification is not yet good enough to analyze an individual blastomere of preimplantation embryos. Furthermore, poly(A) length affects efficiency of RNA amplification. Although the synthesis of new transcripts essentially ceases after germinal vesicle breakdown, poly(A) tails of some classes of existing transcripts in oocytes are elongated, leading to increased translation and protein levels (Bachvarova 1992). Thus, regulation of the poly(A) tail length is a major mechanism for controlling maternal transcript activity. Unlike the T7-oligo(dT) primers used in the conventional linear RNA amplification procedures, the uniquely designed Full Spectrum MultiStart Primers for in vitro transcription from System Biosciences (Mountain View, CA, USA) initiates cDNA synthesis at multiple points along mRNAs with little or no bias with respect to the length of poly(A) tails. Transcript profiles generated from microarray studies using this modified RNA amplification protocol would provide a more accurate perspective of the global changes in populations of both degrading and stable transcripts during oocyte maturation and ZGA (Su et al. 2007).

‘Tailor-made regenerative medicine’ includes nuclear transfer from a patient’s somatic cell to an enucleated donated oocyte, development of the reconstructed embryo up to the blastocyst stage, and establishment of the patient’s ES cells. ‘Making oocytes’ will be an essential technique to develop ‘tailor-made regenerative medicine’ that needs large quantities of healthy ooplasms. ‘Oocyte-like’ cells were recently grown and isolated by utilizing GFP expression as a selection marker during differentiation of ES cells containing GFP expression cassettes under the Pou5f1 promoter (Hubner et al. 2003). Nobody, however, has succeeded in generating oocytes by manipulating gene expression in ES and somatic cells. Even though forced expression of a set of several transcription factors in ES cells
may allow us to generate an oocyte, there is a problem in the oocyte; its nucleus ought to have genetic abnormalities. In contrast, its ooplasm might contain all the gene products that can support embryonic development after fertilization. Unlike another strategy using iPS cells that cannot avoid transmitting genetic abnormalities, the ooplasm can be safely used for ‘tailor-made regenerative medicine’ or ‘ooplasmic donation’.

Since transcriptional cascades that activate an oocyte-specific developmental program are largely unknown, a set of master genes that drive the cascades have not yet been defined. Oocyte-specific transcription factors, however, are likely to be the critical switches for the differentiation into oocytes and good candidates for manipulation of gene expression. For example, NOBOX binds to the NOBOX binding elements with high affinity and augments transcriptional activity of mouse Pou5f1 and Gdf9 promoters (Choi & Rajkovic 2006). Other examples are factor in germ cell (FIGLA) and SOHLH1 that bind to E-box. They are suggested to increase transcriptional activity of Zp1–3, which have promoters including E-box (Yan et al. 2006; Pangas, 2006 #613).

On the other hand, nobody pays attention to a transcription factor whose knockout showed no distinctive phenotypes. Nonetheless, recent advent in microarray technologies allows us to catch any changes in a gene expression profile of cells transfected with a construct to modify gene expression. If a gene expression profile of ES cells approaches that of oocytes in the PCA coordinate, in spite of no phenotypic change, by upregulation of a certain transcription factor, the transcription factor is likely a candidate gene as a tool to induce the oocyte developmental program (Fig. 8). Further forced expression of another transcription factor in the ES cell may result in a similar gene expression profile to that of oocytes and then may achieve a certain remarkable phenotype including follicles or oocytes. Such synergy between cell biology and bioinformatics will become more important and beneficial to establish an in vitro oocyte-development model to ‘make an oocyte’.

**Figure 8** Explanation of a model to transform ES cells efficiently into oocytes using gene expression profiling as a guide. PC, principle component; PGC, primordial germ cell; ICM, inner cell mass; TE, trophoderm; F, fertilized egg; 2, two-cell embryo; 4, four-cell embryo; 8, eight-cell embryo; M, morula; B, blastocyst. If over-expression of ‘gene A’ makes a global expression profile of ES cells closer to that of oocytes in the PCA coordinate, ‘gene A’ could be a good candidate to promote the oocyte developmental program. Even though no changes in phenotypes of ES cells are observed with over-expression of ‘gene A’, forced expression of ‘gene A’ plus that of ‘gene B’ in ES cells may show a distinctive phenotype including oocytes or follicles.

**Acknowledgements**

The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.

**References**


Johnson MH, Lim A, Fernando D & Day ML 2002 Circadian clockwork genes are expressed in the reproductive tract and conceptus of the early pregnant mouse. Reproductive Biomedicine Online 4: 140–145.


Received 22 September 2007
First decision 31 October 2007
Revised manuscript received 1 February 2008
Accepted 27 February 2008