Genomics Proteomics Bioinformatics **12** (2014) 239–248

HOSTED BY

ELSEVIER

Genomics Proteomics Bioinformatics

www.elsevier.com/locate/gpb
www.sciencedirect.com

ORIGINAL RESEARCH

A Comprehensive Transcriptomic Analysis of Infant and Adult Mouse Ovary

Linlin Pan [1,2], Wei Gong [2], Yuanyuan Zhou [2], Xiaonuan Li [1,2], Jun Yu [2,*], Songnian Hu [2,*]

[1] *James D. Watson Institute of Genome Sciences, Zhejiang University, Hangzhou 310058, China*
[2] *CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China*

Received 26 March 2014; revised 31 July 2014; accepted 2 August 2014
Available online 20 September 2014

Handled by Yongliang Zhao

KEYWORDS

Folliculogenesis;
Postnatal ovary development;
Transcriptome;
RNA-seq

Abstract Ovary development is a complex process involving numerous genes. A well-developed ovary is essential for females to keep fertility and reproduce offspring. In order to gain a better insight into the molecular mechanisms related to the process of mammalian ovary development, we performed a comparative transcriptomic analysis on ovaries isolated from infant and adult mice by using next-generation sequencing technology (SOLiD). We identified 15,454 and 16,646 transcriptionally active genes at the infant and adult stage, respectively. Among these genes, we also identified 7021 differentially expressed genes. Our analysis suggests that, in general, the adult ovary has a higher level of transcriptomic activity. However, it appears that genes related to primordial follicle development, such as those encoding Figla and Nobox, are more active in the infant ovary, whereas expression of genes vital for follicle development, such as *Gdf9*, *Bmp4* and *Bmp15*, is upregulated in the adult. These data suggest a dynamic shift in gene expression during ovary development and it is apparent that these changes function to facilitate follicle maturation, when additional functional gene studies are considered. Furthermore, our investigation has also revealed several important functional pathways, such as apoptosis, MAPK and steroid biosynthesis, that appear to be much more active in the adult ovary compared to those of the infant. These findings will provide a solid foundation for future studies on ovary development in mice and other mammals and help to expand our understanding of the complex molecular and cellular events that occur during postnatal ovary development.

Introduction

The development of the mammalian ovary is a complex process that requires precise coordination of multiple molecular, cellular and histogenetic events. During the mammalian

* Corresponding authors.
E-mail: junyu@big.ac.cn (Yu J), husn@big.ac.cn (Hu S).

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

http://dx.doi.org/10.1016/j.gpb.2014.08.002

ovary development, oocytes are induced to differentiate into somatic ovarian cells, whereas germ cells are dispensable in the development of testis [1]. At birth, all oocytes are either 'naked' or only have a few surrounding granulosa cells. Subsequently, the accretion of somatic cells surrounding the oocytes forms primordial follicles, which preserve the primordial pool. Later, these primordial follicles are recruited and grow to mature preovulatory follicles through a process called folliculogenesis. This process is initially controlled by local intercellular signaling events, but later, it responds to the hypothalamic–pituitary axis that regulates selective ovulation.

Folliculogenesis has been described as a two-phase process: the initial recruitment of the follicle from the primordial pool to preantral follicle and the cyclic recruitment of the growing follicles, involving gonadotropin-dependent stages of rapid growth, during which a preantral follicle will develop into a mature preovulatory follicle. Many of the genes that are directly involved in mouse folliculogenesis have been identified. For instance, *Figla*, *Wnt4* and *TrkB* are essential for primordial follicle formation, whereas *Foxo3*, *Gdf9* and *Bmp15* are required for follicle recruitment and growth [2].

Previous transcriptomic surveys of follicle development and ovary formation have focused on a limited number of genes. More recently, microarray studies using whole ovaries from young and old mice have revealed a number of genes that are differentially expressed depending on the age of the female mouse and the developmental phase of the follicles and oocytes [3,4]. For example, in a microarray study of 15,000 genes expressed in the ovary, the expression level of 2000 of these genes was found to be significantly different between newborn and adult mice [5]. In another study, 307 genes were found to be differentially expressed in the ovary during day 2 through day 4 of postnatal development [6]. However, microarray technology has several limitations. For instance, microarrays are devised based on a number of known genes and have limited power in detecting alternatively spliced genes.

Recently, RNA-seq, which is based on the next generation sequencing technology, has been shown to have advantages in both sensitivity and cost [7]. This technique has been used to survey the transcriptomes of soybean, rice, mice and human tissues [8–10]. To provide basic transcriptomic data for postnatal ovary development, we used RNA-seq to deeply sequence the transcriptomes of infant and adult mouse ovaries. Furthermore, using this data, we have performed a comprehensive analysis of the transcriptome dynamics occurring during this period that may serve as a gene expression profile for ovary development. By identifying the differentially-expressed genes (DEGs), we have also provided a blueprint for further investigation of the intricate and potentially subtle changes in gene expression and function that occur during postnatal ovary development from one to eight weeks of age.

Results and discussion

Global analysis of gene expression

To obtain a broad, unbiased evaluation of the transcriptomic dynamic occurring during postnatal development of the mouse ovary, we performed the ribo-minus RNA-sequencing (rmRNAseq) using the ABI SOLiD system to obtain gene transcription profiles for both infant and adult ovary tissues at single-nucleotide resolution. We obtained 23,866,480 and 26,446,960 mapped reads (50 bp in length) for the infant and adult ovary, respectively (**Table 1**). For the adult library, 17,476,893 reads (66.1% of the total mapped reads for adult) were uniquely mapped to the mouse genomic loci, among which, 29.3%, 39% and 27.6% were distributed in exons, introns and intergenic regions, respectively. In addition, we generated 12,585,638 uniquely-mapped reads (52.7%) from the infant library, 15.8% in exons, 40.4% in introns and 40.3% in intergenic regions.

One of the primary aims of RNA sequencing is to compare the gene expression levels between samples. The adult ovary usually contains more developing follicles and is much more active than the infant ovary [11]. Therefore, we speculate that the transcriptional activity of the adult ovary should be higher than that of the infant ovary. To test this, we used two parameters to estimate the holistic transcriptional activity at both developmental stages: the total number of genes expressed at each stage and the sum of the reads per kilobase transcript per million reads (RPKM) values for all of the expressed genes at each stage. Firstly, we estimated the number of genes expressed in the infant and adult ovary. To minimize the number of false positives, we consider a given gene as expressed if at least two reads were uniquely mapped to the exonic region of the gene. Using this threshold, we identified 15,454 and 16,646 genes expressed in the infant and adult ovary, respectively, indicating that more genes are transcribed in the adult than in the infant. Altogether, we detected a total of 17,102 genes expressed in the mouse ovary at one or both developmental stages. While there was a large overlap in gene transcription between the infant and adult ovaries, there were also a substantial number of genes that were uniquely expressed at a particular stage. Expression of 456 genes was only detected in the infant ovary, whereas 1648 genes were only expressed in the adult ovary (**Figure 1A**). Secondly, we calculated the RPKM values for each expressed gene and summed them to get the holistic transcriptomic activity at each stage. The summed RPKM values for infant and adult ovaries are 51,719 and 92,684, respectively. Gene distribution analysis further showed that there were more genes expressed at low levels in the infant ovary compared to the adult ovary, while more highly expressed genes were present in the adult ovary compared to the infant ovary (**Figure 1B**). Therefore, this global analysis of gene expression validated our hypothesis that the transcriptional activity of the adult ovary is much higher than that of the infant ovary.

Upon further scrutinization of the highly-expressed genes (RPKM $\geqslant$ 20), we found that an appreciable number of genes (229 genes) were highly expressed at both stages. However, more genes were highly expressed in the adult only (459 in adult only *vs.* 64 in the infant only) (**Figure 1C**). For example, *Nr5a1*, *Sox4* and *Inhba* were uniquely expressed at high levels in the adult ovary, while *Zp3*, *Col5a1* and *Maz* were highly expressed in the infant ovary. Among the top 10 most highly expressed genes at the infant and the adult stages, 5 genes are common between the two stages. These include *Bat1a* encoding ATP-dependent RNA helicase UAP56/SUB2, *Nd1* encoding NADH-ubiquinone oxidoreductase chain 1, *Gnb2l1* encoding guanine nucleotide-binding protein subunit beta-2-like 1, *Cox1* encoding cytochrome c oxidase subunit I and *Inha* encoding inhibin alpha chain precursor. The high expression level of *Nd1* and *Cox1* at both stages indicates the great need

Table 1 Summary of read mapping from adult and infant mouse ovaries

Read mapping	Infant	Adult
Raw tags	45,160,053	63,828,998
High-quality reads	43,428,824	56,210,485
Mapped reads	23,866,480	26,446,960
Reads mapped to unique loci	12,585,638	17,476,893
Reads mapped to multiple loci (2–10)	8,280,378	7,082,772
Reads mapped in high redundancy (>10)	3,000,464	1,887,295
Reads mapped to exon–exon junction database	335,617	782,159
Unique reads mapped to exons	2,423,442	5,829,814
Unique reads mapped to introns	5,085,428	6,815,597
Unique reads mapped to intergenic regions	5,076,768	4,831,482

for mitochondrial biogenesis and energy during the activation of cell proliferation and growth in ovary development. Meanwhile, the appearance of *Inha* in the top 10 most expressed genes at both stages suggests an important role for hormone-regulated metabolic processes during ovary development. Several genes with unknown function in the mouse were also among the top 10 expressed genes, such as 2700023E23Rik and 1810032O08Rik (RIKEN cDNA database) in the infant, and 4921521F21Rik in the adult, suggesting their potential involvement during ovary development. The identity and function of these unknown genes require further studies.

Analysis of differentially-expressed genes

We applied Fisher's Exact Test on the read count of each gene in the infant and adult to acquire statistical support for the significant expression differences between the two stages. We used the DEGseq package to perform this statistical analysis [12]. With the significance level of $P < 0.001$, we obtained 7021 DEGs, including 5862 upregulated genes and 1159 downregulated genes (adult *vs*. infant). Many of these DEGs are known to be essential for ovary development, such as *Nobox*, *Figla*, *Gdf9* and *Zp3*. It was striking that 46 genes showed significant changes in expression level but did not have any functional annotation in Ensembl. Detailed information on all the DEGs is listed in Table S1.

Further inspection of the DEG distribution in the mouse genome revealed that number of DEGs correlated well with the gene density in a specific chromosome. Thus, more DEGs were mapped on a chromosome if more expressed genes were present on the chromosome (**Figure 2**). We also found that the highest number of the DEGs were originated from chromosome 11 (569 genes), which is followed by chromosome 2, 7, 4 and 9, respectively (Figure 2). It has also been reported that abnormalities in the X chromosome, the only idiochromosome in females, caused by mutations in X-linked genes are the leading identifiable cause of premature ovarian failure [13]. Our data indicate that 237 X-linked genes were differentially expressed during ovary development, consisting of 209 upregulated genes and 28 downregulated genes in the adult compared to the infant.

DEGs were predicted to be important for the specific developmental stage of the mouse ovary [5]. To gain a better understanding of the biological implications of the DEGs during ovary development, we performed a standard Gene Ontology (GO) classification. As a result, the 7021 DEGs were categorized in three basic functional groups: molecular function, biological process and cellular component groups. The top six functional categories among the three groups included cell part, cell binding, cellular process, organelle, metabolic process and biological regulation (**Figure 3**). Furthermore, we performed a functional enrichment analysis using WebGstalt2 to determine the enriched functional groups in the up- and downregulated genes, respectively [14]. Using all of the genes in Ensembl as a reference and setting a threshold cut-off value of $P \leqslant 0.001$, our GO enrichment analysis suggested extensive functional differences. The downregulated genes reflected a vast repertoire of genes involved in regulating nuclear activity and chromatin assembly during morphogenesis in the infant ovary, while the upregulated genes were related to growth and metabolic processes (Table S2). This finding agrees with the physiological characteristics of the two developmental stages studied here. In the infant, all oocyte nuclei are arrested at dictyate and most of the follicles are dormant in the primordial follicle pool; hence, it is not surprising that genes related to nuclear activity and chromatin assembly are more active at this stage. In contrast, the follicles start to develop into mature follicles in the adult, thus requiring the upregulation of genes associated with growth processes and metabolism.

Analysis of alternatively spliced variants

A large number of mammalian genes have been shown to have alternatively spliced transcripts [15], which cannot be specifically detected in microarray experiments. A study using RNA-seq technology to examine 15 different types of tissues, found that 90% of all human genes are alternatively spliced [16]. Thus, using this technology, we can detect specific splice variants via the reads mapped to exon(s) or exon combination(s) that are unique to a transcript. These unique reads for a transcript can be used as markers and allow us to compare the splice variants between two samples. According to Ensembl annotation (version 55), 8886 genes have multiple alternatively spliced transcripts. In our analysis, we detected alternatively spliced transcripts for 4484 and 5320 genes in the infant and adult ovary, respectively. It is evident that genes tended to have more alternatively spliced transcripts in the adult stage, Moreover, our data also indicate that there are considerable splicing variations between 2 developmental stages of the same tissue. These data suggest that alternative splicing is developmentally regulated in the ovary and increases the transcript diversity throughout development. This finding is very similar to that found for the maize leaf transcriptome using ultra high-throughput sequencing technology [17].

Furthermore, we sought to determine whether our RNA-seq assay could detect novel splice isoforms simply by mapping the unique reads to the junction database built in house. To construct the junction database, we generated all possible combinations of exon-exon junctions as 98-bp sequences with 49 bases from each exon. After mapping, we selected the reads only mapped to novel exon-exon junctions. We found that 5394 and 11,139 reads were mapped to junctions distinct from their known placements (defined by Ensembl annotation) at the infant and adult stages, respectively. In the infant, there were 1313 new junctions with at least 2 mapped reads for

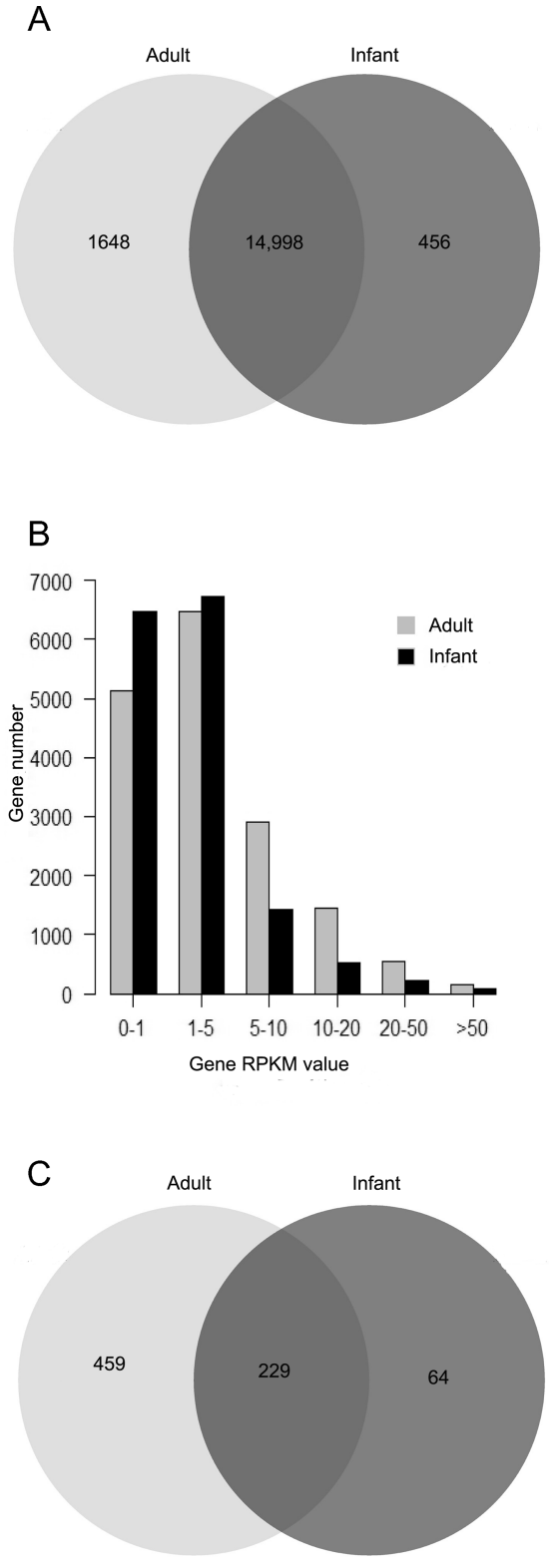

Figure 1 Gene expression profiles in mouse ovaries at infant and adult stages

A. Venn diagram shows the expressed genes in mouse ovaries at infant and adult stages, totally 17,102 genes were expressed. **B.** The RPKM distribution of expressed genes at infant (black bars) and adult (gray bars) stages. **C.** Venn diagram shows the highly-expressed genes (RPKM $\geqslant$ 20).

906 genes. Meanwhile, 2977 new junctions with at least 2 mapped reads were found for 2,025 genes in the adult. In total, we determined unreported splicing forms of 2437 genes during these two developmental stages, with 494 genes supported by the datasets of both stages. Among this list were genes that play essential roles during ovary development, such as *Foxo3*, *Cyp11a1*, *Zfpm2*, *Amhr2*, *Sf1*, *Esr1*, *Cyp17a1*, *Hsd3b1*, *Col4a1* and *Dicer1*.

Expression of genes regulating hormone process genes

One of the two key functions of the ovary is the secretion of steroid hormones that prepare the reproductive tract for fertilization and the establishment of pregnancy. Thus, genes related to steroid secretion and hormone metabolism play important roles during ovary development. Based on our DEG analysis, we found that hormone biosynthesis was greatly activated in the adult ovary. Among the genes associated with steroid hormone biosynthesis, expression of 18 genes was upregulated in the adult ovary compared to the infant ovary. Moreover, none of these genes were significantly higher expressed in the infant ovary. In particular, expression of some genes essential for estrone biosynthesis, such as *Cyp19a1*, *Cyp11a1*, *Cyp17a1*, *Hsd3b1* and *Hsd17b1,* was significantly upregulated in the adult (Figure S1). In addition, we also noticed significantly-upregulated expression of *Fshr*, *Esr1*, *Esr2* and *Lhcgr* in the adult ovary, which encode receptors for follicle stimulating hormone (FSH), estrone and luteinizing hormone (LH), respectively (**Figure 4**). FSH, LH and estrone, along with their receptors, play important roles during folliculogenesis [18,19]. The upregulated expression of genes involved in hormone processing and genes encoding their receptors are likely essential for fertility in the adult mouse.

Expression of genes involved in the MAPK signaling pathway

The MAPK signaling pathway is a basic cellular signal transduction cascade that couples intracellular responses to the binding of growth factors to cell surface receptors and is known to regulate cell proliferation, apoptosis, survival and differentiation [20,21]. Mammals express at least four distinctly-regulated groups of MAPKs: extracellular signal-related kinases (*Erk*)-*1*/*2*, Jun amino-terminal kinases (*Jnk1*/*2*/*3*), p38 proteins (p38alpha/beta/gamma/delta) and *Erk5*, all of which are activated by specific MAPK kinases: *Mek1*/*2* (for *Erk1*/*2*), *Mkk3*/*6* (for the p38), *Mkk4*/*7* (also known as *Jnkk1*/*2;* for the JNKs) and *Mek5* (for *Erk5*). It has been reported that expression of MAPK3/1 (ERK1/2) in ovarian granulosa cells is critical for female fertility as it regulates LH-induced oocyte resumption during meiosis, ovulation and luteinization [22]. Notably, *Erk1* and *Erk2* are co-expressed in all mammalian tissues and have been implicated as key regulators of cell proliferation and differentiation as well as oocyte maturation in culture [23,24]. Loss of *Erk1* causes minimal changes in fertility and embryo viability [25], while mutation of the *Erk2* gene is embryonic lethal in mutant mice [26]. Our data in the current study showed that the MAPK signaling pathway is much more active in the adult ovary, as many of the genes in this pathway had higher expression levels at the adult stage compared to the infant stage (Figure S2). Furthermore, the upregulation of *Egfr*

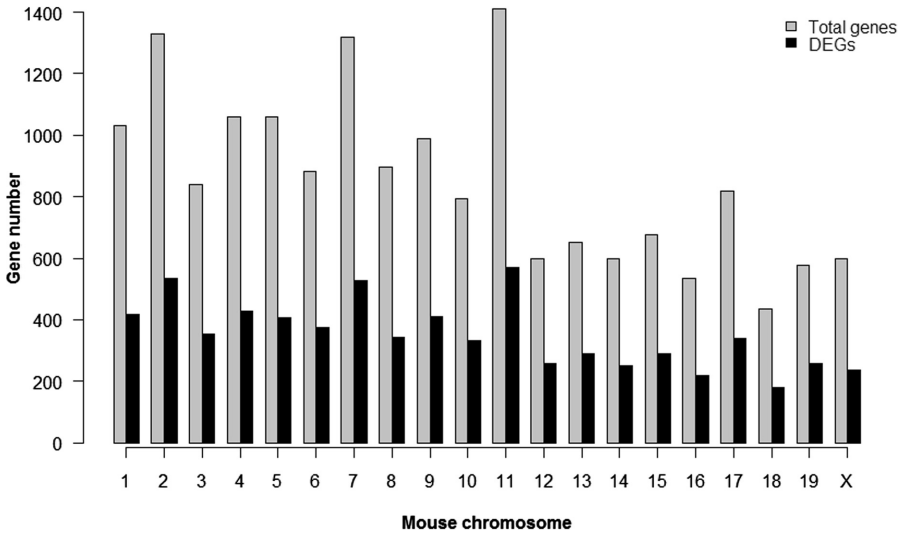

Figure 2 The distribution of DEGs along mouse chromosome
The distribution of expressed genes was presented in respect to their chromosomal location for all expressed genes (gray bars) and DEGs (black bars) stages. DEG, differentially-expressed genes.

observed in the adult likely activated *Mek1/2* (*2k1/2*), *Mapk1* and *Mapk3* in this tissue (Figure S2), suggesting an essential role for *Egfr* gene expression in fertility.

Detection of apoptosis-related genes

Apoptosis is known to be involved in oogenesis, folliculogenesis, oocyte loss/selection and atresia; however, its exact role in ovary development has not been fully characterized [27]. In our study, 42 genes known to be involved in apoptotic processes had significantly altered expression levels, most of which were upregulated in the adult stage (Figure S3). This observation suggests that cellular apoptosis was activated in the adult ovary, resulting in more developing follicles being degraded through apoptotic processes at this stage. Among the genes related to apoptosis, we noted that the genes encoding caspases, the main effector molecules in ovarian apoptosis, were only modestly expressed in the infant ovary and their expression increased in variable degrees in the adult ovary. The significant upregulation of *Caspase1*, *Caspase3*, *Caspase6*, *Caspase8* and *Caspase9* would be expected to greatly activate apoptosis [28]. In addition, members of the *Bcl* gene family, including *Bcl2*, *Bax*, *Bok*, *Bcl2l1*, *Bcl2l2* and *Bad,* were all modestly expressed in the infant and adult stages, although expression of some genes including *Bad* and *Bok* was slightly upregulated in the adult. However, expression of *Mcl1,* another *Bcl* member known as an anti-apoptotic survival factor [29], was greatly upregulated in the adult tissue, although modest in the infant ovary. Meanwhile, expression of some apoptotic inhibitors was also significantly upregulated in the adult. For instance, *Xiap* which encodes the most potent caspase inhibitor, was only modestly expressed in the infant ovary, but was significantly upregulated in the adult ovary. Expression of XIAP in the ovary was upregulated in response to FSH to suppress granulose cell apoptosis and facilitate FSH-induced follicular growth [30]. Our results suggest its role in follicle survival and anti-apoptosis. In summary, genes annotated to participate in cell apoptosis had substantial alterations in expression between the infant and adult stages of development, suggesting the essential role of apoptosis during ovary maturation.

Detection of transcription factors

Transcription factors (TFs) are proteins that control gene expression by binding to specific DNA sequences, thereby controlling gene transcription through the recruitment or dismissal of RNA polymerase and/or other regulatory proteins. There have been 1728 genes identified to encode various TFs in the mouse genome [31]. Our RNA-seq data indicate that at least 1462 genes present in the infant ovary encode TFs, while 1478 TF genes were found in the adult stage, and totally 1515 unique TF genes in this organ (Figure S4). Among these, 16 and 46 genes were highly expressed (RPKM $\geqslant$ 20) in the infant and adult stages, respectively. Interestingly, among the 100 most expressed TF genes in the infant and adult ovary, 59 genes were commonly found at both stages, including *Sox4*, *Foxj1* and *Ddb1*. The consistently high expression of these TFs during ovary development may indicate their vital roles in the maintenance of female fertility. On the other hand, 732 TF genes were differentially expressed between the infant and adult stages, with 632 genes upregulated at the adult stage. TFs that are related to early ovarian follicular development often were downregulated from the infant to the adult, including *Figla* [32,33], *Nobox* [34] and *Pou5f1* [35]. *Figla* is a basic helix-loop-helix transcription factor that is expressed exclusively in oocytes and regulates the transcription of zona pellucida (ZP) genes. A recent comprehensive gene expression study in *Figla* null mice showed that some genes essential for folliculogenesis were downregulated following loss of this gene [36], indicating that *Figla* is a key TF in oocyte-specific gene expression and is indispensable for not only ZP protein synthesis but also for germ cell assembly. Our observation that *Figla* was significantly downregulated in the adult is consistent with

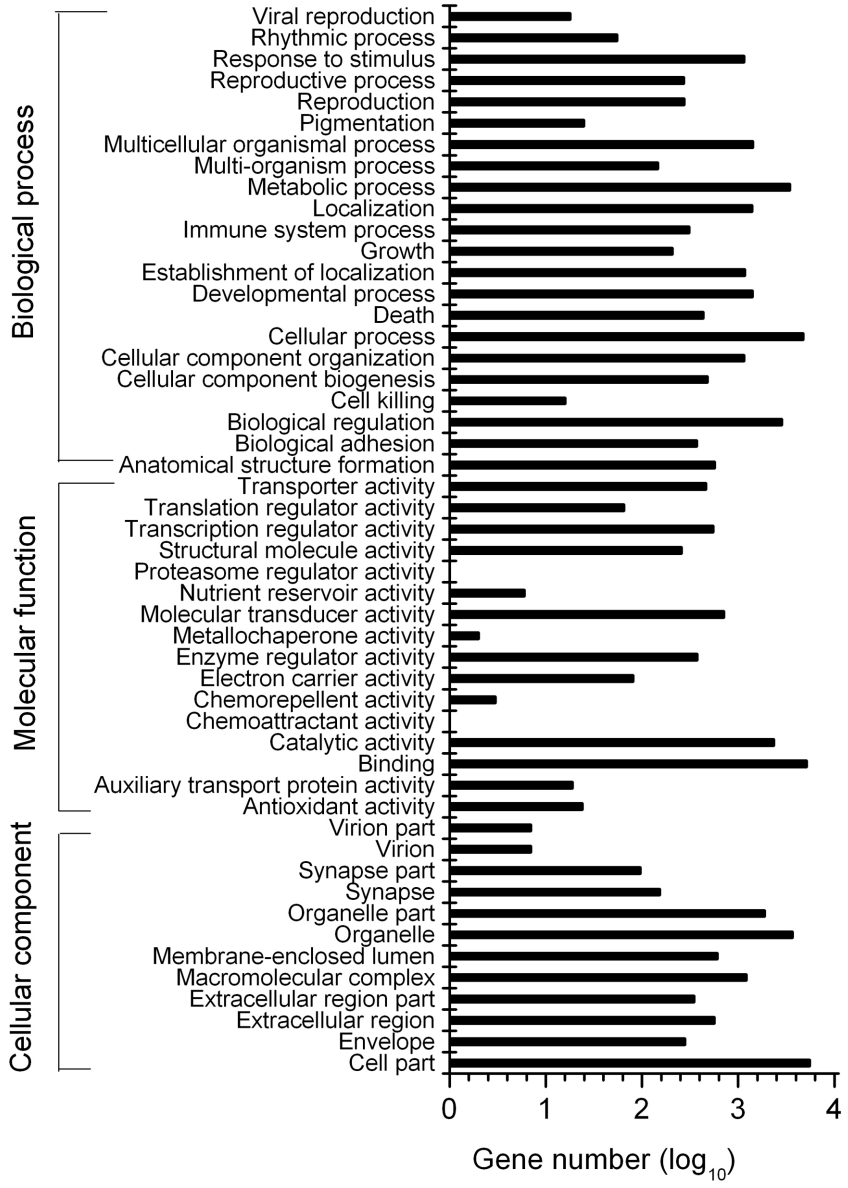

Figure 3 Functional annotation of the DEGs based on GO categorization

Histogram of GO classification of DEGs in ovaries between the infant and adult stages was shown in the three GO functional groups including biological process, molecular function and cellular component.

its known function during ovary development. Alternatively, some TF genes known to be involved in the later stage of follicle growth were upregulated in the adult, including *Foxo1* [37], *Foxl2* [38], *Smad2* [39] and *Gata4* [40]. All the TF gene expression profiles are listed in Table S3.

Expression of TGF-β superfamily genes

The TGF-β superfamily is a structurally-conserved but functionally-diverse group of proteins with at least 35 members in vertebrates. These proteins are widely distributed throughout the body and function as extracellular ligands in numerous physiological processes during both pre-and postnatal life [41]. On the basis of structural characteristics, members of this superfamily have been further categorized into several subfamilies. These include the prototypic TGF-β subfamily, an extensive bone morphogenetic protein (BMP) subfamily, the growth and differentiation factor (GDF) subfamily, the activin/inhibin subfamily (including activins A, AB and B as well as inhibins A and B), the glial cell-derived neurotrophic factor (GDNF) subfamily (including GDNF, artemin and neurturin), as well as several additional members such as anti-Mullerian hormone (AMH) and nodal [42]. Many members of TGF-β superfamily are expressed by ovarian somatic cells and oocytes in a developmental stage-related manner and function as intraovarian regulators of folliculogenesis.

Our results indicate that 28 TGF-β genes were expressed in the infant ovary and 30 were expressed in the adult ovary (Table S4). Consistent with a significant increase in the number of developing follicles during ovary development, we found

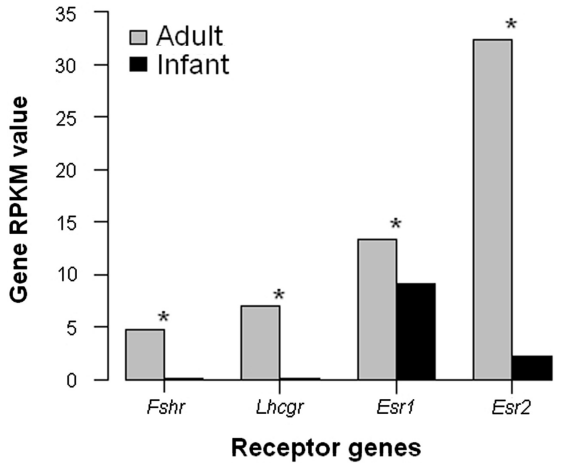

Figure 4 Expression profiles of hormone receptor genes in adult and infant ovaries

Significant differences in expression of genes encoding for hormone receptors between the infant (black bar) and adult (gray bar) stages were indicated with $^{*}(P < 0.001)$. The *P* values were calculated using Fisher's exact test ($P = 2.18E{-}25$ for *Fshr*, 1.56E−39 for *Lhcgr*, 2.73E−12 for *Esr1* and 1.15E−278 for *Esr2*). Fshr, follicle-stimulating hormone receptor; Lhcgr, lutropin–choriogonadotropic hormone receptor; ESR, estrogen receptor.

that a majority of the differentially-expressed TGF-β genes and their receptor genes were substantially upregulated in the adult ovary compared to the infant ovary. Expression of GDFs and BMPs are known to be involved in follicular development. Among the *Gdf*s detected, *Gdf9* and *Gdf11* were upregulated, while *Gdf7* and *Gdf10* were downregulated in the adult. Many reports have suggested that *Gdf9* plays an essential role in the stimulation of granulose cell proliferation and folliculogenesis [43,44]. Furthermore, follicle development in the ovaries of *Gdf9* null mice is arrested at the primary stage [43], indicating that upregulation of *Gdf9* in the adult ovary is quite important for adult mouse fertility. In regard to the BMP subfamily, *Bmp15*, *Bmp3* and *Bmp2* were significantly upregulated from the infant to the adult stage, which was consistent with their reported roles as positive paracrine factors in follicle progression [45]. However, *Bmp7* and *Bmp1* were slightly downregulated from infant to adult, suggesting that they may have unknown roles during the early stage of ovary development. Intriguingly, the *Amh* gene, known to negatively regulate preantral follicle development [46], was also upregulated in the adult even though there are more developing follicles in the adult ovary [11]. Upregulation of the *Amh* gene in the adult ovary may keep the limited primordial follicles silent due to its inhibitory actions on primordial follicle assembly [47].

Expression of oocyte-specific genes

As a specialized cell population, oocytes have been intensively studied to unravel their unique physiology, from the specification of primordial oocytes to the fateful reunion of gametes during fertilization. Elucidating the alterations in oocyte-specific gene expression is essential to understand the unique features of oocytes. It has been reported that gene products expressed solely in oocytes play important roles in folliculogenesis, fertilization and pre-implantation development [48]. In the last decade, many genes have been identified as oocyte-specific based on large-scale high-throughput transcriptome data [49–51]. Here, we have listed the expression profiles of the oocyte-specific genes in infant and adult ovaries (**Figure 5**). The expression patterns of these oocyte-specific genes showed that most of them were more active during the infant stage than in the adult stage.

The ZP is mainly composed of three glycoproteins (ZP1–3), and functions as an extracellular matrix surrounding the ovarian oocyte and ovulated egg that plays important roles during the fertilization process, including acrosome reaction induction, polyspermy prevention and embryo protection [52]. As described previously, *Figla* regulates the transcription of ZP genes, and in this study, the expression pattern of *Zp1* and *Zp3* coincided with that of *Figla*, being drastically downregulated in the adult stage compared to the infant stage. However, *Zp2* expression was slightly upregulated. In addition to the ZP genes, some other key oocyte-specific genes were also downregulated in the adult ovary. For example, *Nobox*, *Sohlh1*, *Gja4,* and *Pou5f1* were all significantly downregulated at the adult stage. *Nobox* (newborn ovary homeobox gene) is necessary for the expression of several key folliculogenesis genes, including *Gdf9* and *Pou5f1* [53,54]. *Gja4* is present in gap junctions between oocyte and granulosa cells, and *Gja4*-deficient mice lack mature follicles, fail to ovulate and develop numerous inappropriate corpora lutea [55,56]. *Sohlh1* (spermatogenesis- and oogenesis-specific bHLH transcription factor 1) is another germ cell-specific gene that encodes a bHLH protein. *Sohlh1* lies upstream of Lhx8, Zp1 and Zp3, and is preferentially expressed in primordial oocytes [57]. In addition to numerous downregulated oocyte-specific genes, we also found that some important oocyte-specific genes showed little to no change during ovary development. For example, the expression of *Oosp1*, *H1foo*, *Lhx8* and *Sohlh2* remained relatively constant between the infant and adult ovaries, suggesting that they have conserved roles during ovary development. *Sohlh2* is preferentially expressed in germ cells of the embryonic ovary, the oocytes of primordial and the primary follicles [58]. A recent study of *Sohlh2*-deficient mice showed that adult females lacking this gene are infertile owing to the misexpression of many oocyte-specific genes, including *Figla*, *Nobox*, *Zp1*, *Zp3* and *Oosp1* [59].

Although many oocyte-specific genes were downregulated or unchanged, some had higher expression in the adult stage compared to the infant stage, including *Kit*, *Kitl*, *Nlrp14* and *Bmp15*. KIT ligand has been reported to be present in oocytes at all stages of folliculogenesis and functions in the transition from primordial to primary follicles [60]. On the other hand, *Nlrp14* is a leucine-rich repeat protein and a member of the NACHT nucleoside triphosphatase family [61], although its exact function during ovary development is still unknown.

Conclusion

In the present study, the gene expression in infant ovaries and adult ovaries has been investigated at the transcriptome level using RNA-seq. We have provided basic gene expression profiles of mouse ovary development. Comparative analysis showed that the adult ovary has a more active transcriptome than that in the infant, in terms of the number of genes

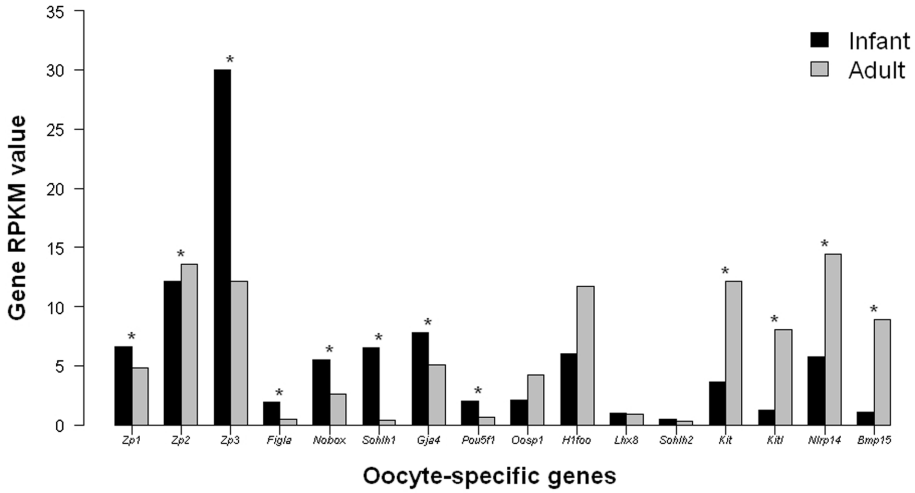

Figure 5 Expression profiles of oocyte-specific genes in adult and infant ovaries
Significant differences in expression of oocyte-specific genes between the infant (black bar) and adult (gray bar) stages were indicated with * ($P < 0.001$). The P values were calculated using Fisher's exact test (see Table S5 for the actual P values).

expressed, the level of gene expression and the number of alternatively spliced transcripts. The expression profiles also suggest vital roles for a large number of regulatory genes in both early and late ovary development, and provide strong evidence for a highly dynamic transcriptome during mouse ovary maturation. Notably, GO analysis reveals that some important functional pathways, including apoptosis, MAPK, as well as steroid and hormone biosynthesis, are more active in adult ovaries than in the infant. Further studies with these differentially-expressed genes identified in the present study will extend our understanding of the complex molecular and cellular events occurring during the ovary development process and help illuminate the metabolism, growth and differentiation of the mammalian ovary.

Materials and methods

Female clean-class Balb/c mice aged 1 week ($n = 40$) and 8 weeks ($n = 15$) were purchased from Vital River Laboratory Animal Technology (Beijing, China). Mice were sacrificed by cervical dislocation. Then all the ovaries of the respective stage were dissected and pooled together before stored in liquid nitrogen prior to use. Adult mice were at the proestrus stage of menstruation at the time of sacrifice. The animal experiments in this study were approved by the Animal Care and Welfare Committee of the Beijing Institute of Genomics, Chinese Academy of Sciences.

RNA-seq library generation and sequencing

Total RNA from each sample was isolated using TRIzol. The RiboM-minus Eukaryote Kit (Invitrogen; Catlog No 10837-08) was then used to deplete the rRNA from the total RNA, which following this procedure, was designated as ribo-minus RNA (rm-RNA). We used 1 μg rm-RNA as the starting material to construct our RNA-seq library using the SOLiD™ Whole Transcriptome Analysis Kit (Applied Biosystems; PN 4425680) as instructed by the manufacturer. Finally, a fraction of the library, in a size range of 100–200 bp, was selected for SOLiD sequencing. Emulsion PCR and SOLiD sequencing were performed according to manufacturer's instructions for the SOLiD™ System. The 50-bp sequences were obtained on a 1/4 SOLiD v3 slide for each library.

Read mapping

After filtering out the low-quality reads (average quality value < 8), we used the Applied Biosystems software (Corona_lite_v4.0r2.0; http://solidsoftwaretools.com/gf/project/corona/) to map the remaining 50-bp reads from the infant and adult ovaries to the mouse reference genome (release on July 09, 2007). We also constructed an exon–exon junction database according to Ensembl gene annotation (version 55). To be specific, for a gene with multiple exons, we used the last 49 nucleotides of one exon as a donor sequence and the first 49 nucleotides of each following exon as an acceptor sequence. We joined the donor sequence and the acceptor sequence into one exon–exon sequence. All the exon-exon sequences generated composed our exon–exon junction database. First, we mapped the full-length 50-bp tags to the reference genome; second, we analyzed the flow-through against our junction database; third, we repeated the first and the second steps for the first 45-, 40-, 35-, 30-, 25-bp truncated tags (after removal of the tag sequences beyond each length). For each length, we used 5-, 5-, 4-, 3-, 3- and 2-bp mismatch options in both the genome and junction mapping steps. Reads mapped to unique positions in the reference sequence were counted according to their mapping location. Then, we calculated the reads number for each Ensembl gene (version 55) based on the mapping information.

Measurement of gene expression

The uniquely-mapped reads were categorized into four groups according to their locations in genome: exonic region, intronic region, intergenic region and junction region. For each gene, we counted the number of unique reads that were located in exonic regions to obtain a raw digital gene expression count.

Genes with at least two uniquely-mapped raw reads in the exonic region were considered as being expressed in the tissue. The genes with raw digital gene expression counts greater than this threshold were then normalized using a slight variation of the RPKM method [62]. The RPKM for each gene was calculated with the equation: $\mathrm{RPKM} = \frac{10^9 C}{NL}$, where C is the number of uniquely-mapped reads in the merged exonic region, N is the total number of uniquely-mapped reads in the library and L is the length of all exons for one gene. Based on the RPLM values, genes were classified as lowly-expressed ($\mathrm{RPKM} < 1$), moderately-expressed ($1 \leqslant \mathrm{RPKM} \leqslant 20$) or highly-expressed ($\mathrm{RPKM} > 20$), respectively.

Identification of DEGs

We used the DEGseq package to identify the DEGs between the infant and adult ovary samples [12]. We used GenMAPP2.0 and KEGG Mapper to visualize the differentially-expressed genes in various KEGG pathways [63,64].

Authors' contributions

LP, JY and SH conceived the project and designed the experiments. LP, XL, WG and YZ performed the experiments; LP, YL and XL analyzed the data. LP and SH wrote the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Grant No. 31271385), the Knowledge Innovation Program of Chinese Academy of Sciences of China (Grant No. KSCX2-EW-R-01-04) and the National High-Tech R&D Program (863 Program, Grant No. 2009AA01A 130) from the Ministry of Science and Technology of China.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.gpb.2014.08.002.

References

[1] McLaren A. Germ and somatic cell lineages in the developing gonad. Mol Cell Endocrinol 2000;163:3–9.
[2] Oktem O, Urman B. Understanding follicle growth *in vivo*. Hum Reprod 2010;25:2944–54.
[3] Eyster KM, Brannian JD. Gene expression profiling in the aging ovary. Methods Mol Biol 2009;590:71–89.
[4] Hasegawa A, Kumamoto K, Mochida N, Komori S, Koyama K. Gene expression profile during ovarian folliculogenesis. J Reprod Immunol 2009;83:40–4.
[5] Herrera L, Ottolenghi C, Garcia-Ortiz JE, Pellegrini M, Manini F, Ko MS, et al. Mouse ovary developmental RNA and protein markers from gene expression profiling. Dev Biol 2005;279:271–90.
[6] Dharma SJ, Modi DN, Nandedkar TD. Gene expression profiling during early folliculogenesis in the mouse ovary. Fertil Steril 2009;91:2025–36.
[7] Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009;10:57–63.
[8] Xue Z, Huang K, Cai C, Cai L, Jiang CY, Feng Y, et al. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature 2013;500:593–7.
[9] Chen QF, Ya HY, Wang WD, Jiao Z. RNA-seq reveals the downregulated proteins related to photosynthesis in growth-inhibited rice seedlings induced by low-energy N beam implantation. Genet Mol Res 2014;13:7029–36.
[10] Shi G, Huang F, Gong Y, Xu G, Yu J, Hu Z, et al. RNA-Seq analysis reveals that multiple phytohormone biosynthesis and signal transduction pathways are reprogrammed in curled-cotyledons mutant of soybean [*Glycine max* (L.) Merr]. BMC Genomics 2014;15:510.
[11] Peters H. The development of the mouse ovary from birth to maturity. Acta Endocrinol (Copenh) 1969;62:98–116.
[12] Wang L, Feng Z, Wang X, Zhang X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 2009;26:136–8.
[13] Zinn AR. The X chromosome and the ovary. J Soc Gynecol Investig 2001;8:S34–6.
[14] Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic Acids Res 2005;33:W741–8.
[15] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature 2001;409:860–921.
[16] Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, et al. Alternative isoform regulation in human tissue transcriptomes. Nature 2008;456:470–6.
[17] Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, et al. The developmental dynamics of the maize leaf transcriptome. Nat Genet 2010;42:1060–7.
[18] Neal P, Challoner S. The development of the mouse ovary and its response to exogenous gonadotropins. J Reprod Fertil 1975;45:449–54.
[19] Hassin S, Claire M, Holland H, Zohar Y. Ontogeny of follicle-stimulating hormone and luteinizing hormone gene expression during pubertal development in the female striped bass, *Morone saxatilis* (Teleostei). Biol Reprod 1999;61:1608–15.
[20] Yang SH, Sharrocks AD, Whitmarsh AJ. Transcriptional regulation by the MAP kinase signaling cascades. Gene 2003;320:3–21.
[21] Xiao Y, Yang FQ, Li SP, Gao JL, Hu G, Lao SC, et al. Furanodiene induces G2/M cell cycle arrest and apoptosis through MAPK signaling and mitochondria-caspase pathway in human hepatocellular carcinoma cells. Cancer Biol Ther 2007;6:1044–50.
[22] Fan HY, Liu Z, Shimada M, Sterneck E, Johnson PF, Hedrick SM, et al. MAPK3/1 (ERK1/2) in ovarian granulosa cells are essential for female fertility. Science 2009;324:938–41.
[23] Su YQ, Wigglesworth K, Pendola FL, O'Brien MJ, Eppig JJ. Mitogen-activated protein kinase activity in cumulus cells is essential for gonadotropin-induced oocyte meiotic resumption and cumulus expansion in the mouse. Endocrinology 2002;143:2221–32.
[24] Chen Z, Gibson TB, Robinson F, Silvestro L, Pearson G, Xu B, et al. MAP kinases. Chem Rev 2001;101:2449–76.
[25] Pages G, Guerin S, Grall D, Bonino F, Smith A, Anjuere F, et al. Defective thymocyte maturation in p44 MAP kinase (Erk 1) knockout mice. Science 1999;286:1374–7.

[26] Aouadi M, Binetruy B, Caron L, Le Marchand-Brustel Y, Bost F. Role of MAPKs in development and differentiation: lessons from knockout mice. Biochimie 2006;88:1091–8.
[27] Hussein MR. Apoptosis in the ovary: molecular mechanisms. Hum Reprod Update 2005;11:162–77.
[28] Hussein MR, Haemel AK, Wood GS. P53-related pathways and the molecular pathogenesis of melanoma. Eur J Cancer Prev 2003;12:93–100.
[29] Leo CP, Hsu SY, Chun SY, Bae HW, Hsueh AJ. Characterization of the antiapoptotic Bcl-2 family member myeloid cell leukemia-1 (Mcl-1) and the stimulation of its message by gonadotropins in the rat ovary. Endocrinology 1999;140:5469–77.
[30] Johnson NC, Dan HC, Cheng JQ, Kruk PA. BRCA1 185delAG mutation inhibits Akt-dependent, IAP-mediated caspase 3 inactivation in human ovarian surface epithelial cells. Exp Cell Res 2004;298:9–16.
[31] Ravasi T, Suzuki H, Cannistraci CV, Katayama S, Bajic VB, Tan K, et al. An atlas of combinatorial transcriptional regulation in mouse and man. Cell 2010;140:744–52.
[32] 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;124:4939–47.
[33] Soyal SM, Amleh A, Dean J. FIGalpha, a germ cell-specific transcription factor required for ovarian follicle formation. Development 2000;127:4645–54.
[34] Rajkovic A, Pangas SA, Ballow D, Suzumori N, Matzuk MM. NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. Science 2004;305:1157–9.
[35] Kehler J, Tolkunova E, Koschorz B, Pesce M, Gentile L, Boiani M, et al. Oct4 is required for primordial germ cell survival. EMBO Rep 2004;5:1078–83.
[36] Joshi S, Davies H, Sims LP, Levy SE, Dean J. Ovarian gene expression in the absence of FIGLA, an oocyte-specific transcription factor. BMC Dev Biol 2007;7:67.
[37] Shi F, LaPolt PS. Relationship between FoxO1 protein levels and follicular development, atresia, and luteinization in the rat ovary. J Endocrinol 2003;179:195–203.
[38] Kuo FT, Bentsi-Barnes IK, Barlow GM, Bae J, Pisarska MD. Sumoylation of forkhead L2 by Ubc9 is required for its activity as a transcriptional repressor of the Steroidogenic Acute Regulatory gene. Cell Signal 2009;21:1935–44.
[39] Kaivo-oja N, Jeffery LA, Ritvos O, Mottershead DG. Smad signalling in the ovary. Reprod Biol Endocrinol 2006;4:21.
[40] Anttonen M, Parviainen H, Kyronlahti A, Bielinska M, Wilson DB, Ritvos O, et al. GATA-4 is a granulosa cell factor employed in inhibin-alpha activation by the TGF-beta pathway. J Mol Endocrinol 2006;36:557–68.
[41] Massague J, Blain SW, Lo RS. TGFbeta signaling in growth control, cancer, and heritable disorders. Cell 2000;103:295–309.
[42] Drummond AE. TGFbeta signalling in the development of ovarian function. Cell Tissue Res 2005;322:107–15.
[43] Dong J, Albertini DF, Nishimori K, Kumar TR, Lu N, Matzuk MM. Growth differentiation factor-9 is required during early ovarian folliculogenesis. Nature 1996;383:531–5.
[44] Knight PG, Glister C. TGF-beta superfamily members and ovarian follicle development. Reproduction 2006;132:191–206.
[45] McNatty KP, Juengel JL, Reader KL, Lun S, Myllymaa S, Lawrence SB, et al. Bone morphogenetic protein 15 and growth differentiation factor 9 co-operate to regulate granulosa cell function. Reproduction 2005;129:473–80.
[46] Durlinger AL, Gruijters MJ, Kramer P, Karels B, Kumar TR, Matzuk MM, et al. Anti-Mullerian hormone attenuates the effects of FSH on follicle development in the mouse ovary. Endocrinology 2001;142:4891–9.
[47] Nilsson EE, Schindler R, Savenkova MI, Skinner MK. Inhibitory actions of Anti-Mullerian Hormone (AMH) on ovarian primordial follicle assembly. PLoS One 2011;6:e20087.
[48] Dean J. Oocyte-specific genes regulate follicle formation, fertility and early mouse development. J Reprod Immunol 2002;53:171–80.
[49] Leo CP, Vitt UA, Hsueh AJ. The Ovarian Kaleidoscope database: an online resource for the ovarian research community. Endocrinology 2000;141:3052–4.
[50] Hsueh AJ, Rauch R. Ovarian Kaleidoscope database: ten years and beyond. Biol Reprod 2012;86:192.
[51] Lin YN, Matzuk MM. High-throughput discovery of germ-cell-specific genes. Semin Reprod Med 2005;23:201–12.
[52] Dean J. Biology of mammalian fertilization: role of the zona pellucida. J Clin Invest 1992;89:1055–9.
[53] Choi Y, Rajkovic A. Characterization of NOBOX DNA binding specificity and its regulation of Gdf9 and Pou5f1 promoters. J Biol Chem 2006;281:35747–56.
[54] Choi Y, Qin Y, Berger MF, Ballow DJ, Bulyk ML, Rajkovic A. Microarray analyses of newborn mouse ovaries lacking Nobox. Biol Reprod 2007;77:312–9.
[55] Simon AM, Goodenough DA, Li E, Paul DL. Female infertility in mice lacking connexin 37. Nature 1997;385:525–9.
[56] Yin BY, Zhang Y, Sun JH, Li JX, Ma YF. Connexin37 mRNA expression in *in vivo* and *in vitro* mouse oocyte. Zygote 2009;17:163–8.
[57] Pangas SA, Choi Y, Ballow DJ, Zhao Y, Westphal H, Matzuk MM, et al. Oogenesis requires germ cell-specific transcriptional regulators Sohlh1 and Lhx8. Proc Natl Acad Sci U S A 2006;103:8090–5.
[58] Ballow DJ, Xin Y, Choi Y, Pangas SA, Rajkovic A. Sohlh2 is a germ cell-specific bHLH transcription factor. Gene Expr Patterns 2006;6:1014–8.
[59] Choi Y, Yuan D, Rajkovic A. Germ cell-specific transcriptional regulator sohlh2 is essential for early mouse folliculogenesis and oocyte-specific gene expression. Biol Reprod 2008;79:1176–82.
[60] Hutt KJ, McLaughlin EA, Holland MK. KIT/KIT ligand in mammalian oogenesis and folliculogenesis: roles in rabbit and murine ovarian follicle activation and oocyte growth. Biol Reprod 2006;75:421–33.
[61] Horikawa M, Kirkman NJ, Mayo KE, Mulders SM, Zhou J, Bondy CA, et al. The mouse germ-cell-specific leucine-rich repeat protein NALP14: a member of the NACHT nucleoside triphosphatase family. Biol Reprod 2005;72:879–89.
[62] Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 2008;5:621–8.
[63] Salomonis N, Hanspers K, Zambon AC, Vranizan K, Lawlor SC, Dahlquist KD, et al. GenMAPP 2: new features and resources for pathway analysis. BMC Bioinformatics 2007;8:217.
[64] Kanehisa M. The KEGG database. Novartis Found Symp 2002;247:91–101; discussion 101–3, 119–28, 244–52.