



## Differential gene expression in pre-laying and laying period ovaries of Sichuan White geese (*Anser cygnoides*)

N. Ding<sup>1\*</sup>, Q. Han<sup>1\*</sup>, X.Z. Zhao<sup>1</sup>, Q. Li<sup>1</sup>, J. Li<sup>1</sup>, H.F. Zhang<sup>1,2</sup>, G.L. Gao<sup>1</sup>, Y. Luo<sup>1</sup>, Y.H. Xie<sup>1</sup>, J. Su<sup>1</sup> and Q.G. Wang<sup>1</sup>

<sup>1</sup>Chongqing Academy of Animal Science, Chongqing, China

<sup>2</sup>College of Animal Science and Technology, Northeast Agricultural University, Harbin, China

\*These authors contributed equally to this study.

Corresponding author: Q.G. Wang

E-mail: wangqigui@hotmail.com

Genet. Mol. Res. 14 (2): 6773-6785 (2015)

Received September 4, 2014

Accepted March 5, 2015

Published June 18, 2015

DOI <http://dx.doi.org/10.4238/2015.June.18.20>

**ABSTRACT.** Geese are an economically important poultry species worldwide. Their superior meat production performance and meat quality make them a popular food. However, they are not bred worldwide because their poor laying capacity increases farming costs. To gain a global view of the genes that are differentially expressed between pre-laying (P) and laying (L) periods and to develop a database for further studies, we performed large-scale transcriptome sequencing of ovarian tissue collected from *Anser cygnoides*. In total, 30,151,422 raw reads, with an average length of 151 bp and a total length of 4,552,864,722 bp, were obtained. After primers and adaptors were removed, 19,167,132 clean reads, with an average length of 134.5 bp and a total length of 2,577,297,281 bp, were obtained, among which 1,268,906,694 bp and 1,308,390,587 bp were from L and P ovarian tissue, respectively. The 16,605 assembled sequences were further functionally annotated by

comparing their sequences to different protein and functional domain databases and assigning gene ontology (GO) terms. Of these, 511 assembled sequences were considered differentially expressed based on the 2-fold method, among which 396 were assigned at least one GO term. Digital expression analysis using the Kyoto encyclopedia of genes and genomes annotation identified 121 genes that were differentially expressed in the P vs L periods. Five of these are of special interest for further investigation of their roles in determining high reproductive performance. This study provides valuable information and sequence resources for uncovering genes determining high egg-laying performance and for future functional genomics analysis of geese.

**Key words:** Geese; Differentially expressed genes; Ovary; Transcriptome

## INTRODUCTION

The goose, an important grass-feeding fowl, is one of the most economically important poultry species worldwide (Stevenson, 1989). With rapid early growth and superior meat production performance, fleshy breeds of geese have high economic benefits; however, poor reproductive performance (low egg production of approximately 30-40 eggs per year) has prevented production on a large scale (Shalev et al., 1991). Laying performance in poultry is closely related to the development of ovarian follicles and the establishment of the follicle hierarchy (Koelkebeck et al., 2001). Compared with other varieties, the Sichuan White goose (*Anser cygnoides*) has more rigorous follicle grades, and this is the main reason for its relatively long duration of peak egg-laying and high egg production (Kang et al., 2009; Pan et al., 2011). Identifying differences in gene expression between distinct follicular stages could help us to understand follicular development, establish a rating mechanism based on gene expression, and elucidate the basis for the higher laying performance yield of Sichuan White geese.

Although entire genomes of model animal species have been sequenced, the goose genome has not. Traditional expressed sequence tag (EST) sequencing through constructing an EST library is time-consuming as well as expensive. Transcriptome sequencing has been proven to be an efficient means of gene discovery, especially with the availability of high-throughput next generation sequencing technology (Kaur et al., 2011; Xie et al., 2012). Here, we performed transcriptome sequencing using RNA extracted from ovaries of pre-laying and laying-period geese using Illumina RNA-seq to discover genes that are differentially expressed in the ovary during these two periods. Using homologies to the *Gallus* genome in combination with various bioinformatic tools, we generated a list of candidate genes that may influence high egg-laying performance in geese.

## MATERIAL AND METHODS

### Ethics statement

All animal work was conducted according to the guidelines for the care and use of experimental animals established by the Ministry of Science and Technology of the People's

Republic of China (approval number 2006-398), and was approved by the Laboratory Animal Management Committee of Chongqing Academy of Animal Sciences.

### **Goose material, RNA extraction, and RNA sequencing**

The Sichuan White geese used in this study were raised at a waterfowl-breeding base in Chongqing, Rongchang County, China. The birds were all maintained under the same environmental conditions and had free access to feed, water, and commercial corn-soybean-based diets. From hatching to 3 weeks of age, the birds received a starter feed [Metabolizable Energy (ME) = 11.76 MJ/kg, Crude Protein (CP) = 20.0%]. From 4 weeks of age to slaughter the birds were fed a grower diet (ME = 12.6 MJ/kg, CP = 17.5%). Two samples were collected from the Sichuan white geese (5 geese in each group), one at 180 days old (pre-laying), the other at 240 days old (laying). In Chongqing, with the specific climatic conditions, the reproductive cycle of geese differs slightly between individuals, but we assume that their individual situations are similar, and ovarian tissue samples were collected from geese with similar development situations.

After the geese were killed, their whole ovaries, including preovulatory follicles, prehierarchical follicles, etc., were collected and put into individual EP tubes. The collected samples were then immediately frozen in liquid nitrogen and stored at -80°C for later use. To achieve statistical validity, each RNA sample included 5 ovaries from 5 geese. Total RNA was extracted using a Biozol total RNA extraction kit (Bioer Technology, Hangzhou, China) according to the manufacturer protocol. The RNA was quantified and its quality checked using a Nano drop ND-1000 (NanoDrop Technologies, Wilmington, DE, USA). Total RNA purity and degradation were also checked on 1% agarose gels before proceeding. A 10 µg sample of total RNA was used for Illumina RNA sequencing.

Polyadenylated mRNA was purified from total RNA using oligo(dT) magnetic beads and Oligotex mRNA kits (Qiagen, USA). mRNAs were fragmented by treating with heat and divalent cations before cDNA synthesis. The cDNA was reverse transcribed with random hexamer primers, end repaired by DNA polymerase, and adapter ligated with T4 DNA ligase, according to the Illumina protocol. Ligated products were PCR-amplified and sequenced from both 5' and 3' ends on an Illumina HiSeq 2000 platform. Raw data of Illumina sequencing were obtained after base calling and stored in fastq format. The raw reads were cleaned by 1) trimming adapter sequences; 2) removing the reads that contained over 10% ambiguous 'N' nucleotides; and 3) filtering the reads with more than 50% bases having a quality score lower than 5. All subsequent analyses were based on the remaining clean reads (Cui et al. 2013).

### **Pipeline of bioinformatic analysis**

#### ***Illumina assembly***

The raw Illumina reads were preprocessed, by removing adaptor sequences, low-quality reads (those with ambiguous bases), and duplicated sequences, and were then assembled using the SOAP de novo software (<http://soap.genomics.org.cn/soapdenovo.html>) with the default settings. First, the clean reads were combined by SOAP de novo based on sequence overlap to form longer fragments without ambiguous bases (contigs). Next, the reads

were mapped back to contigs. Scaffolds were then made using SOAP de novo by connecting the contigs with ambiguous bases to present unknown sequences between each two contigs in the same transcript (Li et al., 2012). Gaps in scaffolds can be filled by paired-end sequence reads to derive sequences having the smallest number of ambiguous bases that cannot be extended from either end. Such sequences are defined as unigenes, and the following analyses are based on these (Li et al., 2012). After excluding short sequences and low-quality sequences containing more than 10% ambiguous 'N' nucleotides or 14 consecutive 'N' nucleotides, unigenes with a minimum length of 200 bp were selected for further analysis (Qin et al., 2011).

### ***Annotation of mRNAs***

We performed comprehensive functional annotation of the unigenes, including gene ontology (GO) term classification and Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment. To identify putative mRNA functions, BLASTx searches of the GenBank nonredundant database hosted by NCBI (<http://www.ncbi.nlm.nih.gov/>) were performed on all unique sequences (e value threshold  $\leq 1e^{-5}$ ; e values less than  $1.0 \times 10^{-5}$  were considered significant).

Additionally, using Blast2GO (Conesa and Gotz, 2008), we extracted GO terms (<http://www.geneontology.org>) from the best hits obtained from the BLASTx search against the nonredundant database (Myhre et al., 2006; Gotz et al., 2008). These results were then sorted by GO categories using in-house Perl scripts. BLASTx was also used to align unique sequences to the Swiss-Prot database ([http://web.expasy.org/docs/swiss-prot\\_guideline.html](http://web.expasy.org/docs/swiss-prot_guideline.html)); the eukaryotic orthologous groups (KOG) database containing proteins from three animals (the nematode *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster*, and *Homo sapiens*), consisting of 4852 clusters of orthologs, including 59,838 proteins, or ~54% of the analyzed eukaryotic 110,655 gene products; and the KEGG, to predict possible functional classifications and molecular pathways (Sui et al., 2011).

## **RESULTS**

### **EST sequence generation and assembly**

We performed Miseq sequencing on cDNA from each of the two groups of ovarian tissue samples taken from female Sichuan White geese. In total, 30,151,422 raw reads, with an average length of 151 bp and a total length of 4,552,864,722 bp were obtained. After primers and adaptors were removed, 19,167,132 clean reads, with an average length of 134.5 bp and a total length of 2,577,297,281 bp remained, among which 1,268,906,694 bp were from L (laying Period) ovarian tissue and 1,308,390,587 bp were from P (pre-laying) ovarian tissue (Table 1).

The ESTs generated in this project were subjected to cluster and assembly analysis. We also combined the data yield from one fourth of 454 GS FLX runs on each of two mixed female Sichuan White female goose tissue samples (Ding et al., 2014). In total, 16,605 assembled sequences were obtained ([Table S1](#)).

**Table 1.** Statistics of goose (*Anser cygnoides*) transcriptome sequencing generated by the MiSeq platform.

	Parameters	Read Number	Data Size (bp)	Average Read Length (bp)
P (pre-laying)	Raw data	17,158,822	2,590,982,122	151
	Clean data	9,437,934	1,268,906,694	135
L (laying Period)	Raw data	12,992,600	1,961,882,600	151
	Clean data	9,729,198	1,308,390,587	134

## Differential gene expression in L and P ovarian tissue

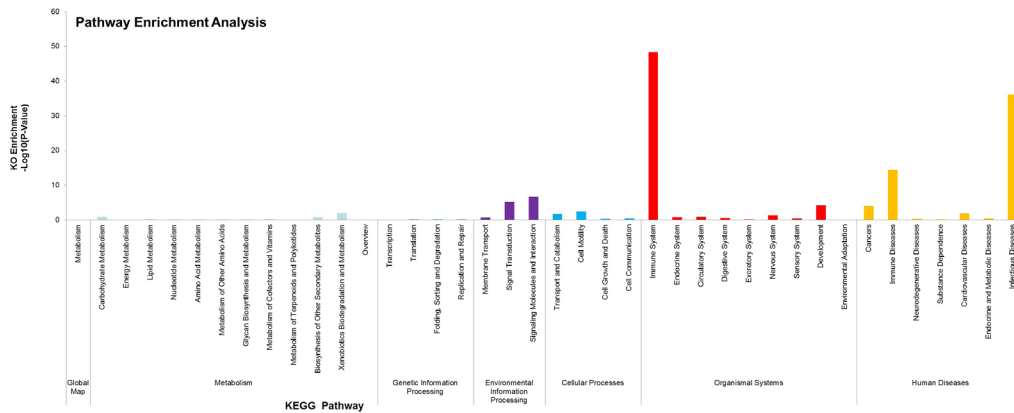
We focused on discovering genes differentially expressed between the two different states of the ovarian tissue. Therefore, in the differential gene expression analysis, we included only the matched unigenes that were merged in the *Gallus* genome predicted genes due to the relative similarities with goose. The rest of the unigenes were used only for functional analysis. Of the 16,605 assembled sequences, 511 were classed as differentially expressed by 2-fold method analysis (Table S2). A number of transcripts showed higher expression levels in the pre-laying period (377 transcripts) than in the laying period, and in the laying period (134 transcripts) of ovary tissues relative to the pre-laying period.

To classify the functions of the genes, GO enrichment analysis was used, based on sequence homology. The 16,605 assembled sequences that had BLAST hits to chicken proteins were categorized into 105 functional groups (Figure 1, Table S2). In each of the three main categories (biological process, cellular component, and molecular function) of the GO classification, “biological process”, “cell”, and “molecular function” terms were dominant, respectively. We also noticed a high percentage of genes from the categories “anatomical structure development”, “cytoplasm”, “intracellular”, “DNA binding”, and “ion binding”, yet only a few genes related to “helicase activity”, “mRNA processing”, “nuclear envelope”, “microtubule organizing center”, “histone binding”, and “RNA binding” (Figure 1). GO terms were further assigned to the obtained unigenes based on their sequence similarities to known proteins in the UniProt database annotated with GO terms, as well as on InterProScan domains they contained. Of the 511 assembled sequences categorized as differentially expressed, 104 of 134 that were more highly expressed in the laying period, and 292 of 377 that more highly expressed in the pre-laying period, were assigned at least one GO term. These unigenes were further classified into different functional categories using a set of GO slims, which are a list of high-level GO terms providing a broad overview of the ontology content (<http://www.geneontology.org/GO.slims.shtml>). The result shows the functional classification of goose unigenes into generic GO slims within the cellular component, molecular function, and biological process categories. Among these categories, genes with functions related to cell adhesion, cell death, cell differentiation, cell motility, cell proliferation, cytoskeletal organization, homeostasis, immunity, locomotion, membrane organization, stress responses, signal transduction, the cytosol, the extracellular region, the extracellular space, the plasma membrane, proteinaceous extracellular matrix, cytoskeletal protein binding, lipid binding, and protein binding and bridging, were the significantly differentially expressed groups, indicating that the ovary was undergoing highly differentiated and intensive metabolic activities.

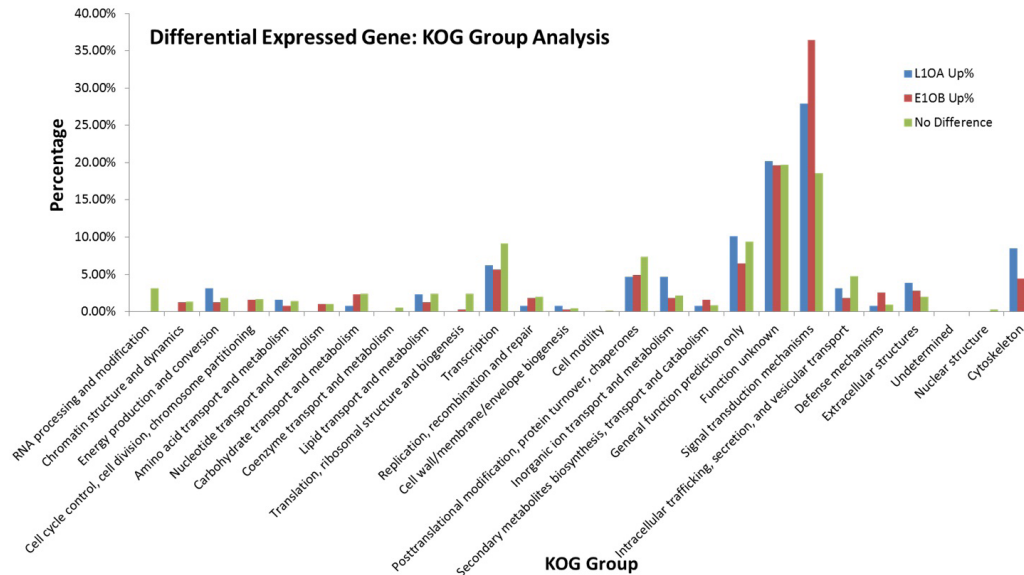
We also performed KOG cluster analysis of KOG IDs based on sequence annotation. The top 26 KOG groups are shown in Figure 2. The groups with the greatest representa-



phatidylinositol phospholipase C, beta (*PLCB*); cytochrome P450, family 19, subfamily A (*CYP19A*); steroid delta-isomerase (3beta-hydroxy-delta5-steroid dehydrogenase, *HSD3B*); and insulin-like growth factor-binding protein 3 (*IGFBP3*). These genes have been extensively studied and are known to be important in mammalian female reproductive organogenesis and reproductive cycle regulation.



**Figure 2.** Differentially Expressed Genes: KOG Group Analysis Note: KOG classification of goose unigenes identified from eukaryotic orthologous groups (KOG).



**Figure 3.** Differentially Expressed Genes: KEGG Pathway Enrichment Analysis Note: KEGG Pathway Enrichment Analysis of differentially expressed genes in geese was processed using KEGG Orthology (KO) annotation.

**Table 2.** Subset of differentially expressed genes upregulated in the ovarian tissue of *Anser cygnoides* during the pre-laying period.

KO1	ID	Gene	Base Mean A2	Base Mean B3	(B/A)Fold+Change	log2Fold+Change	Pval4	padj5	-log10pval	Function distribution
ko:K00070	isotig23801	PLCB; phosphatidylinositol phospholipase C, beta	446.4061	136.1092	0.3049	-1.7135	0.0007463	0.05934	3.1270	Catalyzes the production of the diacylglycerol and inositol 1,4,5-triphosphate from phosphatidylinositol in G-protein-linked receptor-mediated signal transduction.
ko:K07434	isotig20135	PTK2B; FAK2; protein tyrosine kinase 2 beta	817.7254	378.1892	0.4624	-1.1125	0.006087	0.3201	2.2155	Is important to the cell cycle machinery, plays important roles in the processes of oocyte maturation and fertilization.
ko:K10138	isotig31558	HSD3B; 3beta-hydroxy-delta5-steroid dehydrogenase / steroid delta-isomerase	111.0872	261.5241	2.3542	1.2352	0.03605	1	1.4430	Responsible for the oxidation and isomerization of $\Delta 5$ - $\beta$ -hydroysteroid precursors into $\Delta 4$ -ketosteroids, thus catalyzing an essential step in the formation of all classes of active steroid hormones.
ko:K05858	isotig33600	CYP19A; cytochrome P450, family 19, subfamily A	78.1724	16.5275	0.2114	-2.2417	0.04709	1	1.3270	Plays a key role in the normal progress of the menstrual/estrous cycle
ko:K05871	isotig17469	IGFBP3; insulin-like growth factor-binding protein 3	129.6017	40.8327	0.3150	-1.6662	0.04490	1	1.3477	Inhibits IGF-1 and 2, exerts a regulatory role.

KEGG orthology group number; Base Mean of Laying Period; Base Mean of Pre-laying Period; P value; P value adjusted for FDR control.



## DISCUSSION

*Anser cygnoides* has a much bigger body size and a relatively short laying period in comparison with many other poultry species. Geese are also delicious, but produce relatively few eggs (Yin et al., 2008). All the above characteristics make the goose a great model for functional genomic studies of poultry. The goose has a complex genome (amphiploid,  $2n = 78$ ), which would be difficult and expensive to sequence. As a result, progress in goose breeding research and its application has been slow. There are few molecular resources for improved breeding of geese cultivars, such as microsatellite (SSR) markers, amplified fragment-length polymorphism (AFLP) markers, or ESTs. Only 468 goose ESTs are currently in NCBI's GenBank, so it is essential to sequence the goose transcriptome to generate more data. Though sequence information from geese is scarce, fortunately *Gallus* and geese are closely related species, and the complete genome sequence available of *Gallus* has shed light on genomic studies in birds. To discover the genes influencing high egg-laying performance in geese and reconstruct the relevant regulatory network in future studies, we obtained a large EST collection. These ESTs are a valuable functional genomic resource for the goose research community.

We performed read count analyses to identify ESTs that are differentially expressed in two periods of the ovarian breeding cycle. We identified 121 differentially expressed KEGG-annotated genes that were significantly more highly expressed either in pre-laying or in laying period ovarian tissue. Since the individuals under analysis were the same variety, and we specifically sequenced genes expressed in the ovarian tissue of geese, most of the differentially expressed genes are highly likely to be associated with goose reproductive regulation and worthy of special attention in future analyses. However, our differential expression analysis alone is not sufficient to define the specific, relevant reproductive regulatory gene(s). A feasible way to do this would be to examine the co-segregation of the differentially expressed genes with egg-laying performance in the same variety.

Five of the differentially expressed genes (*PTK2B*, *PLCB*, *CYP19A*, *HSD3B*, and *IGFBP3*) are particularly interesting. The oocyte is a highly specialized cell capable of accumulating and storing energy supplies, undergoing meiosis under control of paracrine signals from the follicle, fusing with a single sperm during fertilization, and zygotic development. The oocyte accomplishes this diverse series of events by establishing an array of signal transduction pathway components that include a select collection of protein tyrosine kinases (PTKs) that are expressed at levels significantly higher than in most other cell types, reflecting the highly specialized nature of this germ cell (McGinnis et al., 2011). In addition to the cell cycle machinery, recent studies have discovered that several PTKs play important roles in the processes of oocyte maturation and fertilization (McGinnis et al., 2011). In addition, infertility can result from a wide range of defects, from behavioral, through germ cell development and maturation, to fertilization or embryo development.

Many of the hormones regulating these processes signal via G protein-coupled receptors (GPCRs), which in turn activate a range of plasma membrane enzymes, including phospholipase C (PLC)- $\beta$  isoforms. Many reproductive hormones, notably GnRH, FSH, LH, and oxytocin, signal through GPCRs, and disruption of GPCR signaling can have adverse effects on fertility. GPCR-Gq $\alpha$  signaling is mediated by four different PLC $\beta$  isozymes, PLC $\beta$ 1-4, each the product of a different gene. For most reproductive hormones, it is still not known which PLC- $\beta$  isoform is activated in response to which stimulated receptor (Filis et al., 2009). Among the four PLC $\beta$  isozymes, *PLC $\beta$ 1* appears to play an essential role in mediating GPCR

signaling in reproductive physiology, as mice with disrupted *Plcb2*, *Plcb3*, and *Plcb4* genes all are fertile (Jiang et al., 1997; Ballester et al., 2004). However, female mice with a homozygous null mutation in the phospholipase C b1 (*Plcb1*) gene (*Plcb12/2*, hereafter referred to as KOs), have altered reproductive behavior, ovulation, and pre-implantation embryo development and exhibit an implantation block (Ballester et al., 2004; Filis et al., 2013). It is known that the production of testosterone is a crucial event in masculine differentiation; however, the production of steroid hormones in females is not well understood (Antonio-Rubio et al., 2011).

It has been reported that *CYP17A1* and *CYP19A1* transcripts and CYP17A1 protein increase across the second trimester in the human fetal ovary. Furthermore, pathway analysis of transcripts altered during the onset of primordial follicle formation suggests that many are within interacting pathways that are connected with 17 $\beta$ -estradiol (Fowler et al., 2011). Estrogens are crucial for female and male fertility, as demonstrated by the severe reproductive defects observed when their synthesis (Simpson, 2004) or actions are blocked (Schomberg et al., 1999). In the ovary, locally produced estradiol acts in concert with the gonadotropins secreted from the anterior pituitary to provide for successful folliculogenesis and steroid production. As a secreted hormone, estradiol modulates the structure and function of female reproductive tissues, such as the uterus and oviduct. Estradiol is also one of the principal determinants of pituitary neuron functioning and is critical in enabling these cells to exhibit fluctuating patterns of biosynthetic and secretory activity and to generate the preovulatory surge of the luteinizing hormone (LH). Estradiol also contributes to cyclical variations in sexual female behavior. Therefore, the coordinated and cell-specific expression of the aromatase (*Cyp19a1*) gene in the ovary plays a key role in the normal progress of the menstrual/estrous cycle (Stocco, 2008). In postnatal stages, the gonads of mammals produce three essential groups of steroidal hormones: progestin (progesterone P4), estrogens (estrone E1 and estradiol E2), and androgens (androstenedione A and testosterone T).

These metabolites are produced from cholesterol in the steroidogenic pathway and are synthesized by various enzymes (Simard et al., 2005). The 3 $\beta$ -hydroxysteroid dehydrogenase/D5-D4 isomerase enzyme (*HSD3B*) is important in the synthesis of progesterone and testosterone (Sahmi et al., 2004), and the P450 aromatase (*P450arom*) product of the *Cyp19a1* gene participates in the synthesis of estrone and estradiol (Simpson et al., 1994). *HSD3B* (hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase cluster, 3 $\beta$ -HSD) is responsible for the oxidation and isomerization of  $\Delta^5$ -3 $\beta$ -hydroxysteroid precursors into  $\Delta^4$ -ketosteroids, thus, catalyzing an essential step in the formation of all classes of active steroid hormones. In humans, expression of the type I isoenzyme accounts for the *HSD3B* activity found in placenta and peripheral tissues, whereas the type II *HSD3B* isoenzyme is predominantly expressed in the adrenal gland, ovary, and testis, and its deficiency is responsible for a rare form of congenital adrenal hyperplasia (Simard et al., 2005). In mouse, the *HSD3B* family contains 7 members (*HSD3B1-7*). The different isoforms are expressed in a tissue and developmentally specific manner and fall into two functionally distinct groups. Studies on *HSD3B* in birds are scarce, and mainly concern its expression pattern across various tissues (Sechman et al., 2011).

In the mammalian ovary, insulin-like growth factor 3 (IGFBP-3) in follicular fluid may exert a regulatory role, inhibiting IGF-1 and 2 (Amato et al., 1999). The increased concentration of IGFBP-3 may decrease the level of IGF-1 and 2 in follicular fluid and consequently affect their functions. Previous studies have shown that IGF-1 and IGFBP-3 are associated with oocyte maturation (Oosterhuis et al., 1998; Nardo et al., 2001). Moreover, IGFBP-3 follicular fluid concentration in older women did not differ from that in younger controls (Klein et al., 2000). The five

differentially expressed genes identified in this study are involved in the development and maturation of the ovary, and are likely to play important roles in regulating the goose breeding cycle.

This study is one of several extensive efforts to detect genes related to high reproductive performance that are expressed in developing ovarian tissue or reproductive organs (El-Hanoun et al., 2012; Xu et al., 2013; Zhang et al., 2013; Zhao et al., 2013), yet the critical genes have not yet been discovered. It is noteworthy that low egg-laying performance is common in most species of geese, so the differentially expressed genes are associated with physiological and biochemical characteristics that distinguish geese from other poultry. We conducted differential expression analysis using RNA from different times in the ovarian cycle, hoping to discover differences in functional gene expression. Although some differentially expressed genes were found, involving multiple functional classifications, we cannot establish the direct relationship of these genes with egg-laying performance on the basis of the current data. Moreover, genes known to be important in egg formation and ovarian development appear not to have direct roles in high reproductive performance (Xu et al., 2013). Nevertheless, our study provides some novel insights into the molecular mechanisms of high egg-laying performance determination in geese, as well as a valuable functional genomics resource and a list of candidate genes for functional analysis in the future.

In conclusion, we generated a large EST collection and identified a list of candidate genes that are differentially expressed between the pre-laying and laying periods in goose ovaries. It is noteworthy that we detected 121 differentially expressed annotated genes. In the laying period, 101 of them were downregulated, and 20 of them were upregulated. Five of them are of special interest for further investigation for their roles in determining high reproductive performance. These data will be of considerable interest to the goose research community. Our study also provides an archive for future studies of the molecular mechanism underlying survivable characteristics of large birds, and may also be useful for future *de novo* sequencing of this species.

### Conflicts of interest

The authors declare no conflict of interest.

### ACKNOWLEDGMENTS

The authors would like to acknowledge the members of the Poultry Institute of Chongqing Academy of Animal Science for managing the birds and collecting data, and the work of Shanghai Personal Biotechnology Co., Ltd., in sequencing and data analysis. Research supported by the Application Development Projects of Chongqing Science and Technology (#CSTC2013YYKFC80003); Chongqing Agricultural Development Foundation (#11403); Chongqing Fundamental Research Funds Projects (#2012csk-jbky-00915).

### [Supplementary material](#)

### REFERENCES

- Amato G, Izzo A, Tucker AT and Bellastella A (1999). Lack of insulin-like growth factor binding protein-3 variation after follicle-stimulating hormone stimulation in women with polycystic ovary syndrome undergoing *in vitro* fertilization. *Fertil. Steril.* 72: 454-457.

- Antonio-Rubio NR, Guerrero-Estevez SM, Lira-Romero E and Moreno-Mendoza N (2011). Expression of 3 $\beta$ -HSD1 and P450 Aromatase enzymes during mouse gonad differentiation. *J. Mol. Histol.* 42: 535-543.
- Ballester M, Molist J, Lopez-Bejar M, Sanchez A, et al. (2004). Disruption of the mouse phospholipase C-beta1 gene in a beta-lactoglobulin transgenic line affects viability, growth, and fertility in mice. *Gene* 341: 279-289.
- Conesa A and Gotz S (2008). Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int. J. Plant Genomics* 2008: 619832.
- Cui Z, Li X, Liu Y, Song C, et al. (2013). Transcriptome profiling analysis on whole bodies of microbial challenged *Eriocheir sinensis* larvae for immune gene identification and SNP development. *PLoS One* 8: e82156.
- Ding N, Han Q, Li Q, Zhao XZ, et al. (2014). Comprehensive analysis of Sichuan white geese (*Anser cygnoides*) transcriptome. *Anim. Sci. J.* 85: 650-659.
- El-Hanoun AM, Attia YA, Gad HA and Abdella MM (2012). Effect of different managerial systems on productive and reproductive traits, blood plasma hormones and serum biochemical constituents of geese. *Animal* 6: 1795-1802.
- Filis P, Lannagan T, Thomson A, Murray AA, et al. (2009). Phospholipase C-beta1 signaling affects reproductive behavior, ovulation, and implantation. *Endocrinology* 150: 3259-3266.
- Filis P, Kind PC and Spears N (2013). Implantation failure in mice with a disruption in *Phospholipase C beta 1* gene: lack of embryonic attachment, aberrant steroid hormone signalling and defective endocannabinoid metabolism. *Mol. Hum. Reprod.* 19: 290-301.
- Fowler PA, Anderson RA, Saunders PT, Kinnell H, et al. (2011). Development of steroid signaling pathways during primordial follicle formation in the human fetal ovary. *J. Clin. Endocrinol. Metab.* 96: 1754-1762.
- Gotz S, Garcia-Gomez JM, Terol J, Williams TD, et al. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36: 3420-3435.
- Jiang H, Kuang Y, Wu Y, Xie W, et al. (1997). Roles of phospholipase C beta2 in chemoattractant-elicited responses. *Proc. Natl. Acad. Sci. U.S.A.* 94: 7971-7975.
- Kang B, Guo JR, Yang HM, Zhou RJ, et al. (2009). Differential expression profiling of ovarian genes in prelaying and laying geese. *Poult. Sci.* 88: 1975-1983.
- Kaur S, Cogan NO, Pembleton LW, Shinozuka M, et al. (2011). Transcriptome sequencing of lentil based on second-generation technology permits large-scale unigene assembly and SSR marker discovery. *BMC Genomics* 12: 265.
- Kent WJ (2002). BLAT-the BLAST-like alignment tool. *Genome Res.* 12: 656-664.
- Klein NA, Battaglia DE, Woodruff TK, Padmanabhan V, et al. (2000). Ovarian follicular concentrations of activin, follistatin, inhibin, insulin-like growth factor I (IGF-I), IGF-II, IGF-binding protein-2 (IGFBP-2), IGFBP-3, and vascular endothelial growth factor in spontaneous menstrual cycles of normal women of advanced reproductive age. *J. Clin. Endocrinol. Metab.* 85: 4520-4525.
- Koelkebeck KW, Parsons CM, Douglas MW, Leeper RW, et al. (2001). Early postmolt performance of laying hens fed a low-protein corn molt diet supplemented with spent hen meal. *Poult. Sci.* 80: 353-357.
- Li C, Weng S, Chen Y, Yu X, et al. (2012). Analysis of *Litopenaeus vannamei* transcriptome using the next-generation DNA sequencing technique. *PLoS ONE* 7: e47442.
- McGinnis LK, Carroll DJ and Kinsey WH (2011). Protein tyrosine kinase signaling during oocyte maturation and fertilization. *Mol. Reprod. Dev.* 78: 831-845.
- Myhre S, Tveit H, Mollestad T and Laegreid A (2006). Additional gene ontology structure for improved biological reasoning. *Bioinformatics* 22: 2020-2027.
- Nardo LG, Bellanca SA, Burrello N, Longo G, et al. (2001). Concentrations of insulin-like growth factor (IGF)-I and IGF binding protein-3 in the follicular fluid of women undergoing ovarian hyperstimulation with different gonadotropin preparations. *Gynecol. Endocrinol.* 15: 413-420.
- Oosterhuis GJ, Vermes I, Lambalk CB, Michgelsen HW, et al. (1998). Insulin-like growth factor (IGF)-I and IGF binding protein-3 concentrations in fluid from human stimulated follicles. *Hum. Reprod.* 13: 285-289.
- Pan ZX, Han CC, Wang JW, Li L, et al. (2011). Cloning and expression of stearoyl-CoA desaturase 1 (SCD-1) in the liver of the Sichuan white goose and landes goose responding to overfeeding. *Mol. Biol. Rep.* 38: 3417-3425.
- Qin YF, Fang HM, Tian QN, Bao ZX, et al. (2011). Transcriptome profiling and digital gene expression by deep-sequencing in normal/regenerative tissues of planarian *Dugesia japonica*. *Genomics* 97: 364-371.
- Sahmi M, Nicola ES, Silva JM and Price CA (2004). Expression of 17beta- and 3beta-hydroxysteroid dehydrogenases and steroidogenic acute regulatory protein in non-luteinizing bovine granulosa cells *in vitro*. *Mol. Cell Endocrinol.* 223: 43-54.
- Schomberg DW, Couse JF, Mukherjee A, Lubahn DB, et al. (1999). Targeted disruption of the estrogen receptor-alpha gene in female mice: characterization of ovarian responses and phenotype in the adult. *Endocrinology* 140: 2733-2744.

- Sechman A, Pawlowska K and Hrabia A. (2011). Effect of 3,3',5-triiodothyronine and 3,5-diiodothyronine on progesterone production, cAMP synthesis, and mRNA expression of *STAR*, *CYP11A1*, and *HSD3B* genes in granulosa layer of chicken preovulatory follicles. *Domest. Anim. Endocrinol.* 41: 137-149.
- Shalev BA, Dvorin A, Herman R, Katz Z, et al. (1991). Long-term goose breeding for egg production and crammed liver weight. *Br. Poult. Sci.* 32: 703-709.
- Simard J, Ricketts ML, Gingras S, Soucy P, et al. (2005). Molecular biology of the 3beta-hydroxysteroid dehydrogenase/delta5-delta4 isomerase gene family. *Endocr. Rev.* 26: 525-582.
- Simpson ER (2004). Models of aromatase insufficiency. *Semin. Reprod. Med.* 22: 25-30.
- Simpson ER, Mahendroo MS, Means GD, Kilgore MW, et al. (1994). Aromatase cytochrome P450, the enzyme responsible for estrogen biosynthesis. *Endocr. Rev.* 15: 342-355.
- Stevenson MH (1989). Nutrition of domestic geese. *Proc. Nutr. Soc.* 48: 103-111.
- Stocco C (2008). Aromatase expression in the ovary: hormonal and molecular regulation. *Steroids* 73: 473-487.
- Sui C, Zhang J, Wei J, Chen S, et al. (2011). Transcriptome analysis of *Bupleurum chinense* focusing on genes involved in the biosynthesis of saikosaponins. *BMC Genomics* 12: 539.
- Xie F, Burklew CE, Yang Y, Liu M, et al. (2012). *De novo* sequencing and a comprehensive analysis of purple sweet potato (*Ipomoea batatas* L.) transcriptome. *Planta* 236: 101-113.
- Xu Q, Zhao W, Chen Y, Tong Y, et al. (2013). Transcriptome profiling of the goose (*Anser cygnoides*) ovaries identify laying and broodiness phenotypes. *PLoS ONE* 8: e55496.
- Yin T, Difazio SP, Gunter LE, Zhang X, et al. (2008). Genome structure and emerging evidence of an incipient sex chromosome in *Populus*. *Genome Res.* 18: 422-430.
- Zhang X, Kang B, Zhang LN, Guo JR, et al. (2013). Gene expression profiles of LH, prolactin and their receptors in female Zi geese (*Anser cygnoides*) during development. *Folia Biol. (Krakow)* 61: 59-64.
- Zhao X, Shao T, Wang YQ, Lu XL, et al. (2013). The phytoestrogen daidzein may affect reproductive performance of Zhedong White geese by regulating gene mRNA levels in the HPG axis. *Br. Poult. Sci.* 54: 252-258.