Analysis of expression of candidate genes for polycystic ovary syndrome in adult and fetal human and fetal bovine ovaries†

Menghe Liu[1], Katja Hummitzsch[1], Monica D Hartanti[1,2], Roseanne Rosario[3], Nicole A Bastian[1], Nicholas Hatzirodos[1], Wendy M Bonner[1], Helen F Irving-Rodgers[1,4], Joop S.E. Laven[5], Richard A Anderson[3] and Raymond J Rodgers[1*]

[1] Discipline of Obstetrics and Gynaecology, School of Medicine, Robinson Research Institute, The University of Adelaide, Adelaide, SA 5005, Australia; [2] Faculty of Medicine, Trisakti University, Jakarta 11440, Indonesia; [3] Medical Research Council Centre for Reproductive Health, University of Edinburgh, Edinburgh, EH16 4TJ, United Kingdom; [4] School of Medical Science, Griffith University, Gold Coast Campus, QLD 4222, Australia; [5] Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynaecology, Erasmus University Medical Center, Rotterdam, Netherlands

†Grant support. Funding was provided by Australia Awards Scholarship from the Australian Government, National Health and Medical Research Council of Australia, https://www.nhmrc.gov.au/, and The University of Adelaide, http://www.adelaide.edu.au/ and UK Medical Research Council (grant no G1100357) https://mrc.ukri.org/. Menghe Liu is supported by Adelaide University China Fee Scholarship.

*Correspondence. Prof. Raymond Rodgers, Discipline of Obstetrics and Gynaecology, School of Medicine, Robinson Research Institute, The University of Adelaide, Adelaide, SA 5005, Australia ray.rodgers@adelaide.edu.au

Running title. Expression of PCOS candidate genes in the fetal ovary.

Abstract

Polycystic ovary syndrome (PCOS) appears to have a genetic predisposition and a fetal origin. We compared the expression levels of 25 PCOS candidate genes from adult control and PCOS human ovaries (n = 16) using microarrays. Only one gene was potentially statistically different. Using qRT-PCR, expression of PCOS candidate genes was examined in bovine fetal ovaries from early stages when they first developed stroma through to completion of development (n = 27; 60-270 days of gestation). The levels of *ERBB3* mRNA negatively correlated with gestational age but positively with *HMGA2, FBN3, TOX3, GATA4* and *DENND1A.X1,2,3,4*, previously identified as correlated with each other and expressed early. *PLGRKT* and *ZBTB16*, and less so *IRF1*, were also correlated with *AMH, FSHR, AR, INSR* and *TGFB1I1*, previously identified as correlated with each other and expressed late. *ARL14EP*, *FDFT1*, *NEIL2* and *MAPRE1* were expressed across gestation and not correlated with gestational age as shown previously for *THADA, ERBB4, RAD50, C8H9orf3, YAP1, RAB5B, SUOX* and *KRR1*. *LGCGR* because of its unusual bimodal expression pattern had some unusual correlations with other genes. In human ovaries (n = 15, < 150 days of gestation), *ERBB3.V1* and *ERBB3.VS* were expressed and correlated negatively with gestational age and positively with *FBN3, HMGA2, DENND1A.V1,3,4, DENND1A.V1-7, GATA4* and *FSHR*, previously identified as correlated with each other and expressed early. Thus, the general lack of differential expression of candidate genes in adult ovaries contrasting with dynamic patterns of gene expression in fetal ovaries is consistent with a vulnerability to disturbance in the fetal ovary that may underpin development of PCOS.

Summary sentence

Many PCOS candidate genes are expressed in fetal ovaries in one of three patterns either early, late or across gestation and within these groups, particularly the early and late, the levels of gene expression are correlated with each other.

Keywords: adult ovary, fetal ovary, polycystic ovary syndrome, gene expression, genetics.

Introduction

Polycystic ovary syndrome (PCOS) is one of the most prevalent endocrine disorders among women, with an incidence of about 10% [1, 2]. It is identified with hyperandrogenism often exhibiting male-like body or facial hair [3] and menstrual irregularity or oligomenorrhea and polycystic ovaries [4]. PCOS is also associated with secondary complications, such as impaired glucose tolerance, type 2 diabetes, sleep disorders and possibly adverse cardiovascular diseases [5-7].

The etiology of PCOS is not clearly understood. Studies on twins and families [8, 9] suggested a genetic basis of PCOS and transcriptome and genome-wide association studies (GWAS) have led the studies of PCOS into a new era [10-12]. There is also evidence of a fetal or early postnatal origin of PCOS [13-17]. The prevailing hypothesis is that a female fetus exposed to elevated levels of androgens is at increased risk of developing PCOS in later life. The cause of elevation in androgens is unknown. However, pregnant women with PCOS have elevated levels of androgens [18] possibly due to elevated levels of anti-Mullerian hormone [17] leading to offspring of women with PCOS having an increased predisposition to developing PCOS.

We have investigated a possible link between the genetic and fetal origins of PCOS by studying the expression of PCOS candidate genes in human and bovine fetal ovaries. We initially examined *FBN3* [19] identified by D19S884 microsatellite analyses in familial linkage studies [20, 21]. *FBN3* and other members of the TGFβ signaling pathway were expressed in the fetal ovaries [19] suggesting that this pathway maybe important for development of PCOS, particularly the morphological characteristics of increased stroma and stromal collagen in the PCOS ovary [22, 23]. TGFβ is a well-known stimulator of stromal cell replication and collagen synthesis.

We further examined the expression of other PCOS candidate genes identified from GWAS studies [10, 11] including *DENND1A, LHCGR, FSHB, FSHR, YAP1, INSR, RAB5B, TOX3, HMGA, C9orf3), GATA4, ERBB4, RAD50, THADA, SUOX, KRR1* and *SUMO1P1* and additionally *FBN3, TGFB1I1* and *AR* [24]. Nearly all these genes were expressed in the developing fetal ovary. The pattern of expression of each gene followed one of three patterns, some were expressed only early in gestation, some only late in gestation and some throughout gestation [24]. The mRNA levels of many genes within each group, particularly in the early and late groups, were highly correlated with each other [24]. In other studies of genes selected for association with different cells types in the developing ovary [25, 26] hierarchical clustering [25] also revealed three different patterns of expression across gestation with genes in each group generally being highly correlated with each other. Presumably these different patterns reflect the successive developmental processes of stroma penetrating from the mesonephros into the ovary primoridium and its continued expansion, the

formation of ovigerous cords and their dissolution into follicles and formation of a surface epithelium and the tunica albuginea during the development of the ovary [27, 28].

Subsequently newer GWAS studies identified more loci containing three new genes: plasminogen receptor with a C-terminal lysine (*PLGRKT*), zinc finger and BTB domain containing 16 (*ZBTB16*) and microtubule-associated protein RP/EB family member 1 (*MAPRE1*) [29] and these have not been examined in fetal ovaries. Five other genes, whilst not the closest to the SNP related to PCOS, are also present in these loci and also have not been examined in fetal ovaries. The locus *GATA4* (rs804279) encompasses the promoter region of farnesyl-diphosphate farnesyltransferase 1 (*FDFT1*) [29] which has not been examined in fetal ovaries. Also, interferon regulatory factor 1 (*IRF1*), ADP ribosylation factor like GTPase 14 effector protein (*ARL14EP*), erb-b2 receptor tyrosine kinase 3 (*ERBB3*) and nei like DNA glycosylase 2 (*NEIL2*) are in loci associated with *RAD50*, *FSHB*, *RAB5B* and *GATA4*, respectively [30], and are thus potentially associated with PCOS.

To advance our line of investigation into the genetic and fetal origins of PCOS further, we first examined the expression of PCOS candidate genes using microarray data from human adult ovarian cortex and stroma from control women and women with PCOS [31]. Then using the same bovine and human fetal tissues as previously [24], we analysed the expression of these additional genes (*PLGRKT*, *ZBTB16*, *MAPRE1*, *FDFR1*, *IRF1*, *ARL14EP*, *ERBB3*, *NEIL2*) and related their expression patterns to previous data [24].

Materials and methods

Ethics approvals

The study of human fetal ovaries was approved by the Lothian Research Ethics Committee in Scotland (ref 08/S1101/1) and all methods were performed in accordance with the relevant guidelines and regulations of this approval. There were no ethical issues with the bovine study, because fetal bovine samples were collected from the local abattoir.

Microarray data processing

Microarray data were obtained from a 2004 study conducted by Jansen et al. [31]. That original study was approved by the institutional ethical review boards of Erasmus Medical Center, the Daniel den Hoed Oncology Hospital (both located in Rotterdam), as well as the Flevo Hospital in Almere, The Netherlands [31]. Ovary samples containing cortex and stroma were collected from PCOS patients and normal controls. The controls consisted of women below the age of 40 with

regular menstrual cycles (21 to 35 days) and without polycystic ovaries or ovarian cancer. PCOS subjects (20 to 40 years old) included in this study were women with at least two out of three PCOS criteria (irregular menstrual cycles, clinical or endocrine hyperandrogenism or the presence of polycystic ovaries) with other endocrine conditions excluded as defined previously [32, 33]. Detailed patient information has been published previously [31]. GeneChips HG_U133A and HG_U133B (Affymetrix, High Wycombe, UK) which contained over 45,000 human DNA fragments were hybridised to fragmented biotin-labeled cRNA from 8 different normal ovary samples and 8 different PCOS ovary samples according to Affymetrix protocols. Hybridisation was conducted at the Organon Gene Chip Platform (Newhouse, UK). For microarray data, Affymetrix CEL files were imported into Partek Genomics Suite Software version 7.0 (Partek Incorporated, St Louis, MO, USA). Raw gene expression data was normalized using pre-background adjustment for GC content, Robust Multi-array Average (RMA) background correction with quantile normalization and mean probeset summarisation. For both GeneChips HG_U133A and HG_U133B, mean and standard error of mean (SEM) of non-transformed raw expression data for PCOS candidate genes were calculated.

Collection of human fetal ovaries

Human fetal ovaries (n = 15, 8-20 weeks of gestation) were obtained following medical termination of pregnancy and gestational ages were determined as previously described [24]. Pregnancies were all terminated for social reasons and all fetuses appeared morphologically normal. Maternal informed consent was obtained and the study was approved by the Lothian Research Ethics Committee (ref 08/S1101/1).

Collection of bovine fetal ovaries

Fetuses of pregnant *Bos taurus* cows were collected at Thomas Foods International, Murray Bridge, SA, Australia and transported on ice to the laboratory. The gestational ages of fetuses were estimated with crown-rump length [34]. Gender of young fetuses (smaller than 8 cm) was confirmed as reported previously [24, 27].

Classification of bovine fetal ovaries

As reported previously [35], ovaries were separated into five groups; stage I: ovigerous cord formation (n = 7), stage II: ovigerous cord breakdown (n = 4), stage III: follicle formation (n = 3), stage IV: ovarian surface epithelium formation (n = 8) and stage V: tunica albuginea formation (n = 5).

RNA extraction and cDNA synthesis for human fetal ovaries

As reported previously [24], RNA from fetal ovaries was extracted using the RNeasy Micro Kit (Qiagen, Crawley, UK) according to manufacturer's instructions. Five hundred ng of RNA was reverse transcribed to cDNA using concentrated random primers and Superscript III reverse transcriptase (Life Technologies). cDNA synthesis reaction was diluted 1:20 before proceeding.

RNA extraction and cDNA synthesis for bovine fetal ovaries

Fetal ovaries were homogenized in 1 ml Trizol® (Thermo Fisher Scientific, Waltham, MA, USA) and RNA was extracted according to manufacturer's instructions. Ten µg RNA of each sample was treated with DNase I (Promega/Life Technologies Australia Pty Ltd, Tullmarine, Vic, Australia) for 20 min at 37°C. Two hundred ng of DNAse-treated RNA was used for cDNA synthesis as previously described [24, 27].

Quantitative real-time PCR of human samples

Primers for human fetal ovary samples were designed and primer pair efficiencies were tested as described previously [24] and the details are shown in Supplementary Table S1. *ERBB3* is alternatively spliced and *ERBB3* variant 1 (*ERBB3.V1*; Accession Number: NM_001982.4) is the longer variant. *ERBB3* variant s (*ERBB3.VS*; Accession Number: NM_001005915.1) shares exon 1, exon 2 and 186 bases in exon 3 with *ERBB3.V1*, and lacks other 3' exons found in variant 1 and contains an alternate 3' exon of its own (exon 3). Primers for *ERBB3.V1* were designed to span exon 3 to exon 4, and they were unique to this variant. However, the primers for *ERBB3.VS* were designed to span exon 2 to exon 3, and it could also detect variant 1. The amplifications and data analysis were performed as described previously [24].

Quantitative real-time PCR of bovine samples

Based on available RNA sequences in NCBI, PCR primers for bovine fetal ovary samples were designed to span introns using Primer3 plus and Net primer (PREMIER Biosoft Palo Alto, CA, USA) software and primers are listed in Supplementary Table S1. To test primer combinations, cDNA was diluted from 1 in 4 to 1 in 1000 for establishing 5 standards and generating a standard curve. The combinations which generated a single sharp peak and achieved an amplification efficiency of 0.9-1.1 and an R^2 values ≥ 0.98 were used (Supplementary Table S1). The reactions, cycling conditions, Ct values and gene expression analyses were performed as previous work [24] and the same housekeeping genes, ribosomal protein L32 *(RPL32)* and peptidylprolyl isomerase A *(PPIA)* were used.

Statistical analyses

All statistical analyses were carried out using Microsoft Office Excel 365 (Microsoft Redmond, WA, USA) and GraphPad Prism version 8.00 (GraphPad Software Inc., La Jolla, CA, USA). To determine significant differences in the expression of PCOS candidate genes between normal and PCOS individuals, a one-way ANOVA and Tukey's post hoc tests were conducted. The $2^{-\Delta Ct}$ data for fetal ovarian sample were plotted in scatter plots and bar graphs to describe their levels of expression during ovarian development. Difference between the level of mRNA expression of each gene was analysed using one-way ANOVA with Holm-Sidak *post-hoc* test. Network graphs were plotted and illustrated as previously described [24, 36].

Results

Expression of PCOS candidate genes in adult human ovaries

The microarray intensities from human ovaries of 25 genes were determined on U133A and U133B chips and are shown in Table 1. We identified that only *RAD50* was differentially expressed and upregulated in PCOS ovaries compared with the control ovaries ($P < 0.05$), however, only one of the three probes for *RAD50* on these arrays (Table 1) was statistically significantly. Additionally 57 ANOVAs (Table 1) were conducted and at P equal to 0.05 as the cut off, it is very possible that this result could have arisen by chance.

Expression of PCOS candidate genes in bovine fetal ovaries

The levels of mRNA expression of PCOS candidate genes were determined in ovaries from bovine fetuses (n = 27, gestational ages from 60 to 270 days). *ERBB3* was expressed early in gestation and then declined (Figures 1A and 1E). In contrast, mRNA expression of *IRF1* and *ZBTB16* was low in early gestation and gradually increased until the end of gestation (Figures 1B, 1C, 1F and 1H). *PLGRKT* was upregulated in the early stage and reached the highest point at approximately 200 days of gestation and then declined in the later stages (Figures 1C and 1G). Similarly, *NEIL 2* was upregulated and reached the peak at around 150 days, and then declined in the later stages (Figure 2D and 2H). The other genes *ARL14EP*, *FDFT1* and *MAPRE1* were expressed consistently across gestation (Figures 2A, 2B and 2C).

In order to analyse the correlations between expression levels of PCOS candidate genes and gestational age and correlations between different genes, a Pearson correlation matrix was generated (Table 2). The mRNA expression of *ERBB3* ($P < 0.001$) negatively correlated with gestational age, however, *PLGRKT* ($P < 0.05$), *IRF1* ($P < 0.01$), and *ZBTB16* ($P < 0.0001$)

positively correlated with gestational age (Table 2). The other genes *MAPRE1*, *FDFT1*, *NEIL2* and *ARL14EP* were not significantly correlated with gestational age. *ERBB3* positively correlated with *HMGA2* (Figure 3A, $P < 0.0001$), *FBN3* (Figure 3B, $P < 0.001$), *TOX3* ($P < 0.01$), *GATA4* ($P < 0.01$) and *DENND1A.X1,2,3,4* ($P < 0.01$). *MAPRE1* positively correlated with *NEIL2* ($P < 0.001$), and *ARL14EP* positively correlated with *PLGRKT* ($P < 0.001$) (Table 2). *MAPRE1* and *NEIL2* both positively correlated with *RAD50* ($P < 0.01$), *RAB5B* ($P < 0.01$, $P < 0.001$) and *ARL14EP* ($P < 0.001$, $P < 0.0001$) (Table 2). Likewise, *IRF1* and *ZBTB16* both positively correlated with *AR* ($P < 0.01$, $P < 0.001$), *INSR* ($P < 0.001$) and *TGFB1I1* ($P < 0.001$) (Table 2). In addition, *ZBTB16* negatively correlated with *HMGA2* (Figure 3C, $P < 0.0001$) and *FBN3* (Figure 3D, $P < 0.0001$). After correlation values between genes were identified, adjacent matrixes were plotted using qgraph R package (Figure 4). Most of the genes, except *ERBB4* and *FDFT1*, were closely and highly connected with each other and as well as with the gestational age.

Expression of PCOS candidate genes in human fetal ovaries

The levels of gene expression of *ERBB3.V1* and *ERBB3.VS* were determined in ovaries from human fetuses which were younger than 150 gestational days (Figure 5). *ERBB3.V1* and *ERBB3.VS* were highly expressed before 70 days of gestation, then their expression levels declined with a very similar pattern. During this period of gestation, mRNA levels of *ERBB3.V1* and *ERBB3.VS* both negatively correlated with gestational age (Table 3, $P < 0.05$). In addition, *ERBB3.V1* and *ERBB3.VS* positively correlated with *FBN3* (Figure 6A, $P < 0.001$ and $P < 0.01$, respectively), *HMGA2* ($P < 0.001$), *DENND1A.V1,3,4* ($P < 0.05$), *DENND1A.V1-7* ($P < 0.05$), *GATA4* (Figure 6B, $P < 0.0001$) and *FSHR* ($P < 0.001$). *ERBB3.V1* and *ERBB3.VS* were positively correlated with each other (Figure 6C, $P < 0.0001$).

Discussion

This study reports on the expression of PCOS candidate genes in adult PCOS and non-PCOS ovaries and in developing human and bovine fetal ovaries. The original microarray study of adult PCOS ovaries and non-PCOS ovaries [31] preceded the identification of PCOS candidate genes and so it was illuminating to examine these genes. The study of human and bovine fetal ovaries utilized fetal ovaries that had been previously examined [24] and the current data are compared with them in this manuscript. This study thus provides an overview of the expression of PCOS candidate genes in ovaries and indicates that most of the activity occurs during the development of the fetal ovary.

The microarray data had information on twenty-five PCOS candidate genes and there were no differentially expressed genes identified except for *RAD50*, which was elevated in the PCOS ovaries. Given that there were multiple genes examined, we thus conducted multiple ANOVAs in our study. Thus obtaining statistical significance for *RAD50* at $P < 0.05$ could have happened by chance and this result will need to be confirmed. *RAD50* is a double strand break repair protein essential for telomere maintenance and meiosis [37]. In the ovary, like other genome repair genes, it is responsive to environmental toxicants such as bisphenol A and phalates [38, 39] and is expressed in oocytes [40]. RAD51 is another DNA repair protein, and it collaborates with RAD50 to allow cells to survive in a different pathway [41]. Variants in loci containing the related *RAD51* are associated with early menopause [42] and *RAD51* Associated Protein 1 is downregulated in the theca interna during atresia [43].

Both *ERBB3* and *ERBB4* are located in PCOS candidate loci and are receptors in epidermal growth factor receptor (EGFR) signaling pathways, which regulate growth, proliferation and differentiation of mammalian cells [44]. Pathways involving EGF and EGF-like peptides are critical for follicular and oocyte development and ovulation [45] but studies of fetal ovaries are not comprehensive. The EGFR signaling pathways contains four receptors: EGFR, receptor tyrosine-protein kinase erbB-2 (ERBB2), ERBB4 and ERBB3 which have distinct ligand-binding specificity [46]. Although, ERBB3 lacks an active kinase domain, it has a ligand binding domain which functions without protein phosphorylation during signal transfer. ERBB3, however, forms heterodimers with other kinase-active EGFR family members for the activation of cell proliferation or differentiation. Unlike ERBB3, ERBB4 has a phosphotidylinositol-3 kinase binding site and it is capable of conveying signals downstream [46] and it can form heterodimers with ERBB3 [47]. Our current bovine results show that *ERRB3* was highly expressed early in gestation and *ERBB4* was found previously to be constantly expressed throughout the gestation [24]. The levels of *ERBB3* mRNA were positively correlated with *HMGA2, FBN3, TOX3, GATA4* and *DENND1A.X1,2,3,4*. In humans *ERBB3.V1*and *ERBB3.VS* encode isoform 1 and isoform S, respectively. Isoform 1 (Reference sequence: NP_001973.2) is the longer isoform and it contains an extracellular domain, a transmembrane domain and an intracellular domain [48, 49]. However, isoform S (Reference sequence: NP_001005915.1) lacks the intermembrane region and is thus secreted from cells [50, 51]. In human fetal ovaries both *ERBB3.V1* and *ERBB3.VS* were negatively correlated with gestational age and positively correlated with *HMGA2, FBN3, TOX3, GATA4* and *DENND1A* (V1-7 and V1,3,4), which was consistent with our results in the bovine fetal ovary. Therefore, this suggests that EGFR signaling pathway probably plays a vital role throughout ovary development, although the mechanism might differ from early to late gestation. In addition other studies have linked *DENND1A.V1* and *DENND1A.V2* to PCOS and thecal androgen production [52, 53]. Our

previous study found that the expression of *DENND1A.V2* was not detectable in bovine fetal ovaries [24], however, using primers that could potentially amplify a number of potential variants, elevated levels of *DENND1A* were detected in bovine and human fetal ovaries in the early stages [24].

PLGRKT, *IRF1* and *ZBTB16* were expressed in the late stages of bovine fetal ovary development and their expression levels correlated with gestational age. *PLGRKT* and *ZBTB16*, and less so *IRF1*, were also correlated with *AMH*, *FSHR*, *AR*, *INSR* and *TGFB1I1* previously identified as expressed late in gestation [24]. Plg-RKT is the unique transmembrane receptor of plasminogen and it is located on the cell surface, directly interacting with plasminogen activator [54, 55]. In other tissues it has been shown to play an important role in maintaining extracellular matrix and regulating the proliferation and apoptosis of epithelial cells by regulating fibrinolysis [56]. The plasminogen system is essential for ovarian function in adult ovaries. The expression of plasminogen activator increases in granulosa cells at ovulation [57] and the level of plasminogen activator inhibitor-1 decreases [58]. Furthermore, gonadotropins, especially FSH, can stimulate plasminogen activator in granulosa cells [59]. In a recent study the fibrinolytic/proteolytic system was examined in perinatally androgenized mice [60]. Plasminogen activator inhibitor-1 was expressed throughout the ovaries of hyperandrogenized mice and the pattern of expression was different to the peripheral distribution observed in control ovaries [60]. Plasminogen was present in small follicles but only in the ovaries of the hyperandrogenized mice [60]. IRF1 is a transcription factor that regulates type I interferons and positively [61]. It has been found tha*t IRF1* may lead PCOS risk by a developmental mechanism [12] and *IRF1* is highly expressed in ovarian cancers [62]. *ZBTB16* encodes a promyelocytic leukemia zinc finger, a protein that functions as a transcriptional regulator in development [63, 64] and appears to be responsible for many of the features of the metabolic syndrome [63]. The elevated expression levels of *PLGRKT*, *IRF1* and *ZBTB16* suggest that these three genes and corresponding signals and systems may regulate the late stages of development of the fetal ovary. Furthermore, abnormal expression of these signals at late stage may be related to the fetal origin of PCOS.

ARL14EP, *FDFT1*, *NEIL2* and *MAPRE1* were expressed across gestation and their expression levels were not correlated with gestational age, similarly to other PCOS candidate genes *THADA, ERBB4, RAD50, C8H9orf3, YAP1, RAB5B, SUOX* and *KRR1* as shown previously [24]. *LHCGR* because of its unusual bimodal expression pattern has some unusual correlations with other genes. *ARL14EP*, is also known as *C11orf46*, encodes an effector protein that interacts with ADP-ribosylation factor-like 14 (ARL14). The ADP-ribosylation factors and ARF-like proteins are small GTP-binding proteins of the Ras superfamily involved in many intracellular signaling and vesicular trafficking pathways [65, 66]. *FDFT1* encodes farnesyl-diphosphate farnesyltransferase 1, needed

for squalene synthesis in the mevalonate pathway. Its promoter region is encompassed by the locus containing *GATA4* and *NEIL 2* [29]. *MAPRE1* encodes microtubule-associated protein RP/EB family member 1 (EB1). EB1 is involved in the formation and maintenance of microtubules and Golgi [67]. The NEIL family has three members, NEIL1, 2 and 3, and are DNA glycosylases that differ in substrate specificity, catalytic efficiency and tissue distribution and act in repair of oxidative lesions and in epigenetic demethylation [68, 69].

Some loci associated with PCOS contain more than one gene. These include *ARL14EP/FSHB, ERBB3/RAB5B, RAD50/IRF1* and *NEIL2/GATA4*. Interestingly, the expression levels of genes in the same loci were not correlated with each other. Also the promoter of *FDFT1* lies in the same locus as *GATA4* and their expression levels were not correlated with each other either. This could be the evidence of specific regulation of genes within PCOS-associated loci and not a blanket alteration in expression of genes within a locus.

In summary, most of PCOS candidate genes are not differentially expressed in adult ovaries, except perhaps *RAD50*. However, nearly all the genes in loci associated with PCOS are expressed in the developing fetal ovary and they are expressed in three patterns, early, late and across gestation. Within each group the expression levels of most of the candidate genes are correlated with each other. Thus, the general lack of differential expression of candidate genes in adult ovaries contrasting with the dynamic patterns of gene expression in fetal ovaries is consistent with a vulnerability to disturbance in the fetal ovary that may underpin later development of PCOS.

Supplementary Table S1. List of genes and primers used for qRT-PCR.

Acknowledgements

The authors would like to thank Thomas Food International, Murray Bridge, South Australia and Midfield Meat International, Warrnambool, Victoria for providing the bovine tissues for this research.

Author's roles update

ML, KH, MDH, NAB, NH, HFI-R and RJR designed the study on bovine samples. WMB provided bovine samples. ML, KH, MDH, HFI-R and RJR carried out data analysis of bovine samples. RR carried out gene expression assays and data analysis in human samples. NAB and JSEL carried out the analysis of human microarray data. ML drafted the manuscript. RAB, KH, HFI-R, JSEL, RAA and RJR critically reviewed and approved the final version of the manuscript.

Conflict of interest

RAA reports consultancy work for Ferring, Merck, IBSA, Roche Diagnostics, NeRRe Therapeutics and Sojournix Inc. JSEL reports grants and personal fees from Ferring, grants and personal fees from Titus Healthcare, grants and personal fees from Ansh Labs, during the conduct of the study. The other authors of this manuscript have nothing to declare and no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

References

1. March W, Moore V, Willson K, Phillips D, Norman R, Davies M. The prevalence of polycystic ovary syndrome in a community sample assessed under contrasting diagnostic criteria. Hum Reprod 2010; 25:544-551.
2. Bozdag G, Mumusoglu S, Zengin D, Karabulut E, Yildiz BO. The prevalence and phenotypic features of polycystic ovary syndrome: A systematic review and meta-Analysis. Human Reproduction 2016; 31:2841-2855.
3. Lizneva D, Gavrilova-Jordan L, Walker W, Azziz R. Androgen excess: Investigations and management. Best Pract Res Clin Obstet Gynaecol 2016; 37:98-118.
4. Azziz R. Polycystic Ovary Syndrome. Obstetrics & Gynecology 2018; 132:321-336.
5. Fernandez RC, Moore VM, Van Ryswyk EM, Varcoe TJ, Rodgers RJ, March WA, Moran LJ, Avery JC, McEvoy RD, Davies MJ. Sleep disturbances in women with polycystic ovary syndrome: prevalence, pathophysiology, impact and management strategies. Nat Sci Sleep. 2018; 10:45-64.:10.2147/NSS.S127475. eCollection 122018.
6. Fauser BCJM, Tarlatzis BC, Rebar RW, Legro RS, Balen AH, Lobo R, Carmina E, Chang J, Yildiz BO, Laven JSE, Boivin J, Petraglia F, et al. Consensus on women's health aspects of polycystic ovary syndrome (PCOS): the Amsterdam ESHRE/ASRM-Sponsored 3rd PCOS Consensus Workshop Group. Fertility and Sterility 2012; 97:28-38.e25.
7. Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, Piltonen T, Norman RJ. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Hum Reprod 2018; 33:1602-1618
8. Vink JM, Sadrzadeh S, Lambalk CB, Boomsma DI. Heritability of Polycystic Ovary Syndrome in a Dutch Twin-Family Study. The Journal of Clinical Endocrinology & Metabolism 2006; 91:2100-2104.
9. Kahsar-Miller MD, Nixon C, Boots LR, Go RC, Azziz R. Prevalence of polycystic ovary syndrome (PCOS) in first-degree relatives of patients with PCOS. Fertility and Sterility 2001; 75:53-58.
10. Jones MR, Goodarzi MO. Genetic determinants of polycystic ovary syndrome: progress and future directions. Fertil Steril 2016; 106:25-32.
11. Crespo RP, Bachega T, Mendonca BB, Gomes LG. An update of genetic basis of PCOS pathogenesis. Arch Endocrinol Metab. 2018; 62:352-361. doi: 310.20945/22359-3997000000049.
12. Pau CT, Mosbruger T, Saxena R, Welt CK. Phenotype and Tissue Expression as a Function of Genetic Risk in Polycystic Ovary Syndrome. PloS one 2017; 12:e0168870-e0168870.
13. Dumesic DA, Abbott DH, Padmanabhan V. Polycystic ovary syndrome and its developmental origins. Rev Endocr Metab Disord 2007; 8:127-141.
14. Padmanabhan V, Veiga-Lopez A. Animal models of the polycystic ovary syndrome phenotype. Steroids 2013; 78:734-740.
15. Walters KA. Androgens in polycystic ovary syndrome: lessons from experimental models. Curr Opin Endocrinol Diabetes Obes 2016; 23:257-263.
16. Filippou P, Homburg R. Is foetal hyperexposure to androgens a cause of PCOS? Hum Reprod Update 2017; 23:421-432.
17. Tata B, Mimouni NEH, Barbotin AL, Malone SA, Loyens A, Pigny P, Dewailly D, Catteau-Jonard S, Sundstrom-Poromaa I, Piltonen TT, Dal Bello F, Medana C, et al. Elevated prenatal anti-Mullerian hormone reprograms the fetus and induces polycystic ovary syndrome in adulthood. Nat Med. 2018; 24:834-846. doi: 810.1038/s41591-41018-40035-41595. Epub 42018 May 41514.
18. Sir-Petermann T, Maliqueo M, Angel B, Lara HE, Perez-Bravo F, Recabarren SE. Maternal serum androgens in pregnant women with polycystic ovarian syndrome: possible implications in prenatal androgenization. Hum Reprod. 2002; 17:2573-2579. doi: 2510.1093/humrep/2517.2510.2573.
19. Hatzirodos N, Bayne RA, Irving-Rodgers HF, Hummitzsch K, Sabatier L, Lee S, Bonner W, Gibson MA, Rainey WE, Carr BR, Mason HD, Reinhardt DP, et al. Linkage of regulators of TGF-beta activity in the fetal ovary to polycystic ovary syndrome. FASEB J 2011; 25:2256-2265.

20. Ewens KG, Stewart DR, Ankener W, Urbanek M, McAllister JM, Chen C, Baig KM, Parker SC, Margulies EH, Legro RS, Dunaif A, Strauss JF, 3rd, et al. Family-based analysis of candidate genes for polycystic ovary syndrome. J Clin Endocrinol Metab 2010; 95:2306-2315.
21. Stewart DR, Dombroski BA, Urbanek M, Ankener W, Ewens KG, Wood JR, Legro RS, Strauss JF, 3rd, Dunaif A, Spielman RS. Fine mapping of genetic susceptibility to polycystic ovary syndrome on chromosome 19p13.2 and tests for regulatory activity. J Clin Endocrinol Metab 2006; 91:4112-4117.
22. Hughesdon PE. Morphology and morphogenesis of the Stein-Leventhal ovary and of so-called "hyperthecosis". Obstet Gynecol Surv 1982; 37:59-77.
23. Stein IF, Leventhal ML. Amenorrhea associated with bilateral polycystic ovaries. Am J Obstet Gynecol 1935; 29:181-191.
24. Hartanti MD, Rosario R, Hummitzsch K, Bastian NA, Hatzirodos N, Bonner WM, Bayne RA, Irving-Rodgers HF, Anderson RA, Rodgers RJ. Could perturbed fetal development of the ovary contribute to the development of polycystic ovary syndrome in later life? PLOS ONE 2020; 15:e0229351.
25. Hatzirodos N, Hummitzsch K, Irving-Rodgers HF, Breen J, Perry VEA, Anderson RA, Rodgers RJ. Transcript abundance of stromal and thecal cell related genes during bovine ovarian development. PLoS One. 2019; 14:e0213575. doi: 0213510.0211371/journal.pone.0213575. eCollection 0212019.
26. Hummitzsch K, Hatzirodos N, Irving-Rodgers HF, Hartanti MD, Perry VE, Anderson RA, Rodgers RJ. Morphometic analyses and gene expression related to germ cells, GREL cells and granulosa cells during development of the bovine fetal ovary. Plos One 2019; 14:e0214130.
27. Hummitzsch K, Irving-Rodgers HF, Hatzirodos N, Bonner W, Sabatier L, Reinhardt DP, Sado Y, Ninomiya Y, Wilhelm D, Rodgers RJ. A new model of development of the mammalian ovary and follicles. PLoS One 2013; 8:e55578.
28. Hartanti MD, Hummitzsch K, Bonner WM, Bastian NA, Irving-Rodgers HF, Rodgers RJ. Formation of the bovine ovarian surface epithelium during fetal development. J Histochem Cytochem 2020; 68:113-126.
29. Day F, Karaderi T, Jones MR, Meun C, He C, Drong A, Kraft P, Lin N, Huang H, Broer L, Magi R, Saxena R, et al. Large-scale genome-wide meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for different diagnosis criteria. PLoS Genet. 2018; 14:e1007813. doi: 1007810.1001371/journal.pgen.1007813. eCollection 1002018 Dec.
30. Hayes MG, Urbanek M, Ehrmann DA, Armstrong LL, Lee JY, Sisk R, Karaderi T, Barber TM, McCarthy MI, Franks S, Lindgren CM, Welt CK, et al. Genome-wide association of polycystic ovary syndrome implicates alterations in gonadotropin secretion in European ancestry populations. Nat Commun 2015; 6:7502.
31. Jansen E, Laven JSE, Dommerholt HBR, Polman J, van Rijt C, van den Hurk C, Westland J, Mosselman S, Fauser BCJM. Abnormal gene expression profiles in human ovaries from polycystic ovary syndrome patients. Molecular Endocrinology 2004; 18:3050-3063.
32. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). Hum Reprod 2004; 19:41-47.
33. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. Fertil Steril 2004; 81:19-25.
34. Russe I. Oogenesis in cattle and sheep. Biblthca Anat 1983; 24:77-92.
35. Hartanti MD, Hummitzsch K, Irving-Rodgers HF, Bonner WM, KJ C, Anderson RA, McMillen IC, Perry VE, Rodgers RJ. Morphometric and gene expression analyses of stromal expansion during bovine fetal ovarian development. Reprod Fert Devel 2019; 31:482-495.
36. Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D. qgraph: Network Visualizations of Relationships in Psychometric Data. J Stat Softw 2012; 48:1-18.
37. Rojowska A, Lammens K, Seifert FU, Direnberger C, Feldmann H, Hopfner KP. Structure of the Rad50 DNA double-strand break repair protein in complex with DNA. EMBO J. 2014; 33:2847-2859. doi: 2810.15252/embj.201488889. Epub 201482014 Oct 201488827.
38. Ganesan S, Keating AF. Bisphenol A-Induced Ovotoxicity Involves DNA Damage Induction to Which the Ovary Mounts a Protective Response Indicated by Increased Expression of Proteins Involved in DNA Repair and Xenobiotic Biotransformation. Toxicological Sciences 2016; 152:169-180.
39. Liu X, Craig ZR. Environmentally relevant exposure to dibutyl phthalate disrupts DNA damage repair gene expression in the mouse ovary†. Biology of Reproduction 2019; 101:854-867.

40. Zhang Y, Yan Z, Qin Q, Nisenblat V, Chang HM, Yu Y, Wang T, Lu C, Yang M, Yang S, Yao Y, Zhu X, et al. Transcriptome Landscape of Human Folliculogenesis Reveals Oocyte and Granulosa Cell Interactions. Mol Cell. 2018; 72:1021-1034.e1024. doi: 1010.1016/j.molcel.2018.1010.1029. Epub 2018 Nov 1021.
41. Le S, Moore JK, Haber JE, Greider CW. RAD50 and RAD51 define two pathways that collaborate to maintain telomeres in the absence of telomerase. Genetics 1999; 152:143-152.
42. Laven JSE, Visser JA, Uitterlinden AG, Vermeij WP, Hoeijmakers JHJ. Menopause: Genome stability as new paradigm. Maturitas 2016; 92:15-23.
43. Hatzirodos N, Irving-Rodgers HF, Hummitzsch K, Rodgers RJ. Transcriptome profiling of the theca interna from bovine ovarian follicles during atresia. PLoS One 2014; 9:e99706.
44. Wiley HS. Trafficking of the ErbB receptors and its influence on signaling. Experimental Cell Research 2003; 284:78-88.
45. Richani D, Gilchrist RB. The epidermal growth factor network: role in oocyte growth, maturation and developmental competence. Hum Reprod Update 2018; 24:1-14.
46. Black LE, Longo JF, Carroll SL. Mechanisms of Receptor Tyrosine-Protein Kinase ErbB-3 (ERBB3) Action in Human Neoplasia. The American Journal of Pathology 2019; 189:1898-1912.
47. Kaushansky A, Gordus A, Budnik BA, Lane WS, Rush J, MacBeath G. System-wide Investigation of ErbB4 Reveals 19 Sites of Tyr Phosphorylation that Are Unusually Selective in Their Recruitment Properties. Chemistry & Biology 2008; 15:808-817.
48. Cho H-S, Leahy DJ. Structure of the Extracellular Region of HER3 Reveals an Interdomain Tether. Science 2002; 297:1330.
49. Roskoski R. ErbB/HER protein-tyrosine kinases: Structures and small molecule inhibitors. Pharmacological Research 2014; 87:42-59.
50. Sithanandam G, Anderson LM. The ERBB3 receptor in cancer and cancer gene therapy. Cancer Gene Therapy 2008; 15:413-448.
51. Hsieh S-Y, He J-R, Yu M-C, Lee W-C, Chen T-C, Lo S-J, Bera R, Sung C-M, Chiu C-T. Secreted ERBB3 Isoforms Are Serum Markers for Early Hepatoma in Patients with Chronic Hepatitis and Cirrhosis. Journal of Proteome Research 2011; 10:4715-4724.
52. Teves ME, Modi BP, Kulkarni R, Han AX, Marks JS, Subler MA, Windle J, Newall JM, McAllister JM, Strauss JF, 3rd. Human DENND1A.V2 Drives Cyp17a1 Expression and Androgen Production in Mouse Ovaries and Adrenals. Int J Mol Sci 2020; 21.
53. Tee MK, Speek M, Legeza B, Modi B, Teves ME, McAllister JM, Strauss JF, 3rd, Miller WL. Alternative splicing of DENND1A, a PCOS candidate gene, generates variant 2. Mol Cell Endocrinol 2016; 434:25-35.
54. Andronicos NM, Chen EI, Baik N, Bai H, Parmer CM, Kiosses WB, Kamps MP, Yates JR, III, Parmer RJ, Miles LA. Proteomics-based discovery of a novel, structurally unique, and developmentally regulated plasminogen receptor, Plg-RKT, a major regulator of cell surface plasminogen activation. Blood 2010; 115:1319-1330.
55. Miles LA, Parmer RJ. Plasminogen receptors: the first quarter century. Seminars in thrombosis and hemostasis 2013; 39:329-337.
56. Miles LA, Baik N, Bai H, Makarenkova HP, Kiosses WB, Krajewski S, Castellino FJ, Valenzuela A, Varki NM, Mueller BM, Parmer RJ. The plasminogen receptor, Plg-RKT, is essential for mammary lobuloalveolar development and lactation. Journal of Thrombosis and Haemostasis 2018; 16:919-932.
57. Beers WH. Follicular plasminogen and plasminogen activator and the effect of plasmin on ovarian follicle wall. Cell. 1975; 6:379-386. doi: 310.1016/0092-8674(1075)90187-90187.
58. Liu YX. Plasminogen activator/plasminogen activator inhibitors in ovarian physiology. Front Biosci. 2004; 9:3356-73.:10.2741/1487.
59. Strickland S, Beers WH. Studies on the role of plasminogen activator in ovulation. In vitro response of granulosa cells to gonadotropins, cyclic nucleotides, and prostaglandins. J Biol Chem. 1976; 251:5694-5702.
60. Burchall GF, Pouniotis DS, Teede HJ, Ranasinha S, Walters KA, Piva TJ. Expression of the plasminogen system in the physiological mouse ovary and in the pathological polycystic ovary syndrome (PCOS) state. Reproductive Biology and Endocrinology 2019; 17:33.

61. Honda K, Taniguchi T. IRFs: master regulators of signalling by Toll-like receptors and cytosolic pattern-recognition receptors. Nature Reviews Immunology 2006; 6:644-658.
62. Zeimet AG, Reimer D, Wolf D, Fiegl H, Concin N, Wiedemair A, Wolf AM, Rumpold H, Müller-Holzner E, Marth C. Intratumoral interferon regulatory factor (IRF)-1 but not IRF-2 is of relevance in predicting patient outcome in ovarian cancer. International Journal of Cancer 2009; 124:2353-2360.
63. Seda O, Sedova L, Vcelak J, Vankova M, Liska F, Bendlova B. ZBTB16 and metabolic syndrome: a network perspective. Physiol Res. 2017; 66:S357-S365. doi: 310.33549/physiolres.933730.
64. Koubi M, Poplineau M, Vernerey J, N'Guyen L, Tiberi G, Garciaz S, El-Kaoutari A, Maqbool MA, Andrau J-C, Guillouf C, Saurin AJ, Duprez E. Regulation of the positive transcriptional effect of PLZF through a non-canonical EZH2 activity. Nucleic acids research 2018; 46:3339-3350.
65. Pasqualato S, Renault L, Cherfils J. Arf, Arl, Arp and Sar proteins: a family of GTP-binding proteins with a structural device for 'front-back' communication. EMBO reports 2002; 3:1035-1041.
66. Takai Y, Sasaki T, Matozaki T. Small GTP-binding proteins. Physiol Rev. 2001; 81:153-208. doi: 110.1152/physrev.2001.1181.1151.1153.
67. Yang C, Wu J, de Heus C, Grigoriev I, Liv N, Yao Y, Smal I, Meijering E, Klumperman J, Qi RZ, Akhmanova A. EB1 and EB3 regulate microtubule minus end organization and Golgi morphology. Journal of Cell Biology 2017; 216:3179-3198.
68. Han D, Schomacher L, Schüle KM, Mallick M, Musheev MU, Karaulanov E, Krebs L, von Seggern A, Niehrs C. NEIL1 and NEIL2 DNA glycosylases protect neural crest development against mitochondrial oxidative stress. eLife 2019; 8:e49044.
69. Schuermann D, Weber AR, Schär P. Active DNA demethylation by DNA repair: Facts and uncertainties. DNA Repair 2016; 44:92-102.

Figure legends

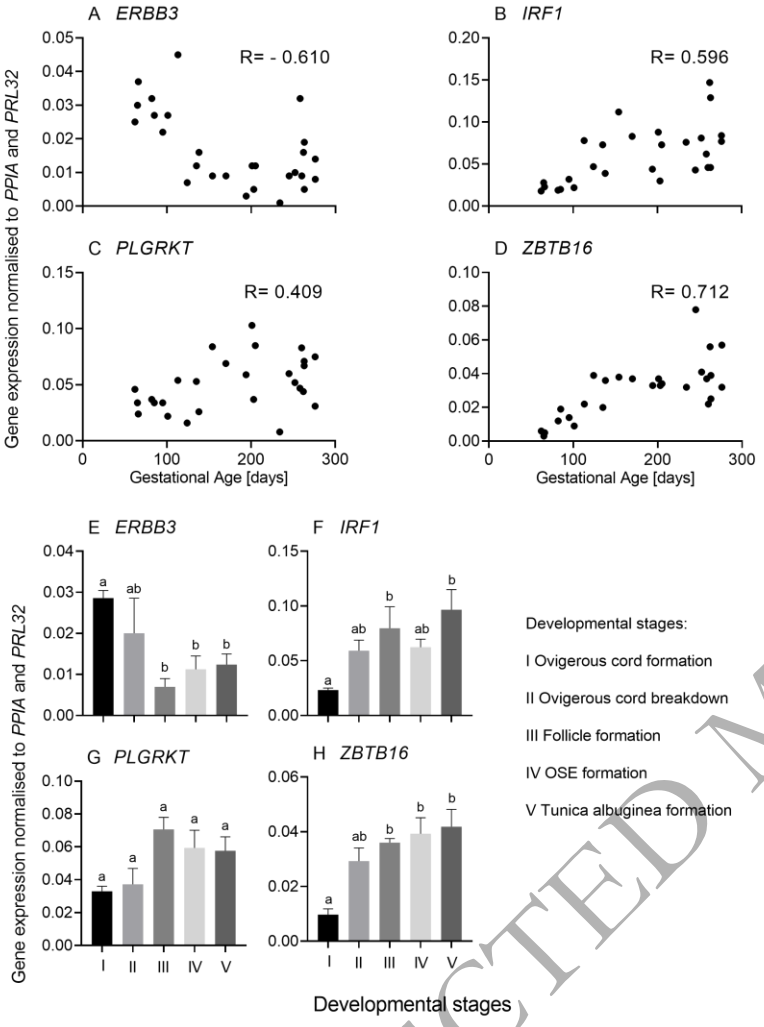

Figure 1. (A-D) Scatter plots of mRNA expression levels of PCOS candidate genes *ERBB3*, *PLGRKT*, *IRF1* and *ZBTB16* in bovine ovaries during gestation (n = 27). Pearson's correlation coefficient (R) tests were used to analyse the data and *P* values are shown in Table 2. (E-H) Differential mRNA expression levels in ovaries grouped into six stages of ovarian development based on their histological morphology: ovigerous cord formation (n = 7, Stage I), ovigerous cord breakdown (n = 4, Stage II), follicle formation (n = 3, Stage III), surface epithelium formation (n = 8, Stage IV) and tunica albuginea formation (n = 5, Stage V). Data are presented as mean ± SEM (normalized to *PPIA* and *RPL32*). One-way ANOVA with *post hoc* Holm-Sidak tests were used to analyse the data. Bars with different letters are statistically significantly different from each other ($P < 0.05$).

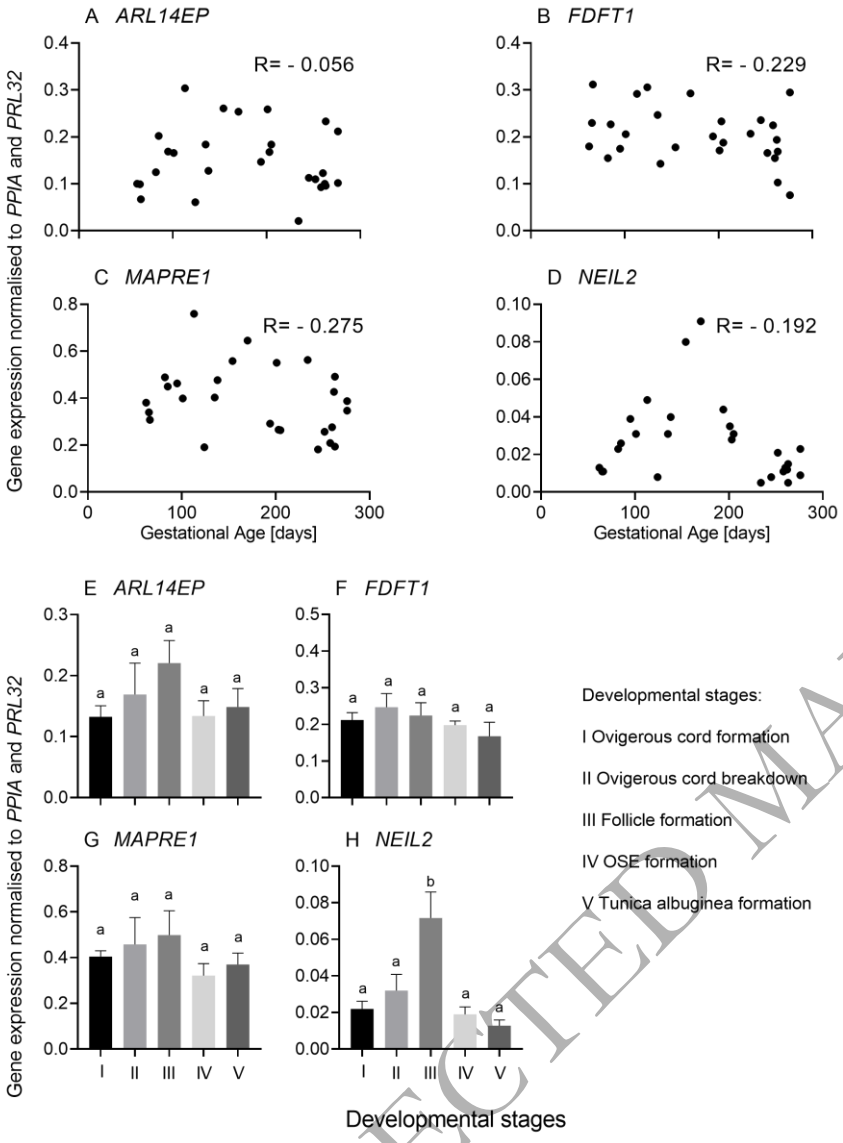

Figure 2. (A-D) Scatter plots of mRNA expression levels of PCOS candidate genes *MAPRE1*, *FDT1*, *NEIL2* and *ARL14WP* in bovine ovaries during gestation (n = 27). Pearson's correlation coefficient (R) tests were used to analyse the data and *P* values are shown in Table 2. (E-H) Differential mRNA expression levels in ovaries grouped into six stages of ovarian development based on their histological morphology: ovigerous cord formation (n = 7, Stage I), ovigerous cord breakdown (n = 4, Stage II), follicle formation (n = 3, Stage III), surface epithelium formation (n = 8, Stage IV) and tunica albuginea formation (n = 5, Stage V). Data are presented as mean ± SEM (normalized to *PPIA* and *RPL32*). One-way ANOVA with *post hoc* Holm-Sidak tests were used to analyse the data. Bars with different letters are statistically significantly different from each other ($P < 0.05$).

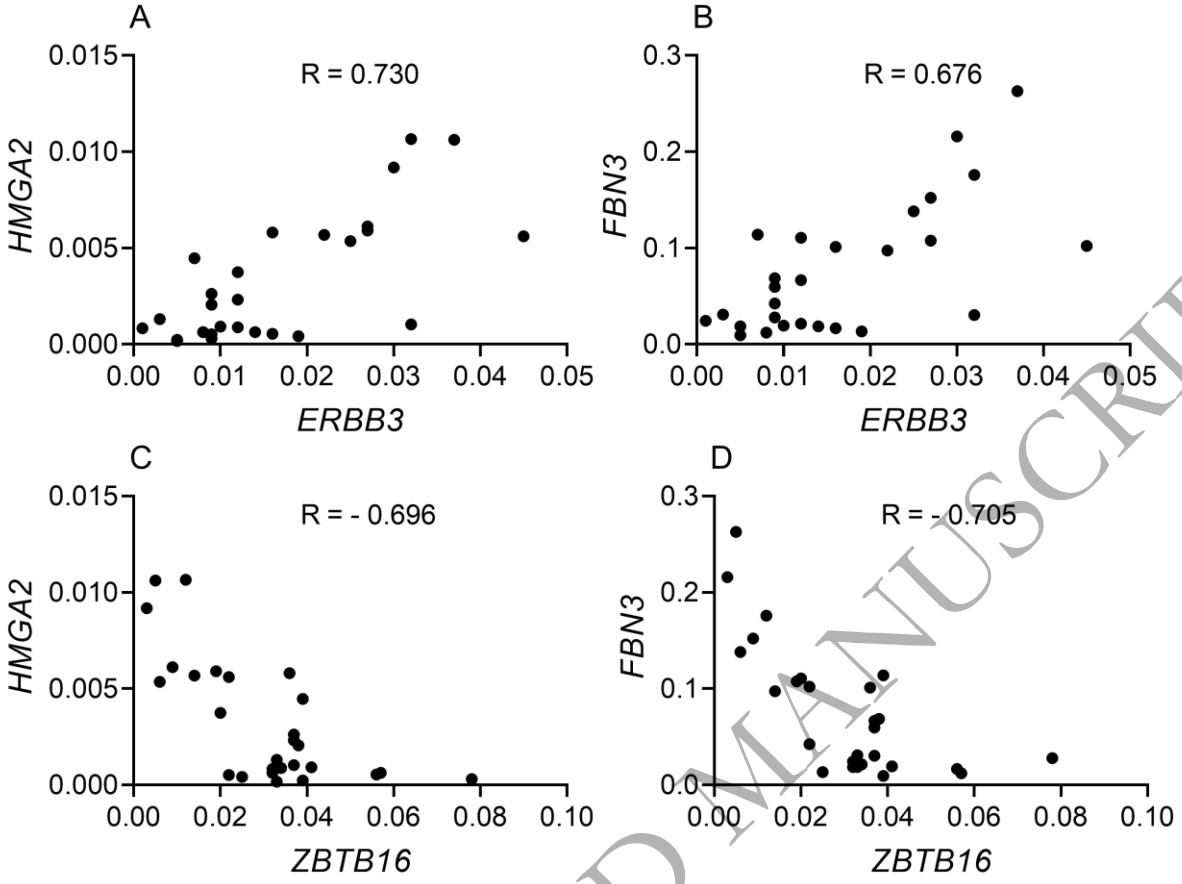

Figure 3. Scatter plots showing related mRNA expression levels of (A) *HMGA2* versus *ERBB3*, (B) *FBN3* versus *ERBB3*, (C) *HMGA2* versus *ZBTB16* and (D) *FBN3* versus *ZBTB16* in whole bovine fetal ovaries (n = 27). Data are presented as normalized gene expression to *PPIA* and *RPL32*. Pearson's correlation coefficient (R) tests were used to analyse the data.

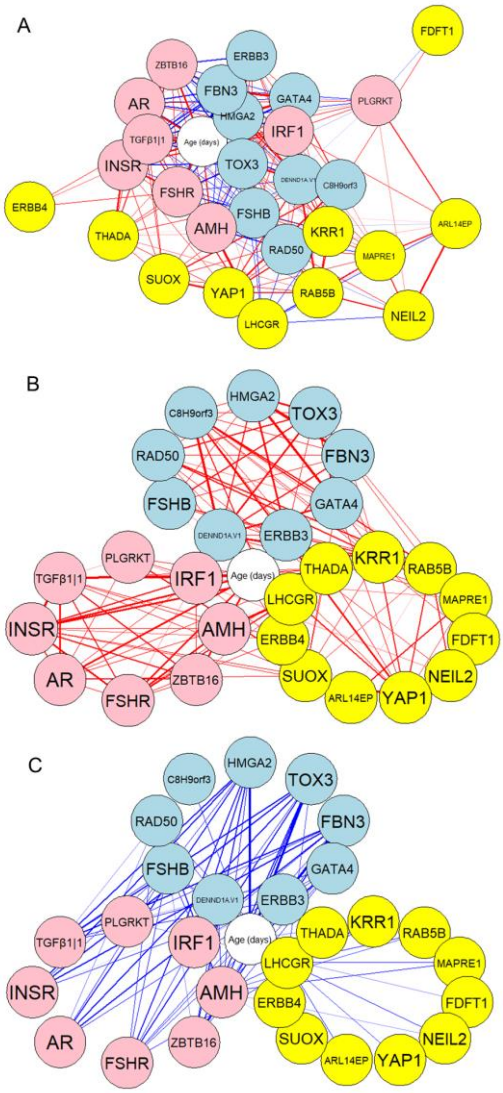

Figure 4. Gene network graphs of correlation coefficients of gene expressions and gestational age in the bovine fetal ovaries. Graphs were plotted using correlation coefficient values from Table 2. (A) Random assembly showing positive and negative associations, (B) grouped positive associations and (C) grouped negative associations were generated with qgraph package in R-program. Nodes represent the genes and age. The thickness of the interconnecting lines represents the strength of the correlations between genes and gestational ages. Red and blue lines represent positive and negative correlations, respectively, whereas blue, yellow and pink nodes represent the genes expressed early, throughout or late in gestation. Age is gestational age.

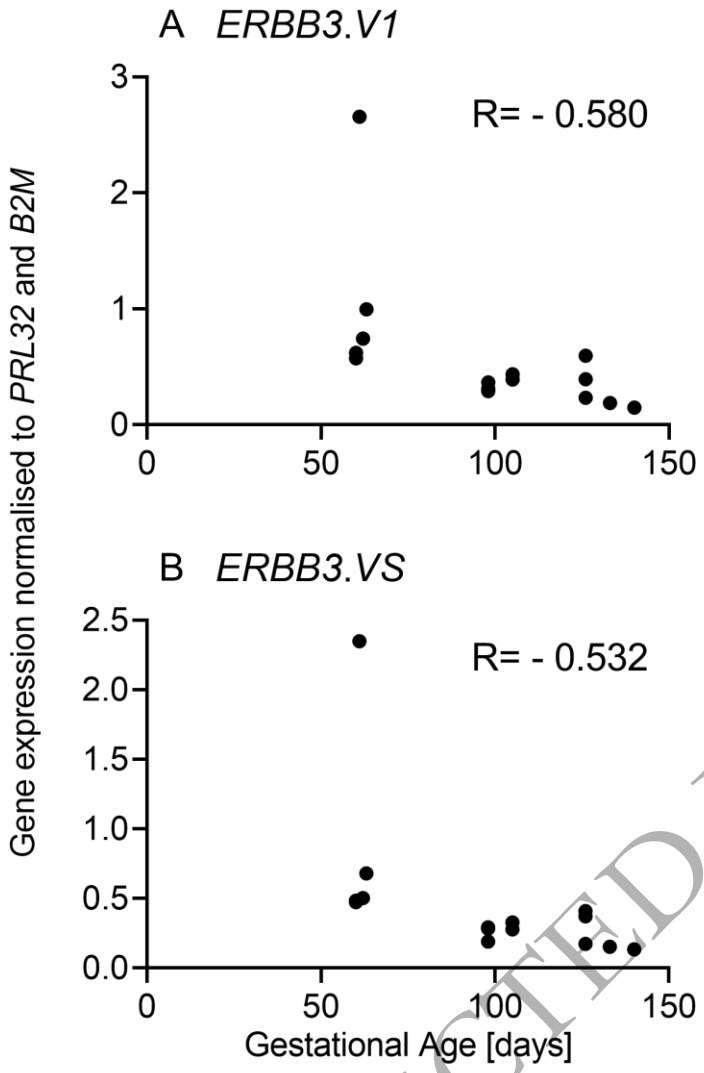

Figure 5. Scatter plot of mRNA expression levels of PCOS candidate gene *ERBB3V1* (A) and *ERBB3VS* (B) in human fetal ovaries during gestation (n = 15, 8 – 20 weeks of gestation). Data are presented as normalized gene expression to *RPL32* and *B2M*. Pearson's correlation coefficient (R) tests were used to analyse the data.

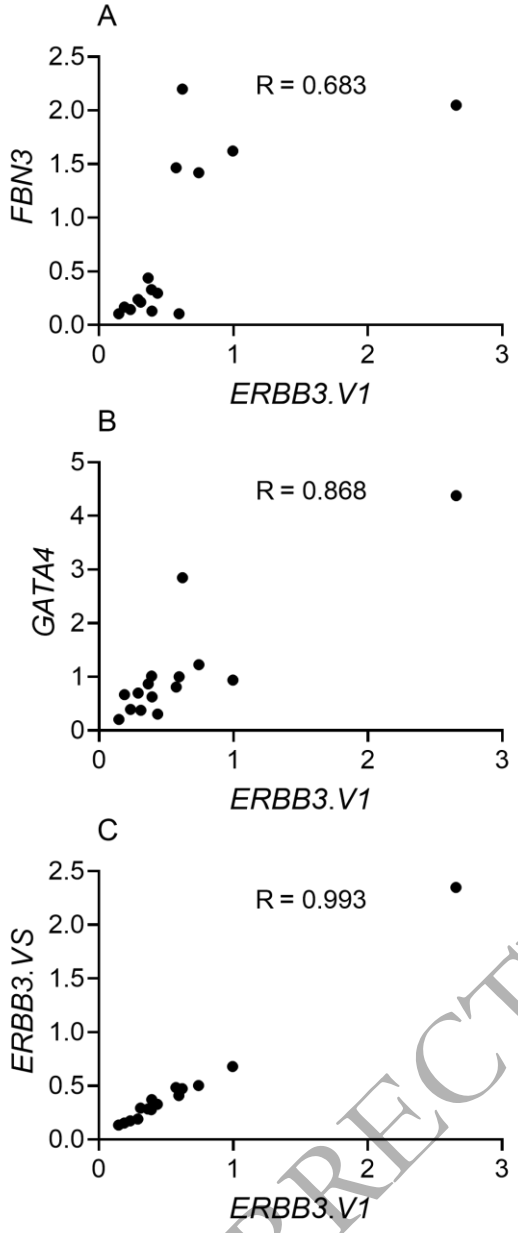

Figure 6. Scatter plots showing related mRNA expression levels of (A) *FBN3* versus *ERBB3.V1*, (B) *GATA4* versus *ERBB3.V1* and (C) *ERBB3.VS* versus *ERBB3.V1* in human fetal ovaries during gestation (n = 15, 8 – 20 weeks of gestation). Data are presented as normalized gene expression to *RPL32* and *B2M*. Pearson's correlation coefficient (R) test was used to analyse the data.

Table 1. Mean ± SEM signal intensities from Affymetrix gene arrays of PCOS-candidate genes in interstitium from control (n = 8) and PCOS ovaries (n = 8) reported from a previous study [31].

Gene Array	Sequence Code	Gene Name	Gene Symbol	Control	PCOS
HGU 133A	208025_s_at	high mobility group AT-hook 2	*HMGA2*	86.6 ± 5.2	91.8 ± 6.1
HGU 133A	214774_x_at	TOX high mobility group box family member 3	*TOX3*	56.5 ± 4.6	58.8 ± 3.0
HGU 133A	215108_x_at	TOX high mobility group box family member 3	*TOX3*	50.3 ± 4.5	50.7 ± 3.5
HGU 133A	216623_x_at	TOX high mobility group box family member 3	*TOX3*	51.8 ± 4.4	54.2 ± 4.4
HGU 133B	230745_s_at	TOX high mobility group box family member 3	*TOX3*	69.5 ± 5.2	58.7 ± 4.0
HGU 133B	240117_at	fibrillin 3	*FBN3*	181.4 ± 12.2	168.2 ± 13.1
HGU 133A	205517_at	GATA binding protein 4	*GATA4*	453.4 ± 62.4	533.8 ± 31.9
HGU 133B	243692_at	GATA binding protein 4	*GATA4*	211.0 ± 21.9	177.5 ± 5.6
HGU 133B	230855_at	GATA binding protein 4	*GATA4*	52.7 ± 1.2	59.6 ± 2.0
HGU 133A	202454_s_at	erb-b2 receptor tyrosine kinase 3	*ERBB3*	151.0 ± 46.6	63.7 ± 5.0
HGU 133A	215638_at	erb-b2 receptor tyrosine kinase 3	*ERBB3*	118.6 ± 7.6	104.3 ± 8.4
HGU 133B	226213_at	erb-b2 receptor tyrosine kinase 3	*ERBB3*	235.3 ± 53.1	178.1 ± 22.7
HGU 133A	219763_at	DENN/MADD domain containing 1A	*DENND1A*	65.3 ± 1.9	63.5 ± 3.4
HGU 133B	226849_at	DENN/MADD domain containing 1A	*DENND1A*	182.1 ± 10.3	185.6 ± 11.4
HGU 133A	208393_s_at	RAD50 homolog, double strand break repair protein	*RAD50*	187.1 ± 4.5	204.8 ± 6.8
HGU 133A	209349_at	RAD50 homolog, double strand break repair protein	*RAD50*	143.7 ± 13.0	150.9 ± 9.5
HGU 133B	238656_at	RAD50 homolog, double strand break repair protein	*RAD50*	115.9 ± 6.5	142.9 ± 6.7*
HGU 133A	203202_at	KRR1, small subunit (SSU) processome component, homolog (yeast)	*KRR1*	134.6 ± 5.9	155.7 ± 12.5
HGU 133A	203203_s_at	KRR1, small subunit (SSU) processome component, homolog (yeast)	*KRR1*	83.2 ± 7.2	90.1 ± 8.3
HGU	232441_at	KRR1, small subunit	*KRR1*	59.6 ± 2.1	73.1 ± 4.8

133B		(SSU) processome component, homolog (yeast)			
HGU 133B	233515_at	KRR1, small subunit (SSU) processome component, homolog (yeast)	*KRR1*	65.6 ± 1.5	70.5 ± 2.9
HGU 133B	235038_at	KRR1, small subunit (SSU) processome component, homolog (yeast)	*KRR1*	56.5 ± 2.6	60.5 ± 2.2
HGU 133B	243930_x_at	KRR1, small subunit (SSU) processome component, homolog (yeast)	*KRR1*	71.1 ± 2.6	72.0 ± 5.6
HGU 133A	201276_at	RAB5B, member RAS oncogene family	*RAB5B*	497.9 ± 21.5	484.5 ± 11.9
HGU 133A	200712_s_at	microtubule-associated protein, RP/EB family, member 1	*MAPRE1*	166.8 ± 11.9	182.7 ± 12.0
HGU 133A	200713_s_at	microtubule-associated protein, RP/EB family, member 1	*MAPRE1*	792.4 ± 29.5	836.5 ± 31.8
HGU 133A	208647_at	farnesyl-diphosphate farnesyltransferase 1	*FDFT1*	799.5 ± 78.2	797.5 ± 52.2
HGU 133A	210950_s_at	farnesyl-diphosphate farnesyltransferase 1	*FDFT1*	731.2 ± 72.6	704.1 ± 45.8
HGU 133B	226585_at	nei-like DNA glycosylase 2	*NEIL2*	126.1 ± 5.4	123.4 ± 4.7
HGU 133A	213342_at	Yes-associated protein 1	*YAP1*	162.8 ± 7.8	168.4 ± 12.5
HGU 133B	224894_at	Yes-associated protein 1	*YAP1*	966.4 ± 64.0	1238.6 ± 109.4
HGU 133B	224895_at	Yes-associated protein 1	*YAP1*	887.3 ± 95.9	1114.9 ± 91.6
HGU 133A	204067_at	sulfite oxidase	*SUOX*	304.7 ± 14.7	268.0 ± 11.6
HGU 133A	206794_at	erb-b2 receptor tyrosine kinase 4	*ERBB4*	45.5 ± 2.6	39.7 ± 1.6
HGU 133A	214053_at	erb-b2 receptor tyrosine kinase 4	*ERBB4*	275.1 ± 134.2	101.7 ± 7.1
HGU 133B	233494_at	erb-b2 receptor tyrosine kinase 4	*ERBB4*	67.3 ± 2.0	72.7 ± 3.1
HGU 133B	233498_at	erb-b2 receptor tyrosine kinase 4	*ERBB4*	149.0 ± 47.7	88.4 ± 2.8
HGU 133A	207240_s_at	luteinizing hormone/choriogonadotropin receptor	*LHCGR*	163.5 ± 61.4	128.0 ± 19.9
HGU 133A	220212_s_at	thyroid adenoma associated	*THADA*	201.8 ± 9.2	188.5 ± 4.2
HGU	54632_at	thyroid adenoma	*THADA*	187.4 ± 11.6	184.7 ± 6.6

133A		associated			
HGU 133A	218992_at	plasminogen receptor, C-terminal lysine transmembrane protein	*PLGRKT*	268.9 ± 38.2	222.9 ± 5.6
HGU 133A	202531_at	interferon regulatory factor 1	*IRF1*	197.7 ± 10.5	172.0 ± 6.9
HGU 133B	238725_at	interferon regulatory factor 1	*IRF1*	117.5 ± 2.6	114.0 ± 5.0
HGU 133A	206516_at	anti-Mullerian hormone	*AMH*	120.9 ± 31.4	124.3 ± 12.1
HGU 133A	205883_at	zinc finger and BTB domain containing 16	*ZBTB16*	393.5 ± 41.9	436.0 ± 82.0
HGU 133A	211201_at	follicle stimulating hormone receptor	*FSHR*	159.9 ± 8.6	160.7 ± 10.7
HGU 133A	211110_s_at	androgen receptor	*AR*	155.9 ± 12.8	158.6 ± 13.9
HGU 133A	211621_at	androgen receptor	*AR*	205.0 ± 34.8	191.5 ± 17.6
HGU 133B	226192_at	androgen receptor	*AR*	809.2 ± 129.9	667.8 ± 24.1
HGU 133B	226197_at	androgen receptor	*AR*	559.1 ± 62.7	522.7 ± 20.6
HGU 133A	207851_s_at	insulin receptor	*INSR*	89.9 ± 11.1	81.6 ± 7.2
HGU 133A	213792_s_at	insulin receptor	*INSR*	961.8 ± 75.0	752.2 ± 84.8
HGU 133B	226212_s_at	insulin receptor	*INSR*	132.1 ± 5.1	134.4 ± 8.2
HGU 133B	226216_at	insulin receptor	*INSR*	209.9 ± 24.5	194.9 ± 15.3
HGU 133B	226450_at	insulin receptor	*INSR*	389.9 ± 19.8	328.5 ± 19.1
HGU 133A	209651_at	transforming growth factor beta 1 induced transcript 1	*TGFB1I1*	1072.0 ± 93.6	1071.0 ± 81.8

* Significantly different from control, $P < 0.05$.

Table 2. Pearson correlation coefficients (R) of mRNA expression levels of PCOS-candidate genes and gestational age in bovine fetal ovaries (n = 27). Data from Hartanti *et al.* [24] are reproduced with permission and combined with new data on *ERBB3*, *KRR1*, *MAPRE1*, *FDFT1*, *NEIL2*, *ARL14EP*, *PLGRKT* and *ZBTB16.* The intensity of the background color indicates the strength of the significance of the correlation. Blue indicates negative correlations and greenindicates positive correlations.

	Age (days)	*HMGA2*	*TOX3*	*FBN3*	*GATA4*	*ERBB3*	*DENND1A.V1*	*FSHB**	*RAD50*	*C8H9orf3*	*KRR1*	*RAB5B*	*MAPRE1*	*FDFT1*	*NEIL2*	*YAP1*	*ARL14EP*	*SUOX*	*ERBB4*	*LHCGR*	*THADA*	*PLGRKT*	*IRF1*	*AMH*	*ZBTB16*	*FSHR*	*AR*	*INSR*
HMGA2	-0.898[d]																										
TOX3	-0.892[d]	0.826[d]																									
FBN3	-0.887[d]	0.959[d]	0.824[d]																								
GATA4	-0.709[d]	0.849[d]	0.700[d]	0.898[d]																							
ERBB3	-0.610[c]	0.730[d]	0.564[b]	0.676[c]	0.586[b]																						
DENND1A.X1,2,3,4	-0.606[c]	0.675[d]	0.731[b]	0.696[d]	0.777[d]	0.494[b]																					
*FSHB**	-0.563[b]	0.443[b]	0.679[c]	0.406[a]	0.182	0.424[a]	0.347																				
RAD50	-0.497[b]	0.434[b]	0.601[c]	0.406[a]	0.359	0.219	0.526[b]	0.563[b]																			
C8H9orf3	-0.393[a]	0.506[b]	0.511[b]	0.543[b]	0.671[c]	0.431[a]	0.841[d]	0.295	0.430[a]																		
KRR1	-0.352	0.463[a]	0.515[b]	0.379	0.368	0.442[a]	0.663[c]	0.505[b]	0.591[b]	0.586[b]																	
RAB5B	-0.314	0.258	0.485[a]	0.246	0.340	0.200	0.714[d]	0.471[a]	0.690[d]	0.644[c]	0.600[c]																
MAPRE1	-0.275	0.23	0.39	0.12	0.17	0.27	0.470[a]	0.39	0.52	0.354	0.35	0.59															

		3	8[a]	7	3	5		6[a]	8[b]		7	3[b]																
FDFT1	-0.229	0.195	0.208	0.276	0.264	0.121	0.234	0.137	0.243	0.396[a]	-0.034	0.223	-0.036															
NEIL2	-0.192	0.007	0.292	-0.038	-0.091	-0.046	0.228	0.281	0.576[b]	0.122	0.214	0.598[c]	0.617[c]	0.031														
YAP1	-0.080	0.217	0.225	0.236	0.491[b]	0.089	0.638[c]	0.064	0.490[b]	0.716[d]	0.590[b]	0.692[d]	0.386[a]	0.073	0.174													
ARL14EP	-0.056	-0.087	0.178	-0.139	-0.200	0.166	0.067	0.317	0.330	0.001	0.188	0.326	0.620[c]	-0.154	0.718[d]	0.063												
SUOX	0.175	0.083	-0.037	0.140	0.403[a]	0.182	0.505[b]	-0.250	0.054	0.639[c]	0.430[a]	0.392[a]	-0.020	0.045	-0.130	0.722[d]	-0.150											
ERBB4	0.305	-0.256	-0.164	-0.240	-0.162	-0.274	0.010	0.049	0.186	0.045	-0.110	0.278	0.203	-0.053	0.130	0.232	0.001	0.073										
LHCGR	0.332	-0.177	-0.382[a]	-0.156	-0.079	-0.121	-0.311	-0.403[a]	-0.577[b]	-0.256	-0.478[a]	-0.459[a]	-0.446[a]	-0.070	-0.451[a]	-0.387[a]	-0.414[a]	0.073	0.162									
THADA	0.380	-0.231	-0.268	-0.264	0.001	-0.165	0.217	-0.176	0.059	0.435[a]	0.334	0.359	0.194	0.039	0.052	0.557[b]	-0.069	0.585[b]	0.225	0.177								
PLGRKT	0.409[a]	-0.467[a]	-0.331	-0.438[a]	-0.338	-0.222	-0.244	-0.321	-0.260	-0.276	-0.240	-0.032	0.080	-0.347	0.349	-0.015	0.612[c]	0.070	0.089	0.135	0.040							
IRF1	0.596[b]	-0.589[b]	-0.403[a]	-0.569[b]	-0.388[a]	-0.336	-0.044	-0.208	-0.160	0.043	0.043	0.219	0.360	-0.132	0.313	0.283	0.428[a]	0.281	0.117	-0.169	0.472[a]	0.499[b]						
AMH	0.682[d]	-0.542[b]	-0.632[c]	-0.525[c]	-0.408[a]	-0.339	-0.476[a]	-0.427[a]	-0.619[c]	-0.311	-0.438[a]	-0.374	-0.439[a]	-0.129	-0.341	-0.306	-0.322	0.153	0.242	0.846[d]	0.313	0.256	0.168					
ZBTB16	0.712[d]	-0.696[d]	-0.630[c]	-0.705[d]	-0.549[b]	-0.611[c]	-0.420[a]	-0.347	-0.283	-0.203	-0.260	-0.064	-0.247	0.152	0.000	-0.033	-0.043	0.088	0.198	0.159	0.386[a]	0.246	0.431[a]	0.401[a]				

FSHR	0.753[d]	-0.640[d]	-0.684[d]	-0.580[b]	-0.436[a]	-0.577[b]	-0.392[a]	-0.531[b]	-0.391[a]	-0.203	-0.432[a]	-0.196	-0.299	-0.004	-0.142	-0.062	-0.216	0.256	0.451[a]	0.675[c]	0.457[a]	0.305	0.347	0.818[d]	0.557[b]			
AR	0.765[d]	-0.654[d]	-0.648[c]	-0.669[c]	-0.420[a]	-0.321	-0.258	-0.452[a]	-0.358	-0.028	-0.040	-0.047	-0.195	-0.092	-0.101	0.137	-0.012	0.420[a]	0.055	0.234	0.491[b]	0.420[a]	0.510[b]	0.568[b]	0.621[c]	0.476[a]		
INSR	0.775[d]	-0.643[d]	-0.632[c]	-0.601[c]	-0.312	-0.449[a]	-0.143	-0.435[a]	-0.283	0.167	-0.070	0.116	-0.071	-0.034	-0.067	0.392[a]	-0.099	0.572[b]	0.369	0.248	0.756[d]	0.272	0.661[c]	0.573[b]	0.636[c]	0.729[d]	0.696[d]	
TGFB1I1	0.853[d]	-0.695[d]	-0.676[c]	-0.664[c]	-0.413[a]	-0.434[a]	-0.188	-0.444[a]	-0.352	0.034	-0.032	0.024	-0.161	-0.176	-0.189	0.292	-0.081	0.518[b]	0.432[a]	0.231	0.584[b]	0.336	0.608[c]	0.537[b]	0.661[c]	0.669[c]	0.772[d]	0.876[d]

[a] $P < 0.05$, [b] $P < 0.01$, [c] $P < 0.001$, [d] $P < 0.0001$; Pearson correlation tests.

*Primer pair from exon 3.

Table 3. Pearson correlation coefficients (R) of mRNA expression levels of PCOS-candidate genes in human fetal ovaries and their gestational ages (n = 15). Data from Hartanti *et al*. [24] are reproduced with permission and combined with new data on *ERBB3.V1* and *ERBB3.VS*. The intensity of the background color indicates the strength of the significance of the correlation. Blue indicates negative correlations and green indicates positive correlations.

	Age	*FBN3*	*HMGA2*	*DENND1A.V1,3,4*	*DENND1A*	*GATA4*	*ERBB3.V1*	*FSHR*	*ERBB3.VS*	*LHCGR*
FBN3	-0.906[d]									
HMGA2	-0.685[b]	0.852[d]								
DENND1A.V1,3,4	-0.674[b]	0.790[c]	0.904[d]							
DENND1A.V1-7	-0.595[a]	0.697[b]	0.864[d]	0.983[d]						
GATA4	-0.593[a]	0.767[c]	0.934[d]	0.808[c]	0.782[c]					
ERBB3.V1	-0.580[a]	0.683[c]	0.827[c]	0.590[a]	0.570[a]	0.868[d]				
FSHR	-0.545[a]	0.575[a]	0.606[a]	0.421	0.445	0.680[b]	0.796[c]			
ERBB3VS	-0.532[a]	0.644[b]	0.811[c]	0.558[a]	0.548[a]	0.868[d]	0.993[d]	0.823[c]		
LHCGR	-0.491	0.421	0.287	0.357	0.291	0.020	0.191	-0.033	0.100	
TOX3	-0.033	-0.134	-0.115	0.232	0.302	-0.078	-0.174	-0.185	-0.211	0.081

[a] $P < 0.05$, [b] $P < 0.01$, [c] $P < 0.001$, [d] $P < 0.0001$; Pearson correlation test.