REPRODUCTION
RESEARCH

In vivo gene expression in granulosa cells during pig terminal follicular development

A Bonnet, K A Lê Cao[1,3], M SanCristobal, F Benne, C Robert-Granié[1], G Law-So, S Fabre[2], P Besse[3], E De Billy, H Quesnel[4], F Hatey and G Tosser-Klopp

INRA, UMR 444, Génétique Cellulaire, F-31326 Castanet-Tolosan Cedex, France, [1]INRA, UR, Station d'Amélioration Génétique des Animaux, F-31326 Castanet-Tolosan Cedex, France, [2]INRA, UMR6175, Physiologie de la Reproduction et des Comportements, CNRS, Université de Tours, Haras Nationaux, F-37380 Nouzilly, France, [3]Université Paul Sabatier, UMR 5219, Institut de Mathématiques, F-31062 Toulouse Cedex 9, France and [4]INRA, UMR INRA/AgroCampus Rennes: Systèmes d'Elevage, Nutrition Animale et Humaine, F-35590 Saint Gilles, France

Correspondence should be addressed to A Bonnet; Email: agnes.bonnet@toulouse.inra.fr

Abstract

Ovarian antral follicular development is clearly dependent on pituitary gonadotrophins FSH and LH. Although the endocrine mechanism that controls ovarian folliculogenesis leading to ovulation is quite well understood, the detailed mechanisms and molecular determinants in the different follicular compartments remain to be clarified. The aim of this study was to identify the genes differentially expressed in pig granulosa cells along the terminal ovarian follicle growth, to gain a comprehensive view of these molecular mechanisms. First, we developed a specific micro-array using cDNAs from suppression subtractive hybridization libraries (345 contigs) obtained by comparison of three follicle size classes: small, medium and large antral healthy follicles. In a second step, a transcriptomic analysis using cDNA probes from these three follicle classes identified 79 differentially expressed transcripts along the terminal follicular growth and 26 predictive genes of size classes. The differential expression of 18 genes has been controlled using real-time PCR experiments validating the micro-array analysis. Finally, the integration of the data using Ingenuity Pathways Analysis identified five gene networks providing descriptive elements of the terminal follicular development. Specifically, we observed: (1) the down-expression of ribosomal protein genes, (2) the genes involved in lipid metabolism and (3) the down-expression of cell morphology and ion-binding genes. In conclusion, this study gives new insight into the gene expression during pig terminal follicular growth *in vivo* and suggested, in particular, a morphological change in pig granulosa cells accompanying terminal follicular growth.

Reproduction (2008) **136** 211–224

Introduction

The growth and development of ovarian follicles leading to ovulation require a series of coordinated events that lead to follicular somatic cell differentiation and oocyte development. It includes two successive periods, preantral (primordial, primary and secondary follicles) and antral (early antral, antral and preovulatory follicles) follicular developments. Follicles undergo morphological and functional changes as they progress towards ovulation. Among all growing follicles, only a small proportion of them (less than 1%) will ovulate. Most of them undergo a degenerative process called atresia that occurs at the different developmental stages. The antral follicular development depends on a complex regulatory network with endocrine regulation by pituitary gonadotrophins follicle-stimulating hormone (FSH) and luteinizing hormone (LH) but also with autocrine and paracrine pathways (Hsueh 1986) including the action of steroids and peptides (Hillier & Miro 1993, Drummond 2006). Those factors control follicular growth either directly or indirectly: for example, bone morphogenetic protein or the insulin-like growth factor (IGF) systems (Monget *et al*. 2002, Mazerbourg *et al*. 2003, Shimasaki *et al*. 2004) modify the sensitivity of follicular cells to FSH, and epidermal growth factor signalling network potentializes LH action (Hattori *et al*. 1995). Follicular development is thus a complex process and requires the coordinate expression of a large number of genes. The mechanisms underlying this development have been intensively studied mainly in granulosa cells because they constitute an important compartment in the mammalian ovarian follicle. They actively participate in the endocrine function of the ovaries by secreting oestradiol or progesterone under FSH or LH stimulation (Duda 1997).

In pigs, *in vitro* experiments or *in situ* hybridization have allowed the individual description of gene expression in granulosa cells during the growth of antral follicles. They referred mainly to the role of FSH, the IGF

ISSN 1470–1626 (paper) 1741–7899 (online)
DOI: 10.1530/REP-07-0312
Online version via www.reproduction-online.org

system and growth factors on gene expression regulation, and the expression of genes involved in steroidogenesis such as *CYP19A* (Chan & Tan 1987) and *STAR* (Balasubramanian *et al*. 1997). Despite continuing progress in the area of ovarian biology, many of the specific mechanisms involved in follicular development, including the initiation of primordial follicle growth, antrum formation, follicular growth/selection and follicular atresia, remain to be elucidated in greater detail. Then, global approaches have been undertaken to identify new genes involved in antral follicle maturation. Our previous data obtained on porcine granulosa cells *in vitro* by suppression subtractive hybridization (SSH) or differential display PCR suggested a role of FSH in extracellular matrix synthesis, chromatin remodelling, regulation of transcription activity and protection against atresia (Clouscard-Martinato *et al*. 1998, Bonnet *et al*. 2006*b*). Moreover, different cDNA libraries obtained from whole follicles of different size classes were generated to create a catalogue of differentially expressed genes along antral follicle development using sequence frequency in each library (Jiang *et al*. 2004). However, despite the development of DNA micro-array techniques in the pig (Tuggle *et al*. 2007), only few transcriptomic data are available concerning the ovarian function.

Thus, using the pig as a model and a transcriptomic approach, the aim of this study was to identify differentially expressed genes along terminal ovarian follicular development, before the onset of preovulatory LH surge. The experiment has focused especially on the time of expression of functional LH receptors in granulosa cells that occur in 4–5 mm follicle size and classically considered as a maturation marker in follicle development (May & Schomberg 1984). Then, we developed first a micro-array using cDNAs from SSH libraries obtained from granulosa cells isolated from three follicle size classes, small (1–2 mm), medium (3–4 mm) and large antral ($\geq$5 mm). In a second step, a transcriptomic analysis using cDNA probes from these three follicle classes identified transcripts differentially expressed along the terminal follicular growth and some of the gene networks associated with this process.

Results

Dedicated porcine micro-array tool construction

Our first goal was to construct a reliable tool to identify genes differentially expressed during terminal follicular development in porcine ovary by a global approach. Four SSH libraries were constructed (cf. Materials and Methods section) and screened leading to the selection of 1697 clones with the best 'differential' potential between small (SF), medium (MF) and large antral follicles (LF). These clones were sequenced, resulting in 1378 good quality sequences from which 441 sequences (35%) corresponded to the whole insert (presence of vector flanking sequences). The average insert sequence length was 640 bp.

All sequences were deposited in EMBL public database (accession numbers: CR939025–CR940296; CT971504–CT971571; CT990474–CT990531) and submitted to an assembly process (Sigenae contig assembly; Pig V3 p.sc.3, 30/01/2006). Our 1378 sequences were part of 345 contigs, 63 of which were new sequences. The contig analysis revealed a global redundancy of 75%. This redundancy was explained by the high proportion of sequences coming from only three genes, *CYP19A*, *GSTA* and *PGFS1* representing 15, 13.5 and 10% of the sequences respectively. There was no contig overlap between the forward and reverse libraries for SF versus LF comparison and around 10–18% for SF versus MF comparison.

The PCR products of the 1697 clones (345 genes/contigs) resulting from SSH experiments were spotted onto nylon membrane along with 1056 already sequenced PCR products (954 genes/contigs) to generate our micro-array platform (GEO accession number GPL3978).

Micro-array analysis

In order to identify genes whose expression differs along the terminal growth and maturation of porcine ovarian follicles, 14 complex cDNA radiolabelled probes synthesized from the three different size class follicles were hybridized onto our specific cDNA micro-array (1275 genes/contigs). Data (GEO accession number GSE5798 dataset) were pre-processed resulting in a list of 1564 expressed cDNAs, corresponding to 515 different genes or contigs of sequences assembled by SIGENAE.

A mixed linear model was applied to the 1564 cDNAs to quantify the hybridization signal intensity measured for each clone and for each complex probe in the function of follicle size class. The mean expression level for all genes was not statistically different between the follicle classes ($P=0.17$). By contrast, some genes had a significantly different expression level than others ($P<0.001$) and the significance of the gene–follicle class interaction ($P<0.001$) indicated some gene differential expression according to the follicle size classes. The biological variation represented 12% of the total variation while the variation between complex probe replicates and experimental variation corresponded respectively to 2 and 3% of the global variation.

The selection of significant differentially expressed genes was made through a *F*-test followed by a false discovery rate (FDR) adjustment on the *P* values (F analysis). The *P* values obtained after the FDR adjustment were very similar to the raw *P* values of the *F*-test, indicating a very low proportion of false positive (643 adjusted *P* value with FDR versus 705 raw F-statistics *P* values <0.002.). At the 0.2% level of significance, F analysis selected 643 cDNAs

corresponding to 75 known genes and 4 unknown contigs (Tables 1 and 2). The unsupervised hierarchical clustering shows that these 643 cDNAs separated the three follicle size classes in only two groups: LF versus MF and SF (Fig. 1). Among the 79 genes/contigs, 25 were overexpressed and 54 were down-expressed in LF group compared with SF/MF. However, when the expression of each selected gene was studied individually, five different expression profiles were observed (Tables 1 and 2), illustrated by the expression of glutathione S-transferase-α (*GSTA1*), *CYP19A*, *TUBA1B*, *EEF1A* and stathmin 1 (*STMN1*) (Fig. 2). Interestingly, numerous genes whose expression increased during the terminal follicular growth were implicated in glutathione metabolism (*GSTA1*, *GSTA2* and *MGST1*) and lipid metabolism (*CYP19A*, *AKR1C3*, *AKR1C4*, *HADHB*, *BDH2*, *CYB5*, *NR5A2* and *RTF1*). By contrast, the terminal follicular growth was notably accompanied with decreased expression of genes implicated in protein translation (16 subunits of ribosomal proteins, *EEF1A*), ion binding (calumenin (*CALU*), *SLC40A1*, calmodulin (*CALM1*) and *S100A11*) and cell shape (*TUBA1B*, *TUB5B*, *TUB7*, *VIM*, *CAPNS1*, *COF1*, smoothelin (*SMTN*), *STMN1*, *RPSA* and *DAG1*).

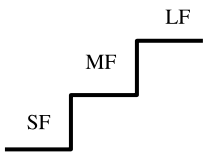
In order to identify a set of predictive genes that could help classify the follicles in their respective class, we have performed the random forest (RF) algorithm. The internal estimation of the generalization error was between 4.76 and 7.14% depending on the forests. A selection of the 120 most important and stable cDNAs with the Mean Decrease Gini importance measure gave an unsupervised hierarchical clustering allowing the separation of the three follicle size classes (Fig. 3). These 120 cDNAs corresponded to 24 known genes and two contigs and included 20 genes already selected by F analysis at the 0.2% level of significance (Tables 1 and 2). Among the six genes selected only by RF analysis, five were found below the 5% level of significance by F analysis. Only the differential expression of the *SFT2D2* gene was considered as non-significant ($P > 0.05$) by F analysis.

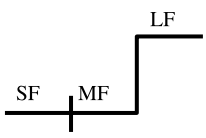
Differential expression validation by quantitative real-time PCR and in situ *hybridization*

In order to validate the micro-array analysis by quantitative real-time PCR, 18 differentially expressed genes selected by F and/or RF analysis were analysed (Table 3). Apart the *MT-CO1* gene, the differential

Table 1 Summary of significant overexpressed genes during pig follicular development.

HUGO name symbol	Gene description	FDR-adjusted *P* value	RF importance ($\times 10^{-3}$)	Maximum fold change	Expression profile
HADHB	Hydroxyacyl-coenzyme A dehydrogenase/ 3-ketoacyl-coenzyme A thiolase/enoyl-coenzyme A hydratase, β-subunit	‡	49.5	3.34	
PSMC2	Proteasome (prosome, macropain) 26S subunit, ATPase, 2	‡	47.4	2.73	
GSTA1*	Glutathione S-transferase A1	‡	222	2.58	
CTSL	Cathepsin L	‡		2.38	
HSPA8*	Heat shock 70 kDa protein 8	‡		2.38	
MGST1*	Microsomal glutathione S-transferase 1	‡		2.19	
GSTA2	Glutathione S-transferase A2	‡		3.01	
ERP29	Endoplasmic reticulum protein 29	‡	52.5	2.99	
TYB9*	Thymosin β9	‡		2.59	
CYP19A*	Cytochrome P450 19A3	‡	139	2.52	
NR5A2*	Nuclear receptor subfamily 5, group A, member 2	‡	55.1	2.45	
GART	Phosphoribosylglycinamide formyltransferase	‡		2.44	
AKR1C4*	Aldo–keto reductase family 1, member C4	‡		2.29	
AKR1C3	Aldo–keto reductase family 1, member C3	‡		2.26	
BX926910.1.p.sc.3*		‡	43.5	2.21	
HSPE1*	Heat shock 10 kDa protein 1	‡		1.98	
TFPI2	Tissue factor pathway inhibitor 2	‡		1.92	
DDX3X	DEAD (Asp–Glu–Ala–Asp) box polypeptide 3	‡		1.9	
CFL2	Cofilin 2	‡		1.84	
HNRPU	Heterogeneous nuclear ribonucleoprotein U	‡		1.82	
RFT1	Putative endoplasmic reticulum multispan transmembrane protein	‡		1.77	
BQ597365.1.p.sc.3		‡		1.56	
PSMD12	Proteasome (prosome, macropain) 26S subunit, non-ATPase, 12	‡		1.53	
CYB5	Cytochrome b-5	‡		1.52	
CCT1	t-Complex 1	‡		1.51	





HUGO name symbol column: *identifies genes with several clones present in the selection. FDR-adjusted *P* value column: ‡$P < 0.002$. Fold change corresponds to the higher expression of the three classes versus the lowest.

Table 2 Summary of significant under-expressed genes during pig follicular development.

HUGO name symbol	Gene description	FDR-adjusted *P* value	RF importance	Maximum fold change	Expression profile
TUBA1B	Tubulin, α-1b	‡	195	2.16	
CALU	Calumenin	‡	46.5	1.98	
CB287006.1.p.sc.3		‡		1.80	
SMTN	Smoothelin	‡	80.5	1.72	
SLC40A1	Solute carrier family 40 (iron-regulated transporter), member 1	‡		1.67	
PKM2	Pyruvate kinase, muscle	‡	73.2	1.63	
CALM1, CALM2	Calmodulin 1	‡	80.1	1.51	
MT-CO1	Mitochondrially encoded cytochrome c oxidase I	*	43.7	1.61*	
CFL1	Cofilin 1	*	40.7	1.39*	
GPRC5C	G protein-coupled receptor, family C, group 5, member C	*	38	1.37*	
SFT2D2	SFT2 domain containing 2	NS	38	1.37*	
FOLR2	Folate receptor 2	†	36.1	1.35*	
GSTO1	Glutathione S-transferase omega 1	*	45.3	1.29*	
RPLP1*	Ribosomal protein, large, P1	‡		2.51	
RPS26*	Ribosomal protein S26	‡	63.2	2.47	
RPS17*	Ribosomal protein S17	‡		2.37	
EEF1A*	Eukaryotic translation elongation factor 1-α1	‡	168	2.34	
RPLP0*	Ribosomal protein, large, P0	‡	38.3	2.20	
RPL37A	Ribosomal protein L37a	‡	137	2.14	
HMGB1	High-mobility group box 1	‡		2.11	
ARL4C	ADP-ribosylation factor-like 4C	‡		2.06	
DAG1	Dystroglycan 1	‡		1.98	
TUBB5	Tubulin, β5	‡		1.97	
GNB2L1	Guanine nucleotide-binding protein (G protein), β-polypeptide 2-like 1	‡		1.94	
TB7*	Tubulin β-7 chain	‡	37.7	1.86	
ITM2A*	Integral membrane protein 2A	‡	38	1.85	
CF179049.1.p.sc.3*		‡	40.2	1.80	
RPS5	Ribosomal protein S5	‡		1.73	
H2AFZ*	H2A histone family, member Z	‡		1.62	
RPS6	Ribosomal protein S6	‡		1.50	
VIM*	Vimentin	‡		2.81	
CAPNS1	Calpain, small subunit 1	‡		2.50	
RPS25*	Ribosomal protein S25	‡		2.43	
SOX4	SRY (sex determining region Y)-box 4	‡		2.39	
RPL11	Ribosomal protein L11	‡		2.36	
STMN1*	Stathmin 1/oncoprotein 18	‡	39.2	2.33	
RPS12*	Ribosomal protein S12	‡		2.23	
RPSA*	Laminin receptor 1	‡		2.21	
BTG2	β-Cell translocation gene 2	‡		2.03	
TMSB10*	Thymosin, β10	‡		2.01	
HIST1H2AC	Histone cluster 1, H2ac	‡		1.91	
GPX3*	Glutathione peroxidase 3	‡		1.89	
EGR1*	Early growth response 1	‡		1.87	
CLTB	Clathrin, light chain	‡		1.86	
RPS8	Ribosomal protein S8	‡		1.85	
BDH2	3-Hydroxybutyrate dehydrogenase, type 2	‡		1.84	
ENTPD1*	Ectonucleoside triphosphate diphosphohydrolase 1	‡		1.82	
PPARG*	PPARG*	‡		1.77	
RPL34*	Ribosomal protein L34, leukemia-associated protein	‡		1.77	
C15ORF21	Chromosome 15 open reading frame 21	‡		1.73	
RPS7	Ribosomal protein S7	‡		1.72	
PABPC1*	Poly(A)-binding protein, cytoplasmic 1	‡		1.71	
MRPL49	Mitochondrial ribosomal protein L49	‡		1.68	
RPL3	Ribosomal protein L3	‡		1.65	
IGFBP2	Insulin-like growth factor-binding protein 2	‡		1.64	
S100A11	S100 calcium-binding protein A11	‡		1.57	
CTGF	Connective tissue growth factor	‡		1.52	
RANBP1	RAN-binding protein 1	‡		1.51	
RPL9	Ribosomal protein L9	‡		1.49	
NONO	Non-POU domain containing, octamer binding	‡		1.44	

HUGO name symbol column: *identifies genes with several clones present in the selection. FDR-adjusted *P* value column: *$P<0.05$; †$P<0.01$; ‡$P<0.002$. Fold change corresponds to the higher expression out of the three classes versus the lowest: *corresponds to genes selected by RF but not with F-analysis.

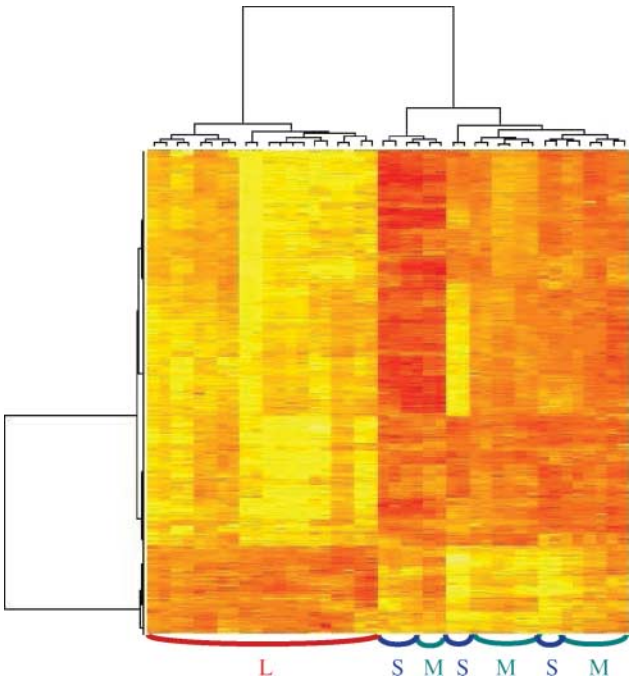


Figure 1 Heat map display of unsupervised hierarchical clustering of 643 cDNAs selected with F-analysis. The cDNAs are displayed in lines and micro-arrays in columns. The yellow colour represents over-expressed cDNAs and red down-expressed cDNAs. L corresponds to large follicles, M to medium follicles and S to small follicles.

expression of all tested genes was coherent when comparing the two approaches. Moreover, the cytochrome P450 11A (*CYP11A*) and the STAR protein (*STAR*) genes, not present on our micro-arrays and whose expression was already known to increase during the terminal development of the follicle (Hatey *et al*. 1992, Conley *et al*. 1994, LaVoie *et al*. 1997) gave the expected results.

GSTA1 gene expression, first ranked (score 222) with the RF analysis as a predictive marker for follicle size class with a relatively high expression fold change (2.58) between LF and SF, was examined by *in situ* hybridization (Fig. 4). *GSTA* mRNA was strongly detected in theca interna cells of small healthy antral follicles (~1 mm in diameter) and not detectable in granulosa cells (Fig. 4A). By contrast, in large follicles (>5 mm), *GSTA* mRNA was strongly detected in granulosa cells (Fig. 4C).

Data integration

Grouping of differentially expressed genes into pathways was achieved by utilizing the Ingenuity Pathway Analysis (IPA). Among the 79 genes/contigs selected by F analysis, 64 genes were taken into account by IPA tool for generating networks. Five highly significant biological networks (score >16: score of 2 have at least 99% confidence of not being generated by chance) encompassing three top biological functions were identified and summarized in Table 4. The first network highlighted the down-expression of ribosomal protein genes during the follicular development as illustrated in Fig. 5. Networks 2–5 (25 focus genes) illustrated changes in the glutathione and lipid metabolism pathways (Supplementary Figure 1, which can be viewed online at www.reproduction-online.org/supplemental). Networks 3–4 (27 focus genes) were related to cellular growth and differentiation. As part of this latter network, we have particularly focused on genes related to the control of cell shape. Using the IPA tool, a specific cell shape network has been constructed. As shown in Fig. 6, this network targeted on the potential role of the actin gene. Thus, in the way to estimate granulosa cell morphological change during terminal follicular development, we have performed actin staining with FITC-conjugated phalloidin on cryosections of pig ovaries (Fig. 7). First, the observation of different sections identified a granulosa cell structure with a huge nucleus and little cytoplasm. Then, we observed different cell shapes between granulosa cells of small–medium follicles that

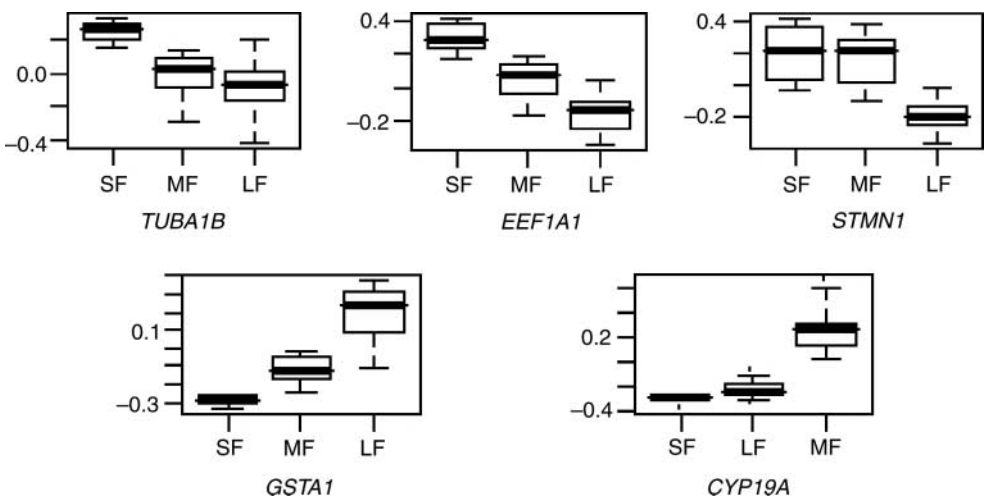


Figure 2 The different gene expression profiles. Y, log scale of gene expression.

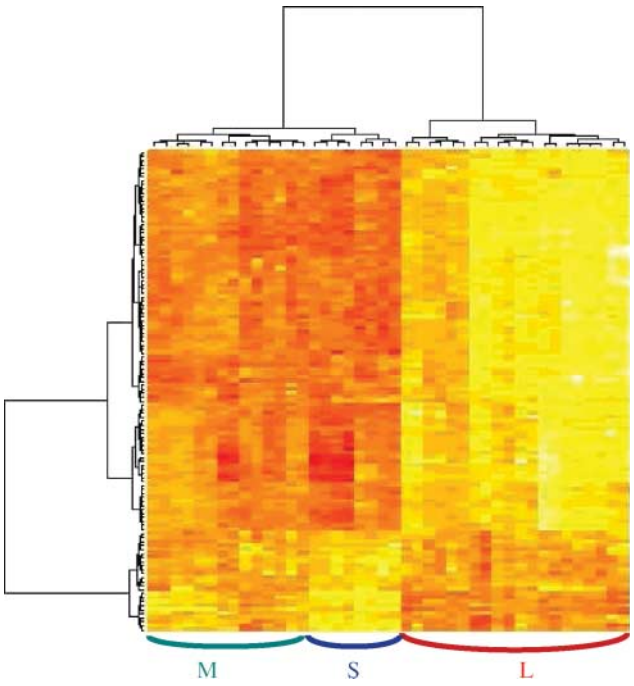


Figure 3 Heat map display of unsupervised hierarchical clustering of 120 cDNAs selected with RFs.

displayed well-defined cell shape and constituted a regular honeycomb network and granulosa cells of large ones, these had less defined borders and an elongated shape. These observations suggested a different organization of actin filaments between two types of granulosa cells. Moreover, the micro-array analysis of *actin* gene expression exhibited no differential expression.

Discussion

The objective of the present study was to identify differentially expressed genes in granulosa cells during terminal follicular development in pigs, and underline gene networks associated with this process. To reach this objective, we have developed a dedicated cDNA micro-array using SSH strategy, further hybridized with RNA probes coming from granulosa cells picked up at different steps of the terminal follicular development (SF, MF and LF).

In order to create a specific porcine micro-array, we applied SSH strategy to obtain enriched cDNA libraries with genes overexpressed in SF compared with LF and MF, or in MF and LF compared with SF. Despite the already described redundancy (Bonnet *et al.* 2006*a*), the sequence data showed a good enrichment with few overlaps between the forward and reverse libraries. For example, there is no overlap between SF/lf and LF/sf libraries, and only 11 genes/contigs (over 170) between SF/mf and MF/sf libraries. Thus, these libraries allowed the design of a valuable micro-array dedicated to the terminal follicular growth in pig species.

After micro-array hybridization and data acquisition, the statistical analyses were performed to (1) identify differentially expressed transcripts using F analysis and (2) select a set of genes that can discriminate the three follicle classes using gene prediction (RFs).

When comparing the *P* values with F analysis and T statistics, we observed a very low proportion of false positive, probably due to the pre-selection by the SSH strategy. The set of 79 selected cDNAs by F analysis

Table 3 Results of qPCR analysis.

		RT-PCR regulation			Micro-array regulation	
Way of selection	**Gene name**	Best comparison	*P* value	Fold change*	*P* value	Fold change
F/RF	BX926910.1.p.sc.3	L/S	*	328.66	‡	2.21
F/RF	NR5A2	L/S	†	5.93	‡	2.45
F/RF	GSTA	L/S	†	5.16	‡	2.58
F	HSPE1	L/S	NS	3.56	‡	1.98
RF	MT-CO1	M/S	NS	1.62	*	−1.61
F	CYB5	M/S	*	1.58	‡	1.52
F	CAPNS1	M/S	NS	−1.47	‡	−2.50
F/RF	TUBA	L/S	NS	−1.67	‡	−2.16
F	RPS5	L/S	‡	−2.15	‡	−1.73
F/RF	RPLP0	L/S	‡	−2.53	‡	−2.20
F/RF	CF179049.1.p.sc.3	L/S	*	−2.90	‡	−1.80
F	SOX4	L/M	NS	−3.25	‡	−2.39
F	HMGB1	L/M	*	−3.34	‡	−2.11
F/RF	STMN1	L/M	‡	−5.86	‡	−2.33
F	GPX3	L/M	†	−6.83	‡	−1.89
F/RF	SMTN	L/S	*	−7.42	‡	−1.72
F/RF	ITM2A	L/M	†	−16.06	‡	−1.85
F	PPARG	L/S	†	−20.04	‡	−1.77
Control +	CYP11A1	L/S	‡	14.02		
Control +	STAR	L/S	†	70.05		

Fold change corresponds to the highest expression by the lowest in relation to the best comparison. As a convention, a minus sign was added for down-regulation during development. *$P<0.05$; †$P<0.01$; ‡$P<0.002$.

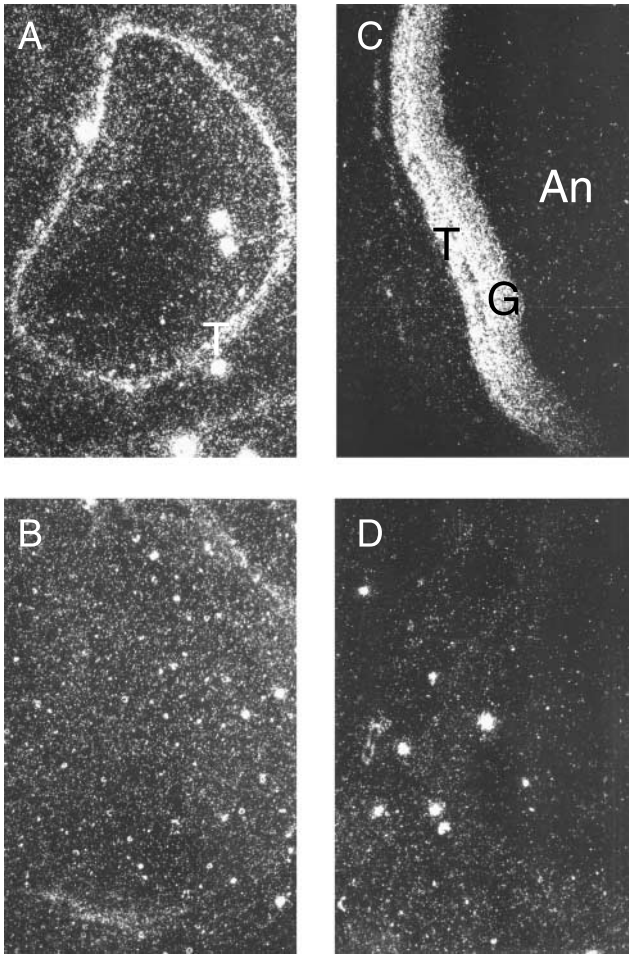


Figure 4 *In situ* hybridization of *GSTA* mRNA. Darkfield views (10× magnification) showing specific (anti-sense probe: A and C) and no specific (sense probe: B and D) hybridizations in cryosections of saw follicles at different stages of development. (A and B) Correspond to small antral follicle (<1 mm). (C and D) Correspond to large antral follicle (5 mm). T, theca cells; G, granulosa cells; An, antrum.

(FDR <0.2%) mostly favoured a hierarchical classification in two clusters, sorting out the LF from the MF and SF (Fig. 1). This observation evidences that main molecular changes occur in pig granulosa cells during the medium to large follicle transition. Interestingly, this is correlated to the appearance of functional LH receptors in granulosa cells (May & Schomberg 1984).

Despite the similarity between SF and MF classes, we tried to discriminate them using an alternative statistical method. RFs were applied to select a predictive set of genes that can classify the follicles in their respective classes (Diaz-Uriarte & Alvarez de Andres 2006, Lê Cao *et al*. 2006). The main advantages of this method are that it can deal with a massive number of correlated input variables (and hence take into account dependencies between cDNAs) and it can also select features using an internal variable importance measure. A weighting procedure is also available to deal with unbalanced classes (biological×technical replicates). RFs selected 26 genes/contigs that could predict the three classes (LF versus MF versus SF, Fig. 3). F and RF selections shared common genes (20) but they did not answer the same biological question: F analysis selected differentially expressed transcripts between one of the three follicle classes, whereas RFs selected transcripts that can predict any of the three classes preferably all together and help classify the follicles.

In order to complete these global analyses, we combined them with the individual study of each selected gene, which revealed five expression profiles (Fig. 2). These expression profiles confirmed that the main significant differences in gene expression during terminal follicular development occurred between the medium and large follicles. Indeed, 49 out of the 79 genes/contigs showed a differential expression only in large follicles (Tables 1 and 2). In addition, 23 genes/contigs exhibited a gradual expression change from SF to LF. Finally, the expression of only 13 genes/contigs decreased during SF to MF transition.

To validate micro-array analysis, gene expression was checked using real-time PCR analysis and *in situ* hybridization. The real-time PCR results were globally in agreement with the micro-array data analyses (Table 3). Discrepancies concern only five genes whose *P* values were considered as non-significant. Among them, only *MT-CO1* displays an opposite regulation between the two approaches, whereas the other four exhibited the same tendency as the one found in micro-array experiments. Non-significance may be attributed to the high variability of gene expression checked by QPCR. A higher number of tested samples would probably confirm the micro-array data for these four genes.

Using IPA, we identified major networks involved in ribosomal protein synthesis, lipid metabolism and cellular growth and proliferation (Table 4). The first network brought together the genes coding for ribosomal proteins and one translational elongation factor (*EEF1A1*), all down-expressed during terminal development (Fig. 5). This network was the most significant (score 26) and included 16 out of the 79 differentially expressed genes. It assumed a decrease in protein synthesis and may be associated with a decreased cellular growth rate. This is in agreement with previous studies describing a decrease in the percentage of proliferating granulosa cells during the final stages of follicular development in pigs and other species (Hirshfield 1986, Fricke *et al*. 1996, Pisselet *et al*. 2000). Our results suggest that the molecular mechanisms leading to granulosa cell proliferation arrest in LF occur as early as the SF to MF transition, as attested by the decreased expression of seven ribosomal genes in MF (Table 2). In addition, this network also suggested the role of the *MYC* family members known to induce ribosomal gene transcription

Table 4 Summary of functional networks for genes selected by F analysis.

Id	Genes[a]	Score	Focus genes	Top functions
1	ANXA2, CDKN2A, ↓**EEF1A1**, EPO, FOS, HRAS, IGF1R, IL3, ITGA2B, JRK, KITLG (includes EG:4254), MDM2 (includes EG:4193), MYC, MYCN, NCL, NGFB, ↓**RPL3**, ↓**RPL9**, ↓**RPL11**, ↓**RPL34**, ↓**RPL37A**, ↓**RPLP1**, RPLP2, ↓**RPLP0** (includes EG:6175), ↓**RPS5**, ↓**RPS6**, ↓**RPS7**, ↓**RPS8**, ↓**RPS12**, ↓**RPS25**, ↓**RPS26**, RPS17 (includes EG:6218), STAU1, TFAP2A, YWHAZ	26	16	Protein synthesis, cancer, cell cycle
2	AMH, ANGPTL4, AQP4, AQP7, BBC3, BCKDHA, ↑**CFL2**, ↑**CYB5A**, CYCS, ↑**DAG1**, E2F1, ↓**ENTPD1**, ↑**ERP29**, F2, ↓**GPX3**, ↓**HIST1H2AC**, ↑**HSPE1**, ↓**IGFBP2**, LAMA2, LAMB3, MAPK13, MDK, ↑**MGST1**, PAPPA, PLA2G4A, ↓**RANBP1**, ↓**RPSA**, SIM1, ↓**SOX4**, ↓**STMN1**, TF, ↑**TFPI2**, TG, THBD, TNF	24	15	Lipid metabolism, molecular transport, small molecule biochemistry
3	ABCB1B, AR, ↓**CALU** (includes EG:813), ↓**CAPNS1**, CCKAR, CIDEC, CSDE1, ↓**CTGF**, ↑**DDX3X**, ↓**EGR1**, ↓**GNB2L1**, ↑**HNRPU**, HSPH1, IFNG, MYC, NCOA4 (includes EG:8031), ↓**NONO**, ↓**PABPC1**, PCBP2, ↓**PPARG**, PRDX2, ↑**PSMC2**, PSMC4, PSMC5, PSMD3, PSMD5, PSMD7, PSMD8, PSMD9, ↑**PSMD12**, PSMD13, ↓**S100A11**, ↓**SMTN**, TERT, WT1	22	14	Cellular growth and proliferation organ development, reproductive system development and function
4	AFP, AGRIN, ↓**ARL4C**, BID, ↓**BTG2**, CCNA1, CDKN2D, CHGB, CP, CST3, ↑**CTSL**, ↑**CYP19A1**, DUSP4, ↓**EEF1A1**, ELN, F12, ↑**GART**, ↓**H2AFZ**, ↑**HADHB**, HAMP, IL6, ↓**ITM2A**, KLKB1, KRAS, LMNA, NGFB, PDIA3, ↓**RPSA**, ↓**SLC40A1**, ↑**TFPI2**, THBD, ↓**TMSB10**, TP53, TPT1, VEGF	20	13	Cellular growth and proliferation organ, genetic disorder, metabolic disease
5	ABCB11, ABCB1B, ACSL1, AFP, AKR1B7, ↑**AKR1C4**, AP3B1, AP3B2, BAAT, CETP, CLTA, ↓**CLTB**, CLTC, ↑**CYP19A1**, CYP21A2, CYP7B1, CYP8B1, ↑**GSTA1**, ↑**GSTA2**, GSTM2, GSTP1 (includes EG:2950), ↓**HMGB1**, HMGB2, HNF4A, HNF4G, ↑**HSPA8**, INS1, MTTP, ↑**NR5A2**, PDK1, ↓**PKM2**, PRLR, SP1, UGT1A9 (includes EG:54600), ↓**VIM**	14	10	Lipid metabolism, molecular transport, small molecule biochemistry

[a]Up and down arrows represent up-regulated and down-regulated genes respectively.

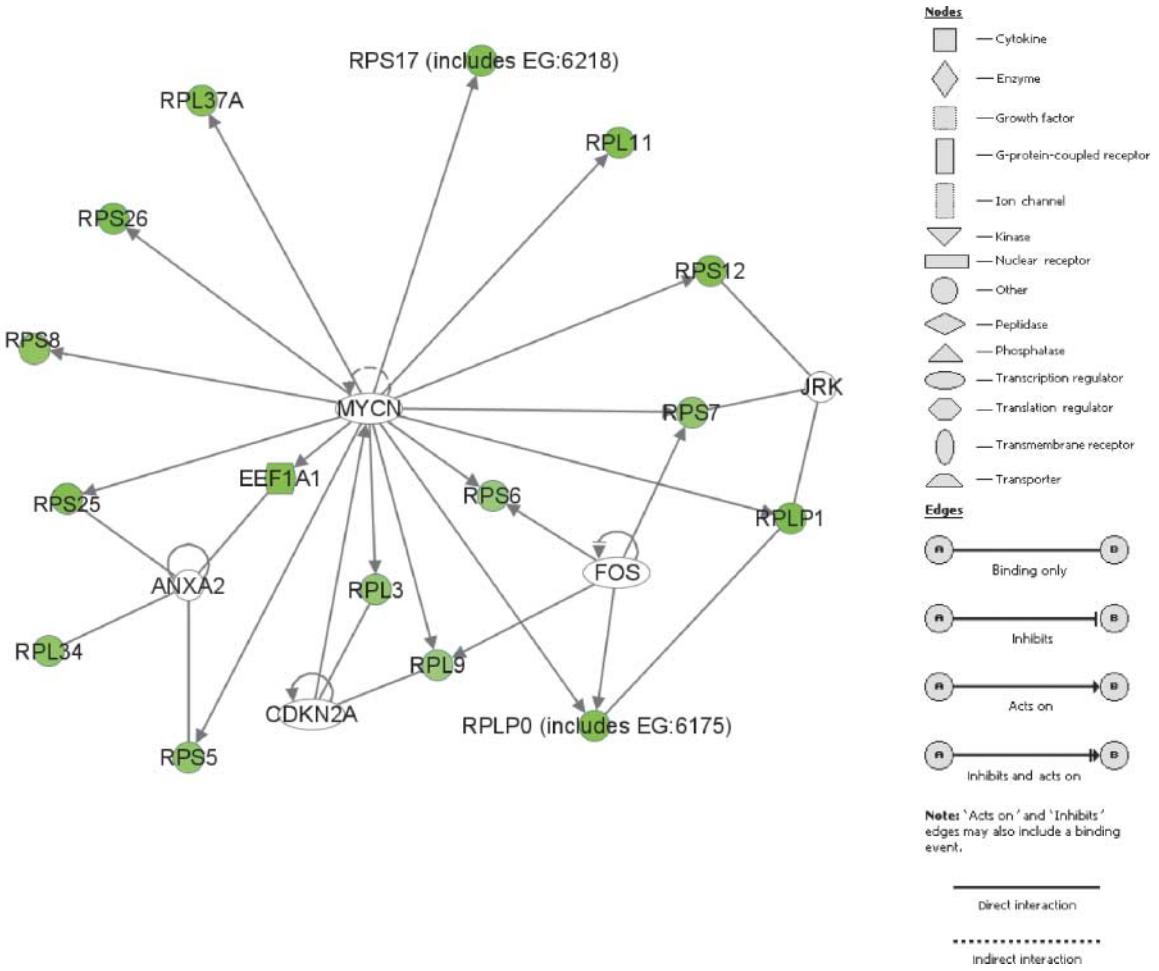


Figure 5 Network 1. Under expression of ribosomal protein genes (green colour) during follicular development.

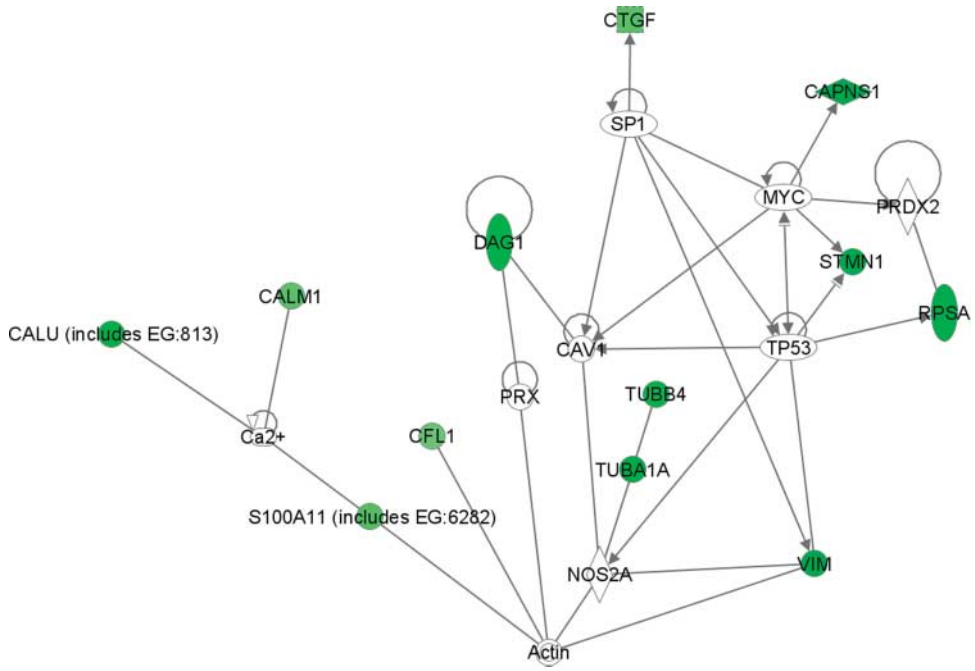


Figure 6 Granulosa cell morphology gene network.

(Boon *et al*. 2001) which needs further investigations in the context of differentiated granulosa cells.

Networks 2 and 5 bring out relations between genes implicated in lipid metabolism (Supplementary Figure 1). In the context of granulosa cells, lipid metabolism is closely related to steroids synthesis, produced *de novo* from cholesterol. We observed an overexpression of different hydroxylase genes as *HADHB*, cytochromes (*CYP19A*, *CYB5*) and *NR5A2*, a factor known to play an important role in the activation of transcription of steroidogenic enzymes like cytochromes (Sirianni *et al*. 2002) and emerging as an important ovarian factor in regulating female reproduction (Saxena *et al*. 2007, Zhao *et al*. 2007). These concomitant overexpressions occurring mainly in LF may favour an increase in steroid synthesis and were in agreement with an increase of oestradiol concentration in follicular fluid during terminal development (Foxcroft & Hunter 1985). This network also underlines a detoxification mechanism that is a consequence of steroid synthesis and allows the transformation of metabolism residues like the lipid hydroperoxydes. Our study highlighted the overexpression of different GST genes (*GSTA1*, *GSTA2* and *MGST*) in LF granulosa cells. This has already been described in steroidogenically active cells (Keira *et al*. 1994, Rabahi *et al*. 1999). In our study, *GSTA1 in situ* experiments fitted in very well with the micro-array analysis for this gene. We underlined also the down-regulation of glutathione peroxidase 3 (*GPX3*; Table 1). The KEGG (http://www.genome.jp/kegg) glutathione pathway (data not shown) suggests two different activities for these enzymes with a specific role of detoxification for GPX3 (catalyses the reduction of peroxides as lipid hydroperoxides (LOOHs)) and an intracellular transport proteins or steroids sequestration function for GSTA. To our knowledge, this is the first evidence of *GPX3* regulation in ovarian cells. Finally, the network 2 highlighted the overexpression of two members of the aldo–keto

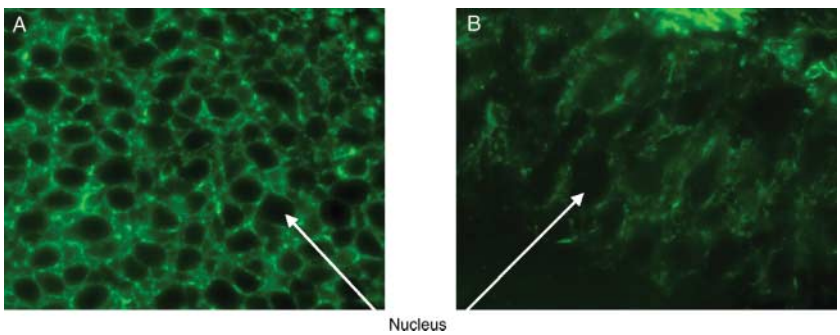


Figure 7 Cell shape. Cell shape was evaluated by actin staining with FITC-conjugated phalloidin on cryosections of pig ovaries including follicles at different stages of development. Representative microscopical field (100× magnification) of granulosa cells from (A) medium follicule, granulosa cells showed a honeycomb shape and (B) large follicle, granulosa cells showed an elongated shape.

reductase superfamily (*AKR1C3* and *AKR1C4*) able to catalyse the conversion of a wide variety of substrates such as aldehydes generated by phospholipid metabolism (Vergnes *et al*. 2003). Moreover, aldo–keto reductases have the ability to modulate the levels of active androgens, oestrogens and progestins (Bauman *et al*. 2004).

Altogether, the overexpression of the genes implicated in lipid metabolism network is in agreement with the differentiation mechanisms leading to fully steroidogenic granulosa cells in LF attested by the overexpression of *CYP19A1* (micro-array), *CYP11A1* and *STAR* genes (QPCR).

Our results notably reveal cell morphology and ion-binding gene regulation. Indeed, along terminal follicular development, we observed a down-regulation of genes coding for cytoskeletal microtubule constituents (*TUB1*, *TUB5* and *TUB7*), intermediate filaments (vimentin) and a gene implicated in their assembly, *STMN1*. This strongly suggests a deep modification of cell architecture of granulosa cells in LF. The network constructed with genes implicated in cell shape and ion binding (Fig. 6) also led us to the hypothesis of actin network implication. We showed up-regulation in LF of cofilin 2 (*CFL2*) involved in actin globular-form sequestration. In contrast, we have observed a decreased expression of *SMTN* that is specifically associated with filamentous actin in stress fibres. In addition, we observed in granulosa cells of LF the down-expression of genes coding for calcium-binding proteins, also implicated in actin cytoskeleton organization: calgizzarin (*S100A11*), *CALM1* and *CALM2* and *CALU*. Altogether, the modifications of these gene regulations let us think of the cell shape changes in LF when compared with earlier stages. This is attested by our *in vivo* observation of actin network in granulosa cells. The observation of actin staining revealed a different organization of actin filaments between the granulosa cells of small/medium and large follicles (Fig. 7). Rounded versus elongated shape may be associated with the *in vivo* differences in morphological cell gene expression between granulosa cells of SF versus LF. Interestingly, in sheep, induced cell rounding was associated with enhanced oestradiol secretion and inhibited proliferation of granulosa cells (Le Bellego *et al*. 2005). In addition, the different studies on luteinization process reported that gonadotrophic regulation of granulosa cell steroidogenesis was associated with cell shape changes (Ben-Ze'ev & Amsterdam 1989). In response to gonadotrophins, several genes coding for microtubule system and intermediate filament have been modulated. This suggests that rearrangement of the cytoskeletal proteins permits better coupling between steroidogenesis involved organelles (Carnegie *et al*. 1988, Sasson *et al*. 2004). Our hypothesis is that one of the microtubule roles tends to facilitate the movement of cholesterol from lipid droplets to mitochondria, possibly by bringing these cellular inclusions closer together (Carnegie *et al*. 1987).

This demonstrates the usefulness of building networks from micro-array experiments: they may highlight new genes that were not analysed in the micro-array experiment but that are relevant to the studied biological phenomenon, even if, in our case, the literature had already pointed these remodelling processes *in vitro*.

Up to now, only three transcriptome analyses on pig folliculogenesis have been performed. Two of them identified genes from the whole follicle whose ovarian expression has been changed as a result of long-term genetic selection for the component of reproduction (Caetano *et al*. 2004, Gladney *et al*. 2004). The third analysis examined gene expression in the whole follicle between preovulatory oestrogenic and luteinized follicles and corresponded to the continuity of our study (Agca *et al*. 2006).

In conclusion, the present study identified 79 regulated genes that may contribute to a better understanding of the mechanism involved in terminal antral follicular growth. Four cDNA sequences were found without significant homologies in databases. Some genes had functions that were already known and their regulation was consistent with the published literature. Some of them have been never associated with folliculogenesis, such as *STMN1*, *SMTN*, *ITM2A* and *GPX3*. Further investigations will be necessary to analyse the spatio-temporal expression pattern of these new genes and their interplay at the RNA and protein levels in the developing ovarian follicle. They may give clues to better understanding of the folliculogenesis process. Integration of the data highlighted gene networks mainly involved in ribosomal protein synthesis; steroid metabolism and cell morphology are perfectly coherent with enhanced steroidogenesis and arrest of growth rate in granulosa cells reaching their final maturation at the end of the terminal follicle development in pig.

Materials and Methods

Collection of ovaries

Oestrous cycles of gilts were synchronized by oral administration of 20 mg/day altrenogest (Regumate, Hoechst-Roussel, Paris, France) for 18 days at INRA experimental farm (animal experimentation authorization B-35-275-32). Ovaries were removed by laparotomy, 24 or 96 h after the last altrenogest feeding. For *in situ* experiment and phalloidin-FITC detection, ovaries were embedded in OCT medium (Miles Laboratories, CML, Nemours, France), frozen in liquid N_2 vapour and stored at −80 °C.

Granulosa cell isolation and RNA extraction

All antral follicles from 1 mm in diameter were isolated carefully using a binocular microscope. The diameter of each follicle was measured and follicles were classified according to

their size class as previously described for pig (Guthrie *et al.* 1993). SF (1–2 mm) and MF (3–4 mm) were recovered 24 h after the last altrenogest feeding of gilts. LF (≥5 mm) were recovered 96 h after the last altrenogest feeding. Granulosa cells were collected from individual follicle in MEM121/F12 (v/v) medium (Gasser *et al.* 1985). A small aliquot fraction was examined using Feulgen colouration to select cells only from healthy follicles (frequent mitosis, no pyknosis), as described previously (Monget *et al.* 1993, Besnard *et al.* 1996).

RNA was extracted from granulosa cells according to the technique described by Chomczynski & Sacchi (1987) with minor modifications (Hatey *et al.* 1995) using pools of granulosa cells from the same follicle size class. Finally, four independent RNA samples were obtained from small healthy follicles, five samples from medium healthy follicles and five samples from large healthy follicles. The quality of each RNA sample was checked through the Bioanalyser Agilent 2100 (Agilent Technologies, Massy, France).

Suppression subtractive hybridization (SSH)

Synthesis of cDNA and SSH were performed as described previously (Bonnet *et al.* 2006*a*), using SMART PCR cDNA synthesis kit and PCR-Select cDNA Subtraction kit respectively (each from Clontech). Briefly, SSH was performed using 300 ng cDNA generated from each RNA sample. After hybridization, the primary PCR amplification was achieved through 27–30 PCR cycles starting with 1 µl of a 23-fold diluted second hybridization reaction. The secondary PCR amplification was achieved through 11–12 PCR cycles starting with 1 µl of a tenfold diluted primary PCR amplification. Resulting SSH products were purified and concentrated, using Amicon Microcon-PCR filters (Millipore, St Quentin-en-Yvelines, France). cDNAs >500 bp were selected by gel filtration on 1 ml sepharose CL2B column, as described in pBluescript II XR Library Construction Kit (Stratagene Europe, Amsterdam, The Netherlands). The first 1 ml fraction was saved and ethanol precipitated with 20 µg glycogen and then diluted in 10 µl water. The selected PCR products were cloned using the pGEM-T Easy Vector (Promega) and electroporated into competent bacteria (DH10 alpha, Clontech).

Two forward SSH libraries were constructed, SF versus LF (SF/lf) and SF versus MF (SF/mf), and two reverse ones, LF versus SF (LF/sf) and MF versus SF (MF/sf). Each library contained 2500 cDNA clones and was respectively enriched with over-expressed genes in SF compared with the LF and MF, and MF and LF compared with SF.

Macro-arrays design and analysis

Bacterial clones of the four different SSH libraries (about 10 000 clones) have been spotted onto nylon filters to generate colony macro-arrays in order to sort out the false-positive clones and select differentially expressed candidate clones between the different follicle classes before sequencing.

Arrays were generated as described (Nguyen *et al.* 1995) and probed in duplicate, using both SMART products (from RNA of small, medium and large follicles) and the four SSH cDNA products (Sf/lf, LF/sf, SF/mf and MF/sf) labelled with α-^{33}P dCTP, as described previously (Hatey *et al.* 1992). After washing, the arrays were exposed 6 or 24 h to storage phosphor screens and scanned thereafter with a phosphor imaging system at a 50 µm resolution (Storm 840; GE Healthcare, Orsay, France).

Image quantification was performed using the XDotReader software and data analysis was performed using the BioPlot software (available at http://biopuce.insa-toulouse.fr). Briefly, after log transformation, the data were normalized by all spot average without background subtraction. The differential cDNA clone's expression is based on following software criterions: overexpressed threshold ratio, 1.5; under-expressed threshold ratio, 0.66 and Student's *t*-test (*P* value threshold, 0.25). A selection of differentially expressed cDNA clones was sequenced and spotted onto micro-array membranes.

Micro-arrays design and hybridization

The micro-arrays contained PCR products from 2849 pig cDNA clones coming from 1697 clones selected after SSH/macro-array experiments, 1056 clones of the AGENAE pig normalized multi-tissue cDNA library (Bonnet *et al.* 2008) and 96 clones used as controls. The multi-tissue library was used to allow a proper normalization of the data. PCR products were spotted in duplicate on two separate fields of the same nylon membrane (18×72 mm, Immobilon-NY+, Millipore) as described (Ferre *et al.* 2007). A detailed description of the resulting micro-array platform is available in the Gene Expression Omnibus database (www.ncbi.nlm.nih.gov/geo).

The quality of spotting and the relative amount of DNA in each spot have been controlled, using a ^{33}P-labelled oligonucleotide corresponding to a vector sequence present in all PCR products (Ferre *et al.* 2007; GEO accession number GSE5797 dataset). After stripping, the arrays were hybridized with ^{33}P-labelled complex probes synthesized from 5 µg of each RNA sample (4 SF, 5 MF and 5 LF), using SuperScript II RNAse H$^-$ reverse transcriptase (Invitrogen). Each complex probe has been hybridized on two individual membranes exposed 6 or 24 h to radioisotopic-sensitive imaging plates (BAS-2025; Fujifilm, Raytest, Courbevoie, France). The imaging plates were scanned thereafter with a phosphor imaging system at 25 µm resolution (BAS-5000, Fujifilm). Hybridization images obtained from oligonucleotide and complex probes were quantified using the semi-automated BZScan software (Lopez *et al.* 2004).

Data management

The experimental design, its implementation and the handling of data comply with MIAME standards (Brazma *et al.* 2001), and all the experimental data were managed using BASE software (Saal *et al.* 2002), adapted by SIGENAE bioinformatics platform (http://www.sigenae.org) to ensure radioactive experiments.

Data analysis

Using the oligonucleotide probe, spots with low signal values (i.e. <2× median of empty spots) were considered as mis-amplified or mis-spotted and were excluded from the

analysis. The data coming from complex probe hybridizations were logarithmically transformed and centred for each membrane. Spots with low signal value (below the average of empty spots +2 SDs) were considered as unexpressed and were excluded from the analysis. Finally, the remaining data were centred for each PCR product. Thereafter, a mixed linear model was fitted on these data, using the MIXED procedure of SAS software (SAS Institute Inc., Cary, NC, USA). Explanatory variables (follicle class, gene and interaction between these factors) were treated as fixed effects and animal, experimental variability (RNA, hybridization) and residual were treated as random effects.

The selection of the differentially expressed genes was performed using the R statistical software system (the Comprehensive R Archive National, http://www.cran.r-project.org). We tested the significance of the follicle classes on the gene expression using *F*-test followed by the Benjamini–Hochberg procedure controlling FDR for each cDNA (Benjamini & Hochberg 1995). Of note, *F*-test selects genes whose intra-class variance is differentially expressed in at least one of the classes. Thereafter, the expression value of each selected gene was compared between each follicle size class, using a Student's *t*-test (*P* value: 5%) to establish the expression profile along terminal follicular development.

A set of predictive genes was identified using the balanced RFs approach (Chen *et al*. 2004). To obtain a stable selection of genes, 1000 of the balanced RFs were launched with each of 15 000 trees and 42 variables randomly sampled as candidates at each split (default value proposed by R). The most important cDNAs that appeared the most frequently (in 90% of the forests) were selected using the Mean Decrease Gini importance measure as feature selection criterion.

Finally, the relevance of the two selections was evaluated via unsupervised hierarchical clustering using the Ward method and Euclidean distance with the R functions *hclust* and *heatmap* (Chipman *et al*. 2003).

Sequence annotation

Each cDNA sequence was compared with Refseq_rna mammalian database using the NCBI blastn program (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi). Blast results with an *e* value inferior to $1e-3$ were parsed and filtered to keep queries matching either a gene, a mRNA or a CDS and possessing at least a global coverage of 70% of the query sequence. Resulting hits were sorted out according to their closeness to the pig genome, their coverage and sequence identity. The selected cDNA sequences were submitted to the Human Genome Organization (HUGO) gene nomenclature committee, using their RefSeq IDs (http://www.genenames.org/). Then, HUGO gene symbols were used to name the genes.

Biological network and pathway analysis

IPA software (Ingenuity Systems Inc., Redwood City, CA, USA) was used to examine molecular pathways. This software combines functional annotations of our differentially expressed genes (focus genes) and the corresponding bibliographic data to generate significant signalling pathways and regulation networks.

Quantitative PCR analysis of gene expression

Total RNA (2 μg) from the same RNA samples used in micro-array experiments was reverse transcribed as described previously (Tosser-Klopp *et al*. 2001). An external standard (plant mRNA: I11a accession number Y10291) was added to each RNA sample (1 pg for 2 μg total RNA sample) before RT to allow quantification of the cDNA production. Primers were designed using 'Primer Express' software (Applied Biosystems, Courtaboeuf, France) and the intron–exon organization of porcine genes has been deduced by comparison with human genome using Iccare software (Muller *et al*. 2004). Translationally controlled tumour protein gene (*TCTP*, accession number BX667045) and *COX3* gene (cytochrome *c* oxidase subunit III, accession number CT971556) were found to be highly expressed but not regulated during follicle development in our micro-array experiment and used as internal controls. All primer sequences are given in Supplementary Table 1, which can be viewed online at www.reproduction-online.org/supplemental. Quantitative real-time PCR was performed using SYBR green fluorescence detection during amplification on an ABI Prism 7900 Sequence Detection System 2.1 (Applied Biosystems), according to the manufacturer's recommendations. Duplicates of each template (120 or 500 pg) were loaded in 384-well plates using a liquid handling robot (TECAN genesis rsp 200X8, Mannedorf/Zurich, Switzerland) with a 10 μl PCR mix SYBR green Power master mix (Applied Biosystems) and 0.5 μM forward and reverse primers (final volume of 13 μl). The PCR amplification conditions were as follows: 50 °C for 30 min, initial denaturation at 95 °C for 10 min and 40 cycles (95 °C for 15 s and 60 °C for 1 min). The last cycle was followed by a dissociation step (ramping to 95 °C). The real-time PCR amplification efficiency has been calculated for each primer pair with four 1:2 dilution points of the calibrator sample (pool of the 14 cDNA samples). After determination of the threshold cycle (Ct) for each sample, the PFAFFL method was applied to calculate the relative changes of each mRNA in each sample (Pfaffl 2001). The relative expression was normalized by the corresponding geometric average of an external control gene and two internal genes using geNorm v3.4 (Vandesompele *et al*. 2002). The significance of expression ratio was tested using *F*-test.

In situ hybridization and actin detection

Frozen ovaries recovered 24 or 96 h after the last altrenogest feeding were serially sectioned at a thickness of 10 μm with a cryostat. For *in situ* hybridization, ^{35}S-labelled cRNA probes (sense and anti-sense) were obtained from *GSTA* gene (accession number X93247 and X91711) by *in vitro* transcription of PCR products generated with the recombinant plasmid using primers containing T3 and T7 promoters at their 5′-end, as described previously (Besnard *et al*. 1996). For actin detection, cryosections were stained with FITC-conjugated phalloidin (Sigma–Aldrich), as previously described (Le Bellego *et al*. 2005), and were analysed using fluorescence microscopy.

Declaration of interest

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

Funding

This work was supported by a grant of the Toulouse Genopole Midi-Pyrénées program.

Acknowledgements

The authors thank Janine Rallières (UMR GC) for technical assistance, H Demay for supplying animals (INRA UMR SENAH), the Centre de Ressources Génotypage Sequençage (Toulouse Genopole Midi-Pyrénées) for technical support and SIGENAE for their support in informatics and statistical improvements. We are grateful to Yvon Tosser for careful reading of the manuscript.

References

Agca C, Ries JE, Kolath SJ, Kim JH, Forrester LJ, Antoniou E, Whitworth KM, Mathialagan N, Springer GK, Prather RS *et al.* 2006 Luteinization of porcine preovulatory follicles leads to systematic changes in follicular gene expression. *Reproduction* **132** 133–145.

Balasubramanian K, Lavoie HA, Garmey JC, Stocco DM & Veldhuis JD 1997 Regulation of porcine granulosa cell steroidogenic acute regulatory protein (StAR) by insulin-like growth factor I: synergism with follicle-stimulating hormone or protein kinase A agonist. *Endocrinology* **138** 433–439.

Bauman DR, Steckelbroeck S & Penning TM 2004 The roles of aldo–keto reductases in steroid hormone action. *Drug News & Perspectives* **17** 563–578.

Le Bellego F, Fabre S, Pisselet C & Monniaux D 2005 Cytoskeleton reorganization mediates alpha6beta1 integrin-associated actions of laminin on proliferation and survival, but not on steroidogenesis of ovine granulosa cells. *Reproductive Biology and Endocrinology* **3** 19.

Benjamini V & Hochberg V 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B* **57** 289–300.

Ben-Ze'ev A & Amsterdam A 1989 Regulation of cytoskeletal protein organization and expression in human granulosa cells in response to gonadotropin treatment. *Endocrinology* **124** 1033–1041.

Besnard N, Pisselet C, Monniaux D, Locatelli A, Benne F, Gasser F, Hatey F & Monget P 1996 Expression of messenger ribonucleic acids of insulin-like growth factor binding protein-2, -4, and -5 in the ovine ovary: localization and changes during growth and atresia of antral follicles. *Biology of Reproduction* **55** 1356–1367.

Bonnet A, Frappart PO, Dehais P, Tosser-Klopp G & Hatey F 2006*a* Identification of differential gene expression in *in vitro* FSH treated pig granulosa cells using suppression subtractive hybridization. *Reproductive Biology and Endocrinology* **4** 35.

Bonnet A, Lê Cao KA, San Cristobal M, Low-So G, Tosser-Klopp G & Hatey F 2006b Transcriptome of pig ovarian cells: discriminant genes involved in pig ovarian development. *Proceedings of the First European Conference on Pig Genomics*, Lodi, Italy, Abstract 39.

Bonnet A, Iannuccelli E, Hugot K, Benne F, Bonaldo MF, Soares MB, Hatey F & Tosser G 2008 A pig multi-tissue normalised cDNA library: large-scale sequencing, cluster analysis and 9K micro-array resource generation. *BMC Genomics* **9** 17.

Boon K, Caron HN, van Asperen R, Valentijn L, Hermus MC, van Sluis P, Roobeek I, Weis I, Voute PA, Schwab M *et al.* 2001 N-myc enhances the expression of a large set of genes functioning in ribosome biogenesis and protein synthesis. *EMBO Journal* **20** 1383–1393.

Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert - C, Aach J, Ansorge W, Ball CA, Causton HC *et al.* 2001 Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nature Genetics* **29** 365–371.

Caetano AR, Johnson RK, Ford JJ & Pomp D 2004 Microarray profiling for differential gene expression in ovaries and ovarian follicles of pigs selected for increased ovulation rate. *Genetics* **168** 1529–1537.

Lê Cao KA, Bonnet A, Besse P, Robert-Granié C & San Cristobal M 2006 Feature selection with random forests for unbalanced multiclass microarray data: application in pig ovarian follicular development. *8th World Congress on Genetics Applied to Livestock Production*, Belo Horizonte, Brazil.

Carnegie JA, Dardick I & Tsang BK 1987 Microtubules and the gonadotropic regulation of granulosa cell steroidogenesis. *Endocrinology* **120** 819–828.

Carnegie JA, Byard R, Dardick I & Tsang BK 1988 Culture of granulosa cells in collagen gels: the influence of cell shape on steroidogenesis. *Biology of Reproduction* **38** 881–890.

Chan WK & Tan CH 1987 Induction of aromatase activity in porcine granulosa cells by FSH and cyclic AMP. *Endocrine Research* **13** 285–299.

Chen C, Liaw A & Breiman L 2004 Using random forest to learn imbalanced data. UC Berkeley.

Chipman H, Hastie T & Tibshirani R 2003 Clustering microarray data. *Statistical analysis of gene expression microarray data*, pp 159–199. New York: Chapman & Hall.

Chomczynski P & Sacchi N 1987 Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Analytical Biochemistry* **162** 156–159.

Clouscard-Martinato C, Mulsant P, Robic A, Bonnet A, Gasser F & Hatey F 1998 Characterization of FSH-regulated genes isolated by mRNA differential display from pig ovarian granulosa cells. *Animal Genetics* **29** 98–106.

Conley AJ, Howard HJ, Slanger WD & Ford JJ 1994 Steroidogenesis in the preovulatory porcine follicle. *Biology of Reproduction* **51** 655–661.

Diaz-Uriarte R & Alvarez de Andres S 2006 Gene selection and classification of microarray data using random forest. *BMC Bioinformatics* **7** 3.

Drummond AE 2006 The role of steroids in follicular growth. *Reproductive Biology and Endocrinology* **4** 16.

Duda M 1997 The influence of FSH, LH and testosterone on steroidsecretion by two subpopulations of porcine granulosa cells. *Journal of Physiology and Pharmacology* **48** 89–96.

Ferre PJ, Liaubet L, Concordet D, SanCristobal M, Uro-Coste E, Tosser-Klopp G, Bonnet A, Toutain PL, Hatey F & Lefebvre HP 2007 Longitudinal analysis of gene expression in porcine skeletal muscle after post-injection local injury. *Pharmaceutical Research* **24** 1480–1489.

Foxcroft GR & Hunter MG 1985 Basic physiology of follicular maturation in the pig. *Journal of Reproduction and Fertility* **33** 1–19.

Fricke PM, Ford JJ, Reynolds LP & Redmer DA 1996 Growth and cellular proliferation of antral follicles throughout the follicular phase of the estrous cycle in Meishan gilts. *Biology of Reproduction* **54** 879–887.

Gasser F, Mulsant P & Gillois M 1985 Long-term multiplication of the Chinese hamster ovary (CHO) cell line in a serum-free medium. *In Vitro Cellular & Developmental Biology* **21** 588–592.

Gladney CD, Bertani GR, Johnson RK & Pomp D 2004 Evaluation of gene expression in pigs selected for enhanced reproduction using differential display PCR and human microarrays: I. Ovarian follicles. *Journal of Animal Science* **82** 17–31.

Guthrie HD, Bolt DJ & Cooper BS 1993 Changes in follicular estradiol-17 beta, progesterone and inhibin immunoactivity in healthy and atretic follicles during preovulatory maturation in the pig. *Domestic Animal Endocrinology* **10** 127–140.

Hatey F, Gasparoux JP, Mulsant P, Bonnet A & Gasser F 1992 P450scc regulation in pig granulosa cells: investigation into the mechanism of induction. *Journal of Steroid Biochemistry and Molecular Biology* **43** 869–874.

Hatey F, Mulsant P, Bonnet A, Benne F & Gasser F 1995 Protein kinase C inhibition of *in vitro* FSH-induced differentiation in pig granulosa cells. *Molecular and Cellular Endocrinology* **107** 9–16.

Hattori MA, Yoshino E, Shinohara Y, Horiuchi R & Kojima I 1995 A novel action of epidermal growth factor in rat granulosa cells: its potentiation of gonadotrophin action. *Journal of Molecular Endocrinology* **15** 283–291.

Hillier SG & Miro F 1993 Inhibin, activin, and follistatin. Potential roles in ovarian physiology. *Annals of the New York Academy of Sciences* **687** 29–38.

Hirshfield AN 1986 Patterns of [3H] thymidine incorporation differ in immature rats and mature, cycling rats. *Biology of Reproduction* **34** 229–235.

Hsueh AJ 1986 Paracrine mechanisms involved in granulosa cell differentiation. *Clinics in Endocrinology and Metabolism* **15** 117–134.

Jiang H, Whitworth KM, Bivens NJ, Ries JE, Woods RJ, Forrester LJ, Springer GK, Mathialagan N, Agca C, Prather RS *et al.* 2004 Large-scale generation and analysis of expressed sequence tags from porcine ovary. *Biology of Reproduction* **71** 1991–2002.

Keira M, Nishihira J, Ishibashi T, Tanaka T & Fujimoto S 1994 Identification of a molecular species in porcine ovarian luteal glutathione S-transferase and its hormonal regulation by pituitary gonadotropins. *Archives of Biochemistry and Biophysics* **308** 126–132.

LaVoie HA, Benoit AM, Garmey JC, Dailey RA, Wright DJ & Veldhuis JD 1997 Coordinate developmental expression of genes regulating sterol economy and cholesterol side-chain cleavage in the porcine ovary. *Biology of Reproduction* **57** 402–407.

Lopez F, Rougemont J, Loriod B, Bourgeois A, Loi L, Bertucci F, Hingamp P, Houlgatte R & Granjeaud S 2004 Feature extraction and signal processing for nylon DNA microarrays. *BMC Genomics* **5** 38.

May JV & Schomberg DW 1984 Developmental coordination of luteinizing hormone/human chorionic gonadotropin (hCG) receptors and acute hCG responsiveness in cultured and freshly harvested porcine granulosa cells. *Endocrinology* **114** 153–163.

Mazerbourg S, Bondy CA, Zhou J & Monget P 2003 The insulin-like growth factor system: a key determinant role in the growth and selection of ovarian follicles? a comparative species study *Reproduction in Domestic Animals* **38** 247–258.

Monget P, Monniaux D, Pisselet C & Durand P 1993 Changes in insulin-like growth factor-I (IGF-I), IGF-II, and their binding proteins during growth and atresia of ovine ovarian follicles. *Endocrinology* **132** 1438–1446.

Monget P, Fabre S, Mulsant P, Lecerf F, Elsen JM, Mazerbourg S, Pisselet C & Monniaux D 2002 Regulation of ovarian folliculogenesis by IGF and BMP system in domestic animals. *Domestic Animal Endocrinology* **23** 139–154.

Muller C, Denis M, Gentzbittel L & Faraut T 2004 The Iccare web server: an attempt to merge sequence and mapping information for plant and animal species. *Nucleic Acids Research* **32** W429–W434.

Nguyen C, Rocha D, Granjeaud S, Baldit M, Bernard K, Naquet P & Jordan BR 1995 Differential gene expression in the murine thymus assayed by quantitative hybridization of arrayed cDNA clones. *Genomics* **29** 207–216.

Pfaffl MW 2001 A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* **29** e45.

Pisselet C, Clement F & Monniaux D 2000 Fraction of proliferating cells in granulosa during terminal follicular development in high and low prolific sheep breeds. *Reproduction, Nutrition, Development* **40** 295–304.

Rabahi F, Brule S, Sirois J, Beckers JF, Silversides DW & Lussier JG 1999 High expression of bovine alpha glutathione S-transferase (GSTA1, GSTA2) subunits is mainly associated with steroidogenically active cells and regulated by gonadotropins in bovine ovarian follicles. *Endocrinology* **140** 3507–3517.

Saal LH, Troein C, Vallon-Christersson J, Gruvberger S, Borg A & Peterson C 2002 BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data. *Genome Biology* **3** SOFTWARE0003.

Sasson R, Rimon E, Dantes A, Cohen T, Shinder V, Land-Bracha A & Amsterdam A 2004 Gonadotrophin-induced gene regulation in human granulosa cells obtained from IVF patients. Modulation of steroidogenic genes, cytoskeletal genes and genes coding for apoptotic signalling and protein kinases. *Molecular Human Reproduction* **10** 299–311.

Saxena D, Escamilla-Hernandez R, Little-Ihrig L & Zeleznik AJ 2007 Liver receptor homolog-1 and steroidogenic factor-1 have similar actions on rat granulosa cell steroidogenesis. *Endocrinology* **148** 726–734.

Shimasaki S, Moore RK, Otsuka F & Erickson GF 2004 The bone morphogenetic protein system in mammalian reproduction. *Endocrine Reviews* **25** 72–101.

Sirianni R, Seely JB, Attia G, Stocco DM, Carr BR, Pezzi V & Rainey WE 2002 Liver receptor homologue-1 is expressed in human steroidogenic tissues and activates transcription of genes encoding steroidogenic enzymes. *Journal of Endocrinology* **174** R13–R17.

Tosser-Klopp G, Bonnet A, Yerle M & Hatey F 2001 Functional study and regional mapping of 44 hormono-regulated genes isolated from a porcine granulosa cell library. *Genetics, Selection, Evolution* **33** 69–87.

Tuggle CK, Wang Y & Couture O 2007 Advances in swine transcriptomics. *International Journal of Biological Sciences* **3** 132–152.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A & Speleman F 2002 Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* **3** RESEARCH0034.

Vergnes L, Phan J, Stolz A & Reue K 2003 A cluster of eight hydroxysteroid dehydrogenase genes belonging to the aldo-keto reductase supergene family on mouse chromosome 13. *Journal of Lipid Research* **44** 503–511.

Zhao H, Li Z, Cooney AJ & Lan ZJ 2007 Orphan nuclear receptor function in the ovary. *Frontiers in Bioscience* **12** 3398–3405.

Received 5 July 2007
First decision 6 September 2007
Revised manuscript received 8 April 2008
Accepted 28 April 2008