Meta-analysis of gene expression profiles in granulosa cells during folliculogenesis

Daulat Raheem Khan¹, Éric Fournier¹, Isabelle Dufort¹, François J Richard¹, Jaswant Singh² and Marc-André Sirard¹

¹Centre de Recherche en Biologie de la Reproduction, Département des Sciences Animales, Faculté des sciences de l’agriculture et de l’alimentation, Université Laval, Quebec City, Québec, Canada, and ²Department of Veterinary Biomedical Sciences, Western College of Veterinary Medicine, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

Correspondence should be addressed to M-A Sirard; Email: marc-andre.sirard@fsaa.ulaval.ca

Abstract

Folliculogenesis involves coordinated profound changes in different follicular compartments and significant modifications of their gene expression patterns, particularly in granulosa cells. Huge datasets have accumulated from the analyses of granulosa cell transcriptomic signatures in predefined physiological contexts using different technological platforms. However, no comprehensive overview of folliculogenesis is available. This would require integration of datasets from numerous individual studies. A prerequisite for such integration would be the use of comparable platforms and experimental conditions. The EmbryoGENE program was created to study bovine granulosa cell transcriptomics under different physiological conditions using the same platform. Based on the data thus generated so far, we present here an interactive web interface called GranulosaIMAGE (Integrative Meta-Analysis of Gene Expression), which provides dynamic expression profiles of any gene of interest and all isoforms thereof in granulosa cells at different stages of folliculogenesis. GranulosaIMAGE features two kinds of expression profiles: gene expression kinetics during bovine folliculogenesis from small (6 mm) to pre-ovulatory follicles under different hormonal and physiological conditions and expression profiles of granulosa cells of dominant follicles from post-partum cows in different metabolic states. This article provides selected examples of expression patterns along with suggestions for users to access and generate their own patterns using GranulosaIMAGE. The possibility of analysing gene expression dynamics during the late stages of folliculogenesis in a mono-ovulatory species such as bovine should provide a new and enriched perspective on ovarian physiology.

Reproduction (2016) 151 R103–R110

Introduction

The ovary is a highly dynamic structure, of which the principal functional unit is the follicle. In foetal ovaries, primordial germ cells proliferate during the first trimester of gestation and develop into primordial follicles by mid-gestation. A primordial follicle is typically 30–40 µm in diameter and each is composed of a partially differentiated oocyte (arrested in prophase-1 of meiosis) enclosed by one layer of specialized somatic cells called follicular or granulosa cells. Further follicle development begins before birth as small cohorts of primordial follicles undergo progressive growth and atresia until puberty. Folliculogenesis progresses in the adult ovary, leading to the formation of a fluid-filled cavity called the antrum and the emergence of a highly specialized type of granulosa cell called cumulus cells, which are in direct contact with the oocyte (Gougeon 1996). In mono-ovulatory species such as cattle, one ovule per reproductive cycle is released from a single dominant follicle, whereas the remaining follicles undergo atresia (Lussier et al. 1987).

In contrast to other somatic tissues, granulosa cells during folliculogenesis undergo very dynamic and highly coordinated changes. During the late stages of folliculogenesis, the changes accelerate in all compartments of the follicle wall (granulosa, cumulus, and theca cells, vascular and inter-cellular stromal components), culminating in the release of a competent oocyte and the formation of a new tissue called the corpus luteum. In the developing follicle, acquisition of oocyte competence involves interplay between a multitude of intrinsic and extrinsic factors, which all act to bring about rapid development of distinct gene expression profiles in different follicular cells (Wigglesworth et al. 2014, Khan et al. 2015). This is particularly apparent in granulosa cells (Sirard 2014), in which gene expression patterns are important not only for the ovulation and luteinization processes but
also for the developmental competence of the oocyte contained therein (Assidi et al. 2008, Hamel et al. 2010). Interestingly, FSH has been implicated in acquisition of oocyte developmental competence, both in vivo and in vitro (Sirard et al. 2007). In bovine, in vivo experiments have shown that ovarian super-stimulation with a FSH support for 5 days (endogenous FSH following removal of dominant follicle (for 2 days) followed by 3 days of FSH injections twice a day) followed by no FSH period (called coasting) for 44–68 h yields the best oocyte quality for subsequent development of embryos (Nivet et al. 2012). The study of granulosa cell transcriptome dynamics in different physiological contexts therefore remains crucial to understand the physiology of ovarian tissue as a whole.

Conventional experimental designs do not provide the overall perspective that is essential in order to understand follicular dynamics. The huge amounts of data that have accumulated remains scattered in database repositories and require integration and meta-analysis in order to chart overall gene dynamics in this tissue. The regular manuscript format allows sharing of 1–2% of the data analysed (i.e. highlighted genes), and access to supplemental data, although possible, is difficult to re-analyse. Other than a recent comparison of cumulus and mural granulosa cell transcriptomes in mice (Wigglesworth et al. 2015), meta-analysis of different ovarian transcriptomic studies remains scarce. An online public collection called the ovarian kaleidoscope database or OKdb (http://okdb.appliedbioinfo.net/) provides information on gene expression in different ovarian cell types and their association with various ovarian functions (Hsueh & Rauch 2012). However, a chronological/dynamic interface of folliculogenesis based on integrated ovarian cell gene expression profiles has yet to be constructed. The principal obstacles to achieving this are incomparable technological platforms and experimental conditions in the different studies (Tseng et al. 2012). The integration of such studies requires vast knowledge of ovarian physiology combined with highly specialized bioinformatics skills.

Based on the availability of several publicly available transcriptomic analyses generated on a single technological platform called “EmbryoGENE”, an online interface called GranulosaIMAGE (Granulosa Integrative Meta-Analysis of Gene Expression) has been developed. GranulosaIMAGE provides easy consultation of the temporal kinetics of gene expression during follicular development from small-diameter (>5 mm) follicles to pre-ovulatory in different physiological contexts, along with dominant follicle gene expression profiles for various post-partum time intervals and metabolic states.

![Figure 1](image-url)  
**Figure 1**  Summary of GranulosaIMAGE workflow. Data from transcriptome studies conducted by EmbryoGENE network scientists on granulosa cells have been deposited in the ELMA database. Granulosa cells were obtained from follicles at various stages of folliculogenesis from cows in different physiological and metabolic states (top panel). This flow diagram provides the working model of data retrieval from ELMA, its normalization, statistical tests and graphical representation on the GranulosaIMAGE web-based resource.
Programme description and methods

GranulosalMAGE provides a view of the dynamics of bovine genes and their isoforms by integrating 74 microarray datasets generated using the EmbryoGENE platform (Robert et al. 2011) and a uniform analysis pipeline. Although these studies were conducted independently, they cover collectively most stages of ovarian follicle development (Fig. 1). GranulosalMAGE (http://emb-bioinfo.isaa.ulaval.ca/GranulosalMAGE/) was thus generated as a web resource for easy consultation of the relative dynamics of practically any gene in bovine granulosa cells.

Studies included

The studies included in GranulosalMAGE database are summarized in Table 1, which provides principally two types of granulosa gene expression pattern for each gene.

1. Time-course gene kinetics during follicular development from the small (6 mm) to pre-ovulatory stage under different hormonal or physiological conditions.
2. Gene expression in granulosa cells recovered post-partum from dominant follicles of cows synchronized using prostaglandins (52 h after synchronization, for uniformity across samples and to precede the endogenous LH surge).

<table>
<thead>
<tr>
<th>Reference</th>
<th>Group</th>
<th>Title of article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Douville &amp; Sirard (2014)</td>
<td>6–9 mm follicles (follicular state effect)</td>
<td>Changes in granulosa cells gene expression associated with growth, plateau and atretic phases in medium bovine follicles</td>
</tr>
<tr>
<td>Girard et al. (2015)</td>
<td>&gt;9 mm follicles (follicular state effect)</td>
<td>Global gene expression in granulosa cells of growing, plateau and atretic dominant follicles in cattle</td>
</tr>
<tr>
<td>Nivet et al. (2012)</td>
<td>Dominant follicle (age effect)</td>
<td>FSH withdrawal improves developmental competence of oocytes in the bovine model</td>
</tr>
<tr>
<td>Gilbert et al. (2012)</td>
<td>Dominant follicle (time to LH surge)</td>
<td>Impact of the LH surge on granulosa cell transcript levels as markers of oocyte developmental competence in cattle GSE69247</td>
</tr>
<tr>
<td>Dias et al. (2013a,b)</td>
<td>Pre-ovulatory follicle (24 h post-LH Age effect)</td>
<td>Effect of duration of the growing phase of ovulatory follicles on oocyte competence in super-stimulated cattle</td>
</tr>
<tr>
<td>Golini et al. (2014)</td>
<td>Dominant follicle (post-partum period effect)</td>
<td>Transcriptome analysis of bovine granulosa cells of pre-ovulatory follicles harvested 30, 60, 90 and 120 days post-partum</td>
</tr>
<tr>
<td>Girard et al. (2015)</td>
<td>Dominant follicle (energy balance effect at 60 ± 5 days post-partum)</td>
<td>The effect of energy balance on the transcriptome of bovine granulosa cells at 60 days post-partum</td>
</tr>
<tr>
<td></td>
<td>Dominant follicle (pre-LH surge effect of vitamin B9 &amp; B12)</td>
<td>Unpublished</td>
</tr>
<tr>
<td>Gagnon et al. (2015)</td>
<td>Dominant follicle (vitamin B9 and B12 effect irrespective of LH surge)</td>
<td>Effects of intramuscular administration of folic acid and vitamin B12 on granulosa cells gene expression in post-partum dairy cows</td>
</tr>
</tbody>
</table>

Statistical methods

Data including intensity files and sample annotation (metadata) from previous studies of granulosa cells using the EmbryoGENE microarray platform (Robert et al. 2011) have been filed in the EmbryoGENE LIMS and Microarray Analysis (ELMA) database (Robert et al. 2011). These data are used here to perform meta-analysis and to generate expression profiles. The working model of GranulosalMAGE is summarized in Fig. 1. Intensity values for any probe (along with their associated metadata) are retrieved from the ELMA database and after logarithmic transformation (base 2) of the raw measured intensity; normalization of the intensity values is performed by subtraction of the background threshold. The background threshold is defined as the mean of the intensities of the negative control spots on an array plus twice the standard deviation of these intensities. The resulting relative intensities are then quantile-normalized using the limma bioconductor package (Ritchie et al. 2015) and plotted on the y-axis for each condition. These normalized intensity values are used to produce the colour scale applied to the different points (the higher the intensity, the darker the point appears) showing the expression level under each condition.

The levels of significant difference for each probe were determined using ad hoc statistical tests comparing specific conditions within each experiment. We used the t-test for experiments with two conditions and a fixed-effect ANOVA for experiments with three or...
more conditions. All $P$-values were then corrected to a unified FDR (false discovery rate). In the resulting graphs, the comparisons with corrected $P$-values <0.05 are highlighted in red. The statistical tests do not take into consideration array effects, dye effects, technical replicates, or secondary experimental factors such as individual animals. For further details, the reference articles should be consulted where the above effects have been considered for each study separately.

Data visualization

GranulosalIMAGE charts the intensity profiles of a given gene in granulosa cells under different conditions. The GranulosalIMAGE homepage (Fig. 2) provides a brief description of the use of the programme along with an introduction to the gene search results panel, links to a help section, a description of the statistical methods used and a list of published results included in the programme, and the search bar on which the sought gene probe ID or gene symbol is entered. A link to the EmbryoGENE microarray annotation file (Robert et al. 2011) is provided for consultation of probe IDs with the corresponding gene names, isoforms and symbols.

GranulosalIMAGE displays the results page in response to submission of a gene symbol. The results page contains several sets of information, which we present along with examples later in this article.

1. Results summary: The summary of results contains the numbers and IDs of the probes corresponding to the entered keyword and also indicates the symbol of the gene corresponding to each probe and whether the probe is in the constitutive region or some untranslated region (UTR) of the genome.

   Examples: We provide two examples. First, by submitting gene symbol AREG (amphiregulin), the results summary shows one probe (EMBV3_31722) that is of constitutive type. Secondly, looking for MAML2 (mastermind-like 2) returns three probes are listed (EMBV3_10398, EMBV3_29461 and EMBV3_34221), the second of which is constitutive, whereas the other two are the alternative 3′-UTR sites.

2. Probe details: By selecting the probe ID, the search is directed to the detailed results of the corresponding probe. A short table summarizes the properties of the probe previously mentioned (probe ID, gene symbol and probe type) along with gene names and options for external links to additional details on the gene of interest and its sequence information.

   Example: Selecting the single probe listed for AREG (EMBV3_31722) directs the user to additional information including the name amphiregulin. Furthermore, the user may choose to consult external links via ‘Entrez Gene’ or ‘Refseq’. This information becomes very helpful when a query returns the probes corresponding to more than one gene. For example, the entry ADAMTS1 (ADAM metallopeptidase with thrombospondin type 1 motif, 1) returns 12 probes, of which only one (EMBV3_20771) corresponds specifically to ADAMTS1. The others correspond to genes with similar names or symbols. This is useful as GranulosalIMAGE thus provides information about the gene families as well as different gene isoforms.
3. **Expression graphs**: Normalized intensity values of any gene are displayed graphically. These graphics are divided into two compartments: “follicles over time” showing gene dynamics from the small to pre-ovulatory follicle stages in different physiological contexts and second dominant follicle gene expression patterns in the post-partum period.

In these graphics, each column contains data from a particular experiment (named at the top), whereas the specific conditions for that experiment are mentioned at the bottom. The red line (at 0) shows the background level. An array value below the red line indicates that the intensity value is smaller than the background or in other words the gene is not expressed under those conditions. These values have not been removed from the analysis, in order to include all the conditions in an experiment regardless of gene expression. The small horizontal lines for each condition refer to the mean relative intensity of the probe for this condition. The panels highlighted in pink indicate that an *ad hoc* statistical test on the given group revealed a significant difference (*P* < 0.05).

**Examples**: *AREG* is a thoroughly studied gene involved in ovarian physiology (Conti et al. 2006). The graphical representation of *AREG* shows that FSH and LH affect expression of this gene (Fig. 3A). FSH stimulation during *in vitro* granulosa culture produced significant differences in *AREG* intensity. Also, when granulosa were obtained from follicles undergoing FSH coasting *in vivo*, significant differences in *AREG* levels were observed. Likewise, *AREG* intensity levels were greatly increased in granulosa cells early (6 h) after the LH surge compared with 2 h before or 22 h after. This information confirms previous findings published in the literature (Conti et al. 2006, Sugimura et al. 2014).
Khan et al. 2015). We note that additional information on AREG dynamics in other contrasts, such as follicular developmental phase (growth, plateau and atresia), animal age, ovarian super-stimulation and post-partum period effects, is seldom described in studies of AREG. Although AREG does not change significantly in such contrasts, grouping these data in one location adds considerable value to GranulosalIMAGE and contributes to understanding the overall physiology of ovarian tissue in all these contexts.

The example of ADAMTS1 further illustrates this notion (Fig. 3B). This gene is related closely to the cleavage of extra-membranous domains and activation of AREG in granulosa cells in response to the LH surge (Sayasith et al. 2013). GranulosalIMAGE shows that the dynamics of its response to FSH stimulation and LH surge are very similar to those of AREG. GranulosalIMAGE provides the additional information that expression of this gene is increased significantly during follicular growth and atresia in small (6 mm) but not large follicles (>9 mm), and more in pre-ovulatory follicles of older than younger cows. This observation of the effect of age on the transcript abundance of a gene that is usually studied in association with ovulatory response and post-LH surge cumulus expansion is very interesting.

Construction of customized profiles

GranulosalIMAGE indicates relative gene expression in terms of mRNA abundance under different conditions of folliculogenesis. This manner of presenting expression kinetics provides a basis for the construction of customized profiles by users in the form of simpler illustrations. We present here two examples: (1) temporal kinetics during different physiological states (growth, plateau and atresia), dominance and relevance to LH surge (Fig. 4) and (2) gene dynamics at different intervals and energy states during the post-partum period (Fig. 5).

1. Gene dynamics during folliculogenesis: First, we present the genes with significant roles in ovarian physiology (e.g. AREG and ADAMTS1) and the FSH and LH receptors (FSHR and LHCGR). The abundance of AREG, FSHR and LHCGR transcripts remains unchanged from the growth to plateau stages, whereas the transition to atresia induces down-regulation of LHCGR (in follicles 6–9 mm in diameter), whereas FSHR and AREG are unaffected. An interesting continuous decrease in ADAMTS1 expression is noted during all of these phases. The illustration shows that the dominance stage is characterized by a relatively greater abundance of

---

**Figure 4** Illustration of gene dynamics during folliculogenesis as represented in GranulosalIMAGE. Dynamics of four genes (FSHR, LHCGR, AREG and ADAMTS1) were consulted initially in GranulosalIMAGE and then represented graphically vs follicular developmental stage, either destined to ovulate (grey line) or undergoing atresia (green line). The x-axis values are arbitrary units. Users may consult the dynamics of any gene of interest and may illustrate these versus their preferred follicular parameters (development, super-stimulation and effect of age or various post-partum metabolic states).
Gene expression profiles of granulosa cells

R109

www.reproduction-online.org

Reproduction (2016) 151 R103–R110

LHCGR and a drop in ADAMTS1. The next major shift in gene expression profiles is due to the LH surge. LHCGR transcripts are very abundant 2 h before the surge and drop continuously during the post-LH period. A transient drop in FSHR mRNA is seen in the period spanning 2 h before the surge to 6 h after, followed by an increase measured 22 h after the LH surge. We note with interest that AREG and ADAMTS1 transcripts mirror FSHR expression, with transient up-regulation 6 h after the LH surge.

Illustrations such as Figs 4 and 5 are helpful for following ovarian physiological response kinetics in a variety of different contexts. Known phenomena such as the loss of LHCGR during atresia (Webb et al. 2003) suggest the arrest of follicular growth and a shift towards follicular atresia. Likewise, the acquisition of LHCGR during dominance and a further increase before the LH surge suggest tight regulation of these events during the pre-ovulatory phase (Webb et al. 2003). Furthermore, the expression of ADAMTS1 and AREG relative to LHCGR affirms their roles in inducing cumulus expansion and ovulation following the LH surge. Presentation of gene dynamics in relationship with follicular development allows the exploration of time-course changes in ovarian physiology and hence the generation of new hypotheses.

2. Gene dynamics during the post-partum period: Our database contains studies covering granulosa gene expression profiles from 30 to 120 days post-partum. In addition, the effects of energy status (BHB level) (Fig. 5A) and vitamin supplementation are analysed at approximately 60±5 days post-partum (Fig. 5B). Here, we illustrate the expression dynamics of cyclin beta-1 (CCNB1), LH receptors (LHCGR), forkhead box O-1 (FOXO1) and pentraxin (PTX3). Transcript abundance shows that LHCGR and FOXO1 activities increase significantly from 30 to 60 days post-partum, whereas CCNB1 and PTX3 remain unchanged. We note that none of these genes was subject to maternal energy status (BHB level) 60 days post-partum. At this time point, increased expression of LHCGR and PTX3 decreased CCNB1 and no effects on FOXO1 are observed in response to vitamin supplementation. The transitions from days 60 to 90 and then days 90 to 120 post-partum denote a steady decline in the expression levels of LHCGR and FOXO1 and relative up-regulation of PTX3, whereas CCNB1 remains unchanged.

The most striking feature of this illustration is the down-regulation of CCNB1 and up-regulation of LHCGR at 60 days post-partum (pre-LH group) in response to vitamin supplementation. In fact, CCNB1 is involved in cell cycle regulation. Its down-regulation in this group indicates a relatively more differentiated cell state in which LH receptors are more abundant. This diagram supports the conclusions drawn by Gagnon et al. (2015), suggesting that vitamin supplementation alters post-partum follicular dynamics.
Conclusions
The ovarian follicle is a remarkable structure having diverse functions and highly complex and dynamic physiology. Understanding ovarian physiology is essential in order to optimize female fertility, and huge amounts of data on ovarian tissues have been generated. However, these data are scattered in databases that are difficult to dig in, creating a need for novel ways of integrating and presenting the information for the purpose of advancing knowledge in this field. We present here GranulosalIMAGE, a web-based interface that provides gene expression profiles of granulosa cells from a new perspective. It is an interactive, easy-to-access resource for researchers in the field of ovarian physiology. This is the first step towards integration of various time points of interest in the reproductive cycle. For the moment, GranulosalIMAGE presents only the transcriptomic data that have been produced using EmbryoGENE platform. Although due to intricate technical constraints GranulosalIMAGE does not include the data produced by various groups using microarray platforms other than the EmbryoGENE, we look forward to include RNAseq data in this tool that could be filtered in a relatively more homogenous manner. It also provides a preliminary basis for comparing different follicular tissues such as theca cells, cumulus cells, and oocytes, which are also becoming increasingly available.

Declaration of interest
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding
This study was funded by Natural Science and Engineering Research Council of Canada (NSERC) as part of EmbryoGENE Network and was conducted in collaboration with Boviteq Inc.

Acknowledgements
The authors acknowledge all the researchers who participated in the EmbryoGENE program and whose work has served to construct GranulosalIMAGE.

References
Sirad MA, Desrosier S & Assidi M 2007 In vivo and in vitro effects of FSH on oocyte maturation and developmental competence. Theriogenology 68 (Supplement 1) S71–S76. (doi:10.1016/j. theriogenology.2007.05.053)