Opportunities and challenges in applying genomics to the study of oogenesis and follicullogenesis in farm animals

Agnès Bonnet, Rozenn Dalbies-Tran, Marc Sirard

REPRODUCTION
REVIEW

Focus on Mammalian Embryogenomics

Opportunities and challenges in applying genomics to the study of oogenesis and folliculogenesis in farm animals

A Bonnet, R Dalbiès-Tran[1,2,3,4] and M A Sirard[5]

INRA UMR Génétique Cellulaire, BP52627 Chemin de Borde Rouge, 31326 Castanet Cedex, France, [1]INRA, UMR85 Physiologie de La Reproduction et des Comportements, F-37380 Nouzilly, France, [2]CNRS, UMR6175, F-37380 Nouzilly, France, [3]Université de Tours, F-37041 Tours, France, [4]Haras Nationaux, F-37380 Nouzilly, France and [5]Centre de Recherche en Biologie de La Reproduction, Pavillon des Services, Université Laval, Québec, G1K 7P4 Canada

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

Abstract

Ovarian oogenesis and folliculogenesis are complex and coordinated biological processes which require a series of events that induce morphological and functional changes within the follicle, leading to cell differentiation and oocyte development. In this context, the challenge of the researchers is to describe the dynamics of gene expression in the different compartments and their interactions during the follicular programme. In recent years, high-throughput arrays have become a powerful tool with which to compare the whole population of transcripts in a single experiment. Here, we review the challenges of applying genomics to this model in farm animal species. The first limitation lies in limited the availability of biological material, which makes the study of the follicle compartments (oocyte, granulosa cells and thecal cells) or early embryo much more difficult. The concept of observing all transcripts at once is very attractive but despite progress in sequencing, the genome annotation remains very incomplete in non-model species. Particularly, oogenesis and early embryo development relate to the high proportion of unknown expressed sequence tags. Then, it is important to consider post-transcriptional and translational regulation to understand the role of these genes. Ultimately, these new inferred insights will still have to be validated by functional approaches. In addition to *in vitro* or *ex vivo* functional approaches, both 'natural mutant' ewe models and RNA interference represent, at the moment, the best hope for functional genomics. Advances in our understanding of reproductive physiology should be facilitated by gene expression data exchange and translation into a better understanding of the underlying biological phenomena.

Reproduction (2008) **135** 119–128

Introduction

In the past decade, the advent of expression genomics has aroused a novel approach for the study of biological systems. Through high-throughput arrays and comparative genomics, it has become possible to detect tens of thousands of transcripts and compare their expression between samples in a single experiment. Thus, scientists can now expect to generate a global molecular footprint of their favourite model. Using custom-developed bioinformatics tools to define signalling cascades, scientists hope to unravel the complex molecular mechanisms underlying a biological process. Reproductive biologists were quick to realise the potential of this approach. In this paper, we will review the recent advances of gene expression profiling in the follicle, oocyte and early embryo of domestic animals, along with the hurdles of applying genomics to these models.

Promises and limitations of genomics

The concept of observing all transcripts at once is very attractive; however, it has faced several limitations, and to this day remains an ideal. Despite recent and spectacular progress in the sequencing and assembly of the genomes of several farm animal species including the cow, genome annotation remains very incomplete in these non-model species as opposed to the mouse, for example. Many transcripts have not been discovered experimentally, and

This article was presented at the 2nd International Meeting on Mammalian Embryogenomics, 17–20 October 2007. The Co-operative Research Programme: Biological Resource Management for Sustainable Agricultural Systems of The Organisation for Economic Co-operation and Development (OECD) has supported the publication of this article. The meeting was also sponsored by Le conseil Régional Ile-de-France, the Institut National de la Recherche Agronomique (INRA), Cogenics-Genome Express, Eurogentec, Proteigene, Sigma-Aldrich France and Diagenode sa.

ISSN 1470–1626 (paper) 1741–7899 (online)
DOI: 10.1530/REP-07-0331
Online version via www.reproduction-online.org

thus the genes themselves can only be predicted based on automated computational sequence analysis and cross-species homology. Predictions, however, are limited in nature: they are not always correct, or they miss alternatively spliced forms, hence there are frequent updates in genome databases. Other genes escape prediction altogether and remain hidden in the genome. Therefore, the creation of a comprehensive pangenomic array is still not possible because it cannot be designed solely on annotated genes. While true in genomics in general, this problem is emphasised in oocyte/embryogenomics as many functionally important genes are 'oocyte-specific' genes that have yet to be characterised, and this will be discussed again later in this review. In addition, transcriptome analysis should not be restricted to protein-coding genes, as confirmed by recent evidence of microRNAs being involved in early development (Murchison *et al.* 2007, Tang *et al.* 2007). Tiling arrays, which cover the entire DNA sequence without considering gene annotation, allow unbiased interrogation of genomic loci, but they are so far available only in human and model species. Apart from limitations associated with the arrays themselves, a second constraint lies in the sensitivity of the technique. Should they indeed be represented within the complex probe, rare or shorter transcripts may not be detected onto microarrays, due to at least two major reasons. First, due to hybridisation kinetics, rare transcripts may not bind to their target. Secondly, even if binding occurs, they may not generate a sufficient signal for detection; the same holds true for shorter transcripts such as microRNAs, as labelling is usually proportional to the probe length.

Microarray approaches describe dynamic changes in transcript levels closely related to regulatory events in gene expression without offering any explanation on how this is all managed by the genome and its interactions with environmental conditions. The common challenge faced by the researchers is to translate these lists of differentially expressed genes into a better understanding of the underlying biological phenomena. As already mentioned, a first step for the transformation of gene expression data into biological information is to improve farm animal genome annotation. There are still numerous genes whose identity and function are unknown and some annotations may also be imprecise and incorrect (Khatri & Draghici 2005). The use of clustering analysis to find clusters of co-expressed genes is a powerful tool to provide new clues for the functional annotation of unknown expressed sequence tags (ESTs) or poorly annotated genes. However, the major challenge remains to associate and integrate information (cellular and chromosome location, molecular function, biological process and disease) on thousands of well-annotated genes into an overall picture. Researchers are faced with abundant information about gene products stored in online searchable databases or literature. Bioinformatic tools are now able to extract and summarise them to provide well-organised and relevant information. In addition, some complementary approaches have already screened the functional relationships among genes and included them into larger networks. These approaches have allowed (1) inference of gene networks and metabolic pathways from available data (Bonnet *et al.*, unpublished observations), (2) discovery of new gene relationships involved in a precise biological process and (3) in combination with genome sequences, promoters and transcription factors, the ability to infer upstream regulatory networks (Werner 2007).

Even though not comprehensive, existing arrays generate massive datasets. This has required elaborating novel algorithms of statistical analysis to evaluate false discovery rates and significance of observed variations; ANOVA has recently been integrated and tools made freely available to the scientific community (Hennequet-Antier *et al.* 2005). After quantitation, the next step is supervised or non-supervised hierarchical clustering of genes based on transcripts co-variation; this is also instrumental in making sense of the data (see below).

Nevertheless, the transition from lists of differentially expressed genes to functional profiles has limitations related to (1) the information available via the integrative databases, which is mainly related to humans and rodents and is required to be reorganised in the context of the specific organ and species of reference for optimal utility; and (2) the problem of non-standardised identifiers from one database to another, which can potentially affect the results. To achieve complete integration of genomic information, it will be useful to include the information provided by genetic models (e.g. hyperprolific pigs and sheep, twin cattle) and transgenic animals. Such information is available through an ovarian database (Ovarian Kaleidoscope Database) which accumulates mice mutant phenotypes, scientific literature, gene sequence and chromosomal position information (Ben-Shlomo *et al.* 2002). Furthermore, meta-analysis should prove to be useful in assembling the large number of related results that have been generated by different experiments investigating similar biological conditions to get an overview of the main variation in profiles across ordered comparisons, transcriptional networks or of a particular biological function (Hennetin & Bellis 2006). Overall, the integration of these databases and information will contribute to a better understanding of the biological processes being studied. Ultimately, these new inferred insights will still have to be validated by functional approaches.

Challenges facing the study of reproduction from folliculogenesis to the early embryo

Folliculogenesis

Ovarian folliculogenesis is a complex biological process difficult to analyse because of its dynamic nature, the involvement of different cell types and their interactions (Fig. 1).

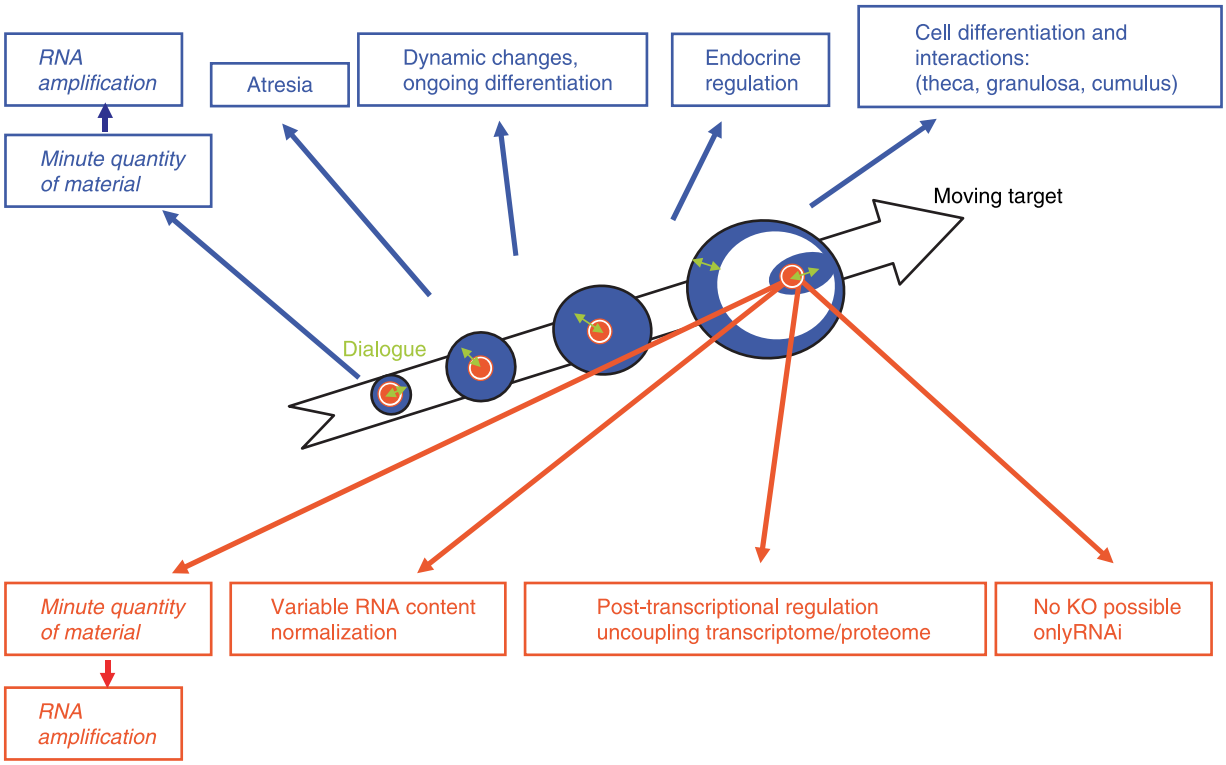

Figure 1 Challenges to the application of genomic tools.

The growth and development of ovarian follicles requires a series of coordinated events that induce morphological and functional changes within the follicle, leading to cell differentiation and oocyte development. The processes of recruitment and selection of the follicles are sequential and under closely coordinated endocrine and paracrine regulation, resulting in the development of a number of ovulatory follicles that are species and breed dependent (Hunter *et al.* 2004). Also, only a few follicles (<1%) will ovulate, while most will undergo a degenerative process called atresia. The fate of follicles during follicular growth is determined by the balance between many processes (proliferation/differentiation/atresia). In this context, the challenge for researchers is to identify the changes associated with each of these processes and the switches in the pathways that will determine the fate of follicles. Transcriptomic approaches should allow the description of gene expression profiles associated with follicle status (healthy/atretic) and different stages of folliculogenesis, as well as a gene signature of the switches involved in this process.

Throughout their dynamic growth, the compartments of a follicle will evolve from the recruitment phase to reach an ovulatory follicle composed of three very distinct compartments: the oocyte, granulosa cells and theca cells. In addition, we can distinguish subcellular populations such as the cumulus granulosa and mural granulosa cells and internal and external theca cells. A comprehensive analysis of folliculogenesis must include the changes in gene expression in all compartments. The experimental limitation of this biological system concerns the small sample quantity available, which makes the study of isolated follicles and their compartments much more difficult, specifically in the early stages of follicular growth. For this reason, most of the work so far has been achieved using the whole ovary for these early stages and either the whole follicle (Agca *et al.* 2006) or granulosa cells (Ndiaye *et al.* 2005) for antral stages. Granulosa cells are in fact the easiest to separate from the follicle at these late stages. Combination of different new technical approaches like microdissection, amplification (see below) and microarray analysis should make it possible to draw a fine description of the molecular mechanisms of the dynamic processes occurring in these different compartments during follicular development.

The follicular programme is also regulated by extra- and intra-ovarian interactions, which adds another level of complexity to this system. The pituitary gonadotrophins, luteinizing hormone (LH) and follicle-stimulating hormone (FSH), provide the primary mechanisms that control selection and dominance via feedback loops with the hypothalamo–pituitary unit. FSH is the main hormone controlling follicular growth in cattle, sheep and pigs and its secretion is in turn controlled via the main secretory products of a large dominant follicle(s), oestradiol and inhibin A (Hunter *et al.* 2004). Within the ovary, cellular interactions between the compartments are essential on the one hand to control the differentiation and proliferation of the follicular cells, and on the other hand to allow

oocyte growth and final differentiation. The interaction between granulosa and theca cells is required for oestrogen biosynthesis. Under influence of LH and FSH respectively, theca cells produce androgens that are used by the granulosa cells to synthesise oestrogens, which in turn protect the dominant follicle(s) from atresia. Intercellular communication between oocytes and granulosa cells occurs via paracrine signalling and gap-junctional exchange of small regulatory molecules. The oocyte secretes soluble paracrine factors that act on its neighbouring granulosa cells, which in turn regulate oocyte development in a bi-directional communication axis (Gilchrist *et al*. 2004). Additionally, interfollicular relationships affect the number of ovulatory follicles present in monotocous or polytocous species. In polytocous species like pigs, the more developed follicles may act to enhance development of the less developed; whereas in monotocous species like the cow, the dominant follicle produces factors that can inhibit follicular development of subordinate follicles (Hunter *et al*. 1992). Once the transcriptome of the different compartments will be characterised, the regulation of genes coding for secreted proteins should be combined with proteomic analysis of the follicular fluid in order to provide a comprehensive picture for the understanding of the interactions involved in folliculogenesis.

Functional approaches

A large number of knockout mice for genes putatively involved in reproductive function have been produced and have aided in the characterising the function of these genes. For example, they have dramatically enhanced the knowledge in gonadotrophin signalling (Rulli & Huhtaniemi 2005), the role of steroids (Drummond 2006) and Smad signalling (Kaivo-oja *et al*. 2006). From primordial germ cell formation through to ovulation and corpus luteum formation, not only stage-specific but also wide-range action genes have been revealed (Barnett *et al*. 2006). However, functional evidence from mouse knockout models cannot be systematically applied to other mammals and thus the production of transgenic large animals would be necessary to assess gene function *in vivo*. Such animals have been produced for more than 20 years in different farm species (Melo *et al*. 2007). However, the fact that true embryonic stem cells do not exist for farm animal species, the technical difficulty and cost of making null transgenics limits the possibility of gene knockouts at this time. These technical problems add to the extended time frame needed to obtain and phenotype adult farm animals with respect to reproductive traits and performance. Remarkably, several natural loss-of-function mutations have been discovered in ewes exhibiting prolific phenotypes, in the BMPR-1B, bone morphogenetic protein 15 (BMP15) or growth and differentiation factor 9 (GDF9) genes (McNatty *et al*. 2005, Bodin *et al*. 2007). These 'natural transgenics' highlight the great advantage of this kind of model in farm animal species. As an alternative, *in vitro* or *ex vivo* functional approaches have been employed in farm animals at the present time.

Since the ovarian follicle is a complex structure composed of different types of cells and the oocyte that are constantly changing and differentiating, *in vitro* culture models are difficult to set up. However, the existing *in vitro* culture models, whether a single cell-type (primary granulosa or cell-line), co-cultures of theca and granulosa cells or whole follicle cultures, are a good way to control the number of variables in an experiment and even more to compare the behaviour of one specific follicular component in different environments. Granulosa cell culture has been an intensively studied model because of their convenient isolation and also because of their sensitivity to FSH, which has been demonstrated *in vivo* to be crucial for ovarian development. In porcine, bovine and ovine species, they constitute a well-standardised model for the *in vitro* study of granulosa cells function, including hormonal regulation. Co-cultures of theca and granulosa cells have also been created mainly for the porcine (Shores *et al*. 2000, Brankin *et al*. 2003) and bovine (Yada *et al*. 1999, Allegrucci *et al*. 2003, Tajima *et al*. 2006) species. These models seem closer to *in vivo* follicles than either granulosa or theca cells cultured alone.

However, it has been long established that somatic cells support oocyte growth and, more recently, that bidirectional interactions between oocyte and somatic cells control folliculogenesis. *In vitro* growth culture systems where oocyte and somatic cells are cultured together have been developed in both pig and bovine species, albeit with limited success (Wu *et al*. 2001, Hirao *et al*. 2004, Hashimoto *et al*. 2007). In these species, culture has been limited to cumulus–oocyte complexes issued from antral follicles, whereas earlier stages can be grown in mouse. As in mouse models, the presence of gap junctions between oocyte and granulosa companion cells are essential and the structure of the follicle must be preserved by adequate culture conditions. To date, progress in culture systems that can support the growth and the development of the whole preantral follicles in domestic species has been also slow for domestic species when compared with mice, for which a complete growth can be achieved resulting in the production of live mice (O'Brien *et al*. 2003). This difference could result from the much longer folliculogenesis in domestic species, which requires a longer culture time when compared with the mouse. In order to get around these difficulties to improve the success rate, more complex alternatives have been set up:

- The activation of primordial follicles *in vitro* using either an organ or a strip culture system has been achieved in several species such as mice and cattle (Wandji *et al*. 1996), but for large animal follicle development rarely progresses to the secondary stage.

- Ovarian cortical pieces including primordial follicles from cattle have been maintained *in ovo*, i.e. grafted beneath the corioallantoic membrane of chick embryos (Cushman *et al*. 2002) and even if spontaneous activation does not occur, this system might be used to study putative stimulators and inhibitors. Nevertheless, some studies have reported the growth from ovine and bovine preantral follicles to antral stage in precise culture conditions (Cecconi *et al*. 1999, Itoh *et al*. 2002).
- Xenotransplantation of ovarian cortical slices from pig or cattle (Kaneko *et al*. 2003, Senbon *et al*. 2003) in immunodeficient SCID mice have also been attempted; some secondary follicles developed to the antral stage and the corresponding oocytes acquired the meiotic competence.

Despite the difficulty in achieving complete maturation for farm animals, culture systems remain useful tools for the study of species-specific mechanisms of folliculogenesis.

The epigenome

At the present time, the study of the transcriptome may not be enough to understand the molecular mechanisms of early development. Recent evidence, much of it related to culture effects and cloning, is pointing at the sensitivity of the gametes and early embryo chromatin to epigenetic modifications. Although it has been known for years that the female environment has an impact on later generation (Barker hypothesis, Finland famines), it is only recently that clues are emerging to associate these external factors with gene expression profiles. In farm animals where there is a defined control of food intake, it becomes important to assess the impact of the maternal metabolic status during late oogenesis and throughout the early foetal development period with an objective to optimise offspring health. The impacts of *in vitro* culture or pregnant mares serum gonadotrophin in the mouse are also known examples of epigenetic cellular responses. Contrary to somatic tissues where most of the epigenetic marking is quite stable and can be defined especially by the DNA methylation status at the CpG islands, the gametes and early embryos have a much less defined pattern. The methylation of DNA is even quite variable according to the species up to the blastocyst stage with high demethylation in mouse compared with larger mammals (Santos & Dean 2006). These epigenetic markings are believed to affect the transcriptome and therefore considerably enhance the complexity of the well-synchronised early development programme that differs between species. At this time, without tiling arrays for farm animals (see above) as well a very limited knowledge of the regulation of gene expression through chromatin modifications, care must be taken in the analysis of follicular or culture conditions affecting later embryo competence.

The oocyte and the early embryo

Transcriptomic analysis of the oocyte presents its own challenges, as summarised in Fig. 1.

Oocyte-specific genes

Early studies in oocyte/embryogenomics have revealed a high proportion of unknown ESTs, primarily those not found in somatic tissues (Rothstein *et al*. 1992, Ko *et al*. 2000, Monk *et al*. 2001). Several factors had contributed to this unexpected outcome. First, oocytes represent only a minor proportion of the cell population within the ovary; therefore, the majority of oocyte-specific transcripts were not represented in whole organ (ovary) libraries. Secondly, the dynamic nature of their regulation allowed them to be detected only at precise stages of folliculogenesis or embryo development. Recently, hundreds of oocyte-specific EST have been discovered using bio-informatics, and many remain uncharacterised despite progress in mouse genome annotation (Paillisson *et al*. 2005). This obviously complicates the construction of causative pathways. It is suspected that such uncharacterised genes might be involved in some unique functions of the oocyte, like reprogramming. They will require special investigation to obtain a complete functional map of the oocyte.

The hypothesis that the so-called oocyte-specific genes would be important for fertility was confirmed by the discovery of reproductive phenotypes in females displaying altered sequence or expression of such genes. In the mouse, functional genomics experiments have delineated the specific stage when the reproductive process becomes impaired through loss of specific gene function (as reviewed in Choi & Rajkovic 2006, Zheng & Dean 2007). A first class of genes is required for normal folliculogenesis. A specific basic helix–loop–helix, best known as factor in the germline alpha (*Figla*), is required to initiate folliculogenesis in the post-natal ovary, while NOBOX oogenesis homeobox (*Nobox*) and *Gdf9* are necessary for follicles to develop beyond the primordial and primary stages respectively. Other oocyte-specific genes are maternal effect genes, i.e. their absence does not affect ovulation nor fertilisation but females are sterile due to preimplantation embryonic failure. Thus, embryos from mothers null for Zygote arrest 1 (*Zar1*), maternal antigen that embryos require (*MATER*; official name NACHT, leucine-rich repeat and PYD containing 5 (*Nalp5*)) are arrested at the one- or two-cell stage respectively. The phenotype is not as early for developmental pluripotency-associated 3 (*Dppa3*), also known as *Stella*, but embryos rarely reach the blastocyst stage, likely due to a defect of the maintenance of methylation involved in epigenetic reprogramming (Nakamura *et al*. 2007). Beyond the mouse model, heterozygous carrier ewes of natural mutations in *BMP15* have an increased ovulation rate and ultimately litter size, while

homozygous mutants are infertile due to impaired oocyte development (Galloway *et al*. 2000, Bodin *et al*. 2007).

Until recently, very few oocyte-specific transcripts had been discovered in non-mouse mammals, and therefore, despite their potential functional importance, most genes were not to be represented in microarrays, whether created 'in house' or commercially. For example, *ZAR1* or *DPPA3* do not appear on the 23K Affymetrix Bovine Genome Array. This has prompted several targeted approaches towards evidencing such genes in bovine. Several orthologues of known mouse oocyte-specific genes were cloned in the cow (Pennetier *et al*. 2004, Dalbies-Tran *et al*. 2005, Uzbekova *et al*. 2006). Libraries of ESTs preferentially represented in oocytes when compared with somatic tissues were generated by suppressive subtractive hybridisation (Pennetier *et al*. 2005, Vallee *et al*. 2005). Alternatively, oocyte preferential EST were revealed through differential hybridisation of a bovine oocyte cDNA library with probes from foetal ovaries versus liver or spleen (Yao *et al*. 2004). Altogether, these efforts have isolated hundreds of bovine oocyte markers, including many novel or non-annotated loci. Evolutionary conservation of these gene sequences and their restricted expression has been investigated. Phylogenetic analysis has indicated that germ cell-specific genes evolve faster than ubiquitous genes (Paillisson *et al*. 2007). Nevertheless, cross-species hybridisation has identified genes conserved between *Xenopus*, mouse and cow; they are thought to be involved in evolutionarily conserved molecular mechanisms (Vallee *et al*. 2005).

Minute amount of biological material

One of the challenges associated with the study of oocytes and embryos in farm animals relates to the very small amount of tissues available. It has been estimated that the oocyte contains ~2 ng RNA and most molecular biology protocols for transcriptome analysis require 1000 times this amount. To circumvent this limitation, amplification protocols have been developed. A first protocol is based on amplification of the entire cDNA pool by PCR. Another method is amplification of RNA through *in vitro* transcription by T7 RNA polymerase; it can be used to amplify RNA in a supposedly more linear fashion. Recent studies have compared these two approaches for amplification distortions and conclude that both methods have a good reliability but may produce amplification disturbances particularly with scarce transcripts (Schneider *et al*. 2004).

Datasets generated from microarrays can be used for cluster analysis to identify transcripts possibly co-regulated during early embryo development, which becomes a powerful tool to study gene function and regulation. However, the period of early embryo development through to the blastocyst stage is characterised by drastically different stage-specific samples. If embryos are recovered just prior to the maternal to embryonic transition (MET), their level of RNA is often very low, several times lower than the germinal vesicle (GV) oocyte (Bilodeau- Goeseels & Schultz 1997). The global amplification methods mentioned above have been validated by assessing the relative abundance between transcripts within the sample (inner sample bias); however, their efficiency to maintain the initial differences between samples bearing natural differences such as early embryos from different developmental stages has not been validated (inter samples bias). Recently, a MET relevant microarray was produced using subtractive suppressive hybridisation reactions to collect candidate ESTs that were printed on glass slides (Sirard *et al*. unpublished data). The microarray was hybridised with T7 amplified probes from GV oocytes and eight-cell embryos. Data analysis using standard methods lead to an apparent over-expression of almost all transcripts found in the eight cells. Not surprisingly, none of these differential levels of expression were confirmed by qPCR, indicating an over-amplification of the eight-cell sample.

To elucidate the cause of this skew, the kinetics of the amplification reactions were investigated using different approaches (Gilbert, Sirard and Robert, unpublished observations). The first amplification round was followed by the quantification of two candidates from aliquots of the reaction mix by qPCR. Since the second amplification round fragments the target sequences that impair the qPCR assessment, the reaction was monitored by a time-course analysis of radioactivity incorporation. Their findings show that the T7 *in vitro* transcription reaction reaches a plateau phase; moreover, the linearity of the reaction was limited to the first 80 min of the 6-h protocol in the first amplification and for about 180 min for the second round. Therefore, the reaction time has now been adjusted to these linear portions to avoid the introduction of the inter-sample bias. This procedure results in variable output of amplified RNA relevant to the differences in the amount of input RNA between the different developmental stage samples. The probes generated using this time-controlled protocol allowed comparisons across embryo developmental stages and cluster analyses. Since the amplified probes contained different amounts of labelled molecules, care was taken to avoid another bias that would have been introduced by the lower signals that would be more prevalent in the eight-cell probe. Therefore, probes were spiked with a control and equal amounts of probes were hybridised on the microarrays. This spiked control served to correct for the over-representation of the eight-cell probe. In addition to the spike, these studies used several internal, external and amplification controls that were assessed by qPCR to validate the procedure. This creative standardisation approach arises from the additional challenges that do not exist in mouse embryos where the MET occurs very early and where the level of transcript may not be as variable as for larger mammals with delayed MET.

The uncoupling between transcription and translation

Vigneault *et al*. (2004) have shown that several transcription factors would display quantitative variations between the one- and eight-cell stages in bovine embryos. When the immunohistochemistry was performed on several factors, it became clear that an inverse relationship was present between transcript abundance and protein abundance. Indeed, the proteins appear as the transcript disappears in contrast with what is most often observed in somatic cells where active transcription is concomitant with active translation. Other examples are *MATER* or *AURORA-A* whose transcripts and proteins also displayed divergent profiles (Pennetier *et al*. 2006; Svetlana Uzbekova personal communication). However, this reverse correlation between transcript and protein abundance during the pre-MET period makes sense considering that maternally stored RNA pools sustain protein synthesis in the absence of embryonic transcription. Even more convincing was the observation that some transcription factors appeared in the embryo cytoplasm two full cell cycles before their movement to the nuclei. In this case, the transcript abundance was at its lowest when the putative protein action was present. These facts must be considered when conclusions are drawn regarding the impact of a given transcript level at a given time. This is especially true before the MET in large mammals.

The study of protein profiles during oocyte maturation or early embryogenesis in the bovine (Coenen *et al*. 2004, Bhojwani *et al*. 2006, Massicotte *et al*. 2006) indicates hundreds of new products before and after fertilisation as well as many proteins that disappear. Since there is no transcription following the germinal vesicle break down and possibly a very low level after fertilisation (Memili & First 1999), these changes must rely on fine regulation of hundreds of transcripts by a mechanism other than transcription. The 3′UTR seems to play a particular role in this regulation but only a subset of the signals are understood today (maternal or embryonic polyadenylation signals, cytoplasmic polyadenylation element (CPE); for a review see Richter 2007).

How does one study the special regulation that maintains some RNAs and send others for processing into proteins? The oocyte, in contrast to the somatic cell, has the capacity to store RNA for long periods. This phenomenon should facilitate the study of these stored RNAs, but on the contrary, this complicates the interpretation since for most genes no apparent link exists between RNA content and protein level or activity. Therefore, it is important to study the regulation of translation in oocyte to achieve an understanding of the role of these genes.

Distinct RNAs bear polyA tails of distinct lengths and many maternal RNAs are regulated during maturation and early embryonic development through shortening/elongation of the polyA tail (Tremblay *et al*. 2005). One of the issues is that the priming reaction for gene expression studies will generate different gene expression profiles based on the length of the polyA tails. Furthermore, the variations in polyA tail length render the physiological interpretation of transcriptome profiling more difficult. Yet how this variable polyA tail can impact genomic data has only recently been taken into account. Scientists have to be careful about using different methods to prepare their RNA for analysis. For example, the use of poly-dT beads or membranes to capture and purify poly-A RNA will often lead to a partial evaluation of the actual transcript content: the longer the poly-A tail, the more likely that the transcript will be retained, creating a bias against short poly-A tail transcripts. Also, preferential amplification of transcripts with a long versus short polyA tail is expected when RT is primed with oligo(dT; as cyclin B, Tremblay *et al*. 2005), whether for cDNA or RNA amplification. On the other hand, if only the long poly-A tails are associated with active translation, then such profiling would favour (if not be strictly limited to) active transcripts in the considered biological situation and possibly a link to function when isolating specific sub-populations of mRNA. Short poly-A tails often suggest the masking of translation and therefore the use of random primers does not produce a footprint of gene activity, but rather a profile of maternal RNA pools, irrespective of whether the RNA is translated at the considered time, stored for later use or destined for degradation. Fractionation of the mRNA population could be instrumental in order to correctly interpret the data being generated from these studies. All current approaches are considered acceptable, but one should be aware of the limitations of these techniques and take it into account when interpreting the data.

In all cases, pre-MET oocytes or embryos should be analysed for protein content and if possible activity before reaching functional conclusions. Because of the possible uncoupling between the transcriptome and the proteome, one must proceed with care before applying the usual tools in functional genomics as gene ontology (GO terms).

The untranslated portion of oocytes mRNA

The 3′UTR contains information on the regulation of translation for the stored transcripts in oocytes. To define this process, it is important to isolate the recurring patterns, perform cluster analysis and identify the regulatory elements involved. It is clear that polyadenylation is influenced by the elements present in the UTR (CPE, ECPE, etc.), but recent studies in the mouse (Svoboda, personal communication) indicates that the 5′ end of the 3′UTR contains sequences showing stretches of similarities with microRNA present in oocytes. This could imply that the regulation of the translation/destruction could depend on something other than the regulatory proteins identified thus far to interact with the 3′UTR such as maskin or CPEB.

The understanding of this regulation will make cluster analysis more functional until the proteomic tools become sensitive enough to allow for protein profiling in small samples such as mammalian early embryos.

Functional genomic in oocytes

Without an adequate culture system for mimicking oogenesis *in vitro*, functional approaches are limited to fully grown oocytes. Since at this stage the oocyte is transcriptionally silent, the range of functional approaches that can be applied to study oogenesis appears limited. At the moment, RNA interference represents the best hope for oocyte functional genomic studies. Since no promoters have been found to be active before the MET, it is not possible to use hairpin RNA; but since Dicer is present and active, and in absence of interferon response, both long and short double-stranded RNA can be used effectively in the bovine species (Paradis *et al*. 2005) as in the mouse. Target genes should be carefully selected as a noticeable effect is expected only for genes, not for already present instable protein form. Also, because of the long delay between oocyte injection and blastocyst formation, some late transcripts might be more difficult to knock down. Despite several publications on bovine oocytes (Paradis *et al*. 2005, Nganvongpanit *et al*. 2006), the resulting embryos are often developmentally compromised.

Folliculogenesis and oogenesis seem integrated at least in the mouse, where knockouts have shown the multiple links between oocyte and follicle survival. However, in the mouse, the competence of the oocyte is mainly a matter of size, while in domestic animals, its late differentiation during the antral growth of the follicle has a real impact on the ability to develop into an embryo after fertilisation. Therefore, the signalling events inducing the changes in oocyte competence seems to come from the follicle and our knowledge at present is limited. In absence of a physiological follicle culture system, most of the data will need to be generated from studies investigating the temporal and spatial gene expression during follicular and oocyte development in terms of competence. The possibility of individual oocyte culture through standardised IVF protocols and assessment of developmental competence will permit the association of follicular characteristics associated with the competent oocyte. Knock-down studies in oocytes would be the appropriate system in which to validate hypotheses coming from these associations.

Conclusion and perspectives

Genomics is a very powerful tool with which to study reproductive physiology and will certainly become a method of choice to reveal the functioning of the ovary and the female gamete. However, the challenges are numerous as described above. First, the tissues under investigation are always changing throughout the process of folliculogenesis. Secondly, the different tissues or compartment are all functionally related and must be studied in the presence of these interactions. Thirdly, the amount of tissue available is limiting, particularly with oocytes and early embryos still dependent upon maternal RNA. Although many technical problems have been solved, there is a need for communication data exchange and integration amongst scientists to develop the right experimental conditions that are essential for making progress in the understanding of reproductive physiology. Undoubtedly, the new genomic tools will be very helpful in this process but only when adapted to the specific conditions present in these reproductive processes.

Declaration of interest

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

Funding

This article is based on research presented at the 2nd International Meeting on Mammalian Embryogenomics, which was sponsored by the Organisation for Economic Co-operation and Development (OECD), Le conseil Régional Ile-de-France, the Institut National de la Recherche Agronomique (INRA), Cogenics-Genome Express, Eurogentac, Proteigene, Sigma-Aldrich France and Diagenode sa. A B and R D-T are members of an INRA-funded institute. All authors received funding from the OECD to attend the meeting.

Acknowledgements

The authors thank Gwenola Tosser-Klopp, Laurence Liaubet, François Hatey and Claude Robert for their major contribution to this 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.

Allegrucci C, Hunter MG, Webb R & Luck MR 2003 Interaction of bovine granulosa and theca cells in a novel serum-free co-culture system. *Reproduction* **126** 527–538.

Barnett KR, Schilling C, Greenfeld CR, Tomic D & Flaws JA 2006 Ovarian follicle development and transgenic mouse models. *Human Reproduction Update* **12** 537–555.

Ben-Shlomo I, Vitt UA & Hsueh AJ 2002 Perspective: the ovarian kaleidoscope database-II. Functional genomic analysis of an organ-specific database. *Endocrinology* **143** 2041–2044.

Bhojwani M, Rudolph E, Kanitz W, Zuehlke H, Schneider F & Tomek W 2006 Molecular analysis of maturation processes by protein and phosphoprotein profiling during *in vitro* maturation of bovine oocytes: a proteomic approach. *Cloning Stem Cells* **8** 259–274.

Bilodeau-Goeseels S & Schultz GA 1997 Changes in ribosomal ribonucleic acid content within *in vitro*-produced bovine embryos. *Biology of Reproduction* **56** 1323–1329.

Bodin L, Di Pasquale E, Fabre S, Bontoux M, Monget P, Persani L & Mulsant P 2007 A novel mutation in the bone morphogenetic protein 15 gene causing defective protein secretion is associated with both increased ovulation rate and sterility in Lacaune sheep. *Endocrinology* **148** 393–400.

Brankin V, Mitchell MR, Webb B & Hunter MG 2003 Paracrine effects of oocyte secreted factors and stem cell factor on porcine granulosa and theca cells *in vitro*. *Reproductive Biology and Endocrinology* **1** 55–69.

Cecconi S, Barboni B, Coccia M & Mattioli M 1999 *In vitro* development of sheep preantral follicles. *Biology of Reproduction* **60** 594–601.

Choi Y & Rajkovic A 2006 Genetics of early mammalian folliculogenesis. *Cellular and Molecular Life Sciences* **63** 579–590.

Coenen K, Massicotte L & Sirard MA 2004 Study of newly synthesized proteins during bovine oocyte maturation *in vitro* using image analysis of two-dimensional gel electrophoresis. *Molecular Reproduction and Development* **67** 313–322.

Cushman RA, Wahl CM & Fortune JE 2002 Bovine ovarian cortical pieces grafted to chick embryonic membranes: a model for studies on the activation of primordial follicles. *Human Reproduction* **17** 48–54.

Dalbies-Tran R, Papillier P, Pennetier S, Uzbekova S & Monget P 2005 Bovine mater-like NALP9 is an oocyte marker gene. *Molecular Reproduction and Development* **71** 414–421.

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

Galloway SM, McNatty KP, Cambridge LM, Laitinen MP, Juengel JL, Jokiranta TS, McLaren RJ, Luiro K, Dodds KG & Montgomery GW *et al.* 2000 Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner. *Nature Genetics* **25** 279–283.

Gilchrist RB, Ritter LJ & Armstrong DT 2004 Oocyte–somatic cell interactions during follicle development in mammals. *Animal Reproduction Science* **82–83** 431–446.

Hashimoto S, Ohsumi K, Tsuji Y, Harauma N, Miyata Y, Fukuda A, Hosoi Y, Iritani A & Morimoto Y 2007 Growing porcine oocyte–granulosa cell complexes acquired meiotic competence during *in vitro* culture. *Journal of Reproduction and Development* **53** 379–384.

Hennequet-Antier C, Chiapello H, Piot K, Degrelle S, Hue I, Renard JP, Rodolphe F & Robin S 2005 AnovArray: a set of SAS macros for the analysis of variance of gene expression data. *BMC Bioinformatics* **6** 150.

Hennetin J & Bellis M 2006 Clustering methods for analyzing large data sets: gonad development, a study case. *Methods in Enzymology* **411** 387–407.

Hirao Y, Itoh T, Shimizu M, Iga K, Aoyagi K, Kobayashi M, Kacchi M, Hoshi H & Takenouchi N 2004 *In vitro* growth and development of bovine oocyte–granulosa cell complexes on the flat substratum: effects of high polyvinylpyrrolidone concentration in culture medium. *Biology of Reproduction* **70** 83–91.

Hunter MG, Biggs C, Faillace LS & Picton HM 1992 Current concepts of folliculogenesis in monovular and polyovular farm species. *Journal of Reproduction and Fertility* **45** 21–38.

Hunter MG, Robinson RS, Mann GE & Webb R 2004 Endocrine and paracrine control of follicular development and ovulation rate in farm species. *Animal Reproduction Science* **82–83** 461–477.

Itoh T, Kacchi M, Abe H, Sendai Y & Hoshi H 2002 Growth, antrum formation, and estradiol production of bovine preantral follicles cultured in a serum-free medium. *Biology of Reproduction* **67** 1099–1105.

Kaivo-oja N, Jeffery LA, Ritvos O & Mottershead DG 2006 Smad signalling in the ovary. *Reproductive Biology and Endocrinology* **4** 21–33.

Kaneko H, Kikuchi K, Noguchi J, Hosoe M & Akita T 2003 Maturation and fertilization of porcine oocytes from primordial follicles by a combination of xenografting and *in vitro* culture. *Biology of Reproduction* **69** 1488–1493.

Khatri P & Draghici S 2005 Ontological analysis of gene expression data: current tools, limitations, and open problems. *Bioinformatics* **21** 3587–3595.

Ko MS, Kitchen JR, Wang X, Threat TA, Hasegawa A, Sun T, Grahovac MJ, Kargul GJ, Lim MK, Cui Y *et al.* 2000 Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development. *Development* **127** 1737–1749.

Massicotte L, Coenen K, Mourot M & Sirard MA 2006 Maternal housekeeping proteins translated during bovine oocyte maturation and early embryo development. *Proteomics* **6** 3811–3820.

McNatty KP, Galloway SM, Wilson T, Smith P, Hudson NL, O'Connell A, Bibby AH, Heath DA, Davis GH, Hanrahan JP *et al.* 2005 Physiological effects of major genes affecting ovulation rate in sheep. *Genetics, Selection, Evolution* **37** (Supplement 1) S25–S38.

Melo EO, Canavessi AM, Franco MM & Rumpf R 2007 Animal transgenesis: state of the art and applications. *Journal of Applied Genetics* **48** 47–61.

Memili E & First NL 1999 Control of gene expression at the onset of bovine embryonic development. *Biology of Reproduction* **61** 1198–1207.

Monk M, Holding C & Goto T 2001 Isolation of novel developmental genes from human germ cell, oocyte and embryo cDNA by differential display. *Reproduction, Fertility, and Development* **13** 51–57.

Murchison EP, Stein P, Xuan Z, Pan H, Zhang MQ, Schultz RM & Hannon GJ 2007 Critical roles for Dicer in the female germline. *Genes and Development* **21** 682–693.

Nakamura T, Arai Y, Umehara H, Masuhara M, Kimura T, Taniguchi H, Sekimoto T, Ikawa M, Yoneda Y, Okabe M, Tanaka S, Shiota K & Nakano T 2007 PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nature Cell Biology* **9** 64–71.

Ndiaye K, Fayad T, Silversides DW, Sirois J & Lussier JG 2005 Identification of downregulated messenger RNAs in bovine granulosa cells of dominant follicles following stimulation with human chorionic gonadotropin. *Biology of Reproduction* **73** 324–333.

Nganvongpanit K, Muller H, Rings F, Hoelker M, Jennen D, Tholen E, Havlicek V, Besenfelder U, Schellander K & Tesfaye D 2006 Selective degradation of maternal and embryonic transcripts in *in vitro* produced bovine oocytes and embryos using sequence specific double-stranded RNA. *Reproduction* **131** 861–874.

O'Brien MJ, Pendola JK & Eppig JJ 2003 A revised protocol for *in vitro* development of mouse oocytes from primordial follicles dramatically improves their developmental competence. *Biology of Reproduction* **68** 1682–1686.

Paillisson A, Dade S, Callebaut I, Bontoux M, Dalbies-Tran R, Vaiman D & Monget P 2005 Identification, characterization and metagenome analysis of oocyte-specific genes organized in clusters in the mouse genome. *BMC Genomics* **6** 76–86.

Paillisson A, Levasseur A, Gouret P, Callebaut I, Bontoux M, Pontarotti P & Monget P 2007 Bromodomain testis-specific protein is expressed in mouse oocyte and evolves faster than its ubiquitously expressed paralogs BRD2, -3, and -4. *Genomics* **89** 215–223.

Paradis F, Vigneault C, Robert C & Sirard MA 2005 RNA interference as a tool to study gene function in bovine oocytes. *Molecular Reproduction and Development* **70** 111–121.

Pennetier S, Uzbekova S, Perreau C, Papillier P, Mermillod P & Dalbies-Tran R 2004 Spatio-temporal expression of the germ cell marker genes MATER, ZAR1, GDF9, BMP15,and VASA in adult bovine tissues, oocytes, and preimplantation embryos. *Biology of Reproduction* **71** 1359–1366.

Pennetier S, Uzbekova S, Guyader-Joly C, Humblot P, Mermillod P & Dalbies-Tran R 2005 Genes preferentially expressed in bovine oocytes revealed by subtractive and suppressive hybridization. *Biology of Reproduction* **73** 713–720.

Pennetier S, Perreau C, Uzbekova S, Thelie A, Delaleu B, Mermillod P & Dalbies-Tran R 2006 MATER protein expression and intracellular localization throughout folliculogenesis and preimplantation embryo development in the bovine. *BMC Developmental Biology* **6** 26–34.

Richter JD 2007 CPEB: a life in translation. *Trends in Biochemical Sciences* **32** 279–285.

Rothstein JL, Johnson D, DeLoia JA, Skowronski J, Solter D & Knowles B 1992 Gene expression during preimplantation mouse development. *Genes and Development* **6** 1190–1201.

Rulli SB & Huhtaniemi I 2005 What have gonadotrophin overexpressing transgenic mice taught us about gonadal function? *Reproduction* **130** 283–291.

Santos F & Dean W 2006 Using immunofluorescence to observe methylation changes in mammalian preimplantation embryos. *Methods in Molecular Biology* **325** 129–137.

Schneider J, Buness A, Huber W, Volz J, Kioschis P, Hafner M, Poustka A & Sultmann H 2004 Systematic analysis of T7 RNA polymerase based *in vitro* linear RNA amplification for use in microarray experiments. *BMC Genomics* **5** 29–37.

Senbon S, Ota A, Tachibana M & Miyano T 2003 Bovine oocytes in secondary follicles grow and acquire meiotic competence in severe combined immunodeficient mice. *Zygote* **11** 139–149.

Shores EM, Picton HM & Hunter MG 2000 Differential regulation of pig theca cell steroidogenesis by LH, insulin-like growth factor I and granulosa cells in serum-free culture. *Journal of Reproduction and Fertility* **118** 211–219.

Tajima K, Orisaka M, Yata H, Goto K, Hosokawa K & Kotsuji F 2006 Role of granulosa and theca cell interactions in ovarian follicular maturation. *Microscopy Research and Technique* **69** 450–458.

Tang F, Kaneda M, O'Carroll D, Hajkova P, Barton SC, Sun YA, Lee C, Tarakhovsky A, Lao K & Surani MA 2007 Maternal microRNAs are essential for mouse zygotic development. *Genes and Development* **21** 644–648.

Tremblay K, Vigneault C, McGraw S & Sirard MA 2005 Expression of cyclin B1 messenger RNA isoforms and initiation of cytoplasmic polyadenylation in the bovine oocyte. *Biology of Reproduction* **72** 1037–1044.

Uzbekova S, Roy-Sabau M, Dalbies-Tran R, Perreau C, Papillier P, Mompart F, Thelie A, Pennetier S, Cognie J & Cadoret V *et al.* 2006 Zygote arrest 1 gene in pig, cattle and human: evidence of different transcript variants in male and female germ cells. *Reproductive Biology and Endocrinology* **4** 12–25.

Vallee M, Gravel C, Palin MF, Reghenas H, Stothard P, Wishart DS & Sirard MA 2005 Identification of novel and known oocyte-specific genes using complementary DNA subtraction and microarray analysis in three different species. *Biology of Reproduction* **73** 63–71.

Vigneault C, McGraw S, Massicotte L & Sirard MA 2004 Transcription factor expression patterns in bovine *in vitro*-derived embryos prior to maternal–zygotic transition. *Biology of Reproduction* **70** 1701–1709.

Wandji SA, Srsen V, Voss AK, Eppig JJ & Fortune JE 1996 Initiation *in vitro* of growth of bovine primordial follicles. *Biology of Reproduction* **55** 942–948.

Werner T 2007 Regulatory networks: linking microarray data to systems biology. *Mechanisms of Ageing and Development* **128** 168–172.

Wu J, Emery BR & Carrell DT 2001 *In vitro* growth, maturation, fertilization, and embryonic development of oocytes from porcine preantral follicles. *Biology of Reproduction* **64** 375–381.

Yada H, Hosokawa K, Tajima K, Hasegawa Y & Kotsuji F 1999 Role of ovarian theca and granulosa cell interaction in hormone production and cell growth during the bovine follicular maturation process. *Biology of Reproduction* **61** 1480–1486.

Yao J, Ren X, Ireland JJ, Coussens PM, Smith TP & Smith GW 2004 Generation of a bovine oocyte cDNA library and microarray: resources for identification of genes important for follicular development and early embryogenesis. *Physiological Genomics* **19** 84–92.

Zheng P & Dean J 2007 Oocyte-specific genes affect folliculogenesis, fertilization, and early development. *Seminars in Reproductive Medicine* **25** 243–251.

Received 19 July 2007
First decision 3 October 2007
Accepted 12 November 2007