The use of microarrays to define functionally-related genes that are differentially expressed in the cycling pig uterus

J.A. Green, J-G. Kim, K.M. Whitworth, C. Agca and R.S. Prather

University of Missouri-Columbia, Division of Animal Sciences; Columbia, Missouri, USA

In swine and other livestock, the uterine endometrium exhibits dramatic morphological and secretory changes throughout the oestrous cycle and during pregnancy. Such physiological changes are a reflection of extremely complex interactions between gene products (RNA and protein). The recent development of genomics and proteomics methods, as well as associated bioinformatics tools, has provided the means to begin characterising such interactions. Indeed, the analysis of the transcriptome and proteome of cells and tissues now comprises a new field of study known as 'systems biology'. Currently, the most powerful technique available to the systems biologist is the microarray. These platforms represent oligonucleotide or cDNA fragments spotted in a specified high-density pattern on a solid support. Hybridisation of fluorescently-tagged cDNAs from different tissue sources permits the measurement of thousands of RNAs in parallel. The method permits the identification of genes that are present at different amounts between the two tissues and, more importantly, it permits the identification of groups of genes (clusters) that are expressed in comparable patterns. Results from a recent expression profiling experiment are described. The goal of the profiling experiment was to define genes that are differentially expressed in endometrium during the oestrous cycle. The experiment used an in-house cDNA microarray with > 14,000 distinct cDNAs cloned from reproductive tissues. Total RNAs from cyclic endometrium (Days 0, 3, 6, 10, 12, 14 and 18 post-oestrus) were reverse transcribed into cDNAs, labelled with fluorescent dye and hybridised to the arrays along with cDNAs derived from a reference RNA pool. A total of 4,827 genes were found to differ significantly at some time during the oestrous cycle. Clustering methods were able to define numerous groups of similarly expressed genes. These data will help to define the complex patterns of endometrial genes acting in concert to create the environments required for fertilisation, embryo growth and conceptus development in swine.

E-mail: greenjo@missouri.edu

Introduction

The uterine endometrium has long fascinated researchers because of the pronounced changes in morphology and secretory capacity that it undergoes during the oestrous cycle, pregnancy and in response to experimental hormone treatments (Basha *et al.*, 1979; Bazer and Roberts, 1983; Croy *et al.*, 1988; Fazleabas *et al.*, 1982; Trout *et al.*, 1992; Yu *et al.*, 1993). During the first three weeks of pregnancy in livestock species, the endometrium nourishes the developing embryo and, during the 'window of implantation', provides an altered uterine epithelium conducive to the attachment and continued growth of the conceptus (Bazer and Roberts, 1983; Burghardt *et al.*, 2002; Geisert *et al.*, 1982; Gray *et al.*, 2002; MacIntyre *et al.*, 2002). It is also known that the endometrium secretes factors that can regulate conceptus development and gene expression (Ezashi and Roberts, 2004; Imakawa *et al.*, 1997). The coordination of these developmental events is due to the integration of endocrine, paracrine and autocrine signals from the ovary, conceptus and the uterus itself. Together, they reflect tightly regulated changes within the uterine transcriptome (Roberts *et al.*, 1993; Robinson *et al.*, 1999; Spencer and Bazer, 1995; Spencer *et al.*, 2004).

Many published reports have described expression patterns of specific genes during the oestrous cycle and pregnancy in the porcine uterus. From these efforts, a myriad of differentially transcribed RNAs have been identified. Examples include spermine/spermidine N1-acetyltransferase (Green *et al.*, 1998), integrins (Burghardt *et al.*, 2002), extracellular matrix proteins (Johnson *et al.*, 2003; Johnson *et al.*, 2001), growth factors (Geisert *et al.*, 2001; Gupta *et al.*, 1998; Moussad *et al.*, 2002) and numerous progesterone-responsive genes (Clawitter *et al.*, 1990; Malathy *et al.*, 1990; Stallings-Mann *et al.*, 1994). Presumably, these gene products are participating in the establishment and maintenance of pregnancy. However, in studying uterine physiology on a gene-by-gene basis, it is has become quite clear that the uterine transcriptome is exceedingly complex. The recent development of genomic and proteomic methods for large-scale expression analysis, along with associated statistical and bio-informatics tools, have provided the means to analyse genes and their products on a global scale. Indeed, if reproductive physiologists are to gain a thorough understanding of uterine physiology, it will be necessary to identify additional uterus-specific transcripts and to define the global patterns of uterine gene and protein expression during the oestrous cycle and pregnancy.

Systems biology

Each end-product of a gene transcription event (protein or non-coding RNA) does not exist in isolation. Rather, the change in the expression of a given gene product is generally an outcome of earlier transcription in other genes as well as numerous concurrent alterations. Together all these coordinated changes are a reflection of regulatory networks and pathways that control cellular homeostasis and differentiation. The study of such networks has developed into a distinct discipline known as 'systems biology' and its goal is to understand how changes in these genetic networks give rise to cellular and tissue phenotypes (de Bivort *et al.*, 2004; Hieronymus and Silver, 2004).

As recently as 10-12 years ago, the ability to identify and characterise gene expression patterns was limited to only a small number of genes at a time. The methods in place at that time included Northern blotting, reverse-transcription PCR (RT-PCR) and real-time PCR (Holland *et al.*, 1991; Lee *et al.*, 1993; Sambrook *et al.*, 1989). These techniques are still used routinely, but they do have limitations in the number of expression patterns that can realistically be characterised. The recent development of several high-throughput technologies has allowed a more extensive assessment of gene expression.

Such high-throughput techniques include: increased sequence capabilities for Expressed Sequence Tags (EST) projects, Serial Analysis of Gene Expression (SAGE), the development of platforms for monitoring the abundance of thousands of distinct RNAs in parallel (microarrays) and the ability to identify hundreds of proteins and their modifications via protein separation and mass spectrometry (Baldwin *et al.*, 2001; Belghazi *et al.*, 2001; Hegde *et al.*, 2000; Tan *et al.*, 2003; Velculescu *et al.*, 1995). These technologies have begun a revolution in biology. Biologists are approaching the point at which they will have the means to study biological phenomena as they really exist – as networks of overlapping and interacting genes, RNAs and proteins.

Both transcriptome and proteome analysis are linked. However, large-scale proteomic techniques and informatics tools are lagging somewhat behind that for analysis of the transcriptome. Therefore, only an overview of transcriptional profiling (and an example of an ongoing profiling experiment) will be provided in this short chapter.

Transcriptome analysis

The use of microarrays to measure the presence of thousands of RNAs concurrently represents one of the first steps toward the ambitious goals of systems biology. Indeed, the power of microarrays is in two areas. The first is the ability to demonstrate the involvement of genes in a biological process in which they had not previously been implicated. The second is in the ability to group or 'cluster' genes based on expression profiles that are closely associated.

The development of a microarray platform by necessity requires a sequence database from which to draw. Most of the time this database consists of complementary DNA (cDNA) sequences arising from EST projects. In regard to the use of transcriptional analysis in porcine reproduction and development, several published reports describe cDNA libraries produced from reproductive organs and used to generate ESTs to define genes expressed in reproductive tissues and early embryos of pigs (Caetano *et al.*, 2003; Fahrenkrug *et al.*, 2002; Jiang *et al.*, 2004; Tuggle *et al.*, 2003; Whitworth *et al.*, 2004). These, and other, publicly available sequence resources have been used to create microarrays for expression profiling experiments (Caetano *et al.*, 2004; Gladney *et al.*, 2004; Whitworth *et al.*, 2005). Such resources have been, and will continue to be, useful tools for the biologist focused on swine reproduction.

Microarray production (cDNA versus oligonucleotide)

A microarray is generally composed of thousands of distinct DNA strands that are placed in a defined order on a silicon or glass support. The spotted DNAs are chosen to be specific to a particular target gene. The principle behind microarray technology is relatively simple. It is based on the fact that complementary sequences present in the RNA population and immobilised on the support will be able to bind to each other under the proper hybridisation conditions, while non-complimentary sequences will not (Kuo *et al.*, 2004). When a sample contains many copies of a particular mRNA, multiple binding events can take place with the complementary DNA on the array, producing a strong signal that is reflective of a gene that is highly expressed (Kuo *et al.*, 2004).

There are two main types of microarrays: cDNA and oligonucleotide arrays (Churchill, 2002; Dobbin and Simon, 2002). Oligonucleotide microarrays are comprised of spotted DNA strands that can vary from as little as ten to as many as 100 nucleotides in length. Often, the oligonucleotides will have been chosen to be similar in regard to hybridisation parameters (e.g. hybridisation temperature, binding affinity) and are spotted on the array in roughly equimolar amounts (Emrich *et al.*, 2003). Consequently, oligonucleotide microarrays can be used to mea-

sure individual samples to provide an absolute measurement for each RNA molecule. However, rarely does this absolute measurement correspond exactly with the specific concentration of the RNA in the sample. In contrast, cDNA microarrays are comprised of full-length cDNAs or cDNA fragments that are typically 500 to 4000 base pairs in length. Because it is generally impractical to accurately spot cDNAs in an equimolar manner and because each cDNA will have its own hybridisation characteristics, experiments involving cDNA microarrays are most typically designed to measure two samples at a time. The resulting data are then used to define the relative abundance of each RNA molecule in the respective samples rather than to provide an absolute measurement (Butte, 2002; Churchill, 2002; Yang and Speed, 2002). An example is illustrated in Fig. 1; it demonstrates the relative RNA abundance of several genes in samples collected from porcine endometrium on Day 0 and Day 3 of the oestrous cycle.
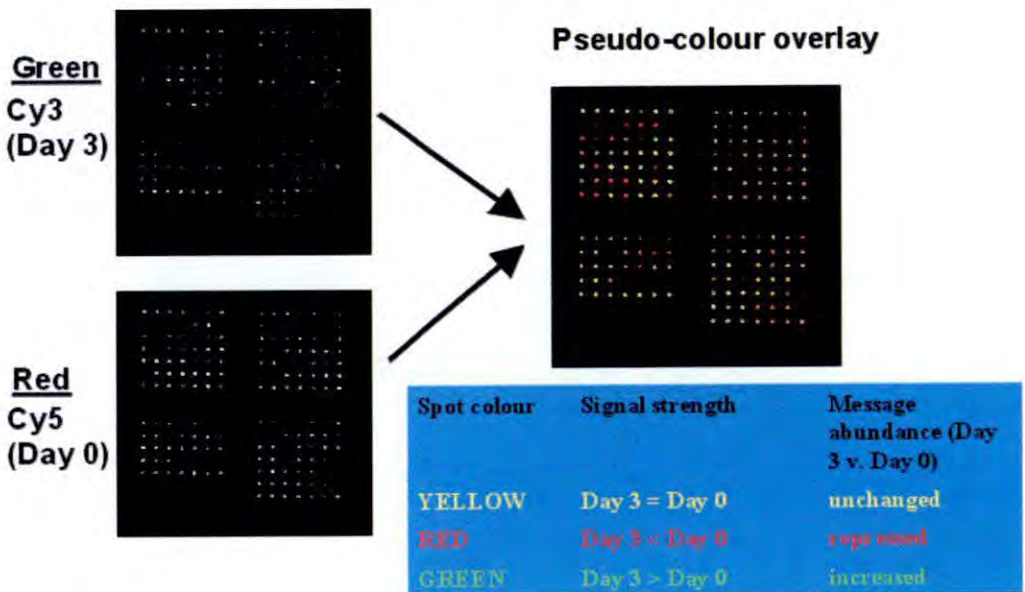

Spot colour	Signal strength	Message abundance (Day 3 v. Day 0)
YELLOW	Day 3 = Day 0	unchanged
RED	Day 3 < Day 0	repressed
GREEN	Day 3 > Day 0	increased

Fig. 1 An example of the use of cDNA arrays to assess relative RNA abundance between tissues at different stages. The panels on the left are fluorescent signals obtained from Cy3- and Cy5-labelled cDNAs from Day 3 and Day 0 endometrium, respectively. The pseudo-coloured overlay of the fluorescent signals is shown in the right panel. Equal fluorescence from the green and red dyes (representing equal abundance of the particular RNA in each sample) is shown as a yellow colour. The presence of a red colour in a spot represents an RNA that is in excess in the Day 0 sample and the presence of a green colour represents an RNA that is in excess in the Day 3 sample.

On overview of transcription profiling with microarrays

Transcriptional profiling experiments with microarrays begin with the collection of tissues or organs of interest. Total RNA (and sometimes polyA RNA) is then extracted from these samples (Sambrook *et al.*, 1989). RNA is unstable and degrades rapidly. Therefore, reverse transcriptase is used to generate a cDNA copy of the mRNA. The cDNA is much more stable than the template RNA and can be easily labelled with a fluorescent dye either during synthesis (by incorporating dye-conjugated nucleotides), after synthesis (by chemical cross-linking) or even after hybridisation of the cDNAs to the array (Hegde *et al.*, 2000; Manduchi *et al.*, 2002; Wrobel *et al.*, 2003).

After hybridisation and washing to remove cDNAs lacking a complementary sequence on the array, the fluorescent signal intensity at each spot is read by a scanner and recorded as a numerical value. Those samples containing many copies of cDNA able to bind to the array (as measured by a strong fluorescent signal) are believed to reflect higher mRNA amounts in the sample, whereas a weak or absent signal reflects low mRNA abundance for the genes corresponding to those spots.

When performing a microarray analysis, it will quickly become apparent that the most time-consuming aspect of such experiments is neither in the collection and processing of the experimental samples nor in hybridisation and data acquisition. Rather, it is in the data analysis and follow-up work needed to understand the biological relevance of the gene expression profiles (Butte, 2002; Quackenbush, 2002).

Reports of microarray data nearly always represent transformed results that have been subjected to some form of pre-processing. Generally, these steps involve 'filtering' to remove signals for bad spots that possess signal-to-noise ratios that are too low to be useful and 'normalisation' to correct for normal fluctuations in fluorescent intensity due to technical and experimental variation (Quackenbush, 2002). Such steps improve the ability to monitor differences in biological samples and permit the comparison of gene expression levels between experiments. It is not possible to provide an adequate overview of normalisation procedures in this short chapter. Interested readers are encouraged to refer to some of the numerous excellent review articles on microarray normalisation and data analysis for additional details (Butte, 2002; Churchill, 2002; Hegde *et al.*, 2000; Quackenbush, 2002).

Due to differences in array format, data acquisition and data analysis, it is important that information arising from microarrays be reported in a standardised way. The development of accepted formats for submission of microarray data was established by the Microarray Gene Expression Data Society (MGED; www.mged.org), an organisation of individuals making use of microarrays for gene expression profiling. The use of standardised reporting procedures allows researchers to share results and make comparisons between microarray experiments. The goal of the organisation is in the establishment of standards for microarray data release and publication; the standards are summarised by the acronym "MIAME" to describe the "minimum information about a microarray experiment" (Stoeckert Jr *et al.*, 2002).

Expression profiling data from microarrays can typically be found in web-based supplements to journal articles as well as in a number of recently established public gene expression databases. Examples include the NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) and the European Bioinformatics Institute's ArrayExpress (http://www.ebi.ac.uk/arrayexpress/). Several university-affiliated websites also provide useful databases and analysis tools for microarray-related research. Some examples include the Yale Microarray Database (http://info.med.yale.edu/microarray/), the Stanford Microarray Database (http://genome-www5.stanford.edu/), the Duke University Center for Applied Genomics and Technology (http://mgm.duke.edu/genome/dna_micro/work/) and the ChipDB database at the Whitehead Institute (http://staffa.wi.mit.edu/chipdb/public/).

The goal of most microarray experiments is the identification of differentially expressed genes that change in response to a given treatment or developmental state. The identification of such genes can be obtained by analysing one gene at a time or several genes concurrently. A gene by gene analysis can be quite useful and is a common approach (Churchill, 2002). Typically, these methods focus on the analysis of fold-differences between samples, calculating p values by using approaches similar to t-Tests and by analysis of variance (Kuo *et al.*, 2004). As was mentioned earlier, most of the time it is a group of genes, acting together, that is responsible for the biological phenomenon being investigated and not simply a single gene acting alone. To identify concurrently expressed genes, a multigene analytical approach is

typically also performed. Such approaches include the use of class discovery, hierarchical cluster-ing, self-organising maps and other clustering methods (Dobbin and Simon, 2002; Hsu et al., 2003; Qin et al., 2003). All of these approaches are useful for identifying groups of genes expressed in similar patterns. However, expanding this information to identify functional relationships between such clustered genes (and their biological relevance) is difficult, although tools are becoming available for this problem as well (Hosack et al., 2003).

Due to the statistical and analytical issues surrounding microarray technology, as well as the complexity in dealing with the vast amount of data generated, it is important that microarray results be confirmed by using independent experimental methods. One popular method is quanti-tative (real time) RT-PCR. Real time (TaqMan) PCR takes advantage of the 5'-nuclease activity of Taq DNA polymerase and its ability to act on a gene-specific DNA oligonucleotide probe possess-ing both quencher and reporter dyes (http://www.appliedbiosystems.com)(Wilhelm and Pingoud, 2003). Upon encountering the probe, the polymerase catalyses removal of the reporter dye, thereby freeing it from suppression by the quencher and allowing reporter fluorescence to be detected. The advantages of TaqMan PCR are that it is rapid and high-throughput. Judicious choosing of the PCR oligonucleotides and labelled probe makes the assay highly specific and, as the name suggests, it permits real-time generation of data without post-PCR manipulation (e.g. gel electrophoresis). The greatest problem with the technology is in the considerable cost of the dye-conjugated probes. This expense can be offset dramatically by using the fluorescent dye, SYBR green to measure amplified products (Wilhelm and Pingoud, 2003). The disadvantage of this approach is the inabil-ity of SYBR green to distinguish specific amplified products from non-specific ones. Alternative confirmation methods include ribonuclease protection assays, reverse-transcription PCR or North-ern blotting. All can be used to confirm microarray results, but these methods do not readily lend themselves to making quantitative assessments of gene expression data. Finally, in situ hybridisation and immunohistochemistry (if an antibody is available) can be used to measure gene expression changes in a cell-specific manner within a heterogeneous tissue. These approaches offer the extra advantage of showing exactly where particular candidate genes are being expressed. This aspect can be important because gene expression changes detected by microarrays may reflect alterations in gene expression arising from only a small proportion of cells in a tissue.

Expression profiling by using porcine cDNA microarrays

Researchers at the University of Missouri-Columbia (UMC) have developed a cDNA microarray for use in expression profiling work (Whitworth et al., 2005). The array is an outcome of two porcine EST projects and consists of 19,968 spotted cDNAs within which are represented 14,129 distinct genes. A list of all the genes on the array can be found on the MU Swine Genome Project website (http://genome.rnet.missouri.edu/Swine/). The cDNAs on the array represent multiple copies of some of the unigene members as well as control spots that can be used for background correction and signal normalisation. The research emphasis within the UMC porcine reproductive biology group is mainly in three areas: ovarian physiology, uterine physiology, and oocyte/em-bryo maturation and development. Initial expression profiling results arising from the ovary and embryo work have been published or are in preparation (Whitworth et al., 2005). Some of the most recent efforts have been on expression profiling of the uterine transcriptome. Selected results from those experiments are described here.

Identification of differentially expressed transcripts in the porcine uterus

One goal of the porcine reproductive physiology group is to define those genes that are differ-entially expressed in normal non-pregnant (NP) endometrium. These data will help to define

the complex patterns of gene expression responsible for creating an environment conducive to sperm transport, early embryonic growth and embryo attachment.

The assayed samples consisted of uterine endometrium that was collected from cycling gilts on days 0, 3, 6, 10, 12, 14 and 18 post-oestrus (n = 3 animals per stage). The total RNA was reverse transcribed and labelled with Cy5. In this work, a reference design was used to permit comparisons across experiments performed at UMC. The reference RNA (labelled with Cy3) was composed of a mixture of RNAs from reproductive (foetus, uterus, ovary and placenta) and non-reproductive (liver, brain, skeletal muscle, heart, kidney, liver and spleen) organs. Each RNA sample was reverse transcribed, labelled and hybridised two times to duplicate array slides to generate two technical replicates of each biological sample. The microarrays were scanned by using a Genepix 4000B scanner and files were loaded into Genespring 6.2.1 for analysis. Spot quality was assessed and 'bad' spots eliminated; LOWESS normalisation was then performed on all good spots. 'Bad' spots are those that are smeared, have irregular signals or whose fluorescent signal is saturating. Typically, bad spots represent only a small proportion of the spots on an array. However, it is worth noting that there can be a subjective aspect to spot-quality assessment and this subjectivity may be a source of experimental variation between research groups.

Comparisons were then made between each of the time points. Gene transcripts that changed in abundance throughout the oestrous cycle were identified by analysis of variance (ANOVA) ($p < 0.05$). The Benjamini and Hochberg False Discovery Rate multiple correction test was performed on each comparison. A total of 4,827 mRNAs were found to differ significantly over the course of the oestrous cycle. To limit the number of genes analysed, only those that were abundant (twice the expression level of the reference RNA signal in at least one of the seven time points) were characterised further. Even with this limitation, numerous genes were found to be up-regulated at each timepoint (Day 0, 118; Day 3, 226; Day 6, 243; Day 10, 508; Day 12, 542; Day 14, 518; and Day 18, 297 genes). Examples of some of the differentially expressed genes are listed in Table 1. Most of those listed have been described previously as being differentially expressed in ungulate endometrium and the relative changes in transcript abundance are consistent with those previous reports.

Identification of genes with similar expression patterns

K-means clustering (standard correlation) of the differentially expressed spots revealed groups of genes possessing similar expression patterns. The clustering (Fig. 2) showed that there were a large number of transcripts that changed in abundance across the developmental stages in a consistent manner. Six distinct clustering patterns are illustrated in Fig. 2. These gene clusters represented transcripts whose abundance increased: **A.** on Day 0 of the oestrous cycle; **B.** on Days 0 and 18; **C.** on Days 3-6; **D.** on Days 3-14; **E.** on Days 10-12; **F.** on Days 10-14. These stages represent periods in which a uterine environment is being established that is conducive to biological phenomena such as sperm maturation and transport (Day 0), blastocyst growth and intra-uterine positioning (Days 3-6), and conceptus growth, apposition and attachment (Days 10-14).

Identification of Functionally-Related Genes by EASE

Expression Analysis Systemic Explorer (EASE; version 2.0) is a software program that helps interpret gene lists from microarray results (http://david.niaid.nih. gov/david/ease.htm) by identifying genes in the data known to be functionally related (e.g components in a multi-subunit complex; involvement in a signal transduction or developmental pathway; etc.). Genbank accession numbers from all significantly up regulated cDNAs from each cluster group were

J.A. Green et al

Table 1. Examples of differentially expressed genes in cyclic porcine endometrium

Day of oestrous cycle Systematic	P-value	0 Ratio*	3 Ratio	6 Ratio	10 Ratio	12 Ratio	14 Ratio	18 Ratio	Description
BC039979	3.01E-12	0.09	0.08	0.10	0.28	0.51	9.64	4.04	Mus musculus tripartite motif-containing 44...
M98553	3.74E-09	1.56	0.20	0.58	4.76	7.75	44.17	32.97	Sus scrofa tartrate-resistant acid phosphatase...
NM_006744	4.36E-09	0.03	0.04	0.05	0.17	0.30	6.08	2.64	Homo sapiens retinol binding protein 4, plasma (RBP4), mRNA...
NM_006804	4.72E-09	0.16	0.05	0.23	0.19	0.67	37.18	107.86	Homo sapiens START domain containing 3 (STARD3), mRNA
NM_002423	4.72E-09	31.89	1.02	1.14	0.37	0.39	1.89	71.58	Homo sapiens (matrilysin, uterine) (MMP7)...
NM_144583	5.25E-09	0.67	0.60	0.67	17.77	19.89	7.53	2.31	Homo sapiens ATPase, H^+ transporting, lysosomal 42kDa...
L14282	5.41E-09	0.12	0.06	0.58	4.72	5.00	9.88	1.90	Pig plasmin trypsin inhibitor mRNA, complete cds
NM_017436	8.66E-09	0.22	0.95	2.06	7.93	5.07	1.21	0.61	Homo sapiens alpha 1,4-galactosyltransferase (A4GALT), mRNA
NM_001901	9.05E-09	0.63	0.67	0.59	2.81	3.03	5.33	1.61	Homo sapiens connective tissue growth factor (CTGF), mRNA
NM_022975	9.14E-09	1.99	0.88	2.03	5.26	4.98	4.63	1.97	keratinocyte growth factor receptor...
NM_014579	9.14E-09	1.47	1.48	1.62	6.79	9.37	25.34	7.60	Homo sapiens solute carrier family 39 (zinc transporter)...
NM_001430	9.14E-09	1.20	0.72	1.37	5.98	6.90	5.26	2.60	Homo sapiens endothelial PAS domain protein 1 (EPAS1), mRNA...
NM_002032	1.42E-08	0.27	1.05	2.67	10.52	7.47	1.74	0.68	Homo sapiens ferritin, heavy polypeptide 1 (FTH1), mRNA
NM_006383	1.58E-08	0.11	0.11	0.07	1.47	2.36	0.36	0.06	Homo sapiens calcium and integrin binding family member 2....
NM_001257	1.75E-08	0.71	1.16	3.11	5.76	7.56	6.29	3.20	Homo sapiens cadherin 13, H-cadherin (heart) (CDH13), mRNA
NM_007014	2.76E-08	0.74	1.26	2.17	5.58	4.55	0.86	0.67	Homo sapiens Nedd-4-like ubiquitin-protein ligase (WWP2)....
NM_006681	5.66E-08	0.12	0.15	0.98	25.08	25.85	16.39	2.06	Homo sapiens neuromedin U (NMU), mRNA
NM_002213	6.10E-08	0.03	0.03	0.02	0.20	0.68	6.54	5.28	Homo sapiens integrin, beta 5 (ITGB5), mRNA

* Ratio represents fluoresent signal relative to the reference RNA signal for each spot.

loaded into the EASE software program and analysed. The program generated an annotation table for each gene list and provided an EASE score (essentially a *P*-value) to identify the most significant biological themes (Table 2).

By using the EASE scores and program, it was possible to quickly identify functionally related genes that changed throughout the oestrous cycle. It was determined that a high propor-
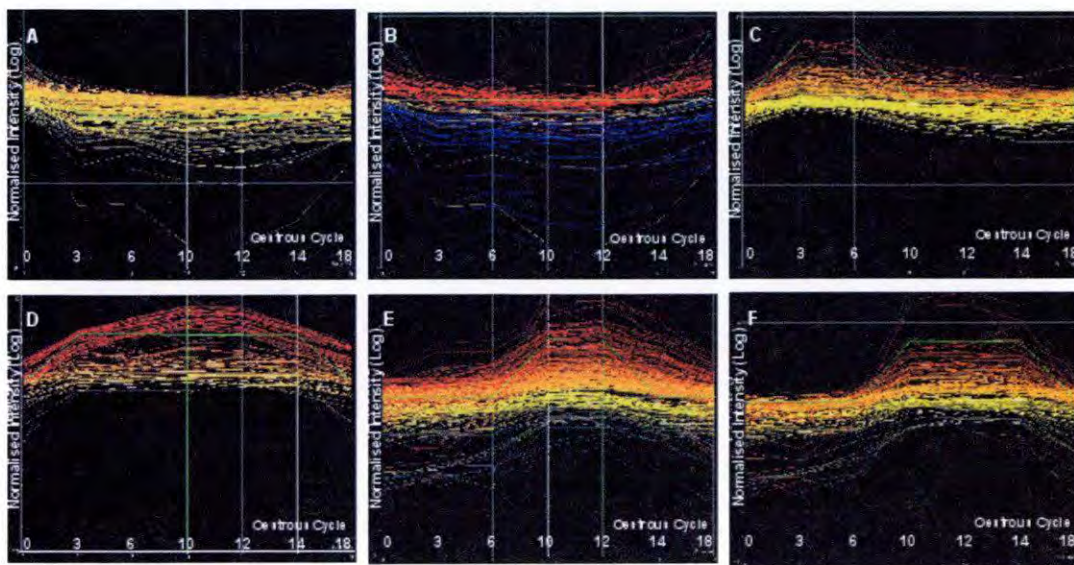
Fig. 2 Gene expression patterns obtained by using K-means clustering. The clustered genes represent those transcripts that were found to fluctuate in a similar pattern at the seven timepoints examined during the oestrous cycle.   **A.** Transcripts maximally expressed on day 0; **B.** Maximally expressed on days 0 and 18; **C.** Maximally expressed on days 3 and 6; **D.** Maximally expressed on days 3-14; **E.** Maximally expressed on days 10-12; **F.** Maximally expressed on days 10-14.

tion of the differentially expressed genes at Day 0 and 18 of the oestrous cycle are involved in immune functions as reflected by changes in the frequency of immune cell markers and cytokine populations. On Days 3-6 and 3-14, biological priorities were focused on oxidative phosphorylation, fatty acid metabolism and sterol biosynthesis. Functionally related genes on Days 10-14 were associated with signal transduction, particularly signalling associated with receptor tyrosine kinase activity (Table 2).

Real-time PCR verification of microarray results

Microarray results were validated by using quantitative gene expression by real-time reverse transcription (RT)-PCR. Seven genes including laminin (LAMC2), vanin 2 (VNN2), ornithine decarboxylase (ODC1), neuromedin U (NMU), uterine plasmin-trypsin inhibitor (UPTI), retinol binding protein (RBP4) and uteroferrin (UF) were selected for validation based on their presence in different clusters and the EASE results (Table 1). Real-time PCR was performed on 0.05 ng/reaction by using the QuantiTect SYBR Green PCR Kit (www.Qiagen.com) on an ABI Prism 7500 system (Applied Biosystems) by following standard protocols. Tyrosine 3-monooxygenase/ tryptophan 5-mon oxygenase activation protein, gamma polypeptide (NM_012479; YWHAG) was chosen as a housekeeping gene. Upon adjusting the expression level relative to the reference RNA sample, it was found that the expression patterns of these transcripts were consistent with the microarray results (Table 3).

Table 2. Gene Ontology EASE comparisons between abundantly expressed genes reflected in the six clustering profiles described here.

Expression pattern	System*	Gene category	EASE score #	Gene name examples
A. Peak at Day 0	BP	Immune response	2.95E-06	Cathepsin S (CTSS); MHC I and II (HLA-DRA, HLA-DRB1, HLA-A, HLA-DMA, HLA-DMB); Proteasome (PSMB8, PSMB9) Complement (C1QB, C1QG, IF)
B. Peak at Days 0&18	BP	Immune response	6.17E-07	Same as in A.
	BP	Cell Adhesion	1.34E-04	Cadherin 11 (CDH11); Collagen (COL12A1); Fibronectin (FN1); Integrin beta 5 (ITGB5); Laminin (LAMC2); Lectin (LGALS1); Osteopontin (SPP1); Discoidin receptor (DDR1)
	CC	Proteasome	3.95E-04	Proteasome (PSMA3, **PSMA6,** PSMB5, PSMB8, PSMB9, PSMB10)
C. Peak at Days 3 - 6	F	Oxidative phosph rylation	3.94E-04	ATP Synthase (ATP5F1, ATP5G1); Ubiquinol-Cytochrome c reductase (UQCR, UQCRH); NADH dehydrogenase (NDUFB1, NUDFC2, NUDFS8)
	BP	Fatty acid metabolism	1.76E-02	stearoyl-CoA desaturase (SCD, SCD4); DCI; ACSL3
D. Peak at Days 3 - 14	BP	Sterol Biosynthesis	2.44E-03	Cytochrome P450 (CYP51A1); Isopentenyl-diphosphate isomerase (ID11); HMGCR; SORL1
E. Peak at Days 10 - 12	MF	ATPase/transmembrane movement of ions	1.00E-03	ATPase transporters (ATP1A1, ATP6V1C1, ATP9A, ATP8A1, ATP2C1, ATP6V1C2)
	MF	Transmembrane receptor tyrosine kinase signaling	5.70E-03	Transducer ERBB (TOB1, 2); Nuclear receptor binding protein (NRBP); Docking protein (DOK4); RUSCI
	BP	Development	3.10E-03	Quaking (QK); Spectrin repeat (SYNE1); Inhibin alpha and beta A (INHA, INHBA); NEUGRIN; PBXIP1; FGFR2
F. Peak at Days 10 - 14	BP	Signal Transduction	7.89E-03 Neuromedin U (NMU)	CREM; CTGF; EPAS1; FGFR2; PTK2B; Inhibin alpha (INHA); AKAP12;

* System reflects Gene Ontology categories: BP (Biological Process); MF (Molecular Function); and CC (Cellular Component).
\# The column labelled 'EASE score' provides the significance of the comparison. Only comparisons with scores less than 0.02 are listed.

Conclusion

The female reproductive tract is the site of the final steps of gamete maturation, oocyte fertilisation, embryo positioning and placental attachment (Bazer and Roberts, 1983; Roberts *et al.*, 1993). The endometrium secr tes an array of proteins that are known to change over the course of the oestrous cycle and pregnancy. Such changes are a reflection of how the uterine endometrium is altered to facilitate the ever-changing needs of the growing conceptus. The preliminary expression profiling experiment described in this chapter illustrated the broad changes in mRNA abundance taking place in the endometrium and has provided a wealth of profiling data that will aid in understanding embryo development and the environment required for its success.

Table 3. Expression of selected genes throughout the oestrous cycle

Genes	Analysis	Day 0	Day 3	Day 6	Day 10	Day 12	Day 14	Day 18
LAMC2	Microarray	2.15 ± 0.46	0.53 ± 0.21	0.62 ± 0.15	0.77 ± 0.08	0.60 ± 0.16	1.46 ± 0.63	3.62 ± 0.68
(NM_005562)	RT-PCR	32.41 ± 3.95	4.43 ± 0.89	2.45 ± 0.33	2.41 ± 0.48	1.42 ± 0.16	9.91 ± 0.74	17.80 ± 1.16
VNN2	Microarray	4.42 ± 0.33	0.19 ± 0.10	0.18 ± 0.10	0.10 ± 0.08	0.19 ± 0.60	4.74 ± 1.81	13.68 ± 2.86
(NM_004665)	RT-PCR	14.53 ± 1.11	0.39 ± 0.03	0.37 ± 0.04	0.09 ± 0.01	0.53 ± 0.07	14.77 ± 2.04	39.32 ± 10.50
ODC1	Microarray	0.08 ± 0.05	5.31 ± 1.86	4.50 ± 2.38	1.35 ± 0.19	0.98 ± 0.14	0.70 ± 0.09	0.90 ± 0.07
(NM_002539)	RT-PCR	1.32 ± 0.28	11.91 ± 1.44	11.64 ± 1.53	2.93 ± 0.37	1.15 ± 0.11	1.23 ± 0.11	1.69 ± 0.20
NMU	Microarray	0.12 ± 0.04	0.15 ± 0.09	0.8 ± 1.92	25.08 ± 6.86	25.85 ± 3.02	16.39 ± 3.28	2.06 ± 2.59
(NM_006681)	RT-PCR	0.33 ± 0.08	1.10 ± 0.19	24.09 ± 4.49	177.68 ± 32.56	67.59 ± 11.22	84.57 ± 16.26	29.18 ± 6.09
UPTI	Microarray	0.12 ± 0.05	0.06 ± 0.02	0.58 ± 0.89	4.72 ± 0.22	5.01 ± 0.65	9.88 ± 2.76	1.90 ± 0.73
(IL14282)	RT-PCR	0.04 ± 0.01	0.01 ± 0.00	1.66 ± 0.22	8.97 ± 0.50	3.86 ± 0.57	6.62 ± 0.39	1.24 ± 0.18
RBP4	Microarray	0.03 ± 0.02	0.04 ± 0.02	0.05 ± 0.02	0.17 ± 0.03	0.30 ± 0.28	6.08 ± 0.76	2.64 ± 1.18
(NM_006744)	RT-PCR	0.05 ± 0.01	0.08 ± 0.01	0.07 ± 0.01	0.39 ± 0.05	0.88 ± 0.10	14.68 ± 1.93	10.27 ± 1.23
Uterferrin	Microarray	1.56 ± 0.15	0.20 ± 0.05	0.58 ± 0.23	4.76 ± 0.49	7.75 ± 1.22	44.17 ± 3.91	32.97 ± 13.85
(M98553)	RT-PCR	3.13 ± 0.14	0.45 ± 0.21	1.45 ± 0.11	8.63 ± 0.57	51.19 ± 16.45	97.56 ± 4.29	86.92 ± 8.09

*Microarray values were normalised to all spots on the microarray by using LOWESS, and Real-time PCR values were normalised to a housekeeping gene. The shaded cells indicate times during the oestrous cycle in which the genes targeted for Real-time PCR analysis exhibited maximal expresssion

A major goal of production animal agriculture is in improving reproductive efficiency. Such efforts can take the form of improved nutrition, heat detection, synchronisation schemes etc. Unfortunately, traditional research approaches into these areas are likely to provide only marginal improvements in reproductive success in livestock. If major improvements are to take place, one area of research likely to be key to such improvements is in manipulating gene expression in the conceptus, uterus or ovary by transgenics and/or somatic cell nuclear transfer. However, knowing how to improve a biological process requires an intimate knowledge of the genetic and protein milieu driving the process. The techniques and tools to define gene- and protein-expression patterns in parallel between experimental treatments and during development have opened up the world of systems biology to the reproductive biologist. In the near term, these approaches will illuminate the mechanics of reproductive physiology in ways undreamt of 15 years ago. A more important outcome, however, will be the opportunity to apply this information to improve reproductive efficiency in swine - a species that is so very important as a world-wide source of nutritional protein.

Acknowledgements

References

Baldwin MA, Medzihradszky KF, Lock CM, Fisher B, Settineri TA and Burlingame AL (2001) Matrix-assisted laser desorption/ionization coupled with quadrupole/ orthogonal acceleration time-of-flight mass spectrometry for protein discovery, identification, and structural analysis *Analytical Chemistry* 73 1707-1720

Basha SMM, Bazer FW and Roberts RM (1979) The Secretion of a Uterine Specific, Purple Phosphatase by Cultured Explants of Porcine Endometrium Dependency upon the State of Pregnancy of the Donor Animal *Biology of Reproduction* 20 431-441

Bazer FW and Roberts RM (1983) Biochemical aspects of conceptus—endometrial interactions *Journal of Expiremental Zoology* 228 373-383.

Belghazi M, Bathany K, Hountondji C, Grandier-Vazeille X, Manon S and Schmitter J-M (2001) Analysis of protein sequences and protein complexes by matrix-assisted laser desorption/ionization mass spectrometry *Proteomics* 1 946–954

Burghardt RC, Johnson GA, Jaeger LA, Ka H, Garlow JE, Spencer TE and Bazer FW (2002) Integrins and extracellular matrix proteins at the maternal-foetal interface in domestic animals *Cells Tissues Organs* 172 202-217

Butte A (2002) The use and analysis of microarray data *Nature Reviews of Drug Discovery* 1 951-960

Caetano AR, Johnson RK and Pomp D (2003) Generation and sequence characterization of a normalized cDNA library from swine ovarian follicles *Mammalian Genome* 14 65-70

Caetano AR, Johnson RK, Ford JJ and 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

Churchill GA (2002) Fundamental of experimental design for cDNA microarrays *Nature Genetics Supplement* 32 490-495

Clawitter J, Trout WE, Burke MG, Araghi S and Roberts RM (1990) A novel family of progesterone-induced, retinol-binding proteins from uterine secretions of the pig *Journal of Biological Chemistry* 265 3248-3255

Croy BA, Waterfield A, Wood W and King GJ (1988) Normal murine and porcine embryos recruit NK cells to the uterus *Cellular Immunology* 115 471-80

de Bivort B, Huang S and Bar-Yam Y (2004) Dynamics of cellular level function and regulation derived from murine expression array data *Proceedings of the National Academy of Science Nucleic Acids Research* 101 17687-17692

Dobbin K and Simon R (2002) Comparison of microarray designs for class comparison and class discovery *Bioinformatics* 18 1438-1445

Emrich SJ, Lowe M and Delcher AL (2003) PROBEmer: a web-based software tool for selecting optimal DNA oligos *Nucleic Acids Research* 31 3746-3750

Ezashi T and Roberts RM (2004) Regulation of Interferon-{tau} (IFN-{tau}) Gene Promoters by Growth Factors that Target the Ets-2 Composite Enhancer: A Possible Model for Maternal Control of IFN-{tau} Production by the Conceptus during Early Pregnancy *Endocrinology* 145 4452-4460

Fahrenkrug SC, Smith TPL, Freking BA, Cho J, White J, Vallet J, Wise T, Rohrer G, Pertea G, Sultana R, Quackenbush J and Keele JW (2002) Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly *Mammalian Genome* 13 475-478

Fazleabas AT, Bazer FW and Roberts RM (1982) Purification and properties of a progesterone-induced plasmin/trypsin inhibitor from uterine secretions of pigs and its immunocytochemical localization in the pregnant uterus *Journal of Biological Chemistry* 257 6886-97

Geisert RD, Renegar RH, Thatcher WW, Roberts RM and Bazer FW (1982) Establishment of Pregnancy in the Pig: I. Interrelationships Between Preimplantation Development of the Pig Blastocyst and Uterine Endometrial Secretions *Biology of Reproduction* 27 925-939

Geisert RD, Chamberlain CS, Vonnahme KA, Malayer JR and Spicer LJ (2001) Possible role of kallikrein in proteolysis of insulin-like growth factor binding proteins during the oestrous cycle and early pregnancy in pigs *Reproduction* 121 719-728

Gladney CD, Bertani GR, Johnson RK and 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

Gray C, Burghardt R, Johnson G, Bazer F and Spencer T (2002) Evidence that absence of endometrial gland secretions in uterine gland knockout ewes compromises conceptus survival and elongation *Reproduction* 124 289-300

Green ML, Chung TE, Reed KL, Modric T, Badinga L, Yang J, Simmen FA and Simmen RCM (1998) Paracrine Inducers of Uterine Endometrial Spermidine/Spermine N1-Acetyltransferase Gene Expression during Early Pregnancy in the Pig *Biology of Reproduction* 59 1251-1258

Gupta A, Ing NH, Bazer FW, Bustamante LS and Jaeger LA (1998) Beta transforming growth factors (TGFs) at the porcine conceptus-maternal interface. Part I: expression of TGFbeta1, TGFbeta2 and TGFbeta3 messenger ribonucleic acids *Biology of Reproduction* 59 905-910.

Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Hughes JE, Snesrud E, Lee N and Quackenbush J (2000) A concise guide to cDNA microarray analysis. *Biotechniques* 29 548-556

Hieronymus H and Silver PA (2004) A systems view of mRNP biology *Genes and Development* 18 2845-2860

Holland PM, Abramson RD, Watson R and Gelfand DH (1991) Detection of specific polymerase chain reaction product by utilizing the 5'——3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proceedings of National Academy of Science USA* 88 7276-7280

Hosack D, Dennis G, Sherman B, Lane H and Lempicki R (2003) Identifying biological themes within lists of genes with EASE *Genome Biology* 4 R70

Hsu AL, Tang S-L and Halgamuge SK (2003) An unsupervised hierarchical dynamic self-organizing approach to cancer class discovery and marker gene identification in microarray data *Bioinformatics* 19 2131-2140

Imakawa K, Carlson KD, McGuire WJ, Christenson RK and Taylor A (1997) Enhancement of ovine tropho-blast interferon by granulocyte macrophage- colony stimulating factor: possible involvement of protein kinase C *Journal of Molecular Endocrinology* **19** 121-30

Jiang H, Whitworth KM, Bivens NJ, Ries JE, Woods RJ, Forrester LJ, Springer GK, Mathialagan N, Agca C, Prather RS and Lucy MC (2004) Large-Scale Genera-tion and Analysis of Expressed Sequence Tags from Porcine Ovary *Biology of Reproduction* **71** 1991-2002

Johnson GA, Burghardt RC, Joyce MM, Spencer TE, Bazer FW, Gray CA and Pfarrer C (2003) Osteopontin Is Synthesized by Uterine Glands and a 45-kDa Cleav-age Fragment Is Localized at the Uterine-Placental Interface Throughout Ovine Pregnancy. *Biology of Reproduction* **69** 92-98

Johnson GA, Bazer FW, Jaeger LA, Ka H, Garlow JE, Pfarrer C, Spencer TE and Burghardt RC (2001) Muc-1, integrin, and osteopontin expression during the implantation cascade in sheep. *Biology of Reproduc-tion* **65** 820-828

Kuo WP, Kim E-Y, Trimarchi J, Jenssen T-K, Vinterbo SA and Ohno-Machado L (2004) A primer on gene ex-pression and microarrays for machine learning re-searchers *Journal of Biomedical Informatics* **37** 293–303

Lee LG, Connell CR and Bloch W (1993) Allelic discrimi-nation by nick-translation PCR with fluorogenic probes. *Nucleic Acids Research* **21** 3761-3766

MacIntyre DM, Lim HC, Ryan K, Kimmins S, Small JA and MacLaren LA (2002) Implantation-associated changes in bovine uterine expression of integrins and extracellular matrix. *Biology of Reproduction* **66** 1430-1436

Malathy PV, Imakawa K, Simmen RCM and Roberts RM (1990) Molecular cloning of the uteroferrin-associ-ated protein, a major progesterone-induced serpin secreted by the porcine uterus, and the expression of its mRNA during pregnancy *Molecular Endocri-nology* **4** 428-440

Manduchi E, Scearce LM, Brestelli JE, Grant GR, Kaestner KH and Stoeckert CJ, JR. (2002) Comparison of dif-ferent labeling methods for two-channel high-den-sity microarray experiments *Physiological Genomics* **10** 169-179

Moussad EE-DA, Rageh MAE, Wilson AK, Geisert RD and Brigstock DR (2002) Temporal and spatial ex-pression of connective tissue growth factor (CCN2;CTGF) and transforming growth factor beta type 1 (TGF-beta1) at the utero-placental interface during early pregnancy in the pig. *Journal of Clinical Pathology: Molecular Pathology* **55** 186-192

Qin J, Lewis DP and Noble WS (2003) Kernel hierarchi-cal gene clustering from microarray expression data *Bioinformatics* **19** 2097-2104

Quackenbush J (2002) Microarray data normalization and transformation *Nature Genetics* **32** Supplement 496-501

Roberts RM, Xie S and Trout WE (1993) Embryo-uterine interactions in pigs during week 2 of pregnancy *Journal of Reproduction and Fertility Supplement* **48** 171-86

Robinson RS, Mann GE, Lamming GE and Wathes DC (1999) The effect of pregnancy on the expression of uterine oxytocin, oestrogen and progesterone re-ceptors during early pregnancy in the cow *Journal of Endocrinology* **160** 21-33

Sambrook J, Fritsch E and Maniatis T (1989) *Molecular Cloning: A laboratory manual* (2nd ed) Cold Spring Harbor Press

Spencer TE and Bazer FW (1995) Temporal and spatial regulation of uterine estrogen receptor and proges-terone receptor gene expression during the estrous cycle and early pregnancy in the ewe *Biology of Reproduction* **53** 1527-1543

Spencer TE, Johnson GA, Burghardt RC and Bazer FW (2004) Progesterone and Placental Hormone Actions on the Uterus: Insights from Domestic Animals *Biol-ogy of Reproduction* **71** 2-10

Stallings-Mann ML, Burke MG, Trout WE and Roberts RM (1994) Purification, characterization, and cDNA cloning of a Kunitz-type proteinase inhibitor secreted by the porcine uterus *Journal of Biological Chemistry* **269** 24090-4

Stoeckert Jr CJ, Causton HC and Ball CA (2002) Microarray databases: standards and ontologies *Na-ture Genetics Supplement* **32** 469-473

Tan Y, Li F, Piao Y, Sun X and Wang Y (2003) Global gene profiling analysis of mouse uterus during the oestrous cycle *Reproduction* **126** 171-182

Trout WE, Hall JA, Stallings-Mann ML, Galvin JM, An-thony RV and Roberts RM (1992) Steroid regulation of the synthesis and secretion of retinol-binding pro-tein by the uterus of the pig *Endocrinology* **130** 2557-2564

Tuggle CK, Green JA, Fitzsimmons C, Woods R, Prather RS, Malchenko S, Soares BM, Kucaba T, Crouch K, Smith C, Tack D, Robinson N, O'Leary B, Scheetz T, Casavant T, Pomp D, Edeal BJ, Zhang Y, Rothschild MF, Garwood K and Beavis W (2003) EST-based gene discovery in pig: virtual expression patterns and comparative mapping to human *Mammalian Ge-nome* **14** 565-579

Velculescu VE, Zhang L, Vogelstein B and Kinzler KW (1995) Serial analysis of gene expression *Science* **270** 484–487

Whitworth K, Springer GK, Forrester LJ, Spollen WG, Ries J, Lamberson WR, Bivens N, Murphy CN, Mathialigan N, Green JA and Prather RS (2004) De-velopmental Expression of 2489 Gene Clusters Dur-ing Pig Embryogenesis: An Expressed Sequence Tag Project *Biology of Reproduction* **71** 1230-1243

Whitworth KM, Agca C, Kim J-G, Patel RV, Springer GK, Bivens NJ, Forrester LJ, Mathialagan N, Green JA and Prather RS (2005) Transcriptional Profiling of Pig Embryogenesis by Using a 15-K Member Unigene Set Specific for Pig Reproductive Tissues and Embryos *Biology of Reproduction* **72** 1437-1451

Wilhelm J and Pingoud A (2003) Real-Time Polymerase Chain Reaction *ChemBioChem* **4** 1120-1128

Wrobel G, Schlingemann J, Hummerich L, Kramer H, Lichter P and Hahn M (2003) Optimization of high-density cDNA-microarray protocols by 'design of experiments' *Nucleic Acids Research* **31** e67

Yang YH and Speed T (2002) Design issues for cDNA microarray experiments. *Nature Reviews Genetics* **3** 579-588

Yu Z, Croy BA, Chapeau C and King GJ (1993) Elevated endometrial natural killer cell activity during early porcine pregnancy is conceptus-mediated *Journal of Reproductive Immunology* **24** 153-64