

# Transcriptional profiling of oocyte maturation and embryonic development elucidates metabolism and control of development

Randall S. Prather, Alana Brown, Lee D. Spate, Bethany K. Redel,  
Kristin M. Whitworth and Jeffrey J. Whyte

*Division of Animal Science, University of Missouri, Columbia, MO 65211, USA*

With the advent of next generation sequencing platforms (RNA-seq), transcriptional profiling permits the characterization of millions of RNAs from even the most limiting samples like early embryos. High-throughput RNA-seq can generate over 600 gigabases (Gb) in a single sequencing run, providing a near-complete record of all of the genes expressed in a sample at the time of collection. Condensing and finding coherence in the immense amount of raw data generated by transcriptional profiling methods such as RNA-seq is a complex task, but necessary if useful information is to be gleaned. Here we review the current technology and describe how transcriptional profiling has been used to improve oocyte maturation and embryo culture conditions, to decrease polyspermy and to improve somatic cell nuclear transfer. Most recently, RNA-seq data has provided a unique framework for understanding metabolism of the early embryo, i.e. the Warburg Effect. Rapidly proliferating cells use glucose for synthesis of nucleotides that are necessary for DNA synthesis. They shunt metabolism away from the tri-carboxylic acid cycle and toward lactic acid production and the pentose phosphate pathway. Pathways identified by RNA-seq data show that early embryos, as it turns out, are quite similar and thrive in conditions that promote proliferation of cancer cells. Application of the Warburg Effect framework to early embryos has, and will continue to contribute to improved culture conditions for embryos *in vitro*.

## Introduction

A comprehensive understanding of factors that control oocyte maturation and pre-blastocyst stage development are needed to reduce reproductive losses and to apply reproductive technologies. Reproductive loss in pigs has been estimated to be as high as 30% through the first month of pregnancy. Any alterations in management or genetic selection that could reduce this loss would have a significant economic impact on production agriculture and may help reduce pregnancy loss in other mammals. In addition to the direct potential impacts on human health, a better understanding of what the oocyte and embryo need for robust growth will enable other technologies that may impact production agriculture. These technologies include *in vitro* embryo production (oocyte maturation, *in vitro* fertilization, and embryo culture) and the technologies that rely on *in vitro* embryo production (cloning by somatic cell nuclear transfer, genetic engineering, sex selection, etc.). To gain an understanding of what regulates

development, the abundance of RNAs that are present within the cell can be estimated. These RNAs can then be used to determine which genes have undergone transcription, which proteins are likely to be produced, and if these proteins are targeted to vesicles, the cytoplasm or to be secreted. Knowledge about these RNAs and proteins can then be used to modify systems of embryo production to improve developmental competence. The estimation of RNA abundance has been termed transcriptional profiling. Below will be described how transcriptional profiling has led to advances in *in vitro* oocyte maturation and embryo culture technologies. But first a description of transcriptional profiling technologies will be provided.

### **Transcriptional profiling-gathering the data**

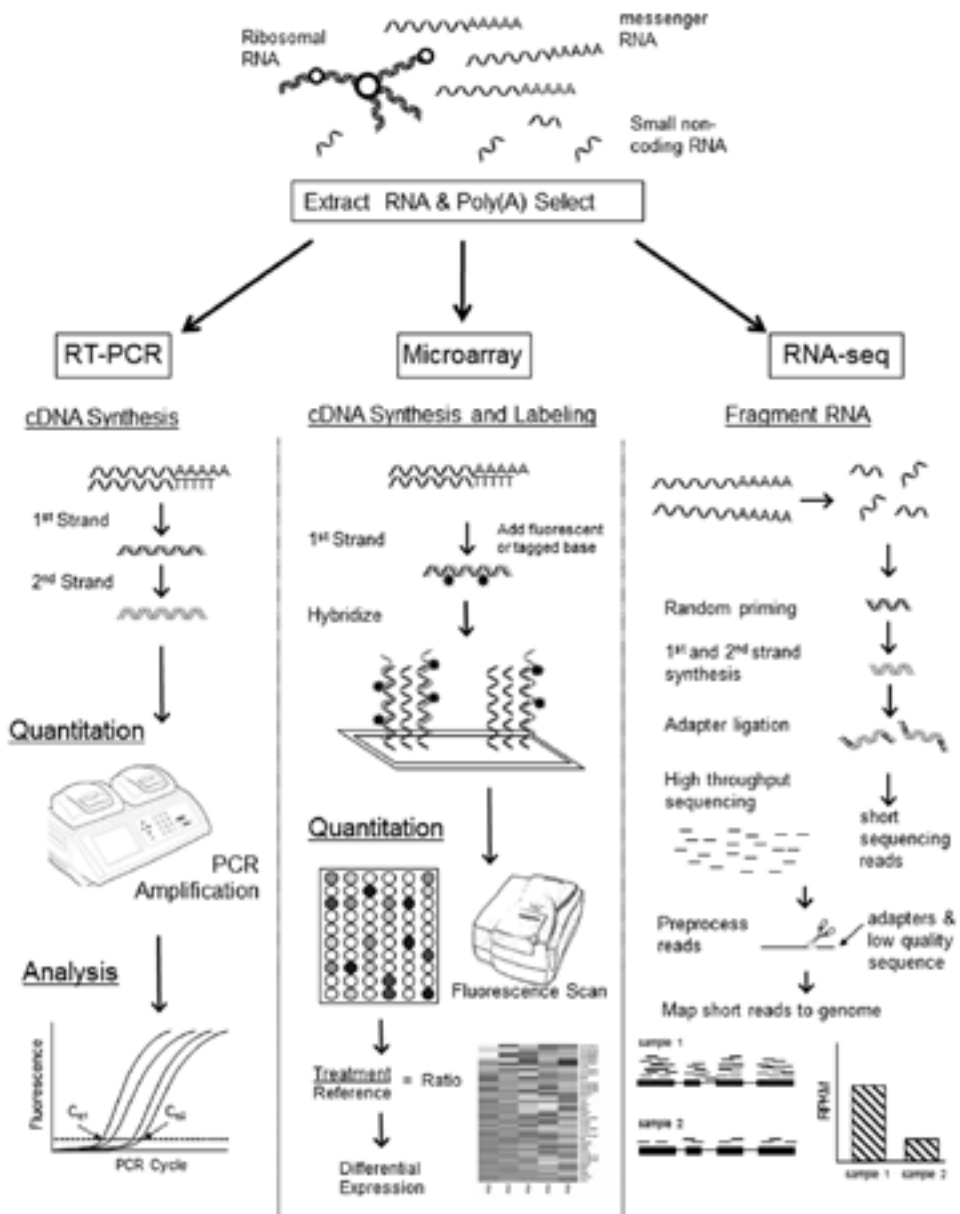
Transcriptional profiling is a phrase that can encompass a number of different approaches. In its broadest sense it includes a characterization of all the RNAs transcribed by a specific cell at a specific time. In the narrowest sense transcriptional profiling may characterize the abundance of only a few transcripts. Many technologies are available for transcriptional profiling (northern blots, sequencing of individual ESTs, quantitative real time (qRT)-PCR, hybridization to cDNA or oligonucleotide microarrays, and deep sequencing (RNA-seq)). While each technology has its limitations, RNA-seq has become the state of the art technology because of the vast amount of information that can be gleaned from a single small sample. Even this technology is rapidly developing, as only a few years ago read length from a single RNA molecule was limited to less than 40 base pairs (bp) with thousands of reads, and is now well over 100 bp with tens of millions reads per sample. The information generated with current technology is so vast that only sophisticated computers with large amounts of memory can handle the storage and analysis of this data. We have developed our own in-house pipelines to process our cDNA microarray data (Whitworth, *et al.* 2005, Whitworth, *et al.* 2010, Whitworth, *et al.* 2011), and our RNA-seq data (Bauer, *et al.* 2010a, Isom, *et al.* 2010, Redel, *et al.* 2012).

The first step to profile the transcripts is to create a library of cDNAs from the tissue(s). This has limitations because in many cases the cDNA is created by using a poly(T) primer which will only anneal to a stretch of poly(A). Generally the poly(A) is located in the 3' untranslated region (UTR) of mRNA. This serves to copy all the mRNA, or protein coding message. The final data that are generated represent a snapshot of the mRNA repertoire in the cell at the time of harvest. This doesn't provide any information about stability, production, or turnover of RNA. Unfortunately, the mRNA represents only a small proportion of the total message in a cell. Thus, these procedures select for mRNA and ignore the rest of the message, some of which is very important for regulating development (Pauli, *et al.* 2011). The different types of RNA include those involved with protein synthesis (messenger RNA: mRNA; ribosomal RNA; transfer RNA; signal recognition particle RNA), post-transcriptional modification or DNA replication (small nuclear RNA; small nucleolar RNA; ribonuclease P; ribonuclease MRP; telomerase RNA), and regulatory RNA (antisense RNA; long noncoding RNA; microRNA: miRNA; piwi-interacting RNA: piRNA; small interfering RNA: siRNA). While much focus has been on coding RNA the importance of non-coding RNA has provided impetus to gain a genome-wide understanding of all RNAs.

### **Transcriptional profiling-making sense of the data**

Condensing and finding coherence in the raw data generated by transcriptional profiling methods such as RNA-seq is a complex task. Until recently, most of the standard techniques for gene expression were limited to a defined number of preselected gene transcripts. With

qRT-PCR, generally a small number of genes are examined, while microarrays can support the analysis of many thousands of transcripts (Chomwisarutkun, *et al.* 2012, Whitworth, *et al.* 2010). Despite the large number of genes that can be analyzed via microarray, the number of probes on a chip is finite. High-throughput RNA-seq can generate over 600 gigabases (Gb) in a single sequencing run, providing a near-complete record of all of the genes expressed (Martin and Wang 2011). Transcriptional profiling, whether performed by qRT-PCR, microarray, or RNA-seq generally begins with isolation of poly(A) RNA sequences from other RNA forms (Fig 1). For qRT-PCR, RNA is first copied to cDNA by reverse transcriptase. Detection of cDNA PCR products amplified from mRNA-specific primers is measured during each cycle by the presence of a fluorescent DNA-binding dye. During initial cycles fluorescence is not detectable, but as PCR product accumulates exponentially, a cycle number is reached where fluorescence signal can be measured. This specific cycle, the quantification cycle, is determined by the amount of target template initially present. Specific guidelines for transcript quantification by qRT-PCR have been established (Bustin, *et al.* 2009). For cDNA microarrays, mRNA is reverse transcribed in the presence of fluorescent dNTPs or a specific tagged base that can be labeled fluorescently. The resulting cDNA is purified and hybridized to the microarray surface, which typically contains thousands of immobilized, known transcript sequences (probes). A fluorescent signal resulting from hybridization of the labeled cDNA target with a specific probe identifies which mRNAs are present in the sample. This can be compared to the fluorescence pattern in a reference sample to estimate up- or down-regulated transcripts. The various methods and related pitfalls for microarray analysis, specifically related to oocyte and embryo gene expression, have been reviewed (Robert 2010). Some of these pitfalls, such as the restriction of microarrays to a limited probe set, have been overcome with RNA-seq. In addition to annotated genes, RNA-seq can reveal new transcripts, new mRNA isoforms and new variants (Trapnell, *et al.* 2012b). The abundance of sequence data has introduced new practical challenges for laboratories (e.g. increased data storage requirements and computing power) and theoretical complications requiring new algorithms to preprocess and map the sequence reads. Sequencing technologies for RNA-seq rely upon massively parallel sequencing by ligation (e.g. 2-base encoding) or sequencing by synthesis (e.g. pyrosequencing; reversible dye termination) (Niedringhaus, *et al.* 2011). This review focuses on sequencing data produced by reversible dye termination (i.e. Illumina sequencing). Data analysis techniques for other technologies are reviewed elsewhere (Kumar, *et al.* 2011, Sandberg, *et al.* 2011). For RNA-seq, short cDNA sequences are required (Fig 1), and can be obtained by fragmenting the isolated mRNA or the synthesized cDNA fragments (Bauer, *et al.* 2010a, Isom, *et al.* 2010, Sanchez-Pla, *et al.* 2012). Then adaptors are ligated to each cDNA fragment (Martin and Wang 2011). Oocyte and early embryonic samples typically offer limited starting quantities of mRNA. This may necessitate PCR amplification of the adaptor-ligated cDNA library prior to sequencing. PCR bias toward AT-rich sequences and sequence duplication can be ameliorated to some extent with modified methods (Aird, *et al.* 2011). Read length produced during sequencing has steadily increased, providing more accurate transcriptome assembly, but this can increase project cost. Alternatively, longer read sequences can be predicted from short reads by sequencing both ends of cDNA fragments (i.e. paired-ends), and computationally filling in the unsequenced gap between the two ends. RNA-seq reads are typically first preprocessed after sequencing to trim reads of low quality bases and contaminating sequence (e.g. adaptors, primers) that could degrade mapping accuracy (Kong 2011, Martin 2011). We have employed the software tool combination of TopHat and Cufflinks to map reads to the swine genome and assemble reads into transcripts, respectively (J.J. Whyte *et al.*, unpublished data). These data can then be visualized, quantitated, and analyzed statistically with the SeqMonk genome browser (described



**Fig. 1** Workflow comparison among three transcriptome analysis methods, qRT-PCR, microarray, and RNA-seq. Generally, the number of transcripts analyzed increases when moving from qRT-PCR to microarray and then to RNA-seq. The steps outlined for each technique are common workflows, but do not represent all possible variations available for each method. Each technique has potential weaknesses and thus each of the three methods can serve an important role in validating and supporting data generated from the other two.

in Chatterjee, *et al.* 2012) loaded with the most recent swine assembly. We have also mapped swine embryonic and oocyte RNA-seq reads to a custom swine transcriptome generated in-house (Bauer, *et al.* 2010a, Isom, *et al.* 2010). Alternative software packages for transcriptome assembly are reviewed by Garber *et al.* (2011). Estimating expression from mapped reads can

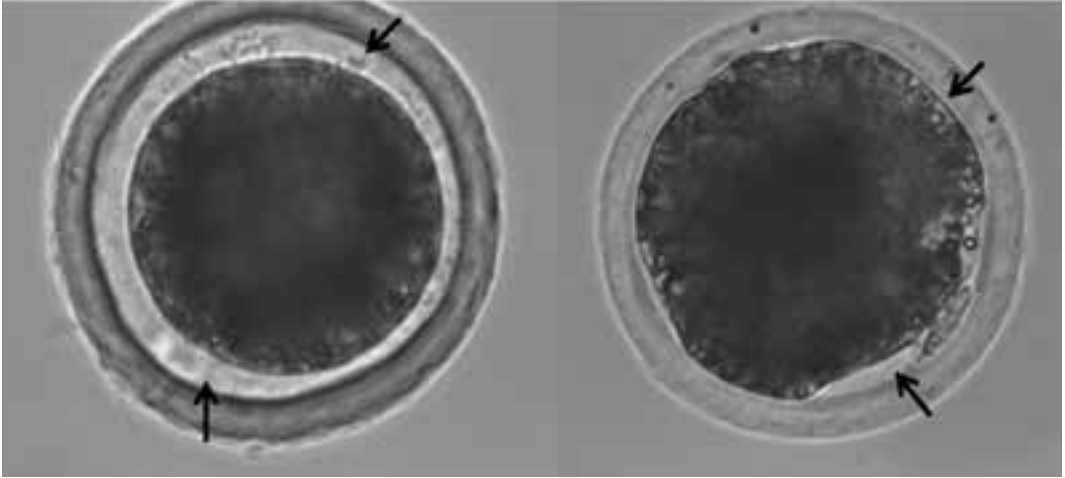
be calculated as fragments (reads) per kilobase of transcript per million mapped fragments. This estimation is computationally complex, due in part to alternative splicing where transcripts share the same exon and cannot be easily distinguished (Trapnell, *et al.* 2012b). As advances in transcriptional profiling and sequence analysis continue, it is important to bear in mind that these advances will likely serve to augment and add to existing methods and data sets, rather than replace or invalidate them. Approaching transcriptional profiling studies with a variety of techniques will strengthen and confirm the final conclusions.

### Oocyte maturation

Oocyte maturation refers to the period of the germinal vesicle (GV) oocyte resuming meiosis and re-arresting at meiosis II. The events that lead to oocyte maturation involve synthesis of the zona pellucida, growth of the oocyte, and accumulation of RNA and proteins. The fully grown oocyte, with its storehouse of RNA and protein, produces very little RNA. Thus one might assume that, barring any degradation, the abundance of various messages at the beginning of oocyte maturation would be quite similar to the abundance found at the end of maturation. Unfortunately there are mechanisms in oocytes that recruit, for example, mRNA that have short poly(A) tails and add a poly(A) tract so that the message can be translated at the appropriate time during meiosis. In addition, some mRNA is actively degraded during maturation. Thus, while all the message is probably present at the GV stage if poly(T) is used to make cDNA, then those mRNA with short poly(A) tails may not be detected. Later during oocyte maturation when those mRNA are polyadenylated they 'appear' when in reality they were always present. One of the controls for extending the poly(A) tail is thought to be the cytoplasmic polyadenylation element (CPE) located in the 3' UTR of the mRNA. One such sequence that is under the control of a CPE is *c-MOS*, a gene that plays a crucial role in regulating meiotic cell division in mammals (Dai, *et al.* 2005). Unfortunately, regulation of the abundance of mRNAs containing these CPEs appears to be more complex than just a simple conserved sequence in the 3'-UTR (Dobbs, *et al.* 2010).

While much focus has been on mRNA, this molecule makes up only a small percentage of the total RNA in a cell. In some systems, small RNA have been studied in detail (He and Hannon 2004, Houwing, *et al.* 2007, Khurana, *et al.* 2010, Kloosterman and Plasterk 2006, Watanabe, *et al.* 2008). One pig report characterized miRNA and piRNA during embryogenesis, and showed a small, but significant, percentage of miRNAs relative to other small RNA (20% in the germinal vesicle oocyte, to 5% at metaphase II, and 10% by the blastocyst stage) (Yang, *et al.* 2012), with a cluster of piRNA on chromosome 6. Some small RNA (i.e. siRNA) mapped to exonic regions for functions like cytoskeleton pathway, while others (i.e. miRNA) paired to 3'-UTR regions associated with genes involved in phosphoprotein, chromosome and acetylation pathways. One of the miRNA that increased during oocyte maturation was *MIR21* while one that decreased was *MIR574*. In *Dicer1*<sup>-/-</sup> and *Ago2*<sup>-/-</sup> mature mouse oocytes, 1,696 and 1,553 genes, respectively, were up-regulated, and 1,571 and 1,121 genes, respectively, were down-regulated compared to wild-type controls, which illustrates the global importance of small RNAs (including microRNAs and endogenous small interfering RNAs) for oogenesis (Tang, *et al.* 2010).

Two studies have profiled different aspects of pig oocyte maturation. The first evaluated *in vitro* maturation versus *in vivo* maturation (Fig 2) and found message representing twelve pathways were overrepresented in the *in vitro* matured oocytes (e.g. cancer, spliceosome, cell cycle, ubiquitin mediated proteolysis) and eight pathways were underrepresented in the *in vitro* matured oocytes (e.g. cytoskeleton regulation, T cell receptor signaling pathway) (Spate, *et al.*



**Fig. 2** Photograph of *in vitro* and *in vivo* matured pig oocytes. Note the difference in the perivitelline space (arrows) with the *in vivo* matured oocyte (left) having a greater space and the *in vitro* matured oocyte (right) having a smaller space.

2012a). The second study evaluated the quality of *in vitro* matured oocytes from prepubertal or cycling gilts and found abundance of 10% of the mRNA was different between the two groups (Paczkowski, *et al.* 2011). Transcripts coding for translation were higher in the prepubertal-derived oocytes, while transcripts for metabolism and regulation of biological processes were increased in oocytes derived from cycling gilts. Similar baseline data illustrated the importance of tumor necrosis factor alpha on maturation (Yuan, *et al.* 2011), as well as the contributions of aberrant redox balance on oocyte developmental competence (Yuan, *et al.* 2012).

### Early embryo development

Early embryo development as defined here will refer to the moment of fertilization until development of the blastocyst stage embryo with an inner cell mass and a trophectoderm. The early pig embryo has some similarities to oocyte maturation in that there is very little, if any, transcription from fertilization until the 4-cell stage (Jarrell, *et al.* 1991, Schoenbeck, *et al.* 1992). Also similar to oocyte maturation, measurement of cDNA provides a snapshot of only the poly(A) mRNA at the time point measured. Even after the 4-cell stage there is some evidence that maternally-derived mRNA is still present and may partially regulate development.

*cDNA Array Analysis of Gene Expression During Preimplantation Development.* One of the first technologies that allowed researchers to evaluate gene expression on a large scale was the use of cDNA and oligonucleotide microarrays. Thousands of cDNAs were printed on glass slides and hybridized to dye-labeled cDNAs from samples of interest. Microarray analysis of early pig embryos revealed differences in gene expression from germinal vesicle oocyte (GV) to *in vivo* 4-cell stage (p4CIVV) to *in vivo* blastocyst stage (pBLIVV) (Whitworth, *et al.* 2005). Additionally, differences in gene expression caused by *in vitro* fertilization and culture were evaluated at the 4-cell (p4CIVP) and blastocyst stage (pBLIVP). This was the first step to use global changes in gene expression to modify *in vitro* culture systems in the pig.

In these initial experiments, pair wise comparisons were performed between GV and p4CIVV and p4CIVP and pBLIVV. There were 3,214 differentially represented transcripts between GV and p4CIVV. When these gene lists were evaluated by Database for Annotation, Visualization

and Integrated Discovery (DAVID, <http://david.abcc.ncifcrf.gov/>), it was shown that GV had a large number of more-abundant transcripts related to cell adhesion receptor activity as well as many transporter activity systems and signal transduction pathways. By the time the oocyte was fertilized and developed to the 4-cell stage, these groups of transcripts were down-regulated and transcripts related to microtubule formation and mRNA binding and RNA splicing increased. As mentioned earlier, in the pig, the embryo actively starts transcribing its own RNAs during the 4-cell stage so this increase in mRNA binding and splicing would be necessary to handle the newly synthesized RNAs. When comparing p4CIVV to pBLIVV, 1,989 transcripts were differentially expressed. Transcripts encoding phosphotransferases and protein kinases were higher in the p4CIVV. In pBLIVV, transcripts encoding ribosomal subunits were up-regulated; which should increase the blastocyst stage embryo's ability to synthesize protein. Hydrogen ion transporter and monovalent inorganic cation transporter transcripts were also up-regulated. There were no differences in gene expression when comparing p4CIVV and p4CIVP and only 10 differentially expressed transcripts between pBLIVV and pBLIVP. The lack of differences could be due to limitations of microarray analysis. The entire pig embryo transcriptome was not represented on the array. Additionally, microarrays may not be sensitive enough to pick up small changes in gene expression. In contrast, transcriptome analysis using RNA-seq was able to identify 1,170 differentially expressed transcripts between *in vitro* and *in vivo* blastocysts, thus illustrating the increased sensitivity of this method (Bauer, *et al.* 2010a) (see below).

*Using Microarrays to Identify Aberrantly Reprogrammed Transcripts in Nuclear Transfer (NT) Embryos.* Microarrays have also been used to identify transcripts for genes that were aberrantly reprogrammed in NT embryos (Whitworth, *et al.* 2011). This study was unique in that all embryo culture after fertilization or activation was performed *in vivo* as all embryos were transferred back into a surrogate after IVF or NT and collected four days later so that observed changes in gene expression were the result of treatment and not simply due to time in culture. Additionally, the expression profile of the donor cell line used to create the NT embryos was examined. Expression profiles of three commonly used NT activation treatments were compared; including: (i) electrical activation/fusion, (ii) electrical activation and fusion followed by treatment with a reversible proteasomal inhibitor MG132, and (iii) electrical fusion followed by chemical activation by thimerosal /dithiothreitol. Interestingly there were no differences in gene expression between the three activation methods. Better characterization of the RNA profile by RNA-seq may yield more clues as to which treatment results in a gene expression pattern more similar to *in vivo* counterparts.

*Deep Sequencing: In Vivo-Derived Embryos.* In a continuing effort to determine global gene expression differences in embryos cultured *in vitro* compared to *in vivo*, RNA-seq was completed on embryos that were fertilized *in vivo* (Bauer, *et al.* 2010a) and cultured either *in vitro* (Porcine Zygote Medium 3 (PZM3- IVC) or *in vivo* (these embryos were derived from the same pig as controls- IVV) for an additional 4 days. Illumina sequencing resulted in 42 bp reads that were tiled against a custom porcine transcriptome comprised of 83,126 members (Isom, *et al.* 2010). A normalization factor was calculated to correct for variations in total expression, sample loading, and differences in cell number between samples. To determine the effect that culture had on the variation of gene expression profiles of embryos within treatment, correlation coefficients were calculated for each biological replicate and compared to the mean for the respective treatment group. The correlation coefficient for the IVV embryos tended ( $P=0.09$ ) to be higher (0.97) than the correlation coefficient for the IVC embryos (0.91), implying that there might be less variation between embryos for those embryos produced *in vivo*.

There were 588 transcripts that were differentially regulated between the two culture groups with a  $p$ -value  $<0.05$  and at least a 2-fold difference between the means. To determine the

biological processes and pathways that were affected by culture, up- or down-regulated genes relative to IVV were uploaded in DAVID (<http://david.abcc.ncifcrf.gov/tools.jsp>; (Dennis, et al. 2003, Huang da, et al. 2009). DAVID Gene Ontology (GO) analysis found gene families involved with mRNA transcription, nucleoside, nucleotide and nucleic acid metabolism, amino acid metabolism, lipid metabolism, and DNA metabolism were enriched in the up-regulated genes. Down-regulated genes belonged to gene families for carbohydrate metabolism, proteolysis, protein glycosylation, cell structure and motility, and lipid, fatty acid, and steroid metabolism. Pathways enriched in up-regulated genes were involved with pyrimidine and purine metabolism. Pathways enriched in down-regulated genes were involved with one-carbon pool by folate, long term potentiation, glyoxylate and dicarboxylate metabolism, and ABC transporters.

To further validate the sequencing results, 11 genes were chosen for qRT-PCR: three pluripotency related genes (*POU5F1*, *SOX2*, *NANOG*), two markers of trophectoderm (*KRT 18*, *CDX2*), one gene associated with DNA methylation (*DNMT3B*), one gene important for polyamine biosynthesis (*ODC1*), and four genes associated with nutrient homeostasis (*RGN*, *STC2*, *SLC6A6*, *SLC7A1*). The qRT-PCR results revealed differences between the IVC and IVV culture groups that validated the sequencing results for all 11 genes. Our findings illustrate the ability to sequence extremely limited samples of approximately 200 cells, to validate the sequencing, and to identify potential genes that can be used to modify our current culture system and find markers of embryo viability.

*Deep Sequencing: In Vitro-Derived Embryos.* A second experiment determined the global gene expression of embryos cultured in low O<sub>2</sub> (5% CO<sub>2</sub>, 5% O<sub>2</sub>, and 90% N<sub>2</sub>) as compared to 5% CO<sub>2</sub> in air. After *in vitro* maturation and fertilization, cleaved embryos were cultured in low O<sub>2</sub> until day 6. On day 6, blastocysts were collected, and total RNA was isolated from pools of 10 embryos per replicate and subjected to Illumina sequencing. The resulting 80 bp reads were aligned to a porcine custom transcriptome (Bauer, et al. 2010a, Isom, et al. 2010). Alignment to the custom transcriptome found 8,920 members that had 5 or more reads aligned.

In the sequencing data for the set of blastocysts cultured in low O<sub>2</sub>, we found: (i) glucose transporters *SLC2A1* and *SLC2A2*, (ii) hexokinases *HK1* and *HK2*, (iii) the fetal form of *PKM2 M2*, (iv) genes involved with the Pentose Phosphate Pathway, *TKT* and *TALDO1*, (v) phosphate dehydrogenase kinases *PDK1*, *PDK2*, and *PDK3* which block the conversion of pyruvate to acetyl-coA and thus inhibits pyruvate from entering the TCA cycle, (vi) a gene involved with alanine production from pyruvate *GPT2*, and (vii) lactate dehydrogenases *LDHA* and *LDHB*. It is hypothesized that oxygen tension can enhance some of these pathways. After culture in either low O<sub>2</sub> (5% CO<sub>2</sub>, 5% O<sub>2</sub>, 90% N<sub>2</sub>) or high O<sub>2</sub> (5% CO<sub>2</sub> in humidified air) qRT-PCR was used to estimate the abundance of the above genes. *TALDO1*, a gene involved with the PPP and *PDK1* which is involved with the blocking of pyruvate into the TCA cycle were both up-regulated in embryos cultured in low O<sub>2</sub> compared to embryos cultured in high O<sub>2</sub> and may illustrate a benefit of low O<sub>2</sub> culture.

A variety of different methods can be used to gain an understanding of the abundance of certain transcripts in the early embryo. These include cDNA microarrays, oligonucleotide arrays and RNA-seq. While each has certain advantages and limitations the real challenge becomes finding beneficial uses of the information that has been generated.

## Applications

Possibly the largest challenge with large scale transcriptional profiling experiments is applying the information generated so that either the embryo culture system can be improved, or a better



understanding of the regulation of development is achieved. As described above RNA-seq of oocytes that were *in vitro* versus *in vivo* matured revealed significant differences. A couple of those differences have been pursued in our laboratories. For example message for both cathepsin B and F were more abundant in oocytes that were matured *in vitro* as compared to those matured *in vivo* (Spate, *et al.* 2012a). Indeed, inhibition of cysteine proteinases promotes improvement in oocyte maturation (Lichtenauer, *et al.* 2012). Similarly, misexpression of members of the Wingless-type MMTV integration site (WNT) pathway were detected with overrepresentation in the *in vitro* matured oocytes. Addition of a WNT ligand, Dickkopf-related protein 1 (DKK1) to the maturation system improved the percent metaphase II and an increased number of nuclei in the resulting blastocyst stage embryos as compared to controls (L.D. Spate and R.S. Prather, unpublished).

*Osteopontin*. In addition to improvements in oocyte maturation, a number of applications to early embryo development have resulted from transcriptional profiling experiments. For example, a 20-fold increase in the number of *SPP1* (osteopontin) ESTs was detected in cDNA libraries produced from oviducts at day 0 of the estrous cycle as compared to day 3. Addition of *SPP1* to the *in vitro* fertilization system as an attempt to address potential polyspermy problems often observed with fertilization resulted in an improved efficiency of fertilization (percentage of zygotes that were monospermic) (Hao, *et al.* 2006).

*Bovine serum albumin (BSA)*. Another project estimated the abundance of message from the GV to 4-cell and blastocyst stages and showed an increase at the blastocyst stage for 65 cell surface receptors of which ligands could be tested to determine the effect on *in vitro* embryo development (Whitworth, *et al.* 2005). Ligands for two receptors were further evaluated included low density lipoprotein receptor (LDLR) and glutamate receptor, ionotropic, N-methyl-D-aspartate 3A (GRIN3A) which were up-regulated from GV to pBLIVV by 5.4 fold and 1.9 fold respectively. When LDL (50  $\mu\text{g/ml}$ ) was added to culture medium that did not contain the protein source BSA, embryos could develop to the blastocyst stage at a high rate (36%) without a reduction in total cell number. Live piglets were born as a result of LDL supplementation replacing BSA for development from the 2-cell to blastocyst stage (Spate, *et al.* 2010). Similarly, GRIN3A is a subunit of the n-methyl-D-aspartic acid (NMDA) receptor. Supplementation of NMDA (0.5 mM) into a BSA free medium also resulted in a high rate of blastocyst development (46.1%) without a reduction in total cell number (Spate, *et al.* 2012b). Interestingly, six embryo transfers did not result in the birth of any live piglets from this treatment as compared to the LDL supplementation.

*Arginine*. DAVID GO analysis identified amino acid metabolism as a family of genes that was up-regulated in the *in vivo* produced, but not in IVC embryos (Bauer, *et al.* 2010b). Amino acids are crucial to support embryo development not only as they are useful for protein biosynthesis but also for the stimulation of genome activation, blastocyst formation and hatching, contribution to energy production, osmoregulation, pH control, cell homeostasis, and signal transduction cascades (Fleming, *et al.* 2004). Specifically, an arginine transporter (*SLC7A1*) was found to be up-regulated about 60 fold in IVC compared to IVV blastocysts. Increasing evidence suggests that arginine plays important roles in placental growth (Gao, *et al.* 2012, Novak, *et al.* 2012, Wu, *et al.* 2010) as well as conceptus growth and development (Gao, *et al.* 2009) as a substrate for synthesis of nitric oxide synthase and polyamines. Characterization of free amino acid concentrations in porcine oviductal and uterine fluid on day 3 and 5 post-fertilization found arginine concentration to be at its highest in day 3 oviductal fluid ( $1.70 \pm 0.52$  mM) and lowest in day 5 uterine fluid ( $0.22 \pm 0.15$  mM) (Li, *et al.* 2007). We hypothesized that our culture system was lacking the optimal concentration of arginine needed by embryos and that adding arginine would improve embryo culture. The current base culture system, PZM3, contains only 0.12 mM arginine and after supplementing arginine to a final concentration of 0.36 mM, *SLC7A1* expression was found by qRT-PCR to be decreased to the same level as IVV blastocysts. Increasing arginine for *in vitro*

produced embryos also showed the correction in expression of *SLC7A1* and an improvement of development of cleaved embryos from 50% in PZM3 to 70% in 1.69 mM arginine (Bauer, *et al.* 2010b). These results illustrate the positive effects that additional arginine is having on porcine embryo development during culture from the 2-cell to the blastocyst stage. Supplementing arginine to a final concentration of 1.69 mM during culture increases development of porcine embryos to blastocyst compared to PZM3 and also corrects expression of *SLC7A1*.

**Folate.** Enriched in the down-regulated transcripts found in the *in vivo* produced embryos that were cultured to the blastocyst stage *in vitro* were 4 genes involved with the one-carbon pool by folate KEGG pathway. This down-regulation of genes involved with folate metabolism may illustrate an impaired folate homeostasis in embryos cultured in the current culture environment. In the cell, folates function as a family of metabolic cofactors that carry and chemically activate single carbons, for a variety of anabolic and catabolic reactions collectively known as folate mediated one-carbon metabolism (Beaudin and Stover 2009). While folate addition had no effect on the percentage of *in vitro* fertilized oocytes that developed to the blastocyst stage, there was a tendency for 0.4 mM folate to have a higher percentage of cleaved embryos develop to the blastocyst stage as compared to 0, 0.2 or 0.8 mM folate. Embryos cultured with 0.2 mM folate had an increase in number of trophectoderm and total cell numbers compared to embryos cultured without folate (Bauer, *et al.* 2010b). Here folate does not hurt development to blastocyst stage and increases cell number and may give rise to a more competent embryo. Embryo transfers still need to be performed to confirm viability.

Looking at oocyte maturation, fertilization and embryo culture, it is evident that transcriptional profiling can be an effective method of identifying ways to improve embryo culture systems and embryo quality. But can the information be used beyond single protein modifications to the culture system?

**Scriptaid Treatment to Improve the Development of Cloned Embryos.** Microarray data was used to compare NT- and *in vivo*-produced blastocyst stage embryos (Whitworth, *et al.* 2011). Twenty up-regulated transcripts were identified in NT embryos compared to the blastocysts and 4 of the 20 are involved in chromatin packaging, including two histones, (*H3F3A* and *H1F0*), a histone deacetylase, sirtuin (*SIRT1*) and the nucleosome assembly protein, SET (myeloid leukemia-associated). Twenty-seven down regulated transcripts were identified and 3 of the 27 function along the mitochondrial respiratory chain. These transcripts include two ATPases, *ATP6V1E1* and *ATP5B*, and NADH dehydrogenase, *NDUF8*. Once these genes for the transcripts that are aberrantly reprogrammed were identified, they were used as markers for improving NT. Zhao *et al.* (Zhao, *et al.* 2010, Zhao, *et al.* 2009) showed that treatment of pig embryos after NT with the histone deacetylase inhibitor, Scriptaid, improved cloning efficiencies. qRT-PCR analysis showed that Scriptaid treatment did correct aberrant gene expression in several (3 of 14 examined) of the aberrantly reprogrammed transcripts that were identified as misregulated by microarray analysis (Whitworth, *et al.* 2011). Interestingly, two of the transcripts that returned to normal *in vivo* levels after Scriptaid treatment are involved in the structural integrity of the embryos including the down-regulated transcript *CAPG* and the up-regulated transcript *SEPT7*. *CAPG* is a member of the actin-regulatory proteins that cap the barbed ends of actin filaments. *SEPT7* is a GDP binding protein that assembles into filamentous cytoskeletal polymers. Subsequent RNA-seq of the same treatments would provide an even more extensive evaluation of aberrant reprogramming following nuclear transfer.

**The Warburg Effect.** While all of the applications listed above deal with a single media component or pathway, major advances may only come from rethinking how development is viewed. One novel approach may be to view early embryo development with the Warburg Effect in mind. In the 1920s, Otto Warburg described a phenomenon in which cancer cells exhibit an

altered glycolytic metabolism (Krisher and Prather 2012). Hallmarks of the WE are (i) increased expression of glucose transporters; (ii) high levels of a cancer-specific hexokinase (*HK2*); (iii) expression of a fetal form of pyruvate kinase, muscle (*PKM2*); (iv) phosphoglycerate mutase (*PGAM*) creating an alternative glycolytic pathway resulting in pyruvate production without producing ATP; (v) high levels of pyruvate dehydrogenase kinases (*PDK*); and (vi) regulation by specific transcription factors (Levine and Puzio-Kuter 2010). The WE is described by an altered metabolic pathway in which there is an up-regulation of glucose by glucose transporters and the metabolism of glucose is shifted towards the Pentose Phosphate Pathway (PPP) and lactic acid, and away from the tricarboxylic acid (TCA) cycle thus producing less ATP. Mining the genes expressed in blastocysts cultured in low O<sub>2</sub> found by RNA-seq created a link suggesting that the WE is functioning in pre-implantation embryos (Redel, *et al.* 2012). Exploitation of this similarity between cancer cells and early embryos will lead to a better understanding of both cell types (Krisher, 2013).

### Conclusion and future

The technology to obtain a complete inventory of gene expression in most any cell type is now available. The current challenge is to understand this information to the extent that pregnancy loss can be reduced and so that meaningful refinements to culture conditions can be made to improve oocyte maturation, fertilization and culture to the blastocyst stage. Overcoming limitations to each technology, our studies have shown that mining the transcriptional profiling data, whether microarray or RNA-seq, has given us a glimpse into the developing oocyte and embryo. Looking at differences in gene expression between treatments has identified targets to potentially augment by adding specific components to culture media and thus promoting oocyte and embryo development. Constant improvements to technologies and bioinformatics have and will continue to assist in improving our culture conditions.

### Acknowledgements

The authors acknowledge funding from Food for the 21<sup>st</sup> Century at the University of Missouri and have no conflict of interest that could be perceived as prejudicing the contents of the review.

### References

- Aird, D, MG Ross, WS Chen, M Danielsson, T Fennell, C Russ, DB Jaffe, C Nusbaum, and A Gnirke** 2011 Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol* **12** R18.
- Bauer, BK, SC Isom, LD Spate, KM Whitworth, WG Spollen, SM Blake, GK Springer, CN Murphy, and RS Prather** 2010a Transcriptional profiling by deep sequencing identifies differences in mRNA transcript abundance in in vivo-derived versus in vitro-cultured porcine blastocyst stage embryos. *Biol Reprod* **83** 791-798.
- Bauer, BK, LD Spate, CN Murphy, and RS Prather** 2010b Arginine supplementation in vitro increases porcine embryo development and affects mRNA transcript expression. *Reproduction, Fertility and Development* **23** 107.
- Beaudin, AE, and PJ Stover** 2009 Insights into metabolic mechanisms underlying folate-responsive neural tube defects: a minireview. *Birth Defects Res A Clin Mol Teratol* **85** 274-284.
- Bustin, SA, V Benes, JA Garson, J Hellemans, J Huggett, M Kubista, R Mueller, T Nolan, MW Pfaffl, GL Shipley, J Vandesompele, and CT Wittwer** 2009 The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* **55** 611-622.
- Chatterjee, A, PA Stockwell, EJ Rodger, and IM Morison** 2012 Comparison of alignment software for genome-wide bisulphite sequence data. *Nucleic Acids Res* **40** e79.
- Chomwisarutkun, K, E Murani, S Ponsuksili, and K Wimmers** 2012 Microarray analysis reveals genes and

- functional networks relevant to the predisposition to inverted teats in pigs. *J Anim Sci* **90** 1-15.
- Dai, Y, B Newman, and R Moor** 2005 Translational regulation of MOS messenger RNA in pig oocytes. *Biol Reprod* **73** 997-1003.
- Dennis, G, Jr., BT Sherman, DA Hosack, J Yang, W Gao, HC Lane, and RA Lempicki** 2003 DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* **4** P3.
- Dobbs, KB, WG Spollen, G Springer, and RS Prather** 2010 The role of cytoplasmic polyadenylation element sequence on mRNA abundance during porcine embryogenesis and parthenogenetic development. *Mol Reprod Dev* **77** 699-709.
- Fleming, TP, WY Kwong, R Porter, E Ursell, I Fesenko, A Wilkins, DJ Miller, AJ Watkins, and JJ Eckert** 2004 The embryo and its future. *Biol Reprod* **71** 1046-1054.
- Gao, H, G Wu, TE Spencer, GA Johnson, and FW Bazer** 2009 Select nutrients in the ovine uterine lumen. VI. Expression of FK506-binding protein 12-rapamycin complex-associated protein 1 (FRAP1) and regulators and effectors of mTORC1 and mTORC2 complexes in ovine uteri and conceptuses. *Biol Reprod* **81** 87-100.
- Gao, K, Z Jiang, Y Lin, C Zheng, G Zhou, F Chen, L Yang, and G Wu** 2012 Dietary L-arginine supplementation enhances placental growth and reproductive performance in sows. *Amino Acids* **42** 2207-2214.
- Garber, M, MG Grabherr, M Guttman, and C Trapnell** 2011 Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat Methods* **8** 469-477.
- Hao, Y, N Mathialagan, E Walters, J Mao, L Lai, D Becker, W Li, J Critser, and RS Prather** 2006 Osteopontin reduces polyspermy during in vitro fertilization of porcine oocytes. *Biol Reprod* **75** 726-733.
- He, L, and GJ Hannon** 2004 MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* **5** 522-531.
- Houwing, S, LM Kamminga, E Berezikov, D Cronembold, A Girard, H van den Elst, DV Filippov, H Blaser, E Raz, CB Moens, RH Plasterk, GJ Hannon, BW Draper, and RF Ketting** 2007 A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell* **129** 69-82.
- Huang da, W, BT Sherman, and RA Lempicki** 2009 Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4** 44-57.
- Isom, SC, WG Spollen, SM Blake, BK Bauer, GK Springer, and RS Prather** 2010 Transcriptional profiling of day 12 porcine embryonic disc and trophectoderm samples using ultra-deep sequencing technologies. *Mol Reprod Dev* **77** 812-819.
- Jarrell, VL, BN Day, and RS Prather** 1991 The transition from maternal to zygotic control of development occurs during the 4-cell stage in the domestic pig, *Sus scrofa*: quantitative and qualitative aspects of protein synthesis. *Biol Reprod* **44** 62-68.
- Khurana, JS, J Xu, Z Weng, and WE Theurkauf** 2010 Distinct functions for the *Drosophila* piRNA pathway in genome maintenance and telomere protection. *PLoS Genet* **6** e1001246.
- Kloosterman, WP, and RH Plasterk** 2006 The diverse functions of microRNAs in animal development and disease. *Dev Cell* **11** 441-450.
- Kong, Y** 2011 Btrim: a fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. *Genomics* **98** 152-153.
- Krisher, RL, and RS Prather** 2012 A role for the Warburg effect in preimplantation embryo development: metabolic modification to support rapid cell proliferation. *Mol Reprod Dev* **79** 311-320.
- Krisher, RL** 2013 Oocyte and embryo metabolomics. In *Control of Pig Reproduction IX* pp 1-18. Eds. H. Rodriguez-Martinez, N.M. Soede and W. Flowers. UK. Context Products Ltd.
- Kumar, S, T Carlsen, BH Mevik, P Enger, R Blaaid, K Shalchian-Tabrizi, and H Kauseerud** 2011 CLOTU: an online pipeline for processing and clustering of 454 amplicon reads into OTUs followed by taxonomic annotation. *BMC Bioinformatics* **12** 182.
- Levine, AJ, and AM Puzio-Kuter** 2010 The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. *Science* **330** 1340-1344.
- Li, R, K Whitworth, L Lai, D Wax, L Spate, CN Murphy, A Rieke, C Isom, Y Hao, Z Zhong, M Katayama, H Schatten, and RS Prather** 2007 Concentration and composition of free amino acids and osmolalities of porcine oviductal and uterine fluid and their effects on development of porcine IVF embryos. *Mol Reprod Dev* **74** 1228-1235.
- Lichtenauer, AM, LD Spate, RS Prather, and JA Green** 2012 Inhibition of cysteine proteinases during in vitro oocyte maturation improves development of porcine cumulus-oocyte-complexes. *Society for the Study of Reproduction Annual Meeting* 309.
- Martin, JA, and Z Wang** 2011 Next-generation transcriptome assembly. *Nat Rev Genet* **12** 671-682.
- Martin, M** 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal* **17** 10-12.
- Niedringhaus, TP, D Milanova, MB Kerby, MP Snyder, and AE Barron** 2011 Landscape of next-generation sequencing technologies. *Anal Chem* **83** 4327-4341.
- Novak, S, F Paradis, JL Patterson, JA Pasternak, K Oxtoby, HS Moore, M Hahn, MK Dyck, WT Dixon, and GR Foxcroft** 2012 Temporal candidate gene expression in the sow placenta and embryo during early gestation and effect of maternal Progenos supplementation on embryonic and placental development. *Reprod Fertil Dev* **24** 550-558.
- Paczkowski, M, Y Yuan, J Fleming-Waddell, CA Bidwell, D Spurlock, and RL Krisher** 2011 Alterations in the transcriptome of porcine oocytes derived from prepubertal and cyclic females is associated with developmental potential. *J Anim Sci* **89** 3561-3571.
- Pauli, A, JL Rinn, and AF Schier** 2011 Non-coding RNAs as regulators of embryogenesis. *Nat Rev Genet* **12** 136-149.
- Redel, BK, AN Brown, LD Spate, KM Whitworth,**

- JA Green, and RS Prather** 2012 Glycolysis in preimplantation development is partially controlled by the Warburg Effect. *Mol Reprod Dev* **79** 262-271.
- Robert, C** 2010 Microarray analysis of gene expression during early development: a cautionary overview. *Reproduction* **140** 787-801.
- Sanchez-Pla, A, F Reverter, MC Ruiz de Villa, and M Comabella** 2012 Transcriptomics: mRNA and alternative splicing. *J Neuroimmunol* **248** 23-31.
- Sandberg, J, B Werne, M Dessing, and J Lundeberg** 2011 Rapid flow-sorting to simultaneously resolve multiplex massively parallel sequencing products. *Sci Rep* **1** 108.
- Schoenbeck, RA, MS Peters, LF Rickords, TT Stumpf, and RS Prather** 1992 Characterization of deoxyribonucleic acid synthesis and the transition from maternal to embryonic control in the 4-cell porcine embryo. *Biol Reprod* **47** 1118-1125.
- Spate, LD, BK Bauer, KM Whitworth, WG Spollen, SM Blake, CN Murphy, JL Dowell, and RS Prather** 2012a Altered mRNA Transcript Expression of In Vitro vs. In Vivo Matured Porcine Oocytes. . *Reproduction Fertility and Development* **24** 210.
- Spate, LD, BK Redel, AN Brown, CN Murphy, and RS Prather** 2012b Replacement of bovine serum albumin with N-methyl-D-aspartic acid and homocysteine improves development, but not live birth. *Mol Reprod Dev* **79** 310.
- Spate, LD, KM Whitworth, KA Walker, BK Bauer, CN Murphy, and RS Prather** 2010 Low-density lipoprotein (LDL) receptor mRNA and protein may enable LDL to replace bovine serum albumin during the in vitro swine embryo development. *Mol Reprod Dev* **77** 298.
- Tang, F, C Barbacioru, E Nordman, B Li, N Xu, VI Bashkurov, K Lao, and MA Surani** 2010 RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat Protoc* **5** 516-535.
- Trapnell, C, A Roberts, L Goff, G Pertea, D Kim, DR Kelley, H Pimentel, SL Salzberg, JL Rinn, and L Pachter** 2012a Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7** 562-578.
- Trapnell, C, A Roberts, L Goff, G Pertea, D Kim, DR Kelley, H Pimentel, SL Salzberg, JL Rinn, and L Pachter** 2012b Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7** 562-578.
- Watanabe, T, H Imai, and N Minami** 2008 Identification and expression analysis of small RNAs during development. *Methods Mol Biol* **442** 173-185.
- Whitworth, KM, C Agca, JG Kim, RV Patel, GK Springer, NJ Bivens, LJ Forrester, N Mathialagan, JA Green, and RS Prather** 2005 Transcriptional profiling of pig embryogenesis by using a 15-K member unigene set specific for pig reproductive tissues and embryos. *Biol Reprod* **72** 1437-1451.
- Whitworth, KM, LD Spate, R Li, A Rieke, P Sutovsky, JA Green, and RS Prather** 2010 Activation method does not alter abnormal placental gene expression and development in cloned pigs. *Mol Reprod Dev* **77** 1016-1030.
- Whitworth, KM, J Zhao, LD Spate, R Li, and RS Prather** 2011 Scriptaid corrects gene expression of a few aberrantly reprogrammed transcripts in nuclear transfer pig blastocyst stage embryos. *Cell Reprogram* **13** 191-204.
- Wu, G, FW Bazer, RC Burghardt, GA Johnson, SW Kim, XL Li, MC Satterfield, and TE Spencer** 2010 Impacts of amino acid nutrition on pregnancy outcome in pigs: mechanisms and implications for swine production. *J Anim Sci* **88** E195-204.
- Yang, CX, ZQ Du, EC Wright, MF Rothschild, RS Prather, and JW Ross** 2012 Small RNA Profile of the Cumulus Oocyte Complex and Early Embryos in the Pig. *Biol Reprod*.
- Yuan, Y, JM Ida, M Paczkowski, and RL Krisher** 2011 Identification of developmental competence-related genes in mature porcine oocytes. *Mol Reprod Dev* **78** 565-575.
- Yuan, Y, MB Wheeler, and RL Krisher** 2012 Disrupted redox homeostasis and aberrant redox gene expression in porcine oocytes contribute to decreased developmental competence. *Biol Reprod* **87** 78.
- Zhao, J, Y Hao, JW Ross, LD Spate, EM Walters, MS Samuel, A Rieke, CN Murphy, and RS Prather** 2010 Histone deacetylase inhibitors improve in vitro and in vivo developmental competence of somatic cell nuclear transfer porcine embryos. *Cell Reprogram* **12** 75-83.
- Zhao, J, JW Ross, Y Hao, LD Spate, EM Walters, MS Samuel, A Rieke, CN Murphy, and RS Prather** 2009 Significant improvement in cloning efficiency of an inbred miniature pig by histone deacetylase inhibitor treatment after somatic cell nuclear transfer. *Biol Reprod* **81** 525-530.

