Transcriptional profiling by RNA-Seq of peri-attachment porcine embryos generated by a variety of assisted reproductive technologies

S. Clay Isom,1,3 John R. Stevens,2 Rongfeng Li,3 William G. Spollen,4 Lindsay Cox,1 Lee D. Spate,3 Clifton N. Murphy,3 and Randall S. Prather3

1Animal Dairy & Veterinary Sciences Department, Utah State University, Logan, Utah; 2Department of Mathematics and Statistics, Utah State University, Logan, Utah; 3Division of Animal Sciences, University of Missouri, Columbia, Missouri; and 4Informatics Research Core Facility; University of Missouri, Columbia, Missouri

Submitted 10 July 2012; accepted in final form 15 May 2013

Isom SC, Stevens JR, Li R, Spollen WG, Cox L, Spate LD, Murphy CN, Prather RS. Transcriptional profiling by RNA-Seq of peri-attachment porcine embryos generated by a variety of assisted reproductive technologies. Physiol Genomics 45: 577–589, 2013. First published May 21, 2013; doi:10.1152/physiolgenomics.00094.2012.—Substantial mortality of in vitro manipulated porcine embryos is observed during peri-attachment development. Herein we describe our efforts to characterize the transcriptomes of embryonic disc (ED) and trophectoderm (TE) cells from porcine embryos derived from in vivo fertilization, in vitro fertilization (IVF), parthenogenetic oocyte activation (PA), and somatic cell nuclear transfer (SCNT) on days 10, 12, and 14 of gestation. The IVF, PA, and SCNT embryos were generated with in vitro matured oocytes and were cultured overnight in vitro before being transferred to recipient females. Sequencing of cDNA from the resulting embryonic samples was accomplished with the Genome Analyzer Ix platform from Illumina. Reads were aligned to a custom-built swine transcriptome. A generalized linear model was fit for ED and TE samples separately, accounting for embryo type, gestation day, and their interaction. Those genes with significant differences between embryo types were characterized in terms of gene ontologies and KEGG pathways. Transforming growth factor-β signaling was downregulated in the EDs of IVF embryos. In TE cells from IVF embryos, ubiquitin-mediated proteolysis and ErkB signaling were aberrantly regulated. Expression of genes involved in chromatin modification, gene silencing by RNA, and apoptosis was significantly disrupted in ED cells from SCNT embryos. In summary, we have used high-throughput sequencing technologies to compare gene expression profiles of various embryo types during peri-attachment development.

We expect that these data will provide important insight into the root causes of (and possible opportunities for mitigation of) suboptimal characteristics in both embryonic disc (ED) and trophectoderm (TE) tissues when compared with in vivo produced embryos at similar stages of development (46). We surmised that ART-induced aberrations in global gene expression patterns would be evident during this period of elongation.

Recent advances in nucleic acid sequencing technologies have made it possible to undertake large-scale cDNA sequencing efforts to characterize relative gene expression patterns, even in extremely small and limiting samples. These so-called RNA-Seq experiments are characterized by the generation of millions of short sequencing “reads,” primarily by utilizing the sequencing platforms produced by Illumina (Genome Ana-
lyzer) or Applied Biosystems (SOLiD). We have applied these techniques to pre- and peri-attachment porcine embryo samples previously (see Refs. 8, 35, 59). Herein, we report our efforts to use high-throughput sequencing technologies to characterize gene expression patterns in ED and TE from day 10, day 12, and day 14 porcine embryos derived from artificial insemination [in vivo fertilization (IVV)], in vitro fertilization (IVF), somatic cell nuclear transfer (SCNT), and parthenogenetic oocyte activation (PA). We hypothesize that the in vitro gamete and embryo manipulations associated with these assisted reproductive technologies will induce enduring changes to gene expression patterns in peri-attachment stage embryos. Furthermore, we expect distinct patterns of gene disruption to occur in the different embryo types generated.

MATERIALS AND METHODS

All chemicals and other bioreagents were purchased from Sigma (St. Louis, MO), unless otherwise indicated. Use and handling of animals were overseen and approved of by the Animal Care and Use Committee at the University of Missouri.

A concerted effort was made throughout this project to use as consistent a genetic background as possible: spermatozoa (IVV and IVF) and karyoplast donor cells (SCNT) were from half-sibling males; in vitro matured oocytes (IVF, SCNT, and PA) were purchased from the same commercial source (ART; Madison, WI), and all embryo recipients (IVV, IVF, SCNT, and PA) were bred and raised on the same swine farm facility at the University of Missouri (Columbia, MO). The IVV, IVF, and SCNT embryos were marked with an enhanced green fluorescent protein (eGFP) transgene derived from breeding stock produced and perpetuated at the University of Missouri (76).

Embryo Production and Sample Collection

IVV. Embryos derived from artificial insemination (AI) served as controls for these experiments. Second- and third-cycle virgin gilts were artificially inseminated at 12 h and 24 h after first observed standing estrus with semen collected from a single, proven eGFP transgenic boar according to standard industry husbandry practices.

IVF. Semen from the same transgenic boar used for AI was frozen as per Wang et al. (72) and used for IVF as described elsewhere (1). Briefly, oocytes were fertilized for 5 h in a Tris-buffered fertilization medium at a concentration of 0.5 × 10⁶ motile spermatozoa/ml and then washed twice before being placed into embryo culture medium.

SCNT. The specific techniques utilized for SCNT are described elsewhere (40). Briefly, metaphase II (MII)-stage oocytes were enucleated by standard micromanipulation techniques. Karyoplast donors were selected from a mixed population of embryonic fibroblast cells expressing eGFP (76) and were placed into the perivitelline space of enucleated oocytes. Incorporation of the donor cell nuclei into the cytoplasm of the enucleated oocytes was accomplished by exposing cytoplasm/karyoplast couplets to two 30 μs pulses of direct electrical current (1.2 kV/cm) in the presence of a calcium-rich fusion/activation medium (0.3 M mannitol, 1.0 mM CaCl₂ · 2H₂O, 0.1 mM MgCl₂ · 6H₂O and 0.5 mM HEPES; pH 7.3). Oocyte activation was accomplished by this same electrical stimulus.

PA. Protocols for production of porcine parthenogenetic embryos have been published previously (34). MI-arrested oocytes were exposed to a brief electrical stimulus (see above) in the presence of a calcium-rich activation medium. These electrical pulses briefly cause voltage-dependent ion channels to open, allowing calcium influx. The resulting surge of free calcium ions serves to induce meiotic resumption and subsequent embryo development.

For IVF, SCNT, and PA production methods, embryos were produced and then cultured in PZM3 culture medium (82) in vitro overnight (15–18 h) at 39°C in 5% CO₂ in air at saturating humidity until transfer into surrogate recipients. After overnight incubation, at least 100 embryos (range 100–159, average 136) were deposited into one oviduct of surrogate recipient females via surgical embryo transfer (ET). Occasionally, more embryos were produced in vitro than were needed for transfer. On these occasions (eight, three, and seven replicates for IVF, SCNT, and PA respectively), the superfluous embryos were cultured for 7 days in vitro under the conditions outlined above, and development to blastocyst was recorded. Fourteen gilts were used to generate AI embryos, and there were 12, 14, and 13 gilts that received embryos derived from IVF, SCNT, and PA, respectively. All embryo recipients were observed in standing estrus within the 24 h prior to ET. All embryo recipients were euthanized 10, 12, or 14 days after embryo transfer. We collected embryos by flushing the uterine horns with 30 ml of HEPES-buffered Tyrode’s Lactate solution [see Lai and Prather (40) for formulation] supplemented with 0.1% (w/v) bovine serum albumin. ED and TE samples were collected from as many individual embryos as possible from each flush. Fine-gauge hypodermic needles were used to manually separate embryonic discs from surrounding tissue. It should be noted here that no effort was made to remove mesoderm or endoderm cells from the TE samples at any stage of embryo development, but for simplicity’s sake, these samples will be referred to simply as TE throughout the rest of this paper. For spherical and ovoid embryos, all of the TE tissue that remained after removing the ED was harvested. For elongated, filamentous embryos, a small (~5 mm) segment of the TE tissue was dissected from the embryo at a site ~2 cm distal from the location of the embryonic disc. ED and TE samples were identified as having come from the same embryo, were snap-frozen in liquid nitrogen, and stored at ~80°C until use.

RNA Extraction and cDNA Synthesis

ED samples from two embryos from each of three distinct recipient flusses were combined in preparation for RNA extraction. The paired TE samples from the embryos used for ED RNA collection were pooled in similar fashion. Total RNA was collected from the pooled samples using the AllPrep DNA/RNA Micro Kit (Qiagen, Gaithersburg, MD) according to the manufacturer’s recommended protocol. First-strand cDNA synthesis was done using 3.5 μl of total RNA in each of the oligo dT-primed reverse transcription (RT) reactions performed for each sample using the SuperScript III First Strand Synthesis System (Invitrogen, Carlsbad, CA). The three RT reactions for each sample type were pooled together in an attempt to normalize any technical variation that might arise during the RT procedure. We used 15 μl of the pooled RT products in each of two second-strand cDNA synthesis reactions to generate double-stranded cDNA. Second-strand reactions were assembled into 150 μl reaction volumes, consisting of the following components: 15 μl of pooled first-strand RT cDNA product, 15 μl 10× buffer 2 (500 mM NaCl, 100 mM Tris-HCl, 100 mM MgCl₂, 10 mM dithiothreitol) from New England Biolabs (Ipswich, MA), 1 μl RNase H (Invitrogen), 5 μl DNA Polymerase I (New England Biolabs), 4 μl 10 mM dNTP Mix (Promega, Madison, WI), 110 μl water. Second-strand synthesis reactions were incubated at 16°C for 4 h. Duplicate second-strand reactions were pooled together, and MinElute PCR purification columns (Qiagen) were used to purify the reactions. Double-stranded cDNA was eluted in 40 μl of 10 mM Tris-Cl (pH 8.5) and submitted to the University of Missouri Genomics Core for cDNA sequencing using the Genome Analyzer IIX (GAIIx) platform from Illumina (San Diego, CA). Prior to library preparation, cDNA was sheared to an average fragment length of 250 base pairs using the Bioruptor from Diagenode (Liege, Belgium) per technical parameters that had been established from previous empirical experimentation and observation (data not shown). We used 10 ng of fragmentalied double-stranded cDNA to prepare sequencing libraries according to protocols provided and validated by Illumina. Upon successful assembly of sequencing libraries, the samples were loaded at a concentration of 7 pM onto

Physiol Genomics • doi:10.1152/physiolgenomics.00094.2012 • www.physiolgenomics.org
individual lanes of GAIIx flow cells and were sequenced per manufacturer-recommended cycling parameters. We sequenced 17 of the 24 samples on a single flow cell lane, while the remaining seven samples were sequenced on two flow cell lanes each, meaning that 31 lanes of data were generated from these 24 samples. These sequencing efforts resulted in millions of sequencing reads, 42 base pairs in length. Read count normalization and alignment were accomplished in a manner similar to what we have reported previously (35), with individual reads being aligned using the software program SOAP (42) to a custom swine transcriptome database (see Ref. 35 for details on database construction). Only reads aligning with two or fewer mismatches to exactly one database member were used in the final analysis.

Statistical Methods

A generalized linear model (6) was fit for the RNA-Seq counts for each discrete sequence within the custom transcriptomic database that had at least one read aligned. This model was fit for each cell type (ED or TE) separately and accounted for embryo type (with levels IVV, IVF, SCNT, PA), gestation days (with levels 10, 12, 14), and their interaction. The statistical analysis was performed by a custom routine written in R (70). All 24 factorial combinations of the three factors (cell type, embryo type, and gestation days) had at least one replicate in the experiment, while seven samples had two replicates, for a total sample size of 31. With additional resources, biological replication would have been preferable to technical, leading to increased statistical power (6, 33). While not optimal for statistical power, the design of this study does lend itself to an appropriate statistical analysis. The generalized linear model here used a Poisson distribution, which has been shown to be valid and most appropriate for technical replicates in RNA-Seq data (24, 45, 60). Across all genes, P values for overall embryo type and interaction effects were converted to q values (66) to control the false discovery rate (10) at 0.10. Post hoc tests between embryo types (but still in each cell type separately) were performed for genes with significant embryo type effects but nonsignificant interaction terms (52). The specific post hoc tests compared IVV vs. IVF (to test for an overall in vitro culture effect), SCNT vs. IVF (to test for an overall cloning effect), and PA vs. IVF (to test for an overall maternal chromosome only effect). Genes showing significance in any of these three post hoc tests (after Bonferroni correction within each gene’s model) were characterized in terms of their associated Gene Ontology (GO) terms (5) and KEGG pathways (37). This test of overrepresentation/underrepresentation was performed by a conditional hypergeometric test (2, 9), as implemented in the GOstats package (23) for Bioconductor (31). P values for the GO terms and KEGG pathways were also converted to q values to control the false discovery rate (10).

In preparation for hierarchical clustering, normalized read counts were log(10)-transformed for all genes shown to be differentially expressed in at least one pairwise comparison between embryo types across all three time points evaluated. These log-transformed data points were then subjected to unsupervised bidirectional hierarchical clustering using the MultiExperiment Viewer (MeV) module of the TM4 Microarray Software Suite (63, 64). Euclidian distance and average linkage metrics were employed for cluster generation. Using the same log-transformed data points for all differentially expressed genes (DEGs), a principal component analysis of cell type, gestation day, and embryo production method was accomplished using the MeV module.

Quantitative PCR

The manufacturer’s recommended protocol was used for quantitative PCR gene expression using the BioMark system from Fluidigm (South San Francisco, CA). This protocol incorporates the use of a DNA binding dye EvaGreen (Biotium, Hayward, CA) for analyzing expression of up to 48 genes for each of 48 samples on Fluidigm’s 48x48 Dynamic Array Integrated Fluidic Circuits (IFC). A panel of 48 genes was selected for analysis by qPCR, in an effort to validate the overall consistency between RNA-Seq and qPCR. Genes were selected for qPCR analysis without regard for differential expression in the RNA-Seq experiment. The genes selected for qPCR analysis fell broadly into one of eight functional categories: housekeeping (EIF4A1, GAPDH, HSP90AA1, RPN1, TAF11), apoptosis (ATM, BCL2L1, CASP3, TP53, XIAP), epigenetic modifiers (ASH2L, DMAP1, DNMT1, DNMT3A, DNMT3B, EZH2, HDAC3, SIRT1), potentially imprinted (GNAS, GRB10, IGF2, IGF2R, NDN, NNAT, PEG10, UBE3A), maternal effect (BMI15, GDF9, MOS, NOBOX, ZAR1, ZP3), pluripotency (KLF4, LIN28, MYC, NANOG, POU5F1, SOX2), sexing (HPR1, SRY), and trophoblast function (ASCL2, CYP17A1, ERF5, HAND1, HSD17B1, KRT8, TEAD4). Gene and PCR primer information for these genes are presented in Supplemental Table S1.1 Primer sets were validated previously (data not shown) and were determined to have similar amplification efficiencies: the coefficient of variation for amplification efficiency across all 48 assays was 0.0916, and only four of the 48 assays had amplification efficiencies that deviated >15% from optimum.

A specific target amplification (STA) was initially performed to enrich each sample prior to quantitative PCR thermal cycling for specific genes of interest. In preparation, a primer mix was prepared by pooling 1 µl of each primer pair (at 20 µM each, from Fluidigm) to form a 200 nM primer mix. For STA thermal cycling, each reaction was composed of 1.25 µl of this primer mix, 2.5 µl of the TaqMan PreAmp Master Mix (Applied Biosystems, Foster City, CA), and 1.25 µl of the cDNA. The reactions were activated at 95°C for 10 min and then amplified for 14 cycles (95°C for 15 s then 60°C for 4 min). Following STA thermal cycling, each reaction was treated with Exonuclease I (Exol, New England Biolabs) to remove any unincorporated primers. The Exonuclease reaction solution was made using 1.4 µl water, 0.2 µl Exol Reaction Buffer, and 0.4 µl Exol for each reaction. Two microliters of Exol reaction solution was added to every STA reaction and then placed in a thermal cycler at 37°C for 30 min to allow the digest to run to completion. The reactions were then inactivated at 80°C for 15 min. Each reaction was then diluted by adding 18 µl of water to the 7 µl reaction (volume of STA reaction + Exol) for a total volume of 25 µl. Reactions were then stored at −20°C until further use.

In preparation for running the Fluidigm chip on the BioMark, a sample Pre-Mix solution was applied to the Fluidigm 48x48 Dynamic Array IFC. This solution consisted of 2.5 µl of the 2x TaqMan Gene Expression Master Mix (Applied Biosystems), 0.25 µl of 20x DNA Dye Binding Sample Loading Reagent (Fluidigm), 0.25 µl EvaGreen DNA Binding Dye (Biotium), and 2 µl of the diluted Exol-treated STA sample. Each sample mix solution was vortexed for 20 s and centrifuged for 30 s. The assay (primer) mix solution was then made for every primer set by using 2.5 µl of the 2x Assay Loading Reagent (Fluidigm), 0.25 µl water, and 2.25 µl of each 20 µM Forward and Reverse Primer Mix. All assay mixes were vortexed for 20 s and centrifuged for 30 s. The IFC chip was primed by injecting 300 µl of control line fluid into each accumulator on the chip and then allowing the IFC Controller MX machine to run the Primer (113x) script. When priming had finished, 5 µl of each assay and 5 µl of each sample were pipetted into their respective inlets on the chip and the chip was returned to the IFC Controller for chip loading. The chip was then placed in the BioMark thermal cycler for quantitative PCR. Cycling parameters included a 5 min initial enzyme activation step, followed by 35 denaturation/extension cycles (95°C for 15 s followed by 60°C for 60 s) and a 3 min final extension cycle. A melt curve analysis was performed following PCR amplification to assess the quality of each amplicon. Quantitative PCR reactions were performed in triplicate for each sample type/primer set combination.

1 The online version of this article contains supplemental material.
Data was analyzed via the Fluidigm Real-Time PCR Analysis Software. This software uses the C(T)_2 value determined through thermal cycling to quantitate the ΔC(T) value by normalizing through a selected housekeeping gene (HSP90AA1). This software also assessed the melting curve to determine the quality of each reaction. We then calculated ΔΔC(T) values by subtracting the ΔC(T) of each experimental sample from the ΔC(T) of a calibrator sample (universal pig reference cDNA sample; Ref. 75) for each gene.

RESULTS

When more in vitro manipulated embryos were produced than were required for embryo transfer, those supernumerary embryos were maintained in culture in the laboratory, to assess development in vitro. Embryos derived from PA were the most successful at developing in vitro, with 30% (87/290, 7 replicates) of activated oocytes developing to blastocyst stage by day 7. Approximately 15.3% (71/464, 8 replicates) of IVF embryos developed to the blastocyst stage. The SCNT embryos were the least successful at developing in vitro, with 11.8% (10/85, 3 replicates) of activated couplets developing into blastocysts by day 7. While in vitro developmental potential may have varied somewhat between embryo production techniques, no obvious differences were noted in the ability of the distinct embryo types to develop successfully after embryo transfer to the prescribed collection days: only seven of the 53 gilts euthanized for embryo collections were open or failed to provide sufficient embryo numbers at the time of flushing (3/14 IVV, 1/12 IVF, 3/14 SCNT, and 0/13 PA). Absolute embryo numbers recovered from flushes were not determined, because of low quality (5.4% of total) or high homology to endogenous “contaminants” such as 18S ribosomal RNA, mitochondrial DNA, etc. (67.8% of total reads). Even after sustaining the loss of these reads, 4.82 × 10⁷ reads remained that aligned with two or fewer mismatches to the putative transcript sequences within our custom swine transcriptome database (an average of 1.55 × 10⁶ reads per sample). Of those reads that aligned to the transcripts in the database, however, ~37.5% matched more than one member of the database with high homology and so were not included in further analyses. Overall, from the 31 sequencing lanes run, 30,061,950 reads, an average of 969,740 reads per lane, aligned to exactly one member of the database and were used in the downstream statistical analyses. Table 1 provides summary data for these alignment statistics for ED and TE samples, respectively. Supplemental Table S2 provides a more thorough documentation of the trimming, purging, and alignment statistics for all samples combined. It should be

Fig. 1. Porcine embryos from days 10, 12, or 14 of gestation. Representative images of embryos flushed from reproductive tracts of pregnant gilts. A, D, G: embryos flushed from gilts on day 10 after embryo transfer; B, E, H: embryos flushed on day 12; C, F, I: embryos flushed on day 14. A, B, C: in vivo-derived embryos on days 10, 12, and 14 of gestation, respectively. D: day 10 parthenogenetic embryos. E: day 12 embryo derived from in vitro fertilization (IVF). F: day 14 embryo derived from somatic cell nuclear transfer (SCNT). G: fluorescence micrograph showing 2 day 10 embryos generated by artificial insemination with semen from a boar hemizygous for an enhanced green fluorescence transgene. Note the bright green fluorescence emanating from the embryonic disc (ED). The ED from the second embryo is not in view. H: ED sample dissected from a day 12 SCNT embryo and stained with Hoechst 33342 to show nuclei of cells. I: highly fluorescent ED sample from a day 14 IVF embryo sitting in close apposition to its corresponding trophectoderm (TE) sample, which displays much lower levels of transgene expression. Fertilization was performed with spermatozoa from the same boar utilized to generate the embryos in G.
Table 1. Short read alignment summary

<table>
<thead>
<tr>
<th>Sample</th>
<th>Reads Aligned Uniquely</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-IVV-ED</td>
<td>710,646</td>
<td>13.8</td>
</tr>
<tr>
<td>10-IVF-ED</td>
<td>673,960</td>
<td>8.6</td>
</tr>
<tr>
<td>10-IVF-TE</td>
<td>1,590,733</td>
<td>41.2</td>
</tr>
<tr>
<td>10-SCNT-ED</td>
<td>870,343</td>
<td>21.2</td>
</tr>
<tr>
<td>10-SCNT-TE</td>
<td>889,254</td>
<td>34.1</td>
</tr>
<tr>
<td>10-PA-ED</td>
<td>195,839</td>
<td>0.9</td>
</tr>
<tr>
<td>10-PA-TE</td>
<td>1,738,151</td>
<td>32.5</td>
</tr>
<tr>
<td>12-IVV-ED</td>
<td>2,179,666</td>
<td>36.4</td>
</tr>
<tr>
<td>12-IVV-TE</td>
<td>1,420,081</td>
<td>42.9</td>
</tr>
<tr>
<td>12-IVF-ED</td>
<td>1,210,916</td>
<td>32.9</td>
</tr>
<tr>
<td>12-IVF-TE</td>
<td>1,646,786</td>
<td>42.2</td>
</tr>
<tr>
<td>12-SCNT-ED</td>
<td>1,836,592</td>
<td>17.8</td>
</tr>
<tr>
<td>12-SCNT-TE</td>
<td>870,940</td>
<td>30.7</td>
</tr>
<tr>
<td>12-PA-ED</td>
<td>2,817,828</td>
<td>31.3</td>
</tr>
<tr>
<td>12-PA-TE</td>
<td>1,378,641</td>
<td>18.6</td>
</tr>
<tr>
<td>14-IVV-ED</td>
<td>764,270</td>
<td>25.9</td>
</tr>
<tr>
<td>14-IVV-TE</td>
<td>1,107,690</td>
<td>35.3</td>
</tr>
<tr>
<td>14-IVF-ED</td>
<td>1,573,189</td>
<td>38.9</td>
</tr>
<tr>
<td>14-IVF-TE</td>
<td>2,073,458</td>
<td>26.7</td>
</tr>
<tr>
<td>14-SCNT-ED</td>
<td>1,301,997</td>
<td>28.3</td>
</tr>
<tr>
<td>14-SCNT-TE</td>
<td>1,197,472</td>
<td>39.0</td>
</tr>
<tr>
<td>14-PA-ED</td>
<td>820,879</td>
<td>20.9</td>
</tr>
<tr>
<td>14-PA-TE</td>
<td>455,250</td>
<td>17.7</td>
</tr>
</tbody>
</table>

Summary of alignment statistics across all lanes of data for all samples analyzed. Presented in the “Reads Aligned Uniquely” column are the total numbers of reads (summed across replicates, if applicable) that aligned with 2 or fewer mismatches to exactly 1 database member. The data in the “%” column show what percentage of the total number of reads (after purging of endogenous “contaminants”) aligned uniquely to the custom database for each sample. Within the “sample” column, “10,” “12,” and “14” refer to embryos collected on gestational days 10, 12, and 14, respectively. IVV, in vivo fertilized; IVF, in vitro fertilized; SCNT, somatic cell nuclear transfer; PA, parthenogenetic activation; ED, embryonic disc; TE, trophodermctoderm. For more detailed information regarding the derivation of these summary statistics, readers are referred to Supplemental Table S2.

Table 2. Summary of statistical and gene set enrichment analyses

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Up In</th>
<th>Versus</th>
<th>Genes</th>
<th>BP</th>
<th>MF</th>
<th>CC</th>
<th>KEGG</th>
</tr>
</thead>
<tbody>
<tr>
<td>ED IVV</td>
<td>165</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>ED IVF</td>
<td>140</td>
<td>289</td>
<td>42</td>
<td>24</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED SCNT</td>
<td>124</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED SCNT</td>
<td>290</td>
<td>206</td>
<td>15</td>
<td>26</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED PA</td>
<td>150</td>
<td>143</td>
<td>4</td>
<td>22</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED PA</td>
<td>54</td>
<td>113</td>
<td>130</td>
<td>14</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TE IVV</td>
<td>185</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TE IVF</td>
<td>349</td>
<td>46</td>
<td>18</td>
<td>36</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TE SCNT</td>
<td>142</td>
<td>311</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TE SCNT</td>
<td>442</td>
<td>11</td>
<td>19</td>
<td>20</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TE PA</td>
<td>167</td>
<td>124</td>
<td>70</td>
<td>15</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TE PA</td>
<td>469</td>
<td>1</td>
<td>45</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Summary of the numbers of “genes” (database members). Gene Ontology (GO) terms, and signaling pathways demonstrated to be significantly enriched in each embryo type, across all gestation days, relative to its most appropriate control sample. BP, GO: Biological Process; MF, GO: Molecular Function; CC, GO: Cellular Component; KEGG, Kyoto Encyclopedia of Genes and Genomes signaling pathways. For example, 140 genes were upregulated in IVF ED samples relative to ED samples from IVV embryos, and a gene set enrichment analysis showed that there were 289 “Biological Process” terms, 42 “Molecular Function” terms, and 24 “Cellular Component” terms enriched for in IVF vs. IVV ED samples. Using the KEGG database, we show no cellular signaling pathways to be significantly affected.
counts (Illumina) and CT values (BioMark) of representative samples are presented in Fig. 5. Scatter plots depicting the inverse relationship between read counts (Illumina) and CT values for all 48 genes within a tissue sample were limited to in vitro oocyte maturation, embryo production, and interpretation of expression profiles and GO analyses (and the discussion of such in the paragraphs that follow) should be undertaken with circumspection. But it should not be inferred that this dataset is definitive in its description of time-, tissue-, and treatment-associated changes in gene expression patterns during peri-implantation porcine embryo development, and interpretation of expression profiles and GO terms and KEGG Pathways enriched in ED for each embryo type relative to its most appropriate control. Sig. Level is the q value or false discovery rate-adjusted P value for the term or pathway.

by embryo type (see Fig. 2). Principal component analysis of all 1,908 differentially expressed transcripts for all samples concurrently, however, showed only the clear separation of samples by tissue type (Fig. 3). No separation by day or by embryo type was readily apparent for the ED samples, although some separation by day was apparent in TE samples, irrespective of embryo type.

Quantitative PCR analysis of 48 genes using the microfluidics-driven BioMark system from Fluidigm yielded results that were consistent with the sequencing data. Images of representative quantitative PCR data are presented in Fig. 4, and the summarized data are presented as Supplemental Tables S3, S4, and S5 (for ED, TE, and ED/TE ratio, respectively). The coefficients of correlation between the number of sequencing reads and the Ct values for all 48 genes within a tissue sample averaged −0.722 and ranged from −0.452 to −0.856. [The least significant correlation value (−0.452) corresponds to the 10-PA-ED sample noted above as being represented by a very low number of aligned reads. If this outlier were removed from consideration, the average correlation coefficient would be −0.735, with the least significant correlation being −0.675.] Scatter plots depicting the inverse relationship between read counts (Illumina) and Ct values (BioMark) of representative samples are presented in Fig. 5.

Table 3. Gene set enrichment analysis results: ED

<table>
<thead>
<tr>
<th>Term/Pathway</th>
<th>Up In</th>
<th>Compared With</th>
<th>Database</th>
<th>Sig. Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromatin modification</td>
<td>IVF</td>
<td>SCNT</td>
<td>GO:BP</td>
<td>0.00005</td>
</tr>
<tr>
<td>Nucleocytoplasmic transport</td>
<td>PA</td>
<td>IVF</td>
<td>GO:BP</td>
<td>0.002</td>
</tr>
<tr>
<td>mRNA metabolic process</td>
<td>PA</td>
<td>IVF</td>
<td>GO:BP</td>
<td>0.002</td>
</tr>
<tr>
<td>Gene silencing by RNA</td>
<td>IVF</td>
<td>RNA</td>
<td>GO:BP</td>
<td>0.004</td>
</tr>
<tr>
<td>Epigenetic regulation of gene</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>expression</td>
<td>IVF</td>
<td>SCNT</td>
<td>GO:BP</td>
<td>0.007</td>
</tr>
<tr>
<td>Dosage compensation by</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X-inactivation</td>
<td>PA</td>
<td>IVF</td>
<td>GO:BP</td>
<td>0.008</td>
</tr>
<tr>
<td>TGFβ signaling</td>
<td>IVV</td>
<td>IVF</td>
<td>KEGG</td>
<td>0.000999</td>
</tr>
<tr>
<td>JNK cascade</td>
<td>IVF</td>
<td>IVF</td>
<td>GO:BP</td>
<td>0.02</td>
</tr>
<tr>
<td>Histone methyltransferase activity</td>
<td>IVF</td>
<td>IVF</td>
<td>GO:MF</td>
<td>0.039</td>
</tr>
<tr>
<td>MAPKKK cascade</td>
<td>IVF</td>
<td>PA</td>
<td>GO:BP</td>
<td>0.04</td>
</tr>
<tr>
<td>BMP signaling pathway</td>
<td>IVV</td>
<td>IVF</td>
<td>GO:BP</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 4. Gene set enrichment analysis results: TE

<table>
<thead>
<tr>
<th>Term/Pathway</th>
<th>Up In</th>
<th>Compared With</th>
<th>Database</th>
<th>Sig. Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein processing in endoplasmic reticulum</td>
<td>IVF</td>
<td>PA</td>
<td>KEGG</td>
<td>0.0004</td>
</tr>
<tr>
<td>Protein catabolic process</td>
<td>IVF</td>
<td>IVV</td>
<td>GO:BP</td>
<td>0.002</td>
</tr>
<tr>
<td>Cellular catabolic process</td>
<td>SCNT</td>
<td>IVF</td>
<td>GO:BP</td>
<td>0.007</td>
</tr>
<tr>
<td>Proteasomal ubiquitin-dependent protein catabolic process</td>
<td>SCNT</td>
<td>IVF</td>
<td>GO:BP</td>
<td>0.0077</td>
</tr>
<tr>
<td>Pyrimidine metabolism</td>
<td>SCNT</td>
<td>IVF</td>
<td>KEGG</td>
<td>0.018</td>
</tr>
<tr>
<td>Chromatin modification</td>
<td>IVF</td>
<td>IVV</td>
<td>GO:BP</td>
<td>0.024</td>
</tr>
<tr>
<td>Neurotrophin signaling</td>
<td>PA</td>
<td>IVF</td>
<td>KEGG</td>
<td>0.025</td>
</tr>
<tr>
<td>ErbB signaling pathway</td>
<td>IVF</td>
<td>IVV</td>
<td>KEGG</td>
<td>0.03</td>
</tr>
<tr>
<td>Gene silencing by RNA</td>
<td>PA</td>
<td>IVF</td>
<td>GO:BP</td>
<td>0.03</td>
</tr>
<tr>
<td>Ubiquitin mediated proteolysis</td>
<td>IVF</td>
<td>IVV</td>
<td>KEGG</td>
<td>0.03</td>
</tr>
<tr>
<td>Germ cell development</td>
<td>PA</td>
<td>IVF</td>
<td>GO:BP</td>
<td>0.031</td>
</tr>
<tr>
<td>mRNA 3’-end processing</td>
<td>IVF</td>
<td>SCNT</td>
<td>GO:BP</td>
<td>0.039</td>
</tr>
<tr>
<td>SAPK signaling cascade</td>
<td>PA</td>
<td>IVF</td>
<td>GO:BP</td>
<td>0.042</td>
</tr>
<tr>
<td>N-glycan biosynthesis</td>
<td>IVV</td>
<td>IVF</td>
<td>KEGG</td>
<td>0.059</td>
</tr>
</tbody>
</table>

GO terms and KEGG Pathways enriched in TE for each embryo type relative to its most appropriate control.

DISCUSSION

To our knowledge, this is the first description of global gene expression patterns in peri-attachment porcine embryos generated by ART. Additionally, this effort is a further development and demonstration of the utility of high-throughput sequencing technologies in carrying out genome-scale experimentation in nontraditional model species.

With the advent of so-called “next-generation” sequencing technologies, it has become possible to generate large numbers of short nucleic acid sequencing reads that relate in identity and abundance to the population of nucleic acids subjected to sequencing. Initially, these technologies were used primarily for whole genome resequencing, characterization of allele frequencies, and identification of single nucleotide polymorphisms. Mortazavi et al. (50), Marioni et al. (45), and Rosenkranz et al. (62) were among the first to pioneer the utilization of high-throughput short-read sequencing technologies for transcriptional profiling purposes. Today, RNA-Seq is finding wider and wider application, with potential advantages and some unique challenges compared with other technologies (i.e., solid-state arrays), which have been reviewed elsewhere (44, 54, 73). It should be noted that for the research presented herein, the number of reads aligned per sample type is somewhat lower than that recommended for RNA-Seq “best practices,” as put forward by the Encyclopedia of DNA Elements Consortium (21). In no way does this invalidate the data presented here, as these experiments should provide valuable insight into poorly understood developmental phenomena. But it should not be inferred that this dataset is definitive in its description of time-, tissue-, and treatment-associated changes in gene expression patterns during peri-implantation porcine embryo development, and interpretation of expression profiles and GO analyses (and the discussion of such in the paragraphs that follow) should be undertaken with circumspection.

The primary purpose of this research was to test the hypothesis that assisted reproductive technologies could have profound effects on patterns of gene expression up to and including the period of peri-attachment development in porcine embryos. The in vitro manipulations used in these experiments were limited to in vitro oocyte maturation, embryo production, and a short period of in vitro embryo culture before transfer into surrogate recipient females. Comparisons in patterns of
relative gene expression between embryo production methods were performed in a pair-wise manner, comparing each experimental method to its most appropriate control: IVF to IVV, SCNT to IVF, and PA to IVF as well. The first of these pairwise comparisons (IVV vs. IVF) allowed for a direct evaluation of genes responsive only to the in vitro manipulations, with the embryo production method (oocyte fertilized by spermatozoon) held constant. The remaining comparisons (SCNT vs. IVF and PA vs. IVF) revealed those genes and pathways that may be altered in response to the different embryo production methods without the potentially confounding effects of those redundant genes altered by in vitro culture.

The potential local and/or systemic effects of surgical embryo transfer on uterine physiology and embryo development are acknowledged. And, although there were no obvious suggestions that some or all of the changes in gene expression between IVV and IVF embryos were in direct response to surgery-induced maternal immunomodulation, that possibility cannot be entirely discounted. These effects were controlled for in the other pairwise comparisons, however, as all of the surrogates for in vitro manipulated embryos shared a common surgical insult upon embryo transfer.

The observation that many transcripts (305 and 534 in ED and TE, respectively) were shown to be significantly altered in embryos fertilized in vitro—compared with in vivo-fertilized embryos at the blastocyst stage (77). Significant differences in the experimental platforms, statistical analyses, and embryo stages might all be invoked as reasons behind this seeming discrepancy. In support of the suggestion that the experimental platform can have profound influences on experimental results even in very similarly conceived experimental

Fig. 2. Heat maps showing relative expression levels of genes differentially expressed between embryo types. For each pairwise comparison between sample types, normalized read counts for the transcripts shown to be differentially expressed across all gestation days were log transformed and subjected to unsupervised bidirectional hierarchical clustering. Within the heat map proper, yellow indicates a high number of aligned reads (high expression), blue indicates very low numbers of reads (low expression), and grey indicates zero reads, or missing values. The dendograms above the heat maps illustrate the calculated relationships between samples and genes, respectively. A: in vivo fertilization (IVV) vs. IVF ED; B: IVF vs. SCNT ED; C: IVF vs. parthenogenetic oocyte activation (PA) ED; D: IVV vs. IVF TE; E: IVF vs. SCNT TE; F: IVF vs. PA TE.

RNA-SEQ OF PERI-ATTACHMENT PORCINE ART EMBRYOS

Physiol Genomics • doi:10.1152/physiolgenomics.00094.2012 • www.physiolgenomics.org
designs, Miles et al. (48) used serial analysis of gene expression to assess relative transcript abundance in porcine blastocysts and found >900 transcripts that were putatively differentially expressed in embryos generated in vitro vs. in vivo. Also, as the genes presented herein as differentially expressed are only those that were shown to be different across all gestation stages, it is highly unlikely that asynchrony between embryo types is responsible for the high number of aberrantly expressed genes. Thus, the larger numbers of DEGs in this report are likely reflective of the relative developmental complexity of peri-attachment embryos compared with blastocyst-stage embryos, as well as the exquisite sensitivity and bias-free nature of the Illumina sequencing platform for transcriptome profiling (i.e., not limited to features on the array).

One pathway shown to be significantly altered in in vitro fertilized embryos was the transforming growth factor (TGF)-β signaling pathway. This pathway is essential for development and patterning during embryogenesis in vertebrates and invertebrates alike (reviewed in Ref. 81). TGF-β signaling has recently been implicated in mediating the interaction of porcine conceptuses with the uterine endometrium at both early and late stages of embryonic attachment (47). Interestingly, though, in our data, the suppression of TGF-β signaling in IVF embryos was observed primarily in the ED samples, and not the TE samples. When evaluated at the pathway level, it is the bone morphogenetic protein (BMP) arm of the TGF-β signaling pathway that is relatively underexpressed in IVF embryos. In fact, every level of the BMP signaling cascade (from the receptor level to the DNA transcription level) is represented in the list of genes expressed at consistently higher levels in ED samples from IVF embryos than from IVF embryos. The BMP signaling family is important in body axis determination, limb development, and organogenesis (reviewed in Refs. 38, 71). It is tempting to speculate that depressed BMP signaling at crucial developmental windows could, at least in part, explain the relatively lower viability rates of in vitro manipulated embryos. Of course, further validation of this hypothesis is necessary.

TE samples from IVF embryos are characterized by an upregulation of the cellular machinery involved in ubiquitin-mediated proteolysis, which may be symptomatic of protein synthesis and/or processing defects, although this too requires

Fig. 3. Principal component analysis. Principal component analysis (PCA) of the 31 sequencing lanes was undertaken using all 1,908 differentially expressed genes as data points. The analysis shows clear separation by tissue type and some separation of TE by day of gestation. No clear distinction of samples by embryo production method was noted. A: samples plotted on PC1 (x-axis) vs. PC2 (y-axis). The blue line surrounds all the ED samples (except for the 2 day 10 ED samples from parthenogenetic embryos, which were consistent outliers in this experiment due to poor sequencing); the green line encircles the trophoblast (TE) samples. B: samples plotted on PC1 (x-axis) vs. PC3 (y-axis). The blue line surrounds the ED samples again. The green line surrounds the day 14 TE samples, and the yellow line surrounds the day 10 and day 12 TE samples (except for the day 12 TE sample from IVF embryos, which was an outlier on this plot). C: PC2 (x-axis) vs. PC3 (y-axis). This principal component was clearly driven by the differences between the day 10 PA ED samples (which didn’t sequence well) and all the rest of the samples; no further clustering was noted on this plot. D: 3-dimensional rendering of the PCA. The axes have been rotated to highlight the separation of the distinct clusters. Blue spheres represent the ED samples; the green spheres represent the “late” TE samples, and the yellow spheres represent the “early” TE samples. Principal components 1, 2, and 3 are represented by the silver, blue, and pink axes, respectively.
reveal three primary themes disrupted by SCNT: epigenetic manipulations of embryos in pigs as well as other model species. Responsible for the altered placental dynamics observed in in vitro manipulated embryos in a number of species (25, 41, 65, 80, 83). To our knowledge, this is the first report of aberrant signaling dynamics. Placental hyperplasia and fetal overgrowth have been commonly observed in pregnancies resulting from in vitro manipulated embryos. Although none of the canonical ErbB receptors (ErbB1–4) were upregulated in IVF embryos, several members of the signaling cascade(s) downstream of the receptors were upregulated, including STAT5b, PAK1, PAK2, SOS1, and mTOR. It is yet too early to understand the exact consequences of hyperactive ErbB signaling in porcine peri-attachment IVF embryos, but it is interesting to note that significant perturbations in placental structure and function are observed in mice that overexpress EGFR, the founding member of the ErbB receptor family (17). Specifically, placental weights from pups heterozygous or homozygous for a hypomorphic EGFR (ErbB1) allele were from 17–55% greater than those from control pups from the same litters. In this mouse model, expression levels of a number of trophoblast-specific genes were altered as well as a consequence of altered ErbB signaling dynamics. Placental hyperplasia and fetal overgrowth have been commonly observed in pregnancies resulting from in vitro manipulated embryos in a number of species (25, 41, 65, 80, 83). To our knowledge, this is the first report of aberrant ErbB signaling in trophoblast cells of in vitro manipulated embryos from any species. It remains to be seen, then, whether this signaling pathway is affected in trophoblast cells of other species at comparable developmental time points and whether or not these signaling defects could be at least in part responsible for the altered placental dynamics observed in in vitro manipulated embryos in pigs as well as other model species.

GO analysis results from the SCNT/IVF (ED) comparison reveal three primary themes disrupted by SCNT: epigenetic control of gene expression, gene silencing by microRNAs, and apoptosis. It comes as no surprise that epigenetic modification of the genome is disrupted in nuclear transfer embryos; this phenomenon has been extensively reviewed elsewhere (3, 18, 51, 84). It is somewhat surprising, however, that genes involved in epigenetic processes were overwhelmingly down-regulated in embryos derived from SCNT. If the epigenome was significantly misprogrammed after nuclear transfer, it would not be unreasonable to expect to see an upregulation of the epigenetic machinery in an attempt to correct the faulty programming. Instead, we observed widespread suppression of the genes that control proper epigenome constitution in SCNT embryos. An alternative explanation for this observation is that the epigenetic machinery of the incipient SCNT embryos was somehow inherently dysfunctional, resulting in faulty embryo reprogramming and poor SCNT success rates. Whatever the explanation, widespread downregulation of genes involved in epigenetic control of gene expression has also been observed in cloned bovine embryos (61). The potential cause-and-effect relationships between faulty genome programming and the epigenetic machinery of the oocyte and early embryo demand further investigations, as additional insight into the causes of and potential strategies for mitigation of faulty epigenetic reprogramming in SCNT embryos is sorely needed.

Very little information is available regarding the status of the cellular machinery responsible for gene silencing by microRNAs in cloned embryos. Cui et al. (16) report aberrant reprogramming of a very limited number of specific microRNAs in cloned mouse embryos, and the same group showed that a histone deacetylase inhibitor (Trichostatin A) can improve the reprogramming of microRNA genes in addition to protein-
coding genes in cloned embryos (15). One recent report suggests that microRNA reprogramming after SCNT is incomplete and inconsistent in elongating bovine embryos (13). Our data support the notion that microRNA-mediated gene silencing may be significantly disrupted in nuclear transfer embryos compared with IVF controls.

An increase in the incidence of apoptosis in SCNT embryos cultured in vitro has been reported previously (22, 32, 43). Our results suggest that increased programmed cell death may not be unique to SCNT embryos maintained in an in vitro culture environment but, rather, is associated with the SCNT process regardless of culture environment. Indeed, it is likely that the cellular stress imposed by widespread gene misregulation (via epigenetic and/or microRNA pathways) would lead to apoptosis in embryos, as this phenomenon has been reported in other model systems (see Refs. 7, 27, 36, 58), including preimplantation-stage bovine embryos (53).

It is interesting to note that even though widespread misregulation of gene expression was also observed in trophoblast tissue from SCNT embryos (as in the ED), neither the same genes nor the same types of genes were shown to be aberrantly expressed in the distinct cell types. Deficiencies in control over gene expression (ED) are instead replaced by dysfunctional metabolic/catabolic pathways and subcellular organizational defects in trophoblast tissue. These observations fit nicely with data previously reported by our laboratory that details gross cytological and histological defects in trophoblast tissue of elongating nuclear transfer embryos (46). It is as yet unclear why, in SCNT embryos, trophoblast cells and cells of the ED have developed such distinct responses to the shared insult of faulty genomic reprogramming during early cleavage-stage development. Whatever the reason, the consequences are clear: placental defects are a key factor in the low embryonic, fetal, and neonatal survival rates after SCNT in all species studied to date (reviewed in Refs. 4, 14, 56). Further insight into this apparent discrepancy between ED and TE regarding epigenomic reprogramming will be crucial to improving the poor performance of SCNT embryos.

Embryos derived from PA were included as a part of this study in an attempt to shed additional light on the phenomenon of genomic imprinting and how faulty imprinting can impact preattachment development. Parthenogenetic embryos are the result of artificially activated oocytes, without any contribution (genetic or otherwise) from spermatozoa. Mammalian parthenogenotes created in this manner are incapable of developing to term, largely as a consequence of improper genomic imprinting (reviewed in Refs. 11, 55). [Parthenogenetic mice have been born that survived into adulthood, but significant genetic engineering to imprinted regions of the genome was necessary to allow this to happen (39).] We hypothesized that a subset of imprinted genes (those expressed exclusively, or preferentially, from the paternally derived allele) would be consistently absent or underepressed when comparing parthenogenotes against IVF embryos. Indeed, when we compared the ED read counts from parthenogenetic samples against those from IVF samples, eight genes (PLAGL1, DIRAS3, SNRPN, NNAT, PEG10, NDN, TMSB4X, KBTBD6) showed a read-count ratio (PA/IVF) of <0.1 for at least two of the three gestation days examined; six of the eight have been previously shown to be preferentially expressed from the paternal allele (in italics above; Ref. 49). For TE samples, 14 genes (PEG10, MEST, AVP, TMEM22, IGFBP7, BGN, APLNR, EHD2, NRN1, CTS1, SLC22A7, TMEM119, CLEC10A, TMSL3) were found that fit that same criteria, yet only two of those were known imprinted genes (in italics above; Ref. 49). Differential gene expression between PA and IVF samples does not automatically qualify a gene as being imprinted. Nevertheless, these are intriguing candidates that we are aggressively investigating at the current time. It has been demonstrated that IVF and SCNT embryos alike are plagued by ART-induced defects in imprinted gene expression.
(see Refs. 12, 26, 67–69, 74). We anticipate that a more complete understanding of genomic imprinting and the role it plays in early embryo development may inform future research regarding the treatment of embryos generated using ART.

The correlations between the quantitative PCR data (C_T values) and the sequencing data (read counts) were compelling. We observed a very few conspicuous discrepancies (DNMT3B, GNAS, IGF2, e.g.) that were likely due to faulty quantitative PCR assays, as the C_T values for these three genes were consistently noted as outliers. Removal of these three genes from the analysis boosted the average correlation coefficient between C_T values and read counts from −0.722 to −0.861. We attribute the remaining minor discordance between C_T values and read counts to the conservative approach we took to aligning reads to our database. Reads were accepted as aligned only if they mapped to exactly one member of our database, thus expressed genes with closely related family members or common and highly conserved motifs may be underrepresented in the read count data but accounted for accurately via quantitative PCR. Notably absent from the quantitative PCR experiments were the day 10 nuclear transfer samples. These were our least abundant samples, and after sequencing, not enough cDNA was left for quantitative PCR analysis.

In summary, this was, to our knowledge, the first large-scale effort to describe differences in global gene expression patterns between peri-attachment pig conceptuses produced using different ART. We have been successful at identifying specific biological processes and signaling pathways that may be impinged upon due to in vitro culture, somatic cell nuclear reprogramming, and parthenogenesis. We have identified a number of new putatively imprinted genes that merit additional consideration. In addition, we have further refined our technical and statistical approaches to global gene expression profiling by high-throughput sequencing. We validated our data with medium-throughput quantitative PCR using the BioMark system from Fluidigm. Significant embryonic losses are incurred in ART embryos at or around the time of embryonic attachment. These data will serve as an invaluable resource to those with an interest in discovering the cause(s) and strategies for mitigation of early embryonic loss in swine embryos generated by ART.

ACKNOWLEDGMENTS

The authors gratefully acknowledge August Rieke, Melissa Samuel, Eric Walters, and David Wax for invaluable contributions to the extensive husbandry efforts required for this project; Kristin Whitworth for molecular biology assistance; and Fluidigm for providing assays for RNA-Seq validation with the BioMark qPCR system.

Current address for R. Li: Center of Metabolic Disease Research, Nanjing Medical University, 140 Han Zhong Rd., Nanjing 210029, China.

GRANTS

This work was funded primarily by the National Center for Research Resources Grant R01 RR-013438 and Food for the 21st Century at the University of Missouri (R. S. Prather) and through Utah Agricultural Experiment Station program projects UTA00151 and UTA01062 for S. C. Isom and J. R. Stevens, respectively.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

ENDNOTE

At the request of the authors, readers are herein alerted to the fact that additional materials related to this manuscript may be found at the institutional website of author S. C. Isom, which at the time of publication is: https://adv.s.usu.edu/htm/faculty-staff/memberID=4042. These materials are not a part of this manuscript and have not undergone peer review by the American Physiological Society (APS). APS and the journal editors take no responsibility for these materials, for the website address, or for any links to or from it.

AUTHOR CONTRIBUTIONS

SCI was responsible for experimental design, sample collection, data generation and analysis, and manuscript preparation; IRS developed and implemented the statistical models for read count analysis; RL performed the manipulations for somatic cell nuclear transfer; WGS oversaw the bioinformatics aspects of this project (short read trimming, parsing, and alignment); LC carried out the quantitative PCR; LDS performed embryo transfers, and assisted with sample collections and in vitro embryo culture; CNM performed the surgical embryo transfers; RSP paid for, directed, and helped to design the experiments detailed herein.

REFERENCES


