Animal Reproduction Science 149 (2014) 39–45

Contents lists available at ScienceDirect

Animal Reproduction Science

journal homepage: www.elsevier.com/locate/anireprosci

Genomic profiling to improve embryogenesis in the pig☆

Randall S. Prather*, Bethany K. Redel, Kristin M. Whitworth, Ming-Tao Zhao

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

ARTICLE INFO

Article history:
Available online 14 May 2014

Keywords:
Transcription
Embryo
DNA methylation
Metabolism
Profiling

ABSTRACT

Over the past decade the technology to characterize transcription during embryogenesis has progressed from estimating a single transcript to a reliable description of the entire transcriptome. Northern blots were followed by sequencing ESTs, quantitative real time PCR, cDNA arrays, custom oligo arrays, and more recently, deep sequencing. The amount of information that can be generated is overwhelming. The challenge now is how to glean information from these vast data sets that can be used to understand development and to improve methods for creating and culturing embryos *in vitro*, and for reducing reproductive loss. The use of ESTs permitted the identification of SPP1 as an oviductal component that could reduce polyspermy. Microarrays identified LDL and NMDA as components to replace BSA in embryo culture media. Deep sequencing implicated arginine, glycine, and folate as components that should be adjusted in our current culture system, and identified a characteristic of embryo metabolism that is similar to cancer and stem cells. Not only will these characterizations aid in improving *in vitro* production of embryos, but will also be useful for identifying, or creating conditions for donor cells that will be more likely to result in normal development of cloned embryos. The easily found targets have been identified, and now more sophisticated methods are being employed to advance our understanding of embryogenesis. Here the technology to study the global transcriptome is reviewed followed by specific examples of how the technology has been used to understand and improve porcine embryogenesis both *in vitro* and *in vivo*.

1. Introduction

The basic premise of understanding a cell is that DNA makes RNA makes protein. Thus for a greater understanding of the early mammalian embryo we must understand its DNA, RNA and protein. The genomes of most species have now been sequenced and working drafts of their genomes are readily available (Groenen et al., 2012) for a perspective on the pig genome (Prather, 2013). From the basic structure of the genome, predictions can be made about the sequence of the RNAs, and the control of their transcription. From the RNAs, predictions can be made both about the proteins that are made and their function. Thus if one were to determine which RNAs were present at a specific stage of development, *e.g.* the blastocyst stage, then it should be possible to predict the genes that were transcribed and the proteins that are made. Extrapolation of that data should permit the description of functional pathways that are present in the blastocyst stage embryo. If this were compared to a blastocyst stage embryo that was created *in vitro*, then it might be possible to predict which culture condition to alter to make the *in vitro* produced embryo less different from the *in vivo* produced embryo.

☆ This paper is part of a special issue entitled: 4th Mammalian Embryo Genomics meeting, Guest Edited by Marc-Andre Sirard, Claude Robert and Julie Nieminen.

* Corresponding author at: 920 East Campus Drive, University of Missouri, Columbia, MO 65211, USA. Tel.: +1 573 882 6414.
E-mail address: PratherR@missouri.edu (R.S. Prather).

http://dx.doi.org/10.1016/j.anireprosci.2014.04.017
0378-4320/

2. Ribonucleic acid

Quantification of the different RNAs in a cell can be challenging. Generally the RNA in focus has been messenger RNA (mRNA) because it is easy to isolate due to its poly(A) tail, and the RNA can be used to predict both the gene and the protein. There are a number of caveats that that should be remembered when working with mRNA. First, one should be careful about interpreting the data from poly(A) isolated RNA, especially if the samples are collected during oocyte. There are numerous examples of preexisting messages being polyadenylated, translated and degraded (Dai et al., 2000, 2005). Such polyadenylation can easily be visualized with an assay that measures the length of the poly(A) tail, and includes transcripts such as *KPNA7*, *H1FOO*, *ID3*, and *PARL* (Dobbs et al., 2010). The question here relates to the isolation of mRNA and does the efficiency of mRNA recovery change when the poly(A) tail lengthens. Secondly, it should also be remembered that mRNA accounts for only 1–3% of the RNA in a typical cell. Often ignored are ribosomal RNA (>80% of total RNA), transfer RNA, long coding RNA, signal recognition particle RNA, small nuclear RNA, small nucleolar RNA, telomerase RNA, micro RNA, Piwi-interacting RNA, and small interfering RNA (reviewed by Prather et al., 2013). Even when the focus includes these diverse RNAs, the RNAs themselves can be edited. Over 100 types of RNA modifications have been identified in all three kingdoms of life. These include changing adenosine to inosine (resulting in an A to G conversion in how the ribosome reads the codon), and methylation of adenosine and cytosine in RNA. For a review of RNA editing (Mallela and Nishikura, 2012). RNA editing has obvious ramifications as an RNA sequence may neither predict the DNA sequence from which it was derived, nor the sequence of amino acids that are translated. Similarly, some of the RNA modifications may affect stability, turnover and translation rate. If this is not complicated enough, proteins can be edited. Inteins are protein sequences that can be spliced out of polypeptides (Elleuche and Poggeler, 2010) and even replace other inteins in cis and trans (Appleby-Tagoe et al., 2011; Aranko et al., 2013). It should be remembered that most of the technologies provide a snapshot of mRNA abundance and do not provide any additional information. Thus the regulation of RNA production, post-transcriptional modification, protein production and post-translational modifications all serve to drive a very complex system. For most technologies analysis of RNA abundance is lethal to the cells or embryos. Efforts to develop technologies that do not harm the embryo include evaluation of the first or second polar body (Klatsky et al., 2010; Jiao and Woodruff, 2013) as this may accurately reflect the abundance of message in the oocyte, and thus predict both the abundance of maternal RNAs in the embryo and hence the developmental quality of the resulting embryo.

3. Profiling technologies

Many technologies are available to quantify RNAs in a sample of cells. Over the past decade the technology to characterize transcription during embryogenesis has progressed from estimating a single transcript to a reliable description of the entire transcriptome. Northern blots were followed by sequencing ESTs, quantitative real time PCR, cDNA arrays, custom oligo arrays, and more recently, deep sequencing (the methods and limitations of these technologies have been recently reviewed from a pig centric viewpoint (Prather et al., 2013). These studies have shown a complex metabolic switch at the transition from maternal control of development to zygotic control of development (MTZ), and these changes are strikingly similar across species (Ostrup et al., 2013). Many of the transcripts that are enriched for prior to the MTZ code for proteins that have a cytoplasmic function; while those enriched for after the MTZ code for proteins that have a nuclear function. Not only is a there a major shift in the transcriptome and hence metabolism of the embryo, when the embryo reaches the blastocyst stage it should be remembered that the early blastocyst stage embryo is composed of at least 3 different cell types that can be defined by their morphology and expression of *Nanog*. This includes the trophectoderm, the inner cell mass composed of epiblast (*Nanog* positive) and the hypoblast (primitive endoderm that is *Nanog* negative: those cells separating the inner cell mass from the blastocoel cavity) each with a different transcriptional signature (Silva et al., 2009; Lanner and Rossant, 2010). When an intact blastocyst stage embryo is used for RNA isolation all three cell types contribute to the final transcript abundance. Thus care should be exercised when interpreting the results as biologically important differences may be masked.

To further exacerbate the problem we tend to think of the pre-blastocyst stage embryos as being uniform, *i.e.* all blastomeres are equal. In reality they may not be equal. An example is the maternal *Trim28* mutant mouse embryo. Trim28 is a protein that is required for protection of the differentially methylated region (DMR) during the global DNA demethylation observed during the cleavage stages (Messerschmidt et al., 2012) and aberrant expression can result in 8-cell stage embryos that have a mosaic DNA methylation pattern (Messerschmidt et al., 2012). Not only might there be mosaic patterns of DNA methylation, the pattern of gene expression within cell types, but between cells has the potential to be quite different. This mosaic expression may be the result of pulsatile expression of specific genes that results in a great deal of variation, or 'noise' in abundance of a transcript (Levine et al., 2013; Sanchez and Golding, 2013). Thus when data from these early embryos is generated it will behoove the reader to be careful of the interpretation, and remember that these embryos may have mosaic epigenetic marks and mosaic transcriptional profiles. Unless single cells are measured the results will be an average of the cells. Thus the oocyte and preimplantation embryo represent a highly dynamic system that on the surface appears to be quite simple, but in reality is quite complex.

This complexity in combination with these technologies now available to describe changes in transcription and DNA structure provide a wealth of information about the embryo. In fact, the amount of information that can now be generated is overwhelming. The challenge is not