

# Proteomic Comparisons of Differentiating Embryonic Stem Cells: An *in vitro* Model of Embryogenesis

Ron Orlando,<sup>1</sup> Peggi M. Angel,<sup>1</sup> D. Brent Weatherly,<sup>2</sup> James A. Atwood<sup>2</sup>

<sup>1</sup>Complex Carbohydrate Research Center & Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA30602

<sup>2</sup>Biolnquire, Athens, GA 30602

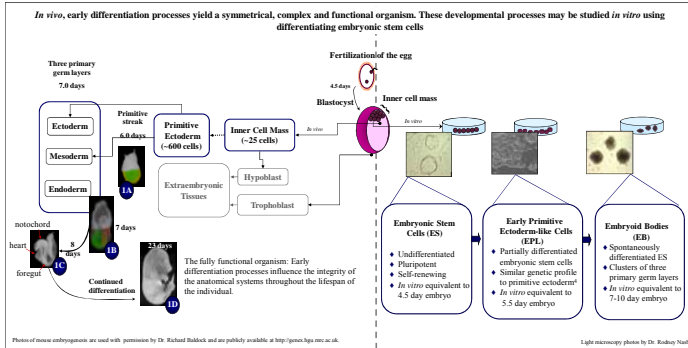
## Overview

Perform a qualitative and quantitative survey of comprehensive protein expression during early development (embryonic days 5 - 8) using differentiating murine embryonic stem cells as an *in vivo* model for early stage mammalian development.

## Introduction

Early mammalian embryogenesis is defined as the time period spanning fertilization of the egg until all major organs of the animal are represented. In mice, at around 4.5 days after fertilization, the developing embryo consists of the blastocyst, which is a collection of ~25 cells called the inner cell mass, surrounded by the trophoblast shell. The trophoblast will form the extra embryonic tissues that nourish and protect the organism during development. The inner cell mass proliferates and differentiates and by ~6 days after fertilization has formed a distinct secondary population of 600 cells called the primitive ectoderm. At approximately 7 days, a column of differentiating cells called the primitive streak begins to extend from this cell population. Differentiating cells disperse from the tip of the primitive streak, to form the tissues, then organs of the developing organism.

The small number of cells that compose each of these developmental stages has limited the study of mammalian development to the investigation of RNA or DNA changes that occur during these time periods. Reports on protein expression are limited to following single protein expression during early development. One way to study these limited developmental populations at a translational level is by using embryonic stem cells to model the early differentiation events. In fact, genetic analysis of mouse embryonic stem cells and partially differentiated embryonic stem cells has been compared to the early mouse embryo, forming a positive association between the *in vitro* and *in vivo* models of early embryogenesis.

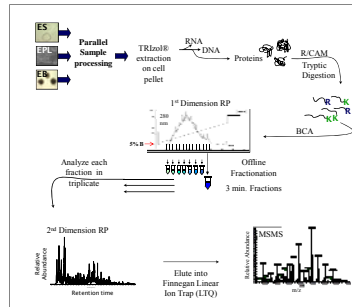


In the current study, three cell types, embryonic stem cells, early primitive ectoderm-like cells, and embryoid bodies are utilized to depict these processes at a translational level. Embryonic stem cells are derived from the *in vivo* 4.5 days post coitum (d.p.c.) inner cell mass. Early primitive ectoderm-like cells are partially differentiated embryonic stem cells with a genetic profile similar to the 5.5 d.p.c. primitive ectoderm. Finally, embryoid bodies are representative of definitive differentiation, with some similarities to the 6-8 d.p.c. embryo. We explore the protein regulation of these cells using shotgun proteomics and spectral counts to correlate the *in vitro* protein expression to the *in vivo* events of early embryogenesis.

## Experimental

R1 murine ES cells were cultured in the absence of feeders on tissue culture grade plastic, pre-coated with 0.1% gelatin-phosphate buffered saline (PBS). ES cell culture medium consisted of Dulbecco's Modified Eagle Medium (DMEM, Gibco BRL) supplemented with 10% fetal calf serum (FCS, Commonwealth Serum Laboratories), 1 mM L-glutamine, 0.1 mM 2-mercaptoethanol, 100 U/ml penicillin, 100 U/ml streptomycin and 1000 U/ml recombinant human leukemia inhibitory factor (LIF) (ESGRO) at 37°C in 10% CO<sub>2</sub>. Spontaneous differentiation of ES cells into embryoid bodies was induced by removing the ES cells from the plates with 0.25% trypsin (Gibco), and replating on 0.2% gelatin-coated plates in non-conditioned complete medium in the absence of LIF. After 2 days cells were trypsinized, washed 3 times with ice-cold PBS and stored at -80°C. R1 mEPL cells were maintained in ES-medium containing 50% MeFl conditioned medium supplemented with 1 mM L-glutamine, 0.1 mM 2-mercaptoethanol, 100 U/ml penicillin, 100 U/ml streptomycin and 1,000 U/ml recombinant LIF at a density of 2 x 10<sup>4</sup> cells/cm<sup>2</sup> in mEPL cells were passaged every 48 hours.

TRIZol® reagent was used to isolate the RNA, DNA and proteins from the cells. The purified protein pellet was reduced and carboxymethylated before being subjected to tryptic digestion. After desalting, BCA quantitation was performed on the peptide population, which was then subjected to a first dimension reverse phase separation. In this step, a total of eight 3-minute fractions were collected from a 4.5 x 150 mm Jupiter C18 column eluted with a linear gradient progressing from 5% B to 65% B. Mobile phase A was 0.1% TFA, mobile phase B was 0.1% TFA/acetonitrile. Fractions were lyophilized to dryness before the second dimension analysis. Second Dimension Reverse Phase: Each fraction was analyzed three times on a 150 μm x 150 mm C18 interfaced with a Finnigan LTQ mass spectrometer. Peptide fractions were eluted using a concentration of mobile phase B (0.1% FA/ACN) that bracketed each collected fraction by ~10% B and eluted over a period of 100 minutes. Each full MS scan (m/z range 450-2000) was followed by nine MS/MS events on the nine most abundant peaks with an exclusion duration of 160 s.



## Dataflow – Data Processing

The experimental approach taken led to several challenges with data processing.

Specifically, three biological populations (ES, EPL, and EB), subjected to a 2-dimensional LC separation where the first dimension produced 10 fractions, each of which subjected to LC-MS/MS. This was repeated 3 times for each cell state, leading to 90 total LC-MS/MS runs. Each of these was processed according to the following steps.

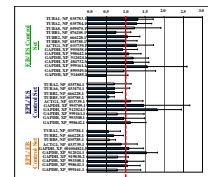
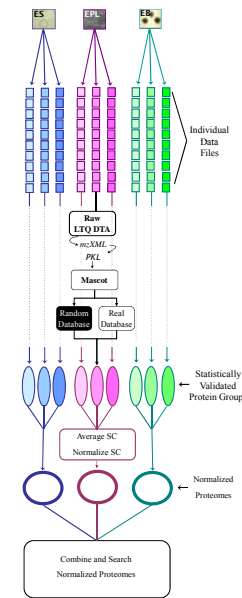
Data files from each LC-MS/MS run were searched against a forward and reversed database using Mascot. **ProteoIQ** was used to cluster peptides to proteins (protein groups) and output lists of proteins having a protein false discovery rate (**Pro FDR**) less than 1%. **ProteoIQ** also kept track of spectral counts from each replicate.

Spectral counts were averaged between runs. The average spectral count for each protein was normalized using Equation 1. This equation normalize the individual protein to the sampling of the cell population as well as normalize for the amount of space each protein occupies in the identified proteome.

$$\text{Equation 1: } N\_SC = \frac{SC_{\text{individual}} / \#AA_{\text{individual}}}{\sum (SC_{\text{Total}} / \#AA_{\text{Total}})}$$

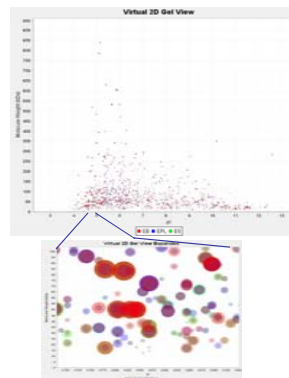
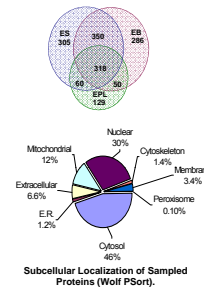
## Normalization Evaluation

39 proteins that have been previously used as loading controls during western blot analysis were used to evaluate the normalization process. As shown by the **bar graph** to the right, the normalization functions appear to have been successful since the expression ratios between the cell types for these controls are all approximately equal to 1.



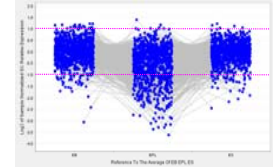
## Protein Distribution and Subcellular Localization

A total of 3328 Proteins were found across these three cell types

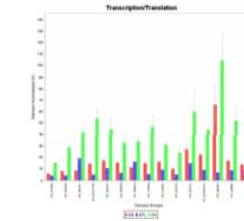


## Analysis of Differentially Expressed Proteins

Proteomic analysis identified 1498 proteins that were contained in all three of these cell types, many of which were found to have the similar expression levels, as shown by the **cluster plot** to the right. Proteins identified to have a 2x change in expression were further investigated to determine their biological function.



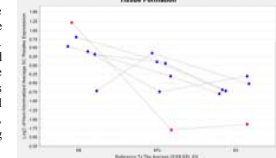
## Up-regulation in ES cells



Many of the proteins found to be up-regulated in ES cells were determined to be transcription/translation factors, which are shown by the normalized spectral counts for all detected transcription/translation factors. Two of these, ELYS and AHCTF1, been shown via knockout mice to be required for survival and development of the inner cell mass. Microarray experiments on wild type mouse embryos showed that AHCTF1 was strongly expressed throughout 6.5 d.p.c while later stages of development showed down regulation of this gene. AHCTF1 has not been reported by other large scale proteomic studies. On the whole, the embryonic stem cell proteins found to be up-regulated in this study appear to correspond with the *in vivo* time point of embryogenesis, prior to differentiation of the inner cell mass to hypoblast and epiblast or can be correlated to previous studies of embryonic stem cells.

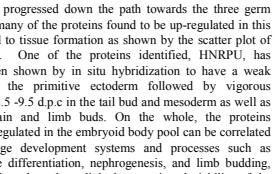
## Up-regulation in EPL cells

Immunocytochemistry has previously demonstrated that many of the proteins up-regulated in early primitive ectoderm-like (EPL) are associated with cell proliferation, migration, and remodeling. Proteins associated with cell migration, such as the well studied SPARC, secreted acidic cysteine rich glycoprotein, were found to be up-regulated in this stage, as shown by the **cluster plot** of proteins with known roles in cell migration. Overall, the profile of these cell migration proteins is reflective of morphogenesis and differentiation, associating this time point with a primitive ectoderm commencing towards differentiation.



## Up-regulation in EB cells

EB cells have progressed down the path towards the three germ layers; hence many of the proteins found to be up-regulated in this cell are related to tissue formation as shown by the scatter plot of these proteins. One of the proteins identified, HNRPU, has previously been shown *in situ* hybridization to have a weak expression in the primitive ectoderm followed by vigorous expression at 8.5-9.5 d.p.c in the tail bud and mesoderm as well as in the forebrain and limb buds. On the whole, the proteins identified up-regulated in the embryoid body pool can be correlated with later stage development systems and processes such as skeletal/muscle differentiation, nephrogenesis, and limb budding, and many of these have been linked to continued viability of the embryo.



## Conclusions:

We have completed a study involving the comprehensive proteomic profiling of murine embryonic stem cells as they progress from an undifferentiated state to a differentiated state. Over 3300 proteins were identified - less than 20% of these proteins have been investigated in terms of embryogenesis.

The functionality of both the single and comparative populations appears to recapitulate *in vivo* differentiation events. Regulation of proteins with known functions follows regulation seen *in vivo* during murine embryogenesis.

The majority of the proteins could not be correlated to known functions, indicating many candidates for future studies on mammalian differentiation.

## Acknowledgement:

Funding for this work was provided by the NIH/NCR Integrated Technology Resource for Biomedical Glycomics, grant P41 RR018502, and the NIH/NCR Resource Center for Integrated Glycotechnology, grant P41 RR05351.

Early Mammalian development



Embryonic stem cells are developed from the inner cell mass, seen *in vivo* at embryonic day 4.5. *In vivo*, the inner cell mass differentiates to a second pluripotent population, primitive ectoderm.



Early primitive ectoderm-like cells, genetically equivalent to the *in vitro* primitive ectoderm. *In vivo* at embryonic day 6.0-6.5, cells begin to migrate and differentiate to form the tissues and organs of the developing organism.



Embryoid bodies represent a state of definite differentiation. Although not the directed differentiation of the implanted embryo, these cell types are often used to represent the 7 day embryo.