ROLES FOR THE CELL CYCLE AND MICRORNAS IN PLURIPOTENT STEM CELLS

by

ROBERT DAVID TROST

(Under the Direction of Stephen Dalton)

ABSTRACT

Pluripotent stem cells (PSCs) are a promising source of material for various cell-based therapies and vital for the understanding of early human development. Maintenance of PSCs and creation of various somatic cell types is orchestrated by the integrations of very complex networks and understanding how their cell-cycle is regulated can improve their maintenance and differentiation *in vitro*. Here I demonstrated, with the utilization of the FUCCI reporter system, a novel signaling event occurring in the G1 phase of human embryonic stem cells (hESCs).

MicroRNAs are major regulators of self-renewal and differentiation in hESCs, allowing for broad and rapid changes in their transcriptional profiles. I identified a microRNA located within a primate-specific microRNA cluster, *miR-520g*, which regulates the inhibitory Smad6. I also show Smad6 being necessary for mesoderm commitment, which reveals a novel role of Smad6 outside of its autoinhibitory activity in Activin/TGFβ signaling.

INDEX WORDS: Stem cells, microRNA, FUCCI, cell-cycle, Smad6, pluripotency

ROLES FOR THE CELL CYCLE AND MICRORNAS IN PLURIPOTENT STEM CELLS

by

ROBERT DAVID TROST

BS, University of North Carolina at Wilmington, 2009

BA, University of North Carolina at Wilmington, 2009

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF PHILOSOPHY

ATHENS, GEORGIA

2016

ROLES FOR THE CELL CYCLE AND MICRORNAS IN PLURIPOTENT STEM CELLS

by

ROBERT DAVID TROST

Major Professor:	Stephen Dalton
Committee:	Shelley Hooks
	Lianchun Wang
	Michael Pierce

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2016

ACKNOWLEDGEMENTS

I would like to first thank my parents, Tim and JoAnn , for their 29 years of unending love and support. I'd also like to thank my siblings, Nick and Alexis, for their support and compassion over the years. Lastly, I'd like to thank Dr. Steve Dalton for his guidance and dedication throughout my graduate career.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

PLURIPOTENT STEM CELLS

Pluripotent stem cells (PSCs) are defined by their potential to give rise to derivatives of all three germs layers and self-renew indefinitely (Chng, Vallier, and Pedersen 2011b). To meet these requirements, PSCs propagated *in vitro* should have the potential to generate the somatic and germ cells of an adult if reintroduced into a blastocyst (Ng and Surani 2011). The two main types of stem cells are embryonic and somatic (i.e. adult). Embryonic stem cells (ESCs) are derived from the inner cell mass (ICM) of the developing embryo. Grown under appropriate conditions (Figure 1.1), ESCs will mimic embryonic development and therefore contain a high therapeutic value in their study of the developing human embryo past implantation (Thomson et al. 1998). Somatic stem cells are generated during fetal development and are maintained through adult life. Although somatic cells are more restricted in differentiation potential than ESCs, their deregulation is the cause of several disease states such Parkinson's disease, diabetes, spinal cord injury, and macular degeneration therefore their understanding is of great medical significance. In 2006, a new type of stem cell was created, known as induced pluripotent stem cells (iPSCs), through the reprogramming of adult cells. This discovery has allowed the derivation of patient specific treatment of disease, further enhancing the utility

of PSCs in regenerative medicine (Takahashi and Yamanaka 2006). iPSCs also give researchers the ability to avoid ethical issues in the use of human embryos for stem cells research.

SIGNALING PATHWAYS REGULATE STEM CELL SPECIFICATION

In our lab, ESCs are directed towards functional, mature cell populations using growth factors or small molecules that recapitulate embryonic development in the culture dish (Smith 2001). Creation of mature, functional somatic cells requires ESCs to first go through multipotent progenitor intermediates. This requires a detailed understanding of embryonic development, which has been gained through the use of animal models. These studies have implemented a complex integration of both extrinsic and intrinsic signaling pathways during embryonic development (Smith 2001; Payne, King, and Hay 2011).

Generation of the three germ layers is one of the most important events during embryogenesis, because these germ layers are the progenitors to all adult tissues (Smith 2001). Germ layer formation initiates during gastrulation of the developing embryo, where a primitive streak (PS) first forms from the ingression of the pluripotent epiblast. Uncommitted cells migrate out of the primitive streak and based on their positioning, become either mesoderm or definitive endoderm through a shared mesendoderm progenitor (Payne, King, and Hay 2011; Murry and Keller 2008). Specification of these cells towards different germ layers is regulated both temporally and spatially, under the control of different signaling environments. Wnt, fibroblast growth factor (FGF), bone morphogenetic protein (BMP), and

Activin/Nodal ligands heavily influence these signaling environments during patterning of the embryo (Beyer et al. 2013).

BMP, Activin, and Nodal ligands belong to the TGF-β signaling superfamily (Massague, Seoane, and Wotton 2005). These factors are secreted, and depending on ligand availability, dosages, and context, exert pleiotropic and sometimes opposing functions, to allow the multi-lineage cell commitment during embryogenesis (Beyer et al. 2013). Nodal and Activin signaling, carried out by their effectors Smad2/3, play important roles in the developing embryo pre- and post-gastrulation by specifying early mesendoderm structures, and by maintaining the epiblast (Massague, Seoane, and Wotton 2005). NODAL-null embryos display hypo-proliferation and reduction of pluripotency markers in the epiblast, consistent with the role of Smad2/3 signaling in maintenance of hESCs *in vitro* (Mesnard, Guzman-Ayala, and Constam 2006; Yu et al. 2011). BMP-dependent Smad1/5/8 signaling is essential during gastrulation and for anterior-posterior axis establishment of the developing embryo. For example, primitive streak formation is abolished in BMP4 deficient mice (Chang et al. 2008)..

Early development has been recapitulated in ESCs by modulation of the before mentioned signaling pathways. Most studies suggest that induction of mesoderm in human ESCs is induced by BMP signaling, while high Activin/Nodal, in the absence of FGF, specifies cells towards endoderm (Beyer et al. 2013). Activation of Wnt signaling in hESCs disrupts self-renewal and supports differentiation towards mesendoderm and mesoderm progenitors. There is a large degree of crosstalk between these signaling pathways in hESCs, creating confusion as to the

role of each pathway during lineage specification (Dalton 2013). For example, Wnt signaling cooperates with Activin during endoderm induction, where β-catenin and Smad2/3 interact with cell-specific transcription factor Eomes to induce transcription of endoderm genes (Dalton 2013). Further identification of the transcriptional landscape during cell commitment will improve the multi-lineage differentiation of hESCs *in vitro*, which is vital for their use in cell therapies.

CELL CYCLE REGULATION

The self-renewal of ESCs in culture is maintained through the use of defined growth factors, cytokines, and small molecules. If grown in self-renewal conditions, ESCs divide symmetrically to create two identical daughter cells (Zwaka and Thomson 2005). Under differentiation conditions, ESCs are specified towards a lineage through asymmetric division events, where one daughter cell remains a stem cell while the other daughter cell sustains a transient proliferating population of a derivative to an adult tissue (Zwaka and Thomson 2005). Cell proliferation occurs through a series of stages collectively termed the cell cycle. The cell cycle is divided into four stages organized around the synthesis (S) phase and mitotic division (M), separated by two intervening gap phases (G1 and G2). The cell cycle is a highly regulated process in ESCs, containing important 'checkpoints', especially at the G2/M and G1/S transitions. These two checkpoints are vital to assess the intrinsic signaling environment and genome integrity of ESCs. A third checkpoint located towards the end of the G1 phase adds the integration of extrinsic signals to the cell, called the restriction (R) point. The R point can signal the stem cell to either

continue towards the G1/S transition, or exit the cell cycle to enter a dormant, or quiescent, state referred to as the G0 phase (Orford and Scadden 2008a). Outcomes of the cell cycle checkpoints are governed by cyclin proteins, cyclin dependent kinases (CDKs), and Retinoblastoma tumour suppressor protein (Rb). Cyclins, with their corresponding CDKs, regulate the phoshphorylation levels of Rb as summarized in Figure 1.3.

Although less is known about the cell cycle of hESCs compared to mESCs, they are both characterized by a shortened cell cycle, a result of an unusually truncated G1 phase with a large percentage of S phase cells (Orford and Scadden 2008b; Singh and Dalton 2009). hESCs and mESCs resist DNA damage, indicating a shared lack of G1 checkpoints in their pluripotent states (Pauklin and Vallier 2013). The characteristics of the cell cycle for these two species diverge in their expression of Cyclin D, where cyclin D2 and CDK4 are upregulated in hESCs upon entry into G1 (Orford and Scadden 2008a). It is well understood that the cell cycle lengthens as ESCs differentiate (also outlined in Figure 1.3), although this shift is still poorly understood (Singh et al. 2013). Recent evidence, including work described in my thesis, suggests a rearrangement in the cell cycle to allow differentiation of PSCS (Calder et al. 2013; Pauklin and Vallier 2013; Singh et al. 2013).

MICRORNAS

Although only 1.2% of the human genome consists of protein coding genes, a large percentage of the genome is transcribed (Mattick and Makunin 2005). A large fraction of the remaining RNA was previously thought to be transcribed from 'junk

DNA', but after the completion of the Human Genome Project approximately 98% of the transcribed genome was confirmed to consist of non-coding RNAs (ncRNAs). Of these ncRNAs, approximately 1000 genes were identified that encode for microRNAs (miRNAs), which account for approximately 3% of the human genome (Ranganathan and Sivasankar 2014).

Significant progress has been made in the past decade in the understanding of miRNA function. They have been identified as important regulators of early cell fate decisions that, like transcription factors, can regulate the expression of multiple genes (Ivey and Srivastava 2010; Greve, Judson, and Blelloch 2013). Currently there are 1,872 annotated human miRNA genes that are processed into ~2,578 mature miRNA sequences (www.mirbase.org). Mature miRNAs are short 20-25-base noncoding RNAs that bind by imperfect matching to the 3'UTR of their target mRNAs, leading to the loss of stability and translational inhibition of the transcripts (Gangaraju and Lin 2009). A few studies have shown select miRNAs can activate translation of their target mRNAs (Gangaraju and Lin 2009; Martinez and Gregory 2010). The sequence on miRNAs responsible for mRNA targeting consists of six to eight bases in the 5' region of a miRNA, referred to as the seed sequence (Greve, Judson, and Blelloch 2013). miRNAs that share a common seed sequence are defined as families, whereas a cluster of miRNAs are expressed on the same locus in the genome and can contain miRNAs from several different families (Greve, Judson, and Blelloch 2013). Individual miRNAs can target hundreds of mRNA targets whereas transcripts can contain hundreds of distinct miRNA binding sites, placing them in large, co-regulated network (Greve, Judson, and Blelloch 2013).

The importance of miRNAs was first understood by deletion of proteins responsible for the multistep miRNA biogenesis pathway (Bernstein et al. 2003). They are transcribed either from individual genes containing their own promoters or within the introns of protein coding genes, which require the use of splicing machinery (Adams, Kasinski, and Slack 2014). MiRNAs are derived from precursor transcripts, called primary miRNAs, by RNA Polymerases II and III. They are first processed in the nucleus into shorter RNA fragments (pre-miRNAs) by the microprocessor complex, containing the Drosha and DGCR8 proteins (Figure 1.4). Pre-miRNAs are then exported to the cytoplasm by the exportin 5-RanGTP shuttle system where they undergo terminal processing into mature miRNAs by enzyme Dicer (Gangaraju and Lin 2009). Dicer loads the miRNA onto the RNA-induced silencer complex (RISC), whose function is supported by a family of proteins called the Argonaute proteins (Gangaraju and Lin 2009).

Initial loss-of-function studies showed important roles for miRNAs in maintaining self-renewal of ESCs as well a loss of pluripotency due to an inability to repress stem cell markers necessary for initial differentiation; however, recent evidence shows miRNAs to have roles in lineage specification of ESCs (Ivey and Srivastava 2010). DGCR8 is exclusively involved in the miRNA pathway, whereas Dicer is involved in both miRNA and siRNA pathways (Gangaraju and Lin 2009). DGCR8 knockout embryos arrest early in development whereas DGCR8-null ES cells show delayed or retarded expression of differentiation markers, most likely as a result of the inability of these cells to repress their self-renewal markers *in vitro*, such as OCT4 and NANOG (Wang et al. 2007). Interestingly, most of the DGCR8-null

ESCs are arrested in the G1 phase, suggesting a large role for miRNAs in promoting the ESC cell G1-S transition (Wang et al. 2007). This is supported by a study that showed miRNAs from the Oct4/Sox2 regulated miR-302 cluster regulated cyclin D1, thus allowing a short G1 phase in hESCs (Card et al. 2008). Ablation of Dicer in ESCs resulting in a similar phenotype to these DGCR8 knockout studies (Bernstein et al. 2003).

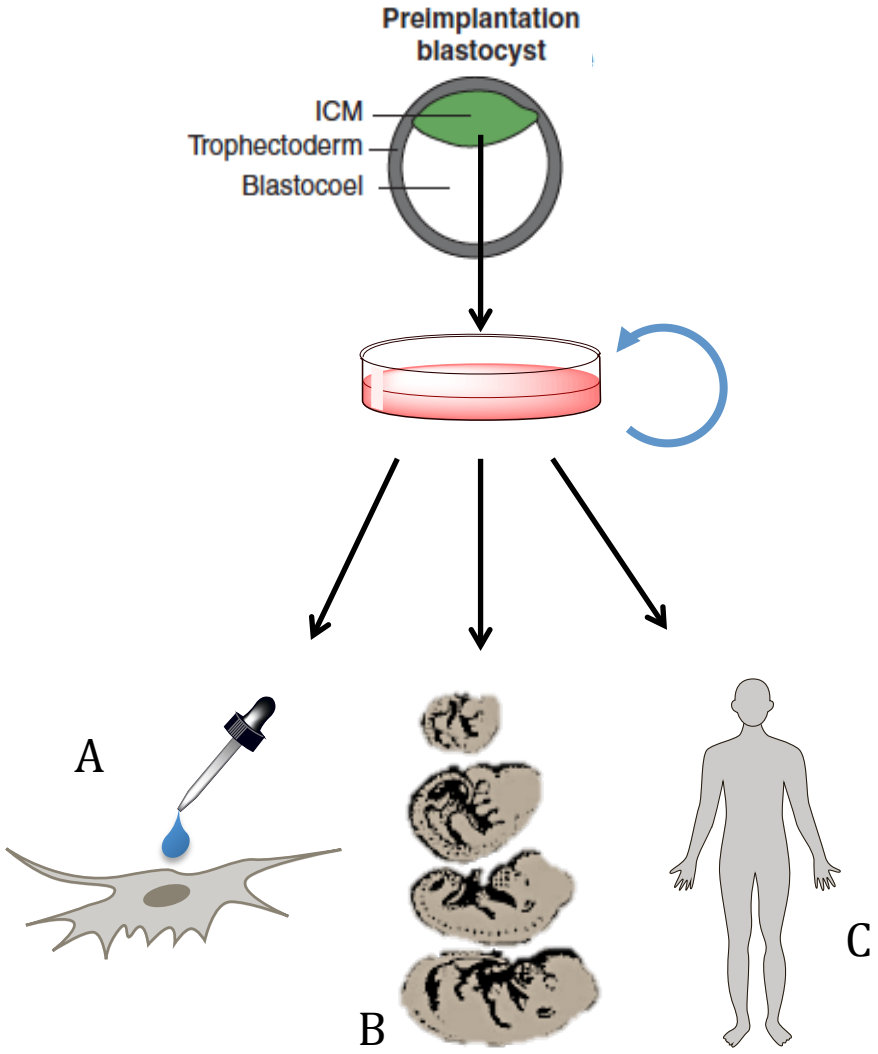

Figure 1.1. Embryonic stem cells (ESCs) are derived from the inner cell mass (ICM) of the developing blastocyst. Cultured ESCs can be used (A) to perform drug toxicity screens, (B) aid in the research of human development, and (C) be used in regenerative medicine. Adapted from (Ohtsuka and Dalton 2008; Chng, Vallier, and Pedersen 2011a).

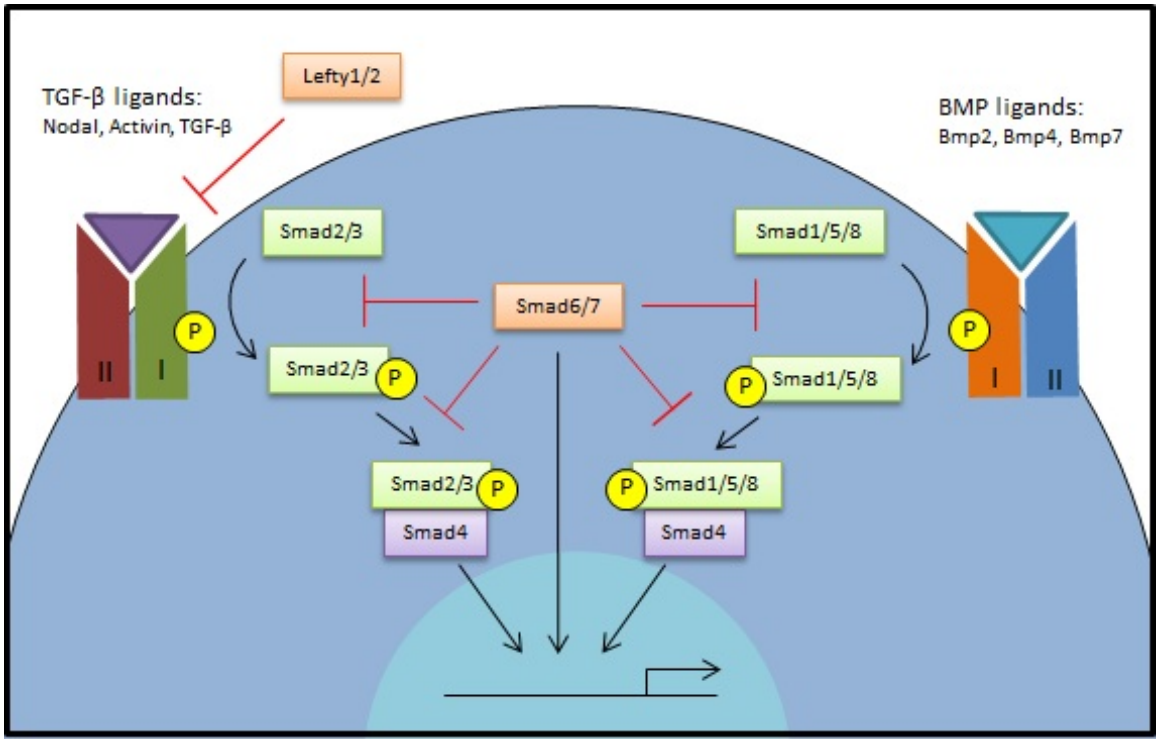

Figure 1.2. Overview of TGF-β signaling. The TGF-β signaling superfamily is divided into 2 branches, TGF-β and BMP. These branches are defined by their effector Smads, Smad2/3 and Smad1/5/8, respectively. Lefty1/2, Smad6, and Smad7 are all inhibitors to TGF-β signaling. Adapted from (Massague, Seoane, and Wotton 2005).

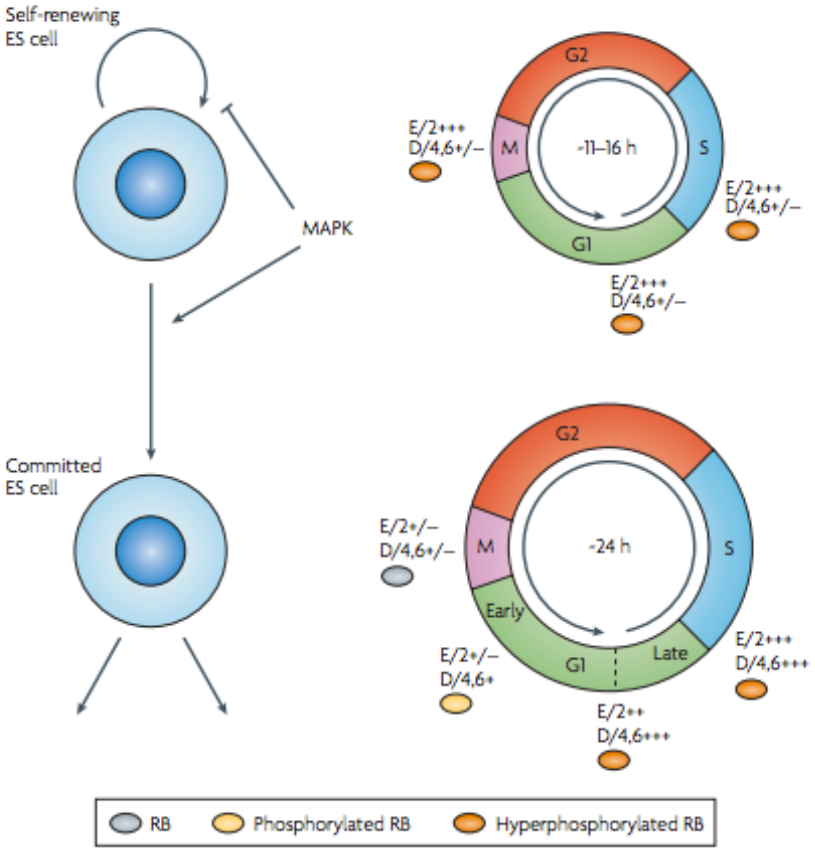

Figure 1.3. Cell cycle in embryonic stem cells. The cell cycle is regulated by cyclins, CDKs, and Rb. The cell division time increases as an ESC becomes more committed. This occurs partly through lengthening of G1 phase. Adapted from (Orford and Scadden 2008a).

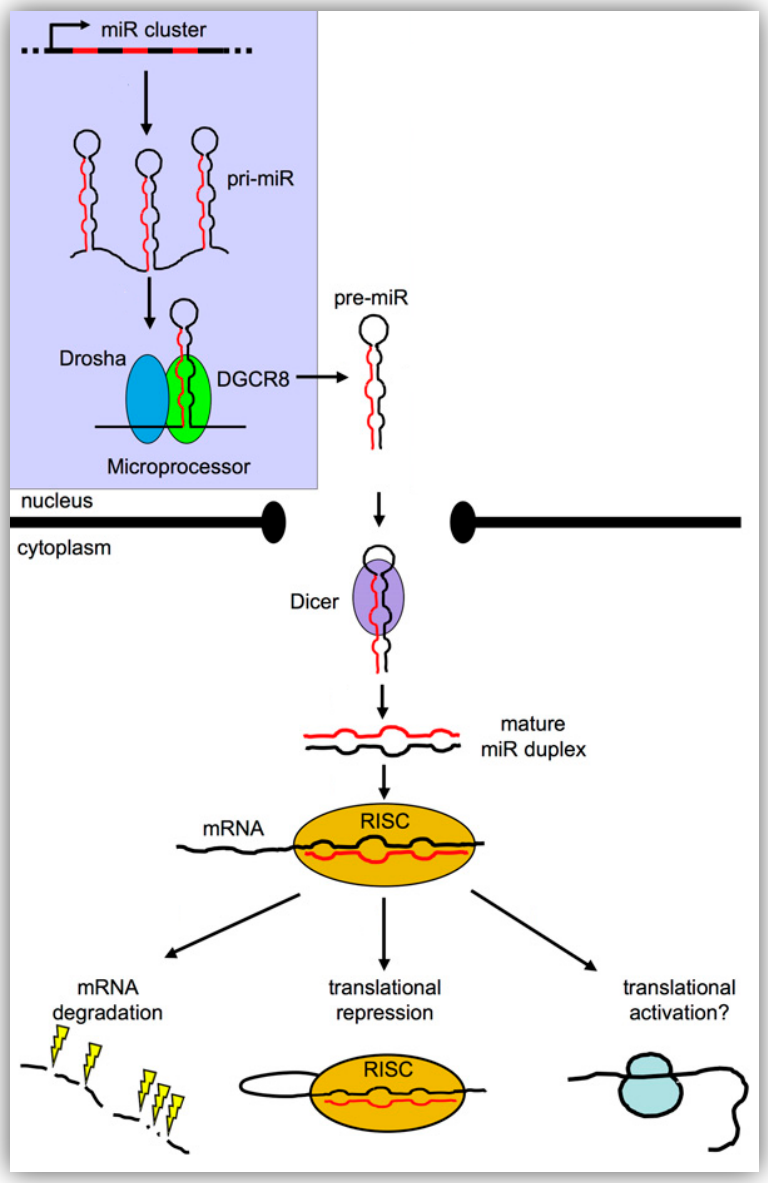

Figure 1.4. Overview of microRNA processing. MicroRNAs are transcribed in the nucleus then exported to the cytoplasm where they can initiate mRNA degradation, translational repression, or even translational activation of their targets. Adapted from (Tiscornia and Izpisúa Belmonte 2010).

REFERENCES

Adams, Brian D, Andrea L Kasinski, and Frank J Slack. 2014. “Aberrant Regulation and Function of microRNAs in Cancer..” *Current Biology : CB* 24 (16): R762–76. doi:10.1016/j.cub.2014.06.043.

Bernstein, Emily, Sang Yong Kim, Michelle A Carmell, Elizabeth P Murchison, Heather Alcorn, Mamie Z Li, Alea A Mills, Stephen J Elledge, Kathryn V Anderson, and Gregory J Hannon. 2003. “Dicer Is Essential for Mouse Development..” *Nature Genetics* 35 (3): 215–17. doi:10.1038/ng1253.

Beyer, T A, M Narimatsu, A Weiss, L David, and J L Wrana. 2013. “The TGFbeta Superfamily in Stem Cell Biology and Early Mammalian Embryonic Development.” *Biochim Biophys Acta* 1830 (2): 2268–79. doi:10.1016/j.bbagen.2012.08.025.

Calder, Ashley, Ivana Roth-Albin, Sonam Bhatia, Carlos Pilquil, Jong Hee Lee, Mick Bhatia, Marilyne Levadoux-Martin, et al. 2013. “Lengthened G1 Phase Indicates Differentiation Status in Human Embryonic Stem Cells..” *Stem Cells and Development* 22 (2): 279–95. doi:10.1089/scd.2012.0168.

Card, Deborah A Greer, Pratibha B Hebbar, Leping Li, Kevin W Trotter, Yoshihiro Komatsu, Yuji Mishina, and Trevor K Archer. 2008. “Oct4/Sox2-Regulated miR-302 Targets Cyclin D1 in Human Embryonic Stem Cells..” *Mol Cell Biol* 28 (20). American Society for Microbiology: 6426–38. doi:10.1128/MCB.00359-08.

Chang, Weise, Zhengshi Lin, Holger Kulessa, Jean Hebert, Brigid L M Hogan, and Doris K Wu. 2008. “Bmp4 Is Essential for the Formation of the Vestibular Apparatus That Detects Angular Head Movements..” Edited by Veronica van

Heyningen. *PLoS Genet* 4 (4): e1000050. doi:10.1371/journal.pgen.1000050.

Chng, Z, L Vallier, and R Pedersen. 2011a. "Activin/Nodal Signaling and Pluripotency." *Vitam Horm* 85: 39–58. doi:10.1016/B978-0-12-385961-7.00003-2.

Chng, Zhenzhi, Ludovic Vallier, and Roger Pedersen. 2011b. *Chapter 3 - Activin/Nodal Signaling and Pluripotency*. *Vitamins and Hormones: Activins and Inhibins*. 1st ed. Vol. 85. Elsevier Inc. doi:10.1016/B978-0-12-385961-7.00003-2.

Dalton, Stephen. 2013. "Signaling Networks in Human Pluripotent Stem Cells." *Curr Opin Cell Biol* 25 (2). Elsevier Ltd: 241–46. doi:10.1016/j.ceb.2012.09.005.

Gangaraju, Vamsi K, and Haifan Lin. 2009. "MicroRNAs: Key Regulators of Stem Cells.." *Nature Reviews. Molecular Cell Biology* 10 (2): 116–25. doi:10.1038/nrm2621.

Greve, Tobias S, Robert L Judson, and Robert Blelloch. 2013. "microRNA Control of Mouse and Human Pluripotent Stem Cell Behavior." *Annu Rev Cell Dev Biol* 29 (1): 213–39. doi:10.1146/annurev-cellbio-101512-122343.

Ivey, K N, and D Srivastava. 2010. "MicroRNAs as Regulators of Differentiation and Cell Fate Decisions." *Cell Stem Cell* 7 (1): 36–41. doi:10.1016/j.stem.2010.06.012.

Martinez, Natalia J, and Richard I Gregory. 2010. "MicroRNA Gene Regulatory Pathways in the Establishment and Maintenance of ESC Identity." *Cell Stem Cell* 7 (1). Elsevier: 31–35. doi:10.1016/j.stem.2010.06.011.

Massague, J, J Seoane, and D Wotton. 2005. "Smad Transcription Factors." *Genes Dev* 19 (23): 2783–2810. doi:10.1101/gad.1350705.

Mattick, John S, and Igor V Makunin. 2005. “Small Regulatory RNAs in Mammals..” *Human Molecular Genetics* 14 Spec No 1 (suppl_1). Oxford University Press: R121–32. doi:10.1093/hmg/ddi101.

Mesnard, D, M Guzman-Ayala, and D B Constam. 2006. “Nodal Specifies Embryonic Visceral Endoderm and Sustains Pluripotent Cells in the Epiblast Before Overt Axial Patterning.” *Development* 133 (13): 2497–2505. doi:10.1242/dev.02413.

Murry, C E, and G Keller. 2008. “Differentiation of Embryonic Stem Cells to Clinically Relevant Populations: Lessons From Embryonic Development.” *Cell* 132 (4): 661–80. doi:10.1016/j.cell.2008.02.008.

Ng, Huck-Hui, and M Azim Surani. 2011. “The Transcriptional and Signalling Networks of Pluripotency..” *Nat Cell Biol* 13 (5): 490–96. doi:10.1038/ncb0511-490.

Ohtsuka, S, and S Dalton. 2008. “Molecular and Biological Properties of Pluripotent Embryonic Stem Cells.” *Gene Ther* 15 (2): 74–81. doi:10.1038/sj.gt.3303065.

Orford, Keith W, and David T Scadden. 2008a. “Deconstructing Stem Cell Self-Renewal: Genetic Insights Into Cell-Cycle Regulation..” *Nat Rev Genet* 9 (2): 115–28. doi:10.1038/nrg2269.

Orford, Keith W, and David T Scadden. 2008b. “Deconstructing Stem Cell Self-Renewal: Genetic Insights Into Cell-Cycle Regulation.” *Nat Rev Genet* 9 (2): 115–28. doi:10.1038/nrg2269.

Pauklin, Siim, and Ludovic Vallier. 2013. “The Cell-Cycle State of Stem Cells Determines Cell Fate Propensity..” *Cell* 155 (1): 135–47. doi:10.1016/j.cell.2013.08.031.

Payne, C, J King, and D Hay. 2011. “The Role of Activin/Nodal and Wnt Signaling in Endoderm Formation.” *Vitam Horm* 85: 207–16. doi:10.1016/B978-0-12-385961-7.00010-X.

Ranganathan, Kannan, and Vaishnavi Sivasankar. 2014. “MicroRNAs - Biology and Clinical Applications..” *Journal of Oral and Maxillofacial Pathology : JOMFP* 18 (2): 229–34. doi:10.4103/0973-029X.140762.

Singh, Amar M, and Stephen Dalton. 2009. “The Cell-cycle and Myc Intersect with Mechanisms That Regulate Pluripotency and Reprogramming..” *Cell Stem Cell* 5 (2): 141–49. doi:10.1016/j.stem.2009.07.003.

Singh, Amar M, James Chappell, Robert Trost, Li Lin, Tao Wang, Jie Tang, Brittany K Matlock, et al. 2013. “Cell-Cycle Control of Developmentally Regulated Transcription Factors Accounts for Heterogeneity in Human Pluripotent Cells..” *Stem Cell Reports* 1 (6): 532–44. doi:10.1016/j.stemcr.2013.10.009.

Smith, A G. 2001. “Embryo-Derived Stem Cells: of Mice and Men.” *Annu Rev Cell Dev Biol* 17: 435–62. doi:10.1146/annurev.cellbio.17.1.435.

Takahashi, K, and S Yamanaka. 2006. “Induction of Pluripotent Stem Cells From Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors.” *Cell* 126 (4): 663–76. doi:10.1016/j.cell.2006.07.024.

Thomson, J A, J Itskovitz-Eldor, S S Shapiro, M A Waknitz, J J Swiergiel, V S Marshall, and J M Jones. 1998. “Embryonic Stem Cell Lines Derived From Human Blastocysts.” *Science* 282 (5391): 1145–47.

Tiscornia, Gustavo, and Juan Carlos Izpisúa Belmonte. 2010. “MicroRNAs in Embryonic Stem Cell Function and Fate..” *Genes Dev* 24 (24). Cold Spring Harbor

Lab: 2732–41. doi:10.1101/gad.1982910.

Wang, Yangming, Rostislav Medvid, Collin Melton, Rudolf Jaenisch, and Robert Blelloch. 2007. “DGCR8 Is Essential for microRNA Biogenesis and Silencing of Embryonic Stem Cell Self-Renewal..” *Nature Genetics* 39 (3): 380–85. doi:10.1038/ng1969.

Yu, Pengzhi, Guangjin Pan, Junying Yu, and James A Thomson. 2011. “FGF2 Sustains NANOG and Switches the Outcome of BMP4-Induced Human Embryonic Stem Cell Differentiation.” *Cell Stem Cell* 8 (3). Elsevier Inc.: 326–34. doi:10.1016/j.stem.2011.01.001.

Zwaka, Thomas P, and James A Thomson. 2005. “Differentiation of Human Embryonic Stem Cells Occurs Through Symmetric Cell Division..” *Stem Cells* 23 (2). John Wiley & Sons, Ltd.: 146–49. doi:10.1634/stemcells.2004-0248.

CHAPTER 2

CHARACTERIZATION OF THE CELL-CYCLE IN HUMAN EMBRYONIC USING THE FUCCI REPORTER SYSTEM

INTRODUCTION

It has been well established that the cell cycle is remodeled during differentiation of ESCs, but less is known about the molecular events involved in lengthening the G1 phase (Sakaue-Sawano et al. 2008). Recent studies have gained insight into these events using the fluorescent ubiquitination-based cell-cycle indicator (FUCCI) system (Calder et al. 2013; Pauklin and Vallier 2013; Sakaue-Sawano et al. 2008). These studies have demonstrated an interaction between the cell cycle and differentiation, consistent with previous reports defining the capability of ESCs to start differentiate from the G1 phase (Chetty et al. 2013; Singh and Dalton 2009). Since gene expression changes in distinct patterns during early differentiation, it is likely that gene expression and cell signaling fluctuates in ESCs in a cell-cycle-dependent manner. To address this question, I utilized the FUCCI system in hESCs in combination with fluorescence-activated cell sorting (FACS) and analyzed gene expression using several different methodologies (Figure 3.2B).

Creation of the FUCCI system by Sukaue-Sawano et al. was based on the fact that the cell cycle is in part controlled by ubiquitin (Ub)-mediated proteolysis. The substrates to E3 ligase activity, Geminin and Cdt1, are involved in “licensing” of

replication origins, where Geminin-mediated inhibition and proteolysis of Cdt1 prevents re-replication. This causes the protein levels of Geminin and Cdt1 to oscillate inversely throughout the cell cycle (Nishitani, Lygerou, and Nishimoto 2004). Cdt1 protein levels are highest during G1, whereas Geminin is highest during the S, G2, and M phases. Based on these data, Sukaue-Sawano et al. developed dual colored probes, wherein they fused Cdt1 and Geminin to red and green emitting fluorescent proteins, the monomeric version of Kusabira Orange (mKO2) and Azami Green (mAG), respectively.

The FUCCI system is best utilized to monitor cell cycle changes in live cells, in culture and *in vivo*. Although the transition from G2 to M phase can be observed by the characteristic morphological changes occurring during cell division, the G1 to S transition is difficult to monitor in live samples. The G1/S transition has been monitored either after nuclear bromoxyuridine (BrdU) staining or by pharmacological synchronization of the cell cycle. Both of these methods fail to represent the normal proliferation and maintenance of self-renewal *in vitro*, therefore creation of FUCCI reporter system to monitor live cell cycle dynamics has created a high-contrast method to tease out cell-cycle fractions in live cells. In this chapter, I demonstrate the use of the FUCCI system in addressing whether hESCs contain heterogeneous expression and signaling events linked to the cell cycle stages.

RESULTS

To construct FUCCI hESC lines, fluorescent reporters were introduced into expression vectors under the control of the constitutive CAGi promoter, linked to either a neomycin (neoR) or puromycin (puroR) selectable marker through an internal ribosome entry site (Figure 2.1A). After drug selection and expansion of WA09 hESCs, we confirmed the authentic cell-cycle-regulated expression of FUCCI indicators by using several approaches in live cells. This system allows FUCCI indicators to only be present during a specific window of time, leading to the identification of cell-cycle position in living cells.

Immunostaining is problematic because both KO2 and Az1 fluorescent proteins undergo rapid photobleaching under fixation conditions; therefore time-lapse microscopy (live imaging) was utilized to first validate the dynamics of the FUCCI system in adherent cells (Figure 2.2A). Images taken from the Viva View imaging system were collated and used to measure the cell cycle length between 14 and 20 hours. The expected pattern of reporter activity in FACS is summarized Figure 2.2C**,** showing an initial double-negative (DN) population indicative of early G1 cells, followed by accumulation of CDT1-KO2 reporter fluorescence throughout the rest of G1 (KO2). As the cells enter S phase, KO2 fluorescence is extinguished and GEMININ-Az1 fluorescence increases until completion of mitosis, wherein cells in S phase express low levels of AzG (AzL) and G2/M cells express high levels (AzH) (Figure 2.2C). Performing a cell size determination assay on a Coulter Counter further validated the FUCCI system. FACS-isolated FUCCI cells were diluted equally in buffer and measured across 2μm increments. As expected, the sorted FUCCI

fractions exhibited a gradual increase in size corresponding to a FUCCI linked progression through the cell cycle (Figure 2.2D).

No global analysis of cell-cycle-dependent genes had been performed in pluripotent cells. To address this we performed RNA-seq analysis on the four FACS-isolated FUCCI cell cycle fractions (DN, KO2, AzL, and AzH) to potentially reveal unseen and significant mechanisms of cell cycle regulation. We identified ~500 transcripts, which followed a reproducible pattern of cell cycle dynamics in three biological replicate experiments (Figure 2.3.A). Using unsupervised cluster analysis, we identified ten cohorts of transcripts, all of which displayed a similar pattern of periodicity, and then performed Gene Ontology (GO) analysis on these data (Figure 2.3.B). Interestingly, the largest group of cell-cycle-dependent transcripts contained genes that are known to regulate development (Figure 2.3.C). This came as a surprise based on previous evaluation of signaling pathways that resulted in no pattern G1-specific expression (Figure 2.4). For example, in our western blot analysis we found no evidence of phospho-dependent regulation of AKT1, ERK1/2, or GSK3b in the different FUCCI fractions (Figure 2.4). In contrast, phosphorylation of ribosomal S6 protein was upregulated in S and G2/M phases, but not G1, as described previously (Shah, Ghosh, and Hunter 2003).

DISCUSSION

We established the FUCCI system as a potent and reliable tool for studying the cell cycle in hESCs. Its use allows for the ability to tease out the effect of the various growth media used in stem cell research, enabling relationships to be

identified between the cell cycle phases and complex signaling events involved in ESC self-renewal and differentiation.

This work has also revealed a novel link between the cell cycle and ESC heterogeneity. Random cell signaling events have been largely attributed to the heterogeneity in ESCs, but we have established a direct link between cell cycle phase and changes in expression of the signaling networks required for development. Through an unsupervised cluster analysis on FUCCI RNA-seq data, we revealed cell-cycle-specific expression and enrichment of transcripts from developmental regulated transcription factors. By contrast, we found no cell cycle regulation for pluripotency factors. Since these developmentally regulated transcription factors are important for the differentiation of ESCs, this suggests that hESCs exist in a "lineage primed" state, possessing a short window of time wherein they can respond to differentiation cues (Figure 2.5). During hESC differentiation, the increase in expression of differentiation markers while the cell cycle specific expression of these markers is maintained would further induce ESCs towards a specific lineage.

EXPERIMENTAL PROCEDURES

Generation and maintenance of human FUCCI hESCs. FUCCI reporter constructs in the pcDNA3 backbone were a gift from Miyawaki and colleagues (Sakaue-Sawano et al. 2008). CDT1-KO2 and GEMININ-AZ1 fusion genes were amplified with primers containing an Eco RI restriction sight using Pfx DNA polymerase (Invitrogen). The resulting amplicons, along with Cag-IRES-PURO and Cag-IRES-NEO, were digested with EcoRI overnight at 37°C. The cut vectors were

then treated with calf intestinal phosphatase CIP (New England Biolabs) for 1 hours at 37°C, and ligated with the digested overnight at 16°C using the DNA Ligation Mighty Mix Kit (Takara). Following ligation, 1 µl of each ligation mixture was transformed into Max Efficiency DH5α Escherichia coli (Invitrogen), grown up under carbenicillin and the plasmids were isolated using the Miniprep Kit (Qiagen). Plasmids were transfected into WA09 hESCs using Lipofectamine 2000 (Invitrogen) and were both selected and maintained in defined media containing puromycin (0.1 µg/ml) and neomycin (200 µg/ml).

Fluorescent activated cell sorting (FACS). To sort cells in a specific cell cycle phase, the FUCCI hESCs were washed with PBS and then put into single-cell suspension using Accutase (Invitrogen). The resulting suspension was centrifuged at 1000 rpm for 5 min, resuspended in fresh medium, and then passed through cell strainer (0.2 µm) to remove any clumps. The strained suspensions were sorted on a Beckman Coulter MoFlo XDP using a 100 µm tip at an excitation of 488 nm for Az1 and 561 nm for KO2.

Cell size determination using Coulter Counter. Cell size analysis of fractions was performed on the Z Series Coulter Counter using a 100 µM aperture. The Coulter Counter is calibrated using the standard 10 µM Instrument Concentration Control solution (Beckman Coulter). Size ranges of 2 µm were set then the sorted FUCCI fractions were diluted (1:500) using a stock diluent, and cell counts were measured for each size group.

Live-cell imaging. A Viva View FL incubator-microscope system (Olympus) was used to validate FUCCI hESCs progressing through the cell cycle. FUCCI hESCs

were grown on a glass-chambered slide inside of a 35 mm plate pre-coated with Geltrex (Life Technologies). Images were taken every 15 min for 24 hours for DIC, green fluorescence, and orange fluorescence at 40x magnification. Images were merged and movie files collated.

RT-qPCR. RNA was isolated from sorted FUCCI fractions using RNA isolation kit (Omega Bio-Tek) and cDNA synthesis was performed using iScript Reverse Transcirption Supermix (Bio-Rad). qRT-PCR was performed on sorted fractions with the use of Taqman assays (Applied Biosystems) on a iCycler (Bio-Rad), according to manufacturer instructions. qRT-PCR assays are normalized to Glyceraldehyde 3-phosphate dehydrogenase (Gapdh) and analyzed according to the ΔΔCT method.

Immunoblotting. Proteins were extracted from cells by adding mammalian cell lysis RIPA buffer (Sigma) containing phosphatase inhibitor cocktail set II, protease inhibitor, and DTT. The cells are left on ice for 30 mins then centrifuged at maximum speed. The protein concentration from the supernatant was quantified using the Bradford assay (BioRad) and mixed with Laemli Buffer for long term storage. Proteins were size-separated by SDS-Page then blotted onto a nitrocellulose membrane and probed for proteins of interest using monoclonal antibodies.

RNA-sequencing. RNA was extracted using TriZol (Invitrogen) from sorted cells and subjected to sequencing with an Illumina HiSeq instrument (Hudson Alpha). For each sample, ~50 M reads were obtained by 2 x 50 nucleotide paired-end sequencing. Alignment of RNA-seq reads was performed using Bowtie version 0.12.7 and TopHat version 1.3.3 (Langmead et al. 2009; Trapnell et al. 2010).

Transcript expression analysis was performed using Cufflinks version 1.2.1 (Trapnell et al. 2010).

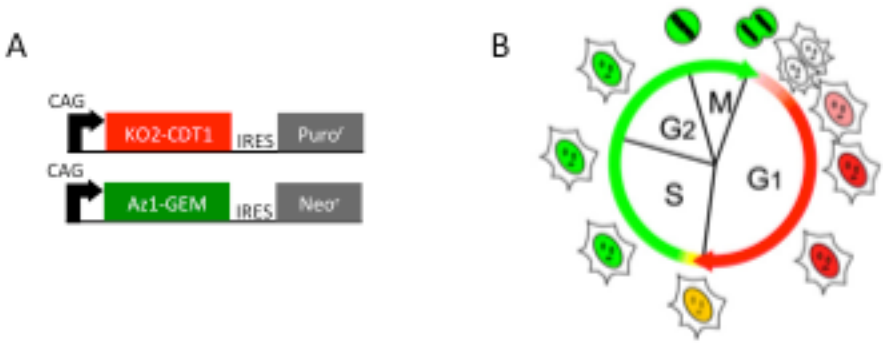

Figure 2.1. FUCCI expression vectors and the cell cycle. (A) FUCCI fluorescent reporters are under the control of the constitutive CAGi promoter, linked to either a neomycin (neoR) or puromycin (puroR) selectable marker through an internal ribosome entry site. (B) The FUCCI system allows isolation of cell cycle fractions based on changes in cell-cell-dependent expression of fluorescent reporters (Sakaue-Sawano et al. 2008).

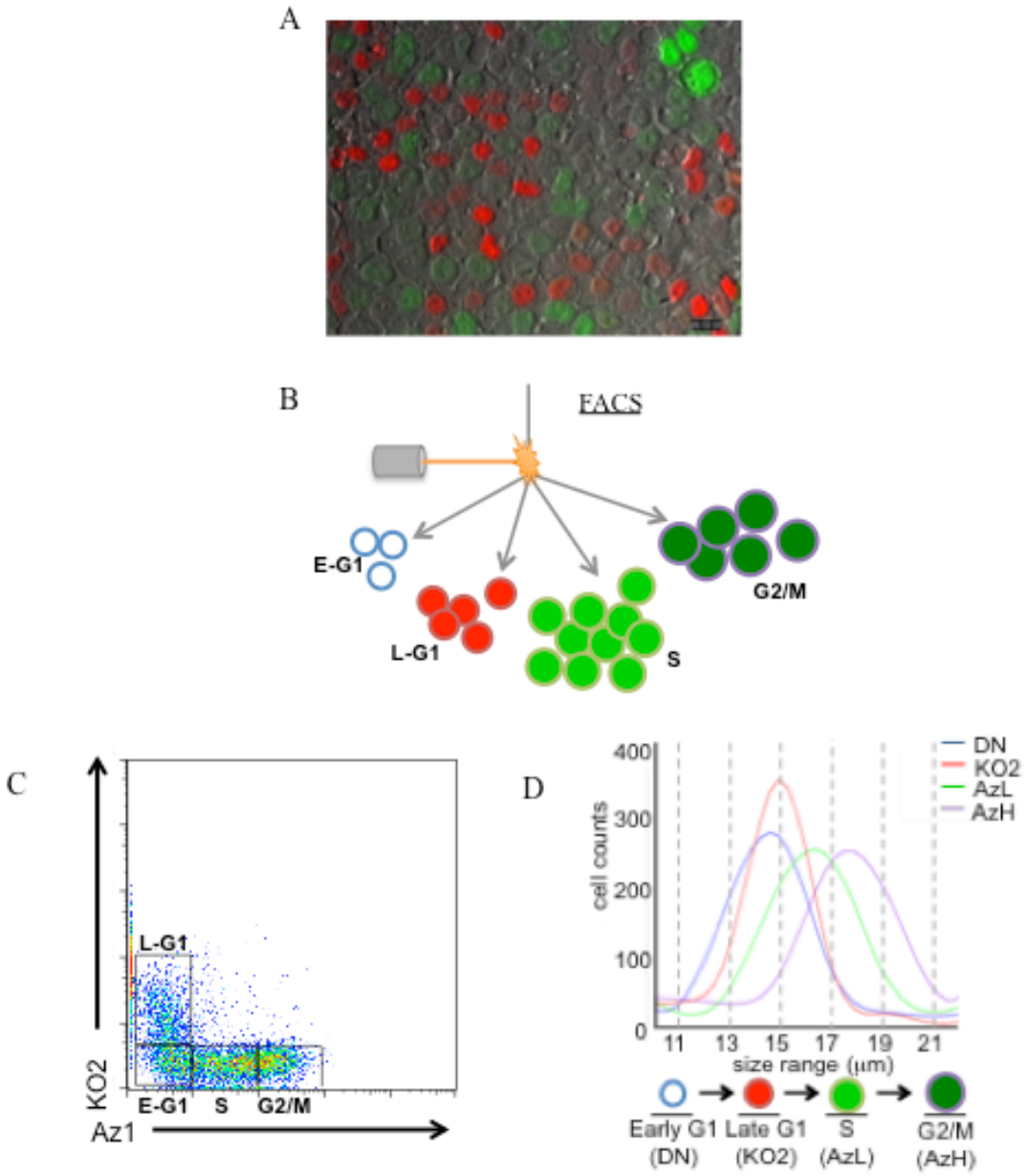

Figure 2.2. Visualizing FUCCI hESCs. (A) FUCCI hESCs on the Viva-view (40x mag.). (B) FUCCI hESCs are sorted with Beckman Coulter MoFlo XDP based on their unique fluorescent profiles (C). (D) Cell size determination assay was performed on FACS-isolated FUCCI hESCs using a Coulter Counter (Singh et al. 2015).

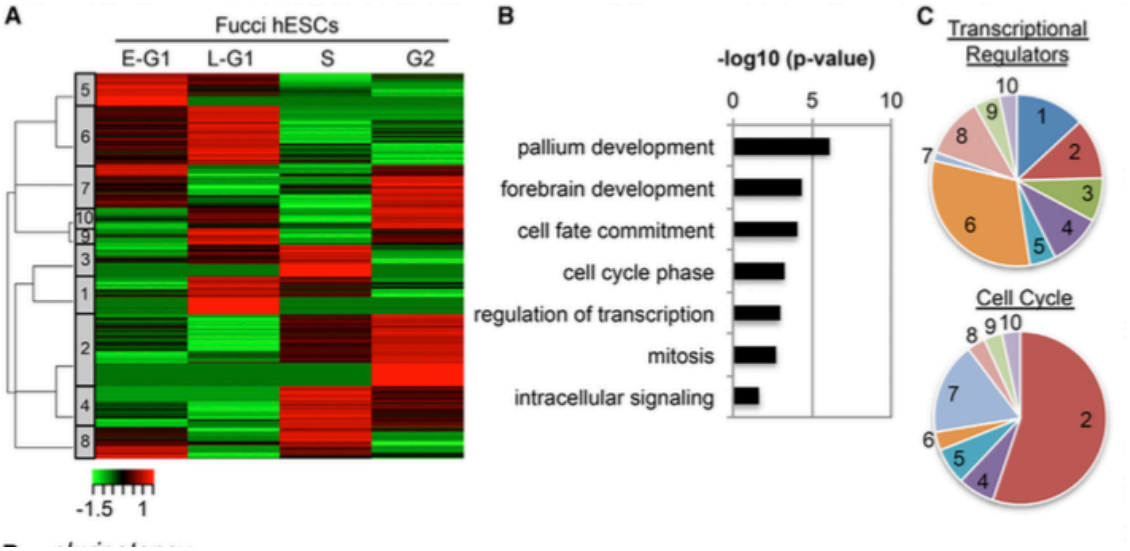

Figure 2.3. Gene-Expression Profiling of Human FUCCI hESCs identified enrichment of developmental markers in G1 phase. (A) Following RNA-seq, cluster analysis was performed on cell-cycle-regulated transcripts in FUCCI hESCs, represented by a heatmap. (B) GO analysis of cell-cycle-regulated genes from human Fucci RNA-seq show an enrichment genes involved in developmental processes. (C) Percentage of genes by GO analysis in the transcriptional regulation or cell-cycle categories according to RNA-seq cluster analysis (Singh et al. 2013).

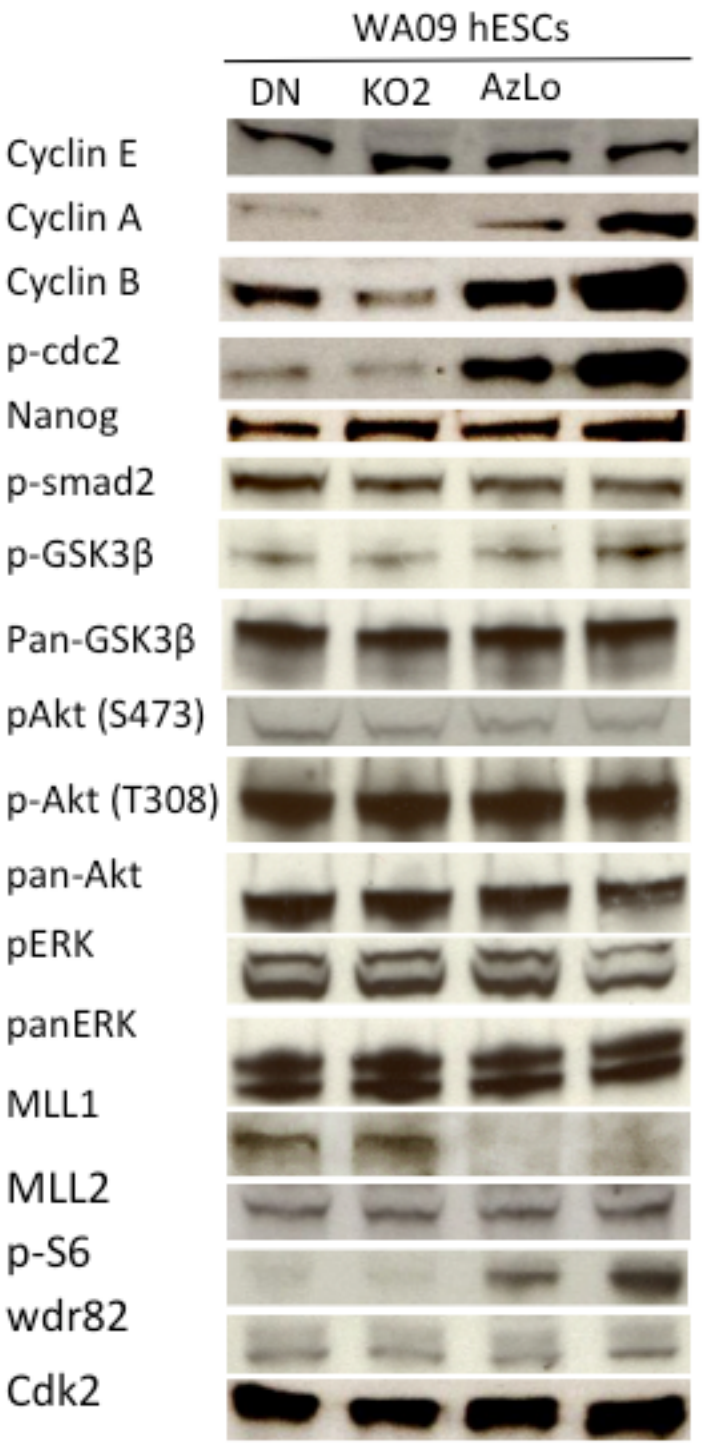

Figure 2.4. Protein Expression of FUCCI hESCs. FACS-isolated FUCCI hESCs were digested using RIPA buffer. The extracted proteins were probed for pluripotency markers and proteins involved in important cell signaling pathways.

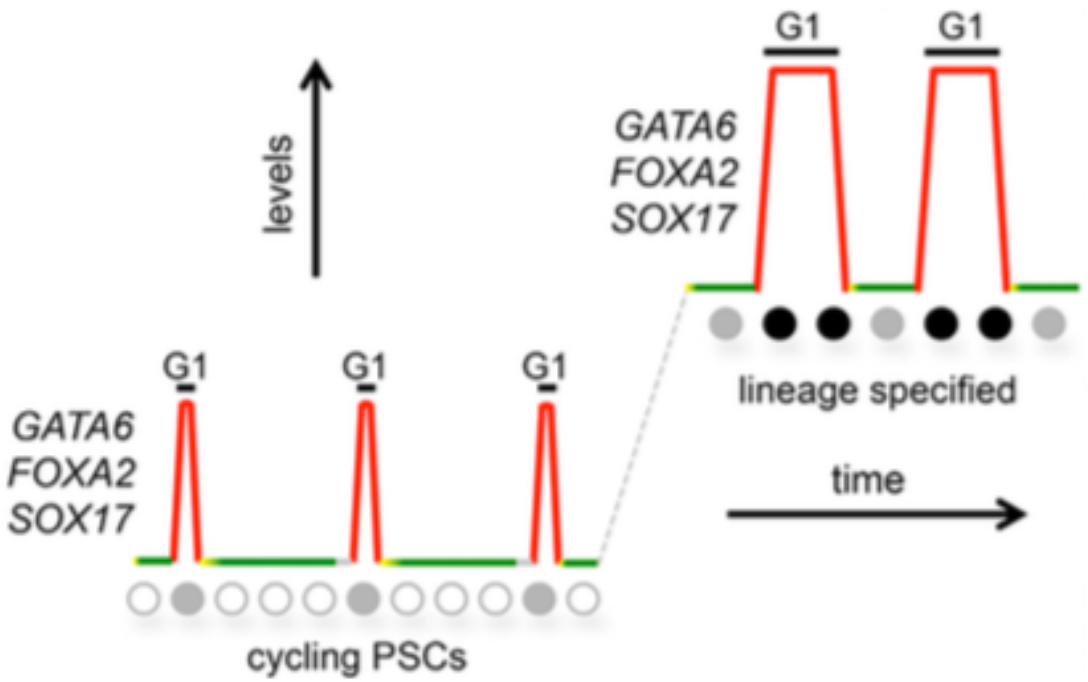

Figure 2.5. The G1 phase expands during differentiation to allow differentiation. Colors are representative of FUCCI fractions (gray, early G1; red, late G1; yellow, G1/S; green, S-G2/M). Circles depict cell-cycle progression, with solid gray circles representing cells that weakly express developmental regulators, and solid black circles indicating cells that strongly express developmental regulators (Singh et al. 2013).

REFERENCES

Calder, Ashley, Ivana Roth-Albin, Sonam Bhatia, Carlos Pilquil, Jong Hee Lee, Mick Bhatia, Marilyne Levadoux-Martin, et al. 2013. “Lengthened G1 Phase Indicates Differentiation Status in Human Embryonic Stem Cells..” *Stem Cells and Development* 22 (2): 279–95. doi:10.1089/scd.2012.0168.

Chetty, Sundari, Felicia Walton Pagliuca, Christian Honore, Anastasie Kweudjeu, Alireza Rezania, and Douglas A Melton. 2013. “A Simple Tool to Improve Pluripotent Stem Cell Differentiation..” *Nature Methods* 10 (6). Nature Publishing Group: 553–56. doi:10.1038/nmeth.2442.

Langmead, Ben, Cole Trapnell, Mihai Pop, and Steven L Salzberg. 2009. “Ultrafast and Memory-Efficient Alignment of Short DNA Sequences to the Human Genome..” *Genome Biology* 10 (3). BioMed Central: R25. doi:10.1186/gb-2009-10-3-r25.

Nishitani, Hideo, Zoi Lygerou, and Takeharu Nishimoto. 2004. “Proteolysis of DNA Replication Licensing Factor Cdt1 in S-Phase Is Performed Independently of Geminin Through Its N-Terminal Region..” *J Biol Chem* 279 (29): 30807–16. doi:10.1074/jbc.M312644200.

Pauklin, Siim, and Ludovic Vallier. 2013. “The Cell-Cycle State of Stem Cells Determines Cell Fate Propensity..” *Cell* 155 (1): 135–47. doi:10.1016/j.cell.2013.08.031.

Sakaue-Sawano, Asako, Hiroshi Kurokawa, Toshifumi Morimura, Aki Hanyu, Hiroshi Hama, Hatsuki Osawa, Saori Kashiwagi, et al. 2008. “Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression..” *Cell* 132 (3): 487–98.

doi:10.1016/j.cell.2007.12.033.

Shah, O Jameel, Sourav Ghosh, and Tony Hunter. 2003. “Mitotic Regulation of Ribosomal S6 Kinase 1 Involves Ser/Thr, Pro Phosphorylation of Consensus and Non-Consensus Sites by Cdc2..” *J Biol Chem* 278 (18). American Society for Biochemistry and Molecular Biology: 16433–42. doi:10.1074/jbc.M300435200.

Singh, Amar M, and Stephen Dalton. 2009. “The Cell Cycle and Myc Intersect with Mechanisms That Regulate Pluripotency and Reprogramming..” *Cell Stem Cell* 5 (2): 141–49. doi:10.1016/j.stem.2009.07.003.

Singh, Amar M, James Chappell, Robert Trost, Li Lin, Tao Wang, Jie Tang, Brittany K Matlock, et al. 2013. “Cell-Cycle Control of Developmentally Regulated Transcription Factors Accounts for Heterogeneity in Human Pluripotent Cells..” *Stem Cell Reports* 1 (6): 532–44. doi:10.1016/j.stemcr.2013.10.009.

Singh, Amar M, Robert Trost, Benjamin Boward, and Stephen Dalton. 2015. “Utilizing FUCCI Reporters to Understand Pluripotent Stem Cell Biology..” *Methods (San Diego, Calif.)*, September. doi:10.1016/j.ymeth.2015.09.020.

Trapnell, Cole, Brian A Williams, Geo Pertea, Ali Mortazavi, Gordon Kwan, Marijke J van Baren, Steven L Salzberg, Barbara J Wold, and Lior Pachter. 2010. “Transcript Assembly and Quantification by RNA-Seq Reveals Unannotated Transcripts and Isoform Switching During Cell Differentiation..” *Nat Biotechnol* 28 (5). Nature Publishing Group: 511–15. doi:10.1038/nbt.1621.

CHAPTER 3

IDENTIFICATION OF MICRORNAS AND THEIR TARGETS IN PLURIPOTENCY AND DURING DIFFERENTIATION

INTRODUCTION

Along with the pivotal roles TGF-β signaling plays in the maintenance of self-renewal and lineage commitment of hESCs, it is also involved in immune responses, tumor suppression, metastasis, and tissue homeostasis (Massague, Seoane, and Wotton 2005). Summarized in Figure 1.2 (Chapter 1), ligands to TGF-β signaling, Activin, Nodal, TGF-β, and BMP, bind to two transmembrane (Type I and Type II) serine/threonine kinases called Activin receptor-like kinases (ALKs). Ligand binding induces heterotetramer ALK formation, with the Type II binding to the Type I receptor, which induces transphosphorylation of the intracellular region of the Type I receptor by Type II. This phosphorylation creates a docking site for the effectors of TGF-β signaling, receptor activated Smads (R-Smads) (Morikawa et al. 2013). Phosphorylation activates the R-Smads to bind the co-Smad (Smad4), primarily as heterotrimers. This activated R-Smad/Smad4-complex can then accumulate in the nucleus to either activate or repress gene transcription through DNA binding (Massague, Seoane, and Wotton 2005).

The TGF-β signaling superfamily of receptors is divided into the TGF-β and BMP branches, which utilize separate R-Smads. BMP signaling occurs through the R-

Smads Smad1/5/8, while Smad2/3 are activated by the TGF-β subclass (Chng, Vallier, and Pedersen 2011). These effectors differ in their DNA binding affinity: Smad2/3 with Smad4 bind to short palindromic Smad-binding elements (SBEs) on DNA, while BMP R-Smads prefer a GC rich sequence (BREs) (Shi et al. 1998). These sequences are found throughout the genome but because of the low binding activity of R-smads, the transcriptional activation of these genes is dependent on the presence of other transcription factors (Nakahiro et al. 2010).

Of the many target genes activated by TGF-β signaling, they also induce expression of the inhibitory Smads (I-Smads), Smad6 and Smad7 (Massague, Seoane, and Wotton 2005). Smad7 is a general inhibitor of TGF-β signaling while Smad6 is thought to mainly regulate BMP signaling. The I-Smads have been shown to inhibit R-Smad activation by blocking their binding to Type I receptors as well as blocking R-Smad/Smad4 complex formation and therefore are categorized as negative feedback inhibitors to the TGF-β signaling superfamily (Lin et al. 2003) (Ishida 2000). Although several studies have shown interaction between I-Smad and TGF-β signaling, the roles of I-Smads (particularly Smad6) are still poorly defined. Similarly, the importance of the TGF-β signaling superfamily in lineage specification is well understood, but less is known about how it regulates cell-commitment at the transcriptional level.

Aberrant Smad6 expression has been shown to play roles in TGF-β associated diseases, supporting its categorization within the TGF-β signaling network (Park 2005). Smad6 is activated as a result of promoter binding by Smad1/5/8, which, along with dual occupation of these R-Smads at its promoter and an upstream

enhancer, is associated with Smad6 expression in the heart, vasculature, and hematopoietic organs, where Smad6 has been shown to play important roles (Morikawa et al. 2011). During hematopoiesis, gain- and loss-of-function studies showed the importance of Smad6 expression for maintaining the pool of HSCs in adults (Kang et al. 2012). As expected, based on its expression during the epithelial-to-mesenchymal transition of the developing heart and its high expression in the adult cardiovascular system, SMAD6-null mice exhibit severe defects in cardiac development (Galvin et al. 2000). Smad6 was also upregulated upon laminar stress of vascular endothelial cells *in vivo*, demonstrating its broad role in maintenance of the cardiovascular system (Topper et al. 1997). In addition to it's expression in mesodermal lineages, Smad6 is also expressed in the lungs, heart, kidneys, immune system, liver, and placenta in humans, suggesting a diverse role of Smad6 in maintenance of several organ systems (Imamura et al. 1997).

Smad6 inhibits TGF-β signaling by binding activated type I receptors through its conserved MH2 domain (N-terminus) (Hata et al. 1998). This same domain has also been shown to regulate the binding of Smad6 partners. Although only a small amount of Smad6 binding partners have been identified thus far, most are categorized as co-repressors therefore Smad6 is also considered to possess co-repressor activity to gene activation (Lin et al. 2003; Bai 2000). Interestingly, Smad6's MH2 domain blocked its DNA binding activity, which showed that co-factor binding is required to confer its gene activation capabilities (Lin et al. 2003). This constraint to direct DNA binding, the necessity of cofactors for gene activation, is shared by the R-Smads (Imamura et al. 1997; Hariharan and Pillai 2008). In another

study, Smad6 participated in a complex with activated glucocorticoid receptors and Hoxc-8 transcription factors on DNA, where they recruited class I histone deacetylases (HDACs) to the promoter elements to repress activation of their target genes (Bai and Cao 2002). Even with these few examples, there is a poor understanding of the mechanism by which Smad6 mediates gene repression. Surprisingly, a recent study showed that post-translational modifications on Smad6 regulated its activity (Zhang et al. 2013). This study, along with the examples showing non-canonical mechanisms for Smad6 signaling in gene repression, suggest a need to be further investigate the roles of Smad6 in mediating cell-fate decisions.

The broad activity of Smad6 suggest cell-specific mechanisms, and raises the question as to what all roles Smad6 plays and which extrinsic signaling conditions regulate Smad6 as a negative feedback inhibitor to TGF-β signaling. Findings outlined in this chapter suggest a novel role for Smad6 in early lineage specification, identified as a result of preliminary screen of miRNA expression during hESC cell-fate commitment.

Several miRNAs that regulate early embryogenesis have been identified, some of which target the TGF-β signaling pathway like the evolutionarily conserved miR-302 family, which is specifically expressed during gastrulation in mammals (Suh et al. 2004). These miRNAs were shown to have important functions in both promoting mesendoderm formation and suppressing ectoderm specification, by targeting components of TGF-β signaling (Rosa, Spagnoli, and Brivanlou 2009). Although researchers are now aware of the critical roles miRNAs play in regulating cell fate decisions during development, further identification and validation of their

expression and mRNA targets will greatly aid in understanding the fine-tuned process of early cell-fate determination (Farazi et al. 2013). The largest human miRNA cluster yet to be identified is a primate-specific microRNA cluster located along a 100kb region of chromosome 19 (C19MC), which is the highest expressed microRNA cluster in both hESCs and placenta (Rippe et al. 2010). Although expression of miRNAs located within C19MC are enriched in hESCs, their roles in maintaining pluripotency are unknown, which we propose partially function to block the induction of differentiation signals by targeting members of the TGF-β family (Liao et al. 2013) The findings outlined in this chapter suggest a relationship between a specific miRNA located in C19MC and an inhibitor to TGF-β signaling, Smad6, through which we suggest a novel mechanism of Smad6 activation in early lineage specification.

RESULTS

This study began by preliminary screening for miRNAs that are differentially expressed during mesoderm induction using a microarray-hybridization approach (Figure 3.1). Most interesting, this screen revealed that expression of miRNAs located in C19MC was repressed during an 8-day mesoderm induction. Approximately 850 putative targets for these miRNAs were then identified *in silico* using Target Scan (www.targetscan.org). Of these potential targets, the most attractive were transcripts inhibitory to TGF-β signaling. Among the possible interactions, regulation of the TGF-β inhibitor Smad6 by *miR520g* was the most intriguing for because of the lack of knowledge regarding Smad6's function.

Smad6's expression was then tested in hESCs which revealed that, while present at low levels in hESCs, it increases during mesoderm and definitive endoderm induction (Figures 3.2A,B,D,E) but not neuroectoderm (Figure 3.2C,F), which supported it's targeting by *miR520g*. To test if *miR520g* inhibited translation of Smad6, the 3'UTR of Smad6 was cloned into a reporter encoding luciferase. Luciferase activity was monitored after co-transfection of *pre-miR520g* with the Smad6-3'UTR reporter in HEK293T cells, which resulted in decreased luciferase activity in the presence of *pre-miR520g*, suggesting *miR520g* blocks Smad6 translation in WAO9 hESCs through binding of to its 3'UTR (Figure 3.3).

To investigate if Smad6 expression regulates mesoderm induction, hESCs were transduced with lentiviral-delivered Smad6 shRNA in order to knockdown Smad6 (Thermo-Scientific, RHS4533-EG3398). The cells were maintained under puromycin selection to obtain stable integration. An approximate 70% knockdown of Smad6 transcript was achieved compared to the control cells transduced with lentiviral-delivered GFP-shRNA. Knockdown of Smad6 was repeated during mesoderm and endoderm induction, by first growing shRNA transduced hESCs in self-renewal conditions for 3 days, then culturing them under differentiation conditions for an additional 3 days. Induction of differentiation markers (Nkx2.5, Gata4, and Gata6) was repressed in mesoderm progenitors under loss of Smad6 (Figure 3.4).

To further define the function of Smad6 in hESCs, I tested its role in TGF-β signaling by performing luciferase assays under Smad6 knockdown (Figure 3.5). The luciferase assays used constructs encoding Activin-response-elements (AREs),

MixL1 promoter and ARE, and resulted in a large decrease in luciferase activity upon Smad6 knockdown in endoderm, suggesting a larger role for Smad6 in the Activin subfamily of TGF-β signaling. To further assess the relationship between Smad6 and Activin signaling, hESCs were stained for Smad6 and Smad3, and performed a Duolink proximity ligation to test if Smad6 may work in complex with Smad3 (Figure 3.5B). Significant signal was observed in hESCs stained for both Smads (Figure 3.5B), which was ablated with addition of the Activin inhibitor, SB-431542, in definitive endoderm (Figure 3.6C).

DISCUSSION

These findings identified primate-specific miRNAs that decreased during the differentiation of hESCs. Several TGF-β signaling family members were putative targets of these miRNAs and these experiments confirmed repression of Smad6 by one of these repressed C19MC miRNAs, *miR520g*. The differentiation of hESCs resulted in the reduction of *miR520g*, allowing for a concomitant increase in Smad6 expression. This relationship is summarized if Figure 3.6A. Preliminary studies also showed genome-wide binding of Smad6 (Figure 3.5D), supporting the proposed role of Smad6 as a transcription factor in hESCs, consistent with previous reports showing Smad6 as a transcriptional co-repressor (Pardali et al. 2005; Ichijo et al. 2005) R-Smad proteins have poor DNA binding activity, and therefore require binding partners to direct them towards specific DNA sequences (Lin et al. 2003; Nakahiro et al. 2010). The results from these studies suggest that Smad3 may work

in tandem with Smad6 in hESCs to elicit its DNA binding activity, which is backed up by the loss of signal seen in SB-431542 treated cells in Figure 3.5C.

The results of this study also suggest novel roles for Smad6 in the regulation of Activin signaling. Smad6 has traditionally been defined as a negative feedback inhibitor to TGF-β signaling, with cytoplasmic expression profiles, but the nuclear localization of Smad6 seen in hESCs and mesoderm imply non-canonical functions (Massague 2005). I propose that Smad6 works during mesendoderm induction to further propagate the activation of differentiation signals, supported by the phenotypic repression of both mesoderm and endoderm markers after Smad6 knockdown (Figure 3.4). If Smad6 were acting as a negative feedback inhibitor of TGF-β signaling during early hESC commitment, the inverse to the results from the luciferase assays in Figure 3.5A would be expected. ARE luciferase activity would increase upon loss of Smad6 based on a predicted activation of Smad2/3 but a significant upregulation under Smad6 knockdown (Figure 3.5A). Based in these results outlined in the chapter, I suggest that Smad6 acts as a co-transcriptional activator in hESCs, targeting early mesendoderm genes (Figure 3.6B). Further work is necessary to address whether Smad6 shares any binding partners in hESCs as seen in other cell types (Lin et al. 2003; Bai 2000).

EXPERIMENTAL PROCEDURES

Isolation of microRNAs from hESCs for microRNA array. hESCs were differentiated towards mesoderm by addition of BMP4 (100ng/ml) and Wnt3a (25 ng/ml) for up to 8 days. The cells were harvested and subjected to Trizol reagent

(Thermo Fisher) to isolate RNA. RNA was quantitated by Nanodrop and sent for sequencing using a microarray-hybridization approach. RNA concentration was validated by RT-qPCR with the use of Taqman® MicroRNA assays (Applied Biosystems) with the Taqman® MicroRNA Reverse Transcription Kit.

Lentiviral production and infection. Lentivirus containing shRNAs vectors (Sigma, pLKO.1) against Smad6 or GFP was prepared in HEK-293T cells and used for infection into WA09 hESCs (MOI: 5). After 24 hours, cells were passaged and re-plated in the presence of puromycin for an additional 3 days and harvested.

Luciferase plasmid construction and assay. Luciferase assays were performed using the Dual Luciferase Reporter Kit (Promega) according to manufacturer instructions and were assayed on a Synergy 2 microplate reader (Bio-Tek). Assays were performed in triplicate and normalized to a Renilla luciferase control. *miR520g* regulation of Smad6 3'UTR was tested by co-transfection of *pre-miR520g* and Smad6 3'UTR in HEK293T cells.

Proximity ligation assay (Duolink®). hESCs were grown on 8-well LabTec slides for 3 days then fixed using 4% paraformaldehyde, blocked in 10% Donkey Serum/PBS, and incubated overnight at 4°C with both primary antibodies in (Smad6/Smad3). The slides were then washed with blocking buffer and secondary antibodies conjugated with PLA probes were added for a 60 min incubation at 37°C, followed by a ligation step at 37°C for 30 min. This was followed by addition of polymerase solution and was incubated at 37°C for 100 min. The slides were then washed and mounted with DAPI and ProLong Gold Anti-Fade Reagent and viewed under the fluorescent microscope.

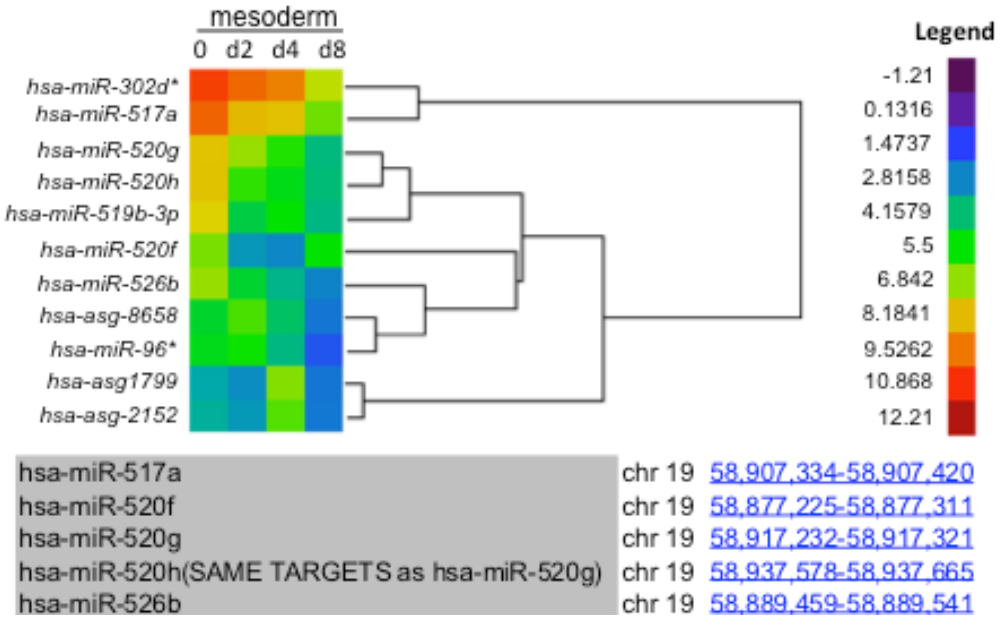

Figure 3.1. Heatmap created from microRNA array performed on an 8-day hESC mesoderm differentiation. The grey box lists miRNAs (and coordinates) that are both significantly reduced during mesoderm differentiation and located within the chromosome 19 microRNA cluster.

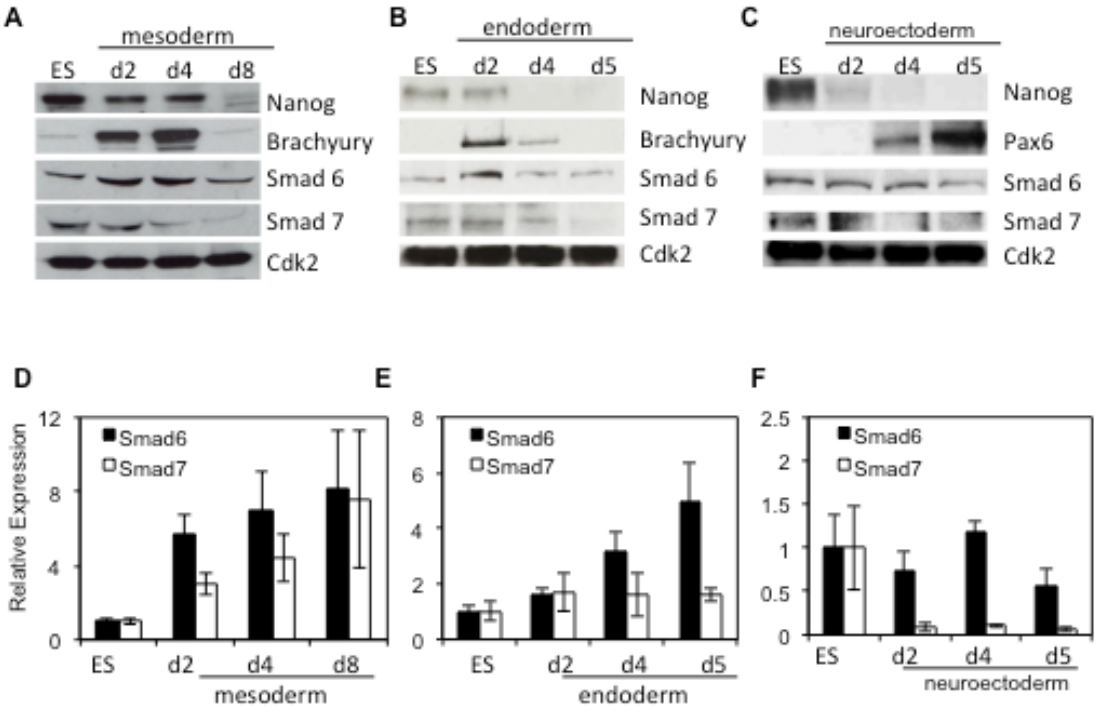

Figure 3.2. Expression of inhibitory Smads during hESC differentiation show upregulation of Smad6. Western blotting was performed on (A) mesoderm, (B) endoderm, and (C) neuroectoderm. RT-qPCR using taqmans for Smad6 and Smad7 were performed on (D) mesoderm, (E) endoderm, and (F) neuroectoderm.

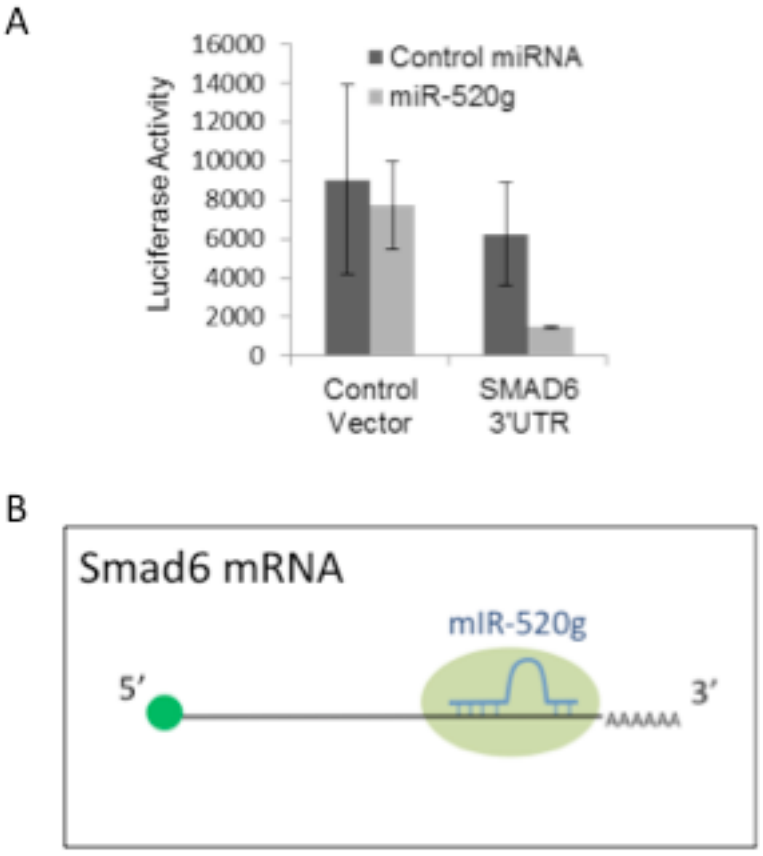

Figure 3.3 miR-520g targets and inhibits Smad6 through its 3'UTR. (A) Smad6 3'UTR luciferase constructs were co-transfected with pre-miR-520g and the resulting luciferase activity was measured. (B) Model showing miR-520g repression of Smad6.

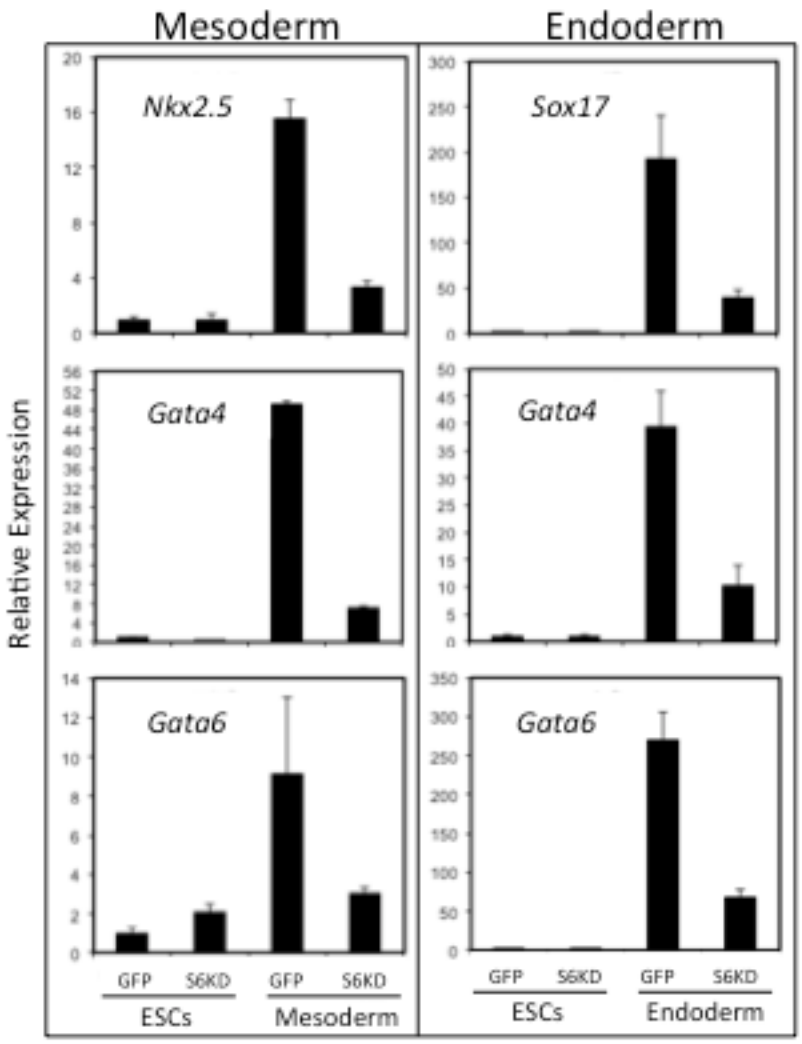

Figure 3.4. Smad6 knockdown showed Smad6 is required for differentiation. Smad6 was knocked down in hESCs with shRNA-delivered Lentivirus. This cells were differentiation and RT-qPCR was performed on developmental genes during differentiation of mesoderm (left) and endoderm (right).

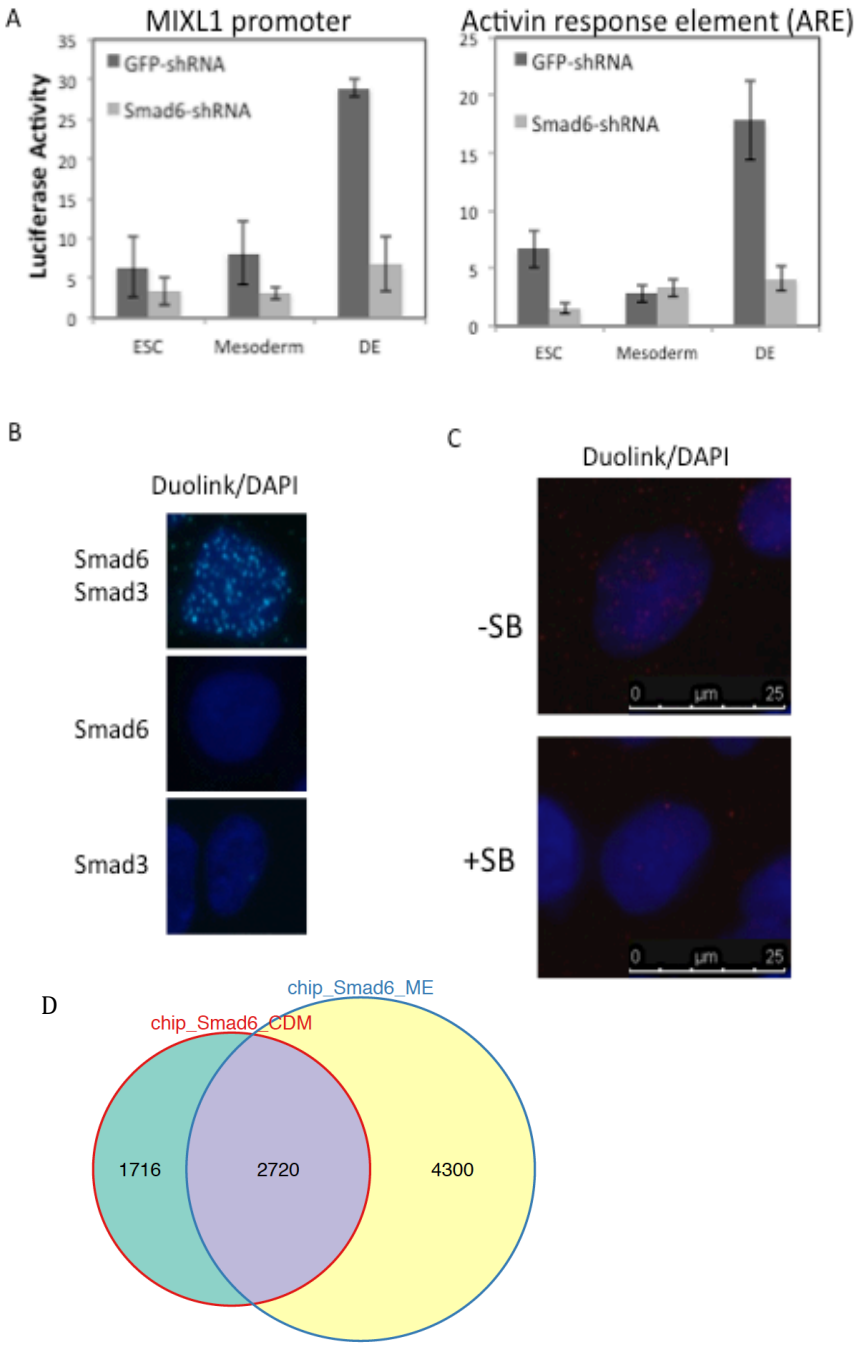

Figure 3.5. Smad6 is required for proper Activin signaling. (A) Luciferase assays were performed during Smad6 knockdown. (B) Duolink assay showed Smad6/Smad3 is in proximity in hESCs. (C) Duolink assay in endoderm cells +/- addition of SB-431542. (D) Smad6 Chip-seq results showed genome binding both in hESCs (CDM) and day-2 mesoderm (ME).

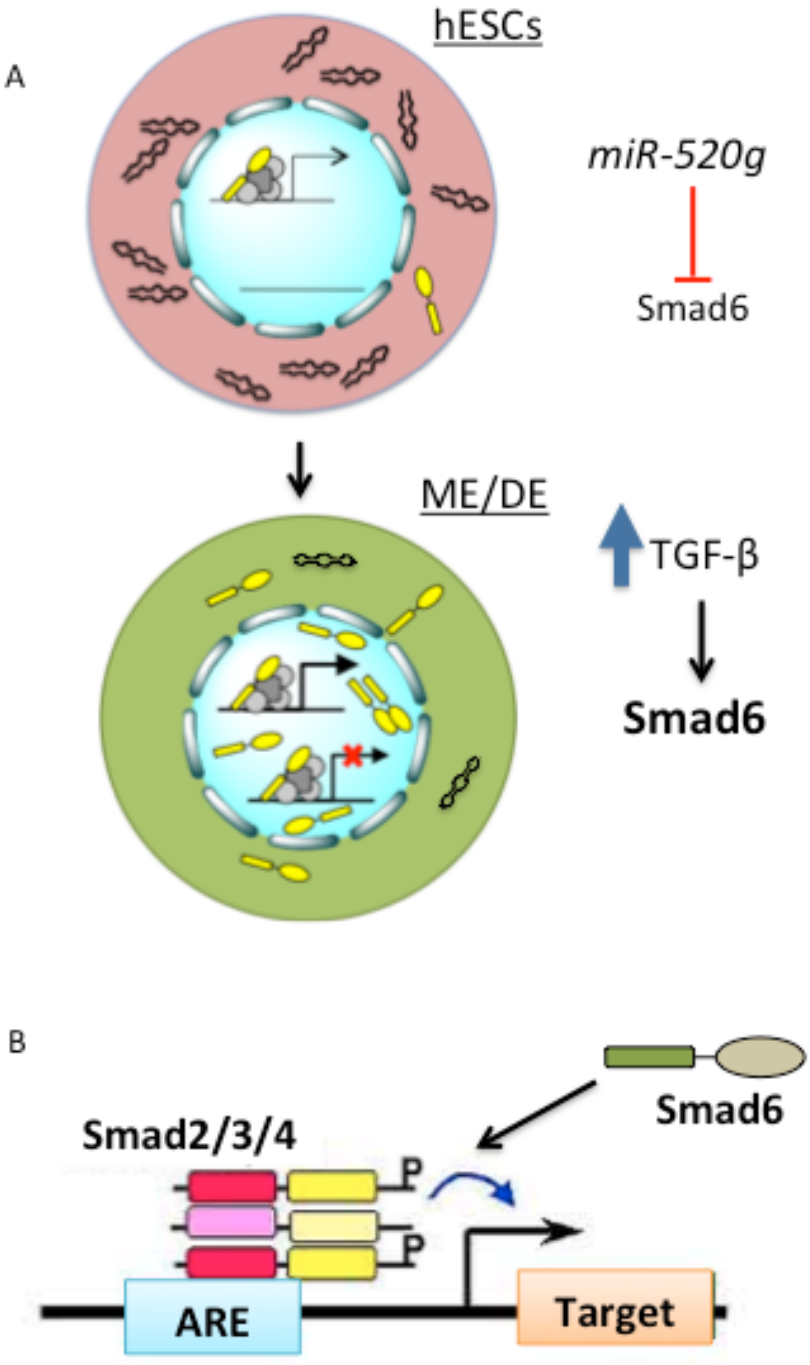

Figure 3.6. Summary of miRNA/Smad6 relationship in hESCs. (A) *miR520g* is high in hESCs, then upon differentiation are suppressed, allowing increase in Smad6 activity. (B) Model for Smad6 acting as a co-transcriptional activator to Activin-responsive gene activation .

REFERENCES

Bai, S. 2000. “Smad6 as a Transcriptional Corepressor.” *Journal of Biological Chemistry* 275 (12): 8267–70. doi:10.1074/jbc.275.12.8267.

Bai, Shuting, and Xu Cao. 2002. “A Nuclear Antagonistic Mechanism of Inhibitory Smads in Transforming Growth Factor-Beta Signaling..” *J Biol Chem* 277 (6): 4176–82. doi:10.1074/jbc.M105105200.

Chng, Z, L Vallier, and R Pedersen. 2011. “Activin/Nodal Signaling and Pluripotency.” *Vitam Horm* 85: 39–58. doi:10.1016/B978-0-12-385961-7.00003-2.

Farazi, T A, J I Hoell, P Morozov, and T Tuschl. 2013. “MicroRNAs in Human Cancer.” *Adv Exp Med Biol* 774: 1–20. doi:10.1007/978-94-007-5590-1_1.

Galvin, K M, M J Donovan, C A Lynch, R I Meyer, R J Paul, J N Lorenz, V Fairchild-Huntress, et al. 2000. “A Role for Smad6 in Development and Homeostasis of the Cardiovascular System..” *Nature Genetics* 24 (2): 171–74. doi:10.1038/72835.

Hariharan, Ramkumar, and M Radhakrishna Pillai. 2008. “Structure-Function Relationship of Inhibitory Smads: Structural Flexibility Contributes to Functional Divergence.” *Proteins* 71 (4): 1853–62. doi:10.1002/prot.21869.

Hata, A, G Lagna, J Massague, and A Hemmati-Brivanlou. 1998. “Smad6 Inhibits BMP/Smad1 Signaling by Specifically Competing with the Smad4 Tumor Suppressor.” *Genes Dev* 12 (2): 186–97.

Ichijo, T, A Voutetakis, A P Cotrim, N Bhattachryya, M Fujii, G P Chrousos, and T Kino. 2005. “The Smad6-Histone Deacetylase 3 Complex Silences the Transcriptional Activity of the Glucocorticoid Receptor: Potential Clinical Implications.” *J Biol Chem* 280 (51): 42067–77. doi:10.1074/jbc.M509338200.

Imamura, T, M Takase, A Nishihara, E Oeda, J Hanai, M Kawabata, and K Miyazono. 1997. “Smad6 Inhibits Signalling by the TGF-Beta Superfamily.” *Nature* 389 (6651): 622–26. doi:10.1038/39355.

Ishida, W. 2000. “Smad6 Is a Smad1/5-Induced Smad Inhibitor. CHARACTERIZATION of BONE MORPHOGENETIC PROTEIN-RESPONSIVE ELEMENT in the MOUSE Smad6 PROMOTER.” *Journal of Biological Chemistry* 275 (9): 6075–79. doi:10.1074/jbc.275.9.6075.

Kang, Y J, J W Shin, J H Yoon, I H Oh, S P Lee, S Y Kim, S H Park, and M Mamura. 2012. “Inhibition of Erythropoiesis by Smad6 in Human Cord Blood Hematopoietic Stem Cells.” *Biochem Biophys Res Commun* 423 (4): 750–56. doi:10.1016/j.bbrc.2012.06.031.

Liao, Xiaoyan, Haipeng Xue, Yu-Chieh Wang, Kristopher L Nazor, Shuren Guo, Neha Trivedi, Suzanne E Peterson, Ying Liu, Jeanne F Loring, and Louise C Laurent. 2013. “Matched miRNA and mRNA Signatures From an hESC-Based in Vitro Model of Pancreatic Differentiation Reveal Novel Regulatory Interactions..” *J Cell Sci* 126 (Pt 17). The Company of Biologists Ltd: 3848–61. doi:10.1242/jcs.123570.

Lin, X, Y Y Liang, B Sun, M Liang, Y Shi, F C Brunicardi, Y Shi, and X H Feng. 2003. “Smad6 Recruits Transcription Corepressor CtBP to Repress Bone Morphogenetic Protein-Induced Transcription.” *Mol Cell Biol* 23 (24): 9081–93.

Massague, J. 2005. “Smad Transcription Factors.” *Genes Dev* 19 (23): 2783–2810. doi:10.1101/gad.1350705.

Massague, J, J Seoane, and D Wotton. 2005. “Smad Transcription Factors.” *Genes Dev*

19 (23): 2783–2810. doi:10.1101/gad.1350705.

Morikawa, M, D Koinuma, K Miyazono, and C H Heldin. 2013. “Genome-Wide Mechanisms of Smad Binding.” *Oncogene* 32 (13): 1609–15. doi:10.1038/Onc.2012.191.

Morikawa, M, D Koinuma, S Tsutsumi, E Vasilaki, Y Kanki, C H Heldin, H Aburatani, and K Miyazono. 2011. “ChIP-Seq Reveals Cell Type-Specific Binding Patterns of BMP-Specific Smads and a Novel Binding Motif.” *Nucleic Acids Res* 39 (20): 8712–27. doi:10.1093/nar/gkr572.

Nakahiro, T, H Kurooka, K Mori, K Sano, and Y Yokota. 2010. “Identification of BMP-Responsive Elements in the Mouse Id2 Gene.” *Biochem Biophys Res Commun* 399 (3): 416–21. doi:10.1016/j.bbrc.2010.07.090.

Pardali, K, M Kowanetz, C H Heldin, and A Moustakas. 2005. “Smad Pathway-Specific Transcriptional Regulation of the Cell Cycle Inhibitor P21(WAF1/Cip1).” *J Cell Physiol* 204 (1): 260–72. doi:10.1002/jcp.20304.

Park, S H. 2005. “Fine Tuning and Cross-Talking of TGF-Beta Signal by Inhibitory Smads.” *J Biochem Mol Biol* 38 (1): 9–16.

Rippe, V, L Dittberner, V N Lorenz, N Drieschner, R Nimzyk, W Sendt, K Junker, G Belge, and J Bullerdiek. 2010. “The Two Stem Cell microRNA Gene Clusters C19MC and miR-371-3 Are Activated by Specific Chromosomal Rearrangements in a Subgroup of Thyroid Adenomas.” *PLoS One* 5 (3): e9485. doi:10.1371/journal.pone.0009485.

Rosa, A, F M Spagnoli, and A H Brivanlou. 2009. “The miR-430/427/302 Family Controls Mesendodermal Fate Specification via Species-Specific Target

Selection." *Dev Cell* 16 (4): 517–27. doi:10.1016/j.devcel.2009.02.007.

Shi, Y, Y F Wang, L Jayaraman, H Yang, J Massague, and N P Pavletich. 1998. "Crystal Structure of a Smad MH1 Domain Bound to DNA: Insights on DNA Binding in TGF-Beta Signaling." *Cell* 94 (5): 585–94.

Suh, Mi-Ra, Yoontae Lee, Jung Yeon Kim, Soo-Kyoung Kim, Sung-Hwan Moon, Ji Yeon Lee, Kwang-Yul Cha, et al. 2004. "Human Embryonic Stem Cells Express a Unique Set of microRNAs.." *Dev Biol* 270 (2): 488–98. doi:10.1016/j.ydbio.2004.02.019.

Topper, J N, J Cai, Y Qiu, K R Anderson, Y Y Xu, J D Deeds, R Feeley, et al. 1997. "Vascular MADs: Two Novel MAD-Related Genes Selectively Inducible by Flow in Human Vascular Endothelium." *Proc Natl Acad Sci U S A* 94 (17): 9314–19.

Zhang, X, J Zhang, A Bauer, L Zhang, D W Selinger, C X Lu, and P Ten Dijke. 2013. "Fine-Tuning BMP7 Signalling in Adipogenesis by UBE2O/E2-230K-Mediated Monoubiquitination of SMAD6." *Embo J* 32 (7): 996–1007. doi:10.1038/emboj.2013.38.