UNIVERSITY OF CALIFORNIA,
IRVINE

The Regulation of BMP Signaling in Mouse Embryonic Stem Cells and Early Mouse Embryogenesis.

DISSERTATION

Submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Mui Nhuc Luong

Dissertation Committee:
Professor Ken W. Y. Cho., Chair
Professor Kavita Arora
Professor Edwin S. Monuki

2015

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Activin receptor II (ActRII)

Activin receptor IIB (ActRIIB)

Asymmetric Initiator Element/Left Side Specific Enhancer (AIE/LSE)

Asymmetric Enhancer (ASE)

BAMBI (BMP and Activin Membrane-Bound Inhibitor)

Basic HLH (bHLH)

BMP response element (BRE)

BMP Type I receptor (BMPRI)

BMP Type II receptor (BMPRII)

Bone morphogenetic proteins (BMPs)

Caudal-related homeobox 2 (Cdx2)

Chromatin imummoprecipitation (ChIP)

Common Smad (co-Smad)

Cycloheximide (CHX)

Cysteine-Rich Secretory Protein (CRISP)

Days post coitum (Dpc)

DNase I hypersensitive (DHS)

Dorsal (Dl)

Drosophila decapentaplegic (Dpp)

Embryonic Stem (ES)

Evolutionarily conserved regions (ECRs)

Extracelluar signal-regulated kinase (Erk)

Gene Ontology (GO)

Glycine/Serine-rich (GS)

Glycogen Synthase Kinase 3 (GSK3)

Growth and differentiation factors (GDFs)

Helix-loop-helix (HLH)

Inhibitors of differentiation (Id)

Inner cell mass (ICM)

kruppel-like zinc finger transcription factor (Klf4)

Leukemia inhibitory factor (LIF)

MAP kinase kinase kinase (MKKK)

Mitogen-Activated Protein Kinases (MAPKs)

Node specific enhancer (NDE)

Normalized Log-Odds (NLOD)

OAZ, (OE/EBF-associated- zinc finger protein Znf423)

ONPG (*ortho*-Nitrophenyl-β-galactoside)

Phosphorylated-Smad (p-Smad)

Posterior Epiblast Element (PPE)

Primitive endoderm (PE)

Protein Analysis Through Evolutionary Relationships (PANTHER)

Quantitative PCR (qPCR)

Reads Per Kilobase of transcript per Million mapped reads (RPKM)

Receptor Smads (R-Smads)

RNA integrity value (RIN)

Screw (Scw)

Sry-box containing gene 2 (Sox2),

TGFβ-activated kinase 1 (Tak1)

Tetracycline (Tc)

Thick Veins (Tkv)

Transcription start site (TSS)

Transforming growth factor-β (TGF-β signaling)

Trophectoderm (TE)

Untranslated Regions (UTRs)

ACKNOWLEDGMENTS

I would like to thank first and foremost, my mentor and committee member chair, Dr. Ken Cho for his mentorship and support throughout my PhD studies. He has been amazingly patient as I learned to become an independent scientist. This dissertation would not have been possible without him.

I would also like to thank my committee members, Dr. Edwin Monuki and Dr. Kavita Arora, for their guidance as well. I am grateful for all their support throughout the years and their constructive feedback on my research. Additionally, I would also like to thank Dr. Chris Hughes and Dr. Pierre Baldi for being on my advancement committee.

I am also thankful to Dr. Bruce Blumberg, Dr. Christine Suetterlin, and Dr. Mike Mulligan for their support of my research.

I would also like to thank Dr. Arul Subramanian, Dr. Ira Blitz, and Dr. Margaret Fish for their mentorship and sharing their knowledge and great appreciation for science.

I am also grateful to all the Cho lab members: Soledad Reyes de Mochel, Anna Javier, William Chiu, Jin Cho, Rebekah Le, and Kitt Parasio for their support of my research. I thank them for their friendship and help throughout the years.

Lastly, I would like to thank my family. They have been my rock throughout my life and I would not be where I am without them. I would like to thank them for their endless support in everything I do. I dedicate my thesis to my family.

CURRICULUM VITAE

Mui Nhuc Luong

EDUCATION:

Ph.D. in Biological Sciences, University of California, Irvine, Department of Developmental and Cell Biology, Winter 2015

B.S. in Biochemistry, University of California, Los Angeles, Department of Chemistry and Biochemistry

Minor in Classics, University of California, Los Angeles, Department of Classics

PUBLICATION:

Reyes de Mochel NS, **Luong M**, Chiang M, Javier AL, Luu E, Toshihiko F, MacGregor GR, Cinquin O, Cho KW. BMP signaling is required for cell cleavage in preimplantation-mouse embryos. Dev Biol. 2014 Oct 14. pii: S0012-1606(14)00503-X.

Blitz IL, **Luong M**, Chiu WTY, Cho KW. Applications of deep sequencing in developmental systems. Principles of Developmental Genetics. Sept 2014. Page 37-48.

Javier AL, Doan LT, **Luong M**, Reyes de Mochel NS, Sun A, Monuki ES, Cho KW. Bmp indicator mice reveal dynamic regulation of transcriptional response. PLoS One. 2012;7(9):e42566.

Boldin MP, Taganov KD, Rao DS, Yang L, Zhao JL, Kalwani M, Garcia-Flores Y, **Luong M**, Devrekanli A, Xu J, Sun G, Tay J, Linsley PS, Baltimore D. miR-146a is a significant brake on autoimmunity, myeloproliferation, and cancer in mice. J Exp Med. 2011 Jun 6;208(6):1189-201.

Kozak KR, Amneus MW, Pusey SM, Su F, **Luong MN**, Luong SA, Reddy ST, Farias-Eisner R. Identification of biomarkers for ovarian cancer using strong anion-exchange ProteinChips: potential use in diagnosis and prognosis. Proc Natl Acad Sci U S A. 2003 Oct 14;100(21):12343-8.

AWARDS:
Fine Tools Travel Award 2011
International Society for Developmental Biology Travel Award 2011

CONFERENCE AND POSTER PRESENTATION:

Luong, M, Blitz IL.,Cho KW. A High-Throughput Genomic Approach to Understanding Transcriptional Control in Oct4 Depleted Mouse Embryonic Stem Cells. University of California, Irvine Department of Developmental and Cell Biology retreat, Society for Developmental Biology Conference, Chicago, IL, 2011.

PROFESSIONAL AND TEACHING EXPERIENCE:

Teaching assistant for Human Development: Conception to Birth; Winter 2009 & Spring 2010 & Winter 2012. Advisor: Dr. Ronald Meyer

Teaching assistant for Developmental Biology; Winter 2011 & Winter 2013. Advisor: Dr. Ken Cho and Dr. David Gardiner

Teaching assistant for Developmental Biology; Summer 2013. Advisor: Dr. Ken Cho

Teaching assistant for Developmental Biology; Spring 2014. Advisor: Dr. David Gardiner and Dr. Kavita Arora

Research Assistant in Dr. Robin Farias-Eisner's Lab in the Department of Obstetrics and Gynecology at the University of California, Los Angeles (UCLA) Sept 2004- Sept 2005

Research Assistant in Dr. David Baltimore's Laboratory at California Institute of Technology (CALTECH) Jan 2006 – Sept 2007

ABSTRACT OF DISSERTATION

The Regulation of BMP Signaling in Mouse Embryonic Stem Cells and Early Mouse Embryogenesis.

By

Mui Nhuc Luong

Doctor of Philosophy in Biological Sciences
University of California, Irvine, 2014
Professor Ken W. Y. Cho

Understanding the molecular mechanisms that govern animal development is one of the major challenges in developmental biology. The TGF-β members such as the bone morophogenetic proteins (BMP) play critical and diverse roles throughout embryonic development in both vertebrates and invertebrates and understanding the mechanisms by which they initiate specific cellular differentiation programs and control of gene expression is important. This dissertation will focus on the characterization of the BMP responsive element (BRE) identified in our lab. Additionally, I will demonstrate that the BRE functions as an activator in BRE-mediated BMP signaling and is responsible for modulating a subset of BMP target genes in mouse embryonic stem (ES) cells. To complement our *in vitro* studies, our lab also investigated the role of BMPs in early mouse embryonic development. Our studies reveal that BMP signaling may be active in the inner cell mass (ICM) during mouse embryogenesis and is playing a role in sustaining pluriopotency by interacting with the *Oct4/Sox2/Nanog* network.

Chapter 1

Introduction

Understanding the molecular mechanisms that govern animal development is one of the major challenges in developmental biology. After the fertilization of an egg, a single cell has the potential to become many diverse cell types and structures to form a functional living organism. Developmental cues from extracellular signaling and secreted molecules are essential for proper embryogenesis. The transforming growth factor-β (TGF-β signaling) superfamily is a class of growth and differentiation factors that are responsible for driving key developmental programs and controlling cell fate and behaviors. The TGF-β superfamily of growth factor is well conserved through Metazoa (Perterson et al., 2008). It consists of 30 family members including TGF-β, Activins, Inhibins, Nodals, Bone Morphogenetic Proteins (BMPs), and growth and differentiation factors (GDFs) (Derynck et al., 1997; Derynck et al., 1998; Feng et al., 2005). These TGF-β members play critical and diverse roles throughout embryonic development in both vertebrates and invertebrates and understanding the mechanisms by which they initiate specific cellular differentiation programs and controls of gene expression is important. Through a combinatorial approach of using different ligand, receptors, and intracellular signaling components, the TGF- β superfamily members can produce a large diversity of transcriptional outputs to regulate a multitude of developmental processes.

General Molecular Mechanism of TGF-β signaling through Smad Proteins:

The members of TGF-β superfamily proteins signal primarily through receptors known as "Type I" and "Type II" receptors. Both receptors are structurally similar and contain serine/threonine kinases. However, the Type I receptors have a conserved Glycine/Serine-rich (GS) sequence domain, while the Type II receptor does not. In the presence of a ligand, the formation of a ligand/receptor complex is initiated allowing the Type II receptor kinases to phosphorylate the GS sequence of the Type I receptor. This in turn activates the Type 1 receptor kinase and enabling it to autophoshorylate itself as well as phosphorylating downstream effectors such as the Smad proteins (Derynck et al., 2003; Shi et al., 2003).

The Smad proteins are essential intracellular effectors of TGF-β signaling. There are 3 distinct classes of Smad proteins: Receptor Smads (R-smads: Smad 1,2,3,5,8), a common Smad (e.g. Smad4), and inhibitory Smads (Smad 6 & 7). Both R-Smads and the common Smad (co-Smad) are made up of a well-conserved polypeptide chain containing a MH1 (N-terminal Mad Homology1) domain and MH2 (C-terminal Mad Homology 2) domain. A key difference between the R-Smads and co-Smad is that the co-Smad (Smad4) lacks a C-terminal SXS motif and therefore cannot be phosphorylated by the Type I receptor. On the other hand, the R-Smads do not contain a C-terminal SXS motif and are phosphorylated by the Type I receptor upon ligand activation. The C-terminal SXS phosphorylation of the R-Smads induces a conformational change, allowing it to dissociate from the Type I receptor and form a heterotrimeric complex containing two R-Smads and co-Smad (Derynck et al., 2003). This complex can then translocate into the nucleus and acts as a transcriptional factor by recruiting other coactivators/corepressors to activate or repress gene transcription (Figure 1.1).

Figure 1.1

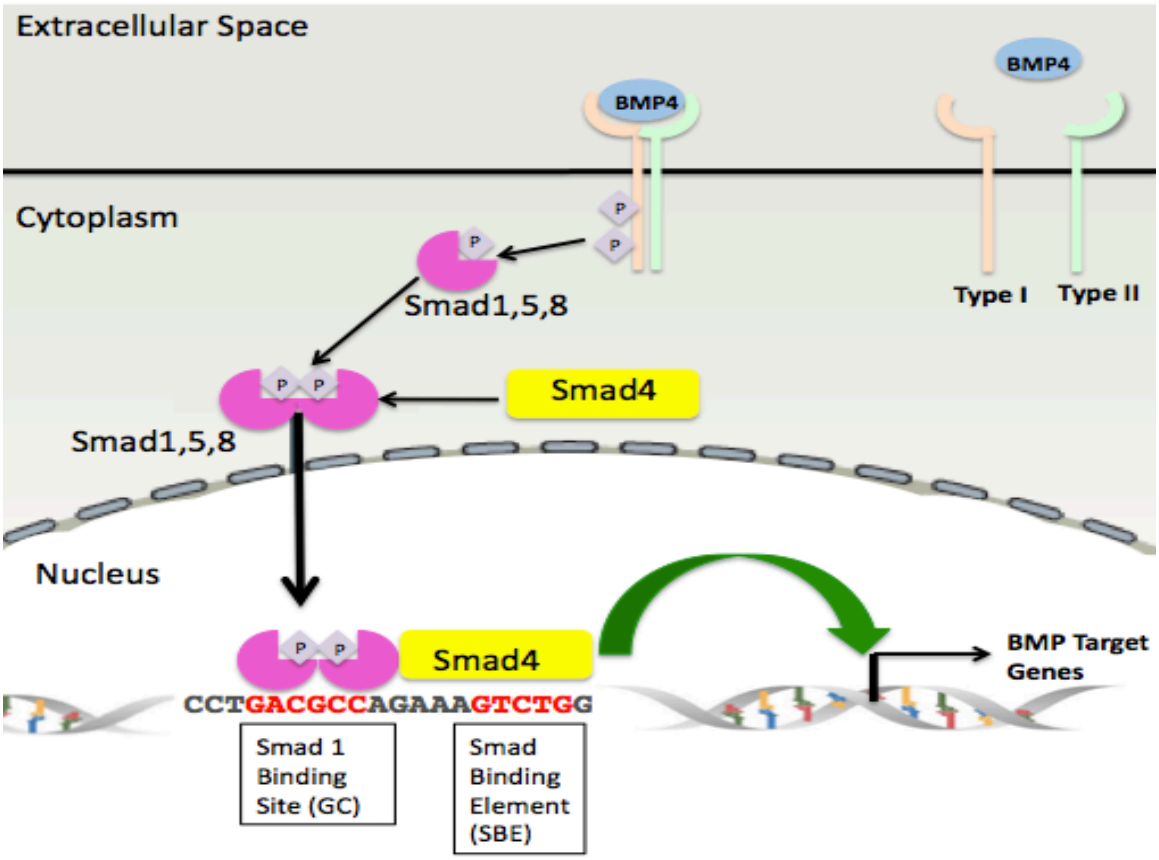


Figure 1.1: Canonical BMP signal transduction cascade. Activation of the canonical BMP signaling pathway begins by BMP receptor activation by the BMP ligand. This enables Smad phosphorylation of the R-Smad (Smad 1/5/8) by the Type-I receptor to form a heteromeric complex with co-Smad (Smad4). The complex then accumulates in the nucleus to activate gene transcription.

Combinatorial Use of Receptors in TGF-β signaling:

Combinatorial use of the Type I and Type II receptors allows for diversity and selectivity in ligand binding as well as intracellular signaling. There are currently seven Type I receptors (ALK1-7) and five Type II receptors identified so far in the human genome. Two different Type1 and Type II homodimer receptors can form a heterotetrameric complex upon ligand activation. However, only specific Type I and Type II receptors can come together to form a ligand-receptor complex upon ligand activation. The molecular basis of selectivity of TypeI-Type II receptor interactions is determined at the interface by specific ligands. TGF-β-1, TGF-β-3, and Activin ligands can bind tightly to TGF-β Type II receptors, Activin receptor II (ACTRII), and Activin receptor IIB (ACTRIIB) without the presence of Type I receptors, even though these ligands can bind to both types of receptors (Boesen et al., 2002; Hart et al., 2002; Greenwald et al., 2004). Conversely, BMP2 and BMP4 ligands have low binding affinity for the Type II receptor (BMPRII) alone but in the presence of the Type I receptor (BMPRIA/ALK3 and BMPRIB/ALK6), ligand binding affinity increases allowing the formation of a heteromeric complex (Kirsch et al., 2000; Keller et al., 2004). Therefore, it is possible that ligands with different affinities for different receptor combinations can be potentially used to modulate a dose-dependent response in the cell.

TGF-β Superfamily Intracellular Signaling:

The Smad signaling pathway downstream of the TGF-β superfamily receptors has been studied extensively. The TGF-β superfamily can be divided into two main branches: 1) BMP/GDF branch signaling through ALK2, 3, 6 via R-Smads 1/5/8 and 2) TGF-β/Nodal/Activin branch signaling through ALK4, 5, 7 via R-Smads 2 and 3 (Massague et al.,

2005) (Table 1.1). However, these interactions are not strictly confined. GDF 8, 9, and 11 can signal through ALK4, 5, 7 (Rebbapragada et al., 2003; Andersson et al., 2006; Gilchrist et al., 2006; Schmierer et al., 2007) and TGF-β can activate both Smad2/3 and Smad1/5/8 in different tissue types (Goumans et al., 2003, Liu et al., 2004; Bharathy et al., 2008; Daly et al., 2008). These findings provide evidence that a particular ligand cannot be unequivocally assigned to only one branch of the TGF-β superfamily signaling, thus lending to the possibility that mixed recruitment of Type I receptors could be an additional mode of regulation in the TGF-β superfamily.

Once phosphorylated, the R-Smad/co-Smad heteromeric complex is translocated to the nucleus and binds to promoter sequences to positively or negatively regulates target gene transcription. Different R-Smad/co-Smad heteromeric complexes have distinct DNA binding specificities to enable the diverse transcriptional output found from different members of the TGF-β superfamily. In the TGF-β activin/nodal signaling pathway, the N-terminal MH1 domain of Smad3 (R-Smad) and Smad4 (co-Smad) recognizes the sequence 5'-GTCT-3'or its reverse complement, 5'-AGAC-3,' known as the Smad Binding Element (SBE) (Dennler et al., 1998). In contrast, the Smad1/5/8 binds to a GC-rich sequence. This GC rich consensus sequence (5'-GRCGNC-3') allows Smad1/5/8 to complex with Smad4, which will independently bind to the SBE. Interestingly, research has shown that Smad1/5/8 binding site and SBE can be spatially restricted by 5 base pairs (Pyrowolakis et al., 2004) through transcriptional regulators such as coactivators and corepressors to stabilize the R-Smad/co-Smad complex since binding of Smads to SBE and the GC-rich site is weak.

Table 1.1

Ligand	Type II Receptor	Type I Receptor	R-Smad
TGF-β	TGF-βRII	ALK4/ACTRIB ALK5/TGF-βRI	Smad 2 Smad 3
BMP & GDF	BMPRII ACTRII ACTRIIB	ALK1 ALK2/ACTRI ALK3/BMPRIA ALK6/BMPRIB	Smad 1 Smad 5 Smad 8
Activin & Nodal	ACTRII ACTRIIB	ALK4/ACTRIB ALK7	Smad 2 Smad 3
AMH	Amhr2	ALK2/ACTRI ALK3/BMPRIA ALK6/BMPRIB	Smad 2 Smad 3

Table 1.1: Summary of the various ligands, receptors, and Smads involved in the TGF-β superfamily signaling cascade.

Although much is known on how Smad-DNA binding can lead to activating or repressing a transcriptional response, its selectivity in choosing which target genes or transcriptional program to initiate is not well understood. Published works have called into question whether the traditional R-Smad/Smad4 complex is even required for regulation since a number of TGF-β induced genes requires only R-Smad or Smad4 for its regulation (Levy et al., 2005; Lim et al., 2006, Descargues et al., 2008). Evidence has shown that Smad2/3 (R-Smads) can interact with IκB kinase in the absence of Smad4 (co-Smad) to regulate Mad1, a Myc antagonist (Descargues et al., 2008). It is known that the commonly studied R-Smad/Smad4 complex regulates most BMP target genes but whether it is always necessary is unclear. A recent finding demonstrated that R-Smads could be regulated by microRNA 21 (mir-21) (Blahna et al., 2012). Taken together, these findings suggest that there are possibly multiple modes of regulation of Smads not yet known, which do not necessarily require the traditional R-Smad/Smad4 complex or a specific DNA-binding sequences.

Canonical BMP Signaling Cascade:

BMPs are the largest subgroup belonging to the TGF-β superfamily and were originally identified as factors that induce the formation of bone and cartilage when ectopically implanted in rats (Urist, 1965; Wozney et al., 1988). BMPs are evolutionarily conserved in all vertebrates and the majority of invertebrates (except for the taxa Choanozoa and Fungi). They can influence a variety of biological processes such as stem cell self-renewal, differentiation, proliferation, tissue morphogenesis, and apoptosis. Members of the BMP subfamily include the BMP-2/4 group (including *Drosophila*

decapentaplegic (Dpp)), the BMP-5/6/7/8 group (OP-1group), GDF-5/6/7 group, and the BMP-9/10 group (Figure 1.2). Active BMPs are composed of 50–100 amino acids with seven cysteines, six of them forming three intramolecular disulfide bonds, known as cysteine knots. The seventh cysteine is used for dimerization with another monomer by forming a covalent disulfide bond, thus forming the biologically active BMP signaling molecule (McDonald et al., 1993; Griffith et al., 1996). BMP family members bind to serine/threonine kinase receptors (i.e. Type I and Type II) and can transduce signals through Smad-dependent and independent mechanisms.

Figure 1.2:

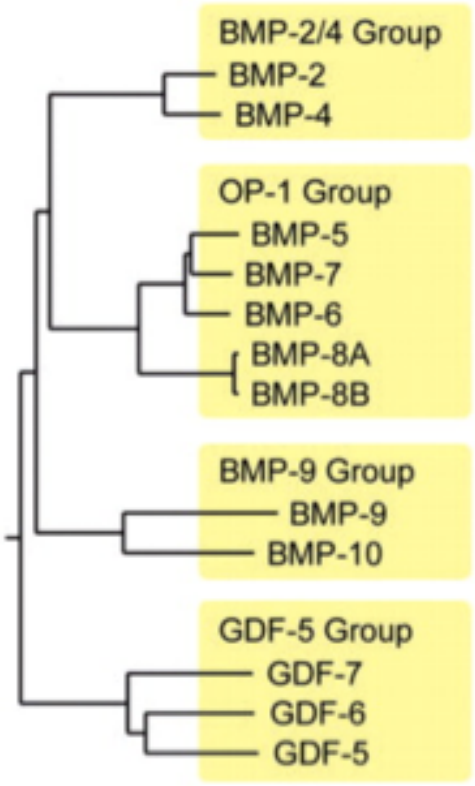


Figure 1.2: Phylogenetic relationship between BMP ligands. The BMP ligands can be grouped into 4 distinct groups: BMP-2/4, OP-1, BMP-9, and GDF-5. Alignment was performed using the C-terminal region containing the cystine knot structure. (Derynck et al., 2008; Miyazono et al., 2010)

Upon BMP receptor activation, the Smad pathway is switched on, which are the major signaling transducers inside the cell. This occurs via phosphorylation of Smad1/5/8 by the Type-I receptor to form a complex with Smad4. Smads 1/5/ 8 is functionally redundant in mouse (Pangas et al., 2008; Huminiecki et al., 2009) and all three receptor Smads have a high degree of amino acid sequence similarity (Hester et al., 2005; Arnold et al., 2006; Pangas et al., 2008; Retting et al., 2009). It has been hypothesized that Smad 1/5/8 may have arose from genome duplication and therefore uncovering their individual function has been difficult. Smad1/5/8 has a preferential binding to GC rich element and Smad4 to SBE as previously stated. It has been speculated that Smad1 preference for GC-rich elements may be due to the highly basic amino acids adjacent to the β-hairpin residues found in all three BMP Receptor Smads, which is not found in Activin receptor smads (Shi et al., 1998; Shi 2001). Furthermore, the promoters of BMP induced genes (e.g. *Bambi*, *Id*, and *Vent* genes) have multiple Smad binding sites to possibly enhance gene transcription. Since the binding affinities of Smads to their binding site are weak (Shi et al., 1998), long concatemers of Smad binding sites may be required to observe transcriptional activation in cells (Zawel et al., 1998; Korchynskyi et al., 2002; Lopez Rovira et al., 2002; Karaulanov et al., 2004). True endogenous BMP signaling may be hard to detect unless Smad1 binding site and SBE are multimerized to amplify BMP signaling.

Transcriptional Response of the BMP Response Element (BRE):

The efficiency of coactivator and corepressor interactions with the Smad complexes is likely to play a major role in determining transcriptional and cellular responses to BMP ligands. The first identified Smad partner, OAZ, (OE/EBF-associated- zinc finger protein

Znf423) was shown to mediate the BMP signaling through the BMP response element (BRE) of the *Xenopus Vent2* gene (Hata et al., 2000). OAZ cooperates with Smad1 by interacting with a specific binding sequence known as the BRE (5'-CAGAC-3' and a 3' flanking sequence, TGGAGC). It is proposed that OAZ functions in a tissue specific manner to regulate BMP signaling (Cheng et al. 2007; Kratitnger et al., 2007) and therefore may not be a universal Smad binding partner. Runx2 is another transcription factor that can regulate BMP target genes. Runx2 and Smad1 complex together to bind to the promoter of genes (*Osteocalcin*, *Osteopontin*, *Col10*, and *Smad6*) involved in the development of osteoblast and chondrocytes (Zhang et al., 2000; Drissi et al., 2003; Wang et al., 2007). It remains unclear whether there is a universal Smad-binding partner to regulate the various BMP-responsive genes.

Schnurri is another well-studied Smad-binding partner of BMP signaling in *Drosophila* and vertebrates. It is a large protein (~275 kDa) composed of eight zinc fingers and interacts with Mad (Smad1/5/8 homolog) and Medea (Smad4 homolog) to regulate DPP (BMP signaling) in *Drosophila*. In *Drosophila,* the Smad1/5/8/-Smad4-Schnurri complex results in *brinker* repression and induces derepression of Dpp (BMP homolog) target genes (Marty et al., 2000; Pyrowolakis et al., 2004; Yao et al., 2008). The regulatory region of the *Brinker* gene contains a silencer element (brkSE) composed of a Mad and Medea consensus-binding site separated by a five base pair spacer sequence (Pyrowolakis et al., 2004; Gao et al., 2005). This five-spacer region is critical for Schnurri docking to mediate *Brinker* repression and therefore relieve repression of Dpp targets (Marty et al., 2000; Yao et al., 2008). Schnurri interacts with the MH2 domains of Smad1/5/8 and Smad4. In contrast to Schnurri's repressive role in *Drosophila*, the vertebrate Schnurri-

Smad1/5/8-Smad4 complex in *Xenopus* can activate *Vent2* gene expression (Yao et al., 2006) through the BRE in *Xenopus* and human.

The Cho lab has shown that the BRE can mediate BMP induction during early *Xenopus* embryogenesis. The BRE was independently found in promoters of three known BMP target genes: *Id3*, *Vent2, Bambi* (Karanulanov et al., 2004; von Bubnoff et al., 2005). A comparison of the *Xenopus*, Human, and Mouse *Id3* promoters shows a high degree of conservation in a 49 bp region (-718/-670) among all three *Id3* promoters. This BRE contains a SBE and GC-rich sequence similar to an OAZ binding site. The BRE of *Xenopus Id3* (*X.Id3*) consensus sequence is "5'-GRCGNC-NNNNN-GTCG-3'. The *Xenopus Id3* BRE shares similarities to a previously published BRE found in the mouse *Id1* promoter (Katagiri et al., 2002; Korchyinskyi et al., 2002; López-Rovira et al., 2002). The *Id1* BRE promoter has a SBE and a GC rich palindromic sequence 5'-GGCGCC-3'. Smad1/5/8 can bind this *Id1* palindromic sequence while Smad4 preferentially binds to the SBE. The palindromic sequence 5'-GACGCC-3' of *Xenopus Id3* BRE is very similar to the *Id1* except with one nucleotide difference (Figure 1.3). All the *X. Id3* BRE, *X. Vent2* BRE, mouse *Id1* BRE are sufficient for BMP response (Katagiri et al., 2002; Korchyinskyi et al., 2002; López-Rovira et al., 2002; von Bubnoff et al., 2005). Although the *X. Id3* BRE is responsive in *Xenopus*, whether this particular BRE is responsiveness to BMPs in a mammalian system (in particular mouse embryonic stem cells) will be addressed in my thesis. It is hypothesized that in mammals, the *X. Id3* BRE acts as an activator of BRE-mediated BMP signaling.

Figure 1.3:

Figure 1.3: Comparison of BRE sequences from different species. Each BRE is aligned for sequence comparison of their Smad1 Binding Site (Red) and the Smad Binding Element (Blue). The SBE is highly conserved across all species. The Smad Binding Site is GC-rich and differs by 2 nucleotides between species.

Non-canonical BMP signaling:

BMPs can also transduce signaling via a non-canonical (Smad-independent) pathway. The best-established example of non-canonical BMP signaling pathway is the activation of the TGFβ-activated kinase 1 (Tak1, a MAP kinase kinase kinase (MKKK) (Yamaguchi et al., 1995; Gunnell et al., 2010). Tak1 can be activated by BMPs or TGF-β and directly phosphorylates MKKK 3/6, leading to the subsequent phosphorylation and activation of p38 pathway by (Moriguchi et al., 1996; and Shirakabe et al., 1997). During *Xenopus* embryonic development, Tak1 is involved in mesoderm induction and patterning. Tak1 deficient mouse embryos have defects in the vasculature of the embryo proper and yolk sac, a phenotype with striking similarities to that exhibited by loss-of-function mutations in the gene encoding for the Type I receptor, ALK1 (Jadrich et al., 2006). This evidence suggests that Tak1 can act through a TGF-β responsive pathway *in vivo* to regulate vascular development. The relative roles of canonical versus non-canonical BMP pathways, and whether they act independently, cooperatively (Stanton et al., 2004; Qiao et al., 2005; Reilly et al., 2005), or antagonistically with each other is still not well understood and more complex than we thought previously.

Spatial Restriction of BMP signaling Activity:

The spatial regulation of BMP activity in a developing embryo is achieved by regulation at multiple levels from ligand expression to Smad activity. Transcription of *BMPs* is spatially regulated when transcription factors required for *BMP* transcriptional activation or repression is localized within the embryo. At the ligand level, BMPs can be spatially regulated through a positive feedback loop to localized BMPs within the embryo.

An good example is the positive regulation of Dpp in wing disc development of the *Drosophila* embryos. In the *Drosophila* wing disc, Hedgehog is expressed in the posterior compartment of the wing disc and can induce Dpp expression anteriorly to its expression domain (Zecca et al., 1995; Nahmad et al., 2009). Another example of spatial regulation of BMP ligands can be seen in zebrafish embryos. The maternal BMP molecule, Radar (GDF6A), and the maternal transcription factor, POU2, are ubiquitously expressed and can induced *BMP2b*, *BMP7a*, and *BMP4* transcription. However, *BMP2b* transcription is repressed in the dorsal side of the embryo by the dorsally-expressed transcriptional repressor, Bozozok (Leung et al., 2003), and FGF signaling (Fürthauer et al., 2004). This repression leads to the enrichment of BMP mRNAs and proteins in the ventral side of the embryo. Spatial regulation of BMP expression through its ligand localization is important for proper embryonic development.

Spatial regulation of BMP activity is also determined by ligand heterodimerization, creating a localized area of high BMP activity. In *Drosophila* embryos, the BMP orthologs Decapentaplegic (Dpp) and Screw (Scw) ligands are required for dorsal-ventral (D/V) axis specification and patterning. In blastoderm stage embryos, Dpp is expressed in the dorsalmost part of the embryo and restricted dorsally through repression by Dorsal (Dl) in ventral cells (Morisato et al., 1995), while Scw is expressed ubiquitously (Arora et al., 1994). Both Dpp and Scw are required to establish the high BMP signaling necessary to specify the amnioserosa. Dpp/Scw ligand heterodimers have high affinity for Short gastrulation (Sog) and Twisted gastrulation (Tsg) and are transported dorsally by Sog and Tsg. This results in high levels of p-Mad accumulation at the dorsal midline just prior to gastrulation (Shimmi et al., 2005). The Dpp/Scw ligand heterodimers are able to

synergistically stimulate phosphorylation of Mad compared to their homodimers alone (Shimmi et al., 2005). Additionally, ligand heterodimerization such as BMP2-BMP7 can rescue *BMP2b/Swirl* mutants in gastrulating zebrafish embryos to specify ventral fates while BMP2-BMP2 homodimer does not (Little et al., 2009). Heterodimerization of BMP ligands can therefore contribute to spatial regulation of BMP activity and subsequently cell fate determination.

Cleavage of BMPs from their inactive to mature form can bring another layer of spatial regulation. In *Drosophila*, Dpp proteins are cleaved at two sites: S1 and S2 (Sopory et al., 2010). In the *Drosophila* wing, both sites are cleaved whereas in the midgut only the S1 site is utilized for cleavage (Sopory et al., 2010). Cleavage at both sites confers higher levels and longer more stable Dpp signaling in the wing disc. In a S2 cleavage mutant BMP4 mouse, only a subset of BMP4 functions is affected suggesting S2/BMP4 cleavage may be tissue specific (Goldman et al., 2006).

Differential expression of BMP receptors can also contribute to spatially defined BMP signaling activity during embryogenesis. One example is the differential expression of BMP Type I Receptor (BMPRI), Thick Veins (Tkv) in *Drosophila*. Tkv is downregulated in and just anterior to the *Dpp* expression domain, resulting in a dip in p-Mad staining close to the anterior/posterior border. However, Tkv expression is upregulated in the posterior compartment (Entchev et al., 2000; Tanimoto et al., 2000; Teleman et al., 2000) of the imaginal disc. The differential expression domains of the BMP Type I receptor are essential to establish BMP gradient activity for proper formation of the wing longitudinal veins. In mouse embryos, it has been shown that *BMPRIA* mRNA was only detected in blastocysts, whereas *BMPRIB* transcripts were detected at all stages from the one-cell

zygote to the 4-cell stage, but not in the compacted morula and blastocyst (Roelen et al., 1997). This suggests that at different developmental stages different receptors are utilized for cell fate determination and specification.

Ligand/receptor affinities can also govern BMP activity within the embryo. It has been shown that BMP receptors possess varying affinities for their ligands (Wrana et al., 1992; Brummel et al., 1994; Koenig et al., 1994; Penton et al., 1994; Wrana et al., 1994; Letsou et al 1995; Nohno et al., 1995; Yamashita et al., 1995; Natsume et al., 1997; Hata et al., 2000; Kirsch et al., 2000; Greenwald et al., 2003; Sebald et al., 2004). Binding of ligands to receptors can occur through two possible binding modes: sequential or cooperative binding. In the sequential mode, the Type II receptor binds ligands with high affinity, followed by the binding of the Type I Receptor, which has a lower affinity to the free ligand (Wrana et al., 1992; Attisano et al., 1993; Ebner et al., 1993; Franzen et al., 1993; Bassing et al., 1994; Brummel et al., 1994; Wrana et al., 1994). Ligand binding to the Type I receptor only occurs when the ligand is first bound to the Type II receptor. In the cooperative mode, the Type I and II receptors bind to ligand with high affinity when expressed together but bind with low affinity when expressed separately (Koenig et al., 1994; ten Dijke et al., 1994; Letsou et al., 1995; Liu et al., 1995; Nohno et al., 1995; Rosenzweig et al., 1995; Yamashita et al., 1995; Saitoh et al., 1996). Both modes of binding possibly could modulate receptor promiscuity and therefore BMP activity. Research suggests that the BMPs bind cooperatively (Massague et al., 1998). BMP ligands display low affinity to both Type I and Type II receptors alone, but high affinity when the two receptors are present. For example, the Activin receptor II and Activin receptor IIB are bona fide activin receptors, which on

their own, bind BMP poorly, if at all. However, ACTRII and ACTRIIB can bind BMP2 and BMP7 in cooperation with BMPRIA or BMPRIB (Yamashita et al., 1995).

Membrane-bound molecules also add another level of regulation, to either promote BMP signaling or inhibit BMP diffusion. A notable example is the membrane-bound inhibitor known as Bambi (BMP and Activin Membrane-Bound Inhibitor). Bambi is a receptor that can form a complex with BMP ligands at the cell surface to prevent BMP diffusion while restricting BMP activity during embryogenesis (Onichtchouk et al., 1999). Bambi can be considered a pseudoreceptor because it sequesters BMP ligands away from the Type I and Type II receptors.

Lastly, Smads are expressed ubiquitously in most cells and their activation is determined by specific receptors. The length of the response to BMPs can be modulated by post-translational modifications of Smads by Mitogen-Activated Protein Kinases (MAPKs) or Wnts. In *Xenopus*, Smad1 linker can be phosphorylated by MAPK and targeted to Smurf1 (an E3 ubiquitin-protein ligase) for degradation, thus resulting in reduced BMP signaling activity in the dorsal part of the embryo (Sapkota et al., 2007). Similarily, Glycogen Synthase Kinase 3 (GSK3) can phosphorylate Smad1, to promote Smad degradation via Smurf1 (Fuentealba et al., 2007). Wnt can inhibit GSK3 and promote BMP activity by maintaining Smad1 expression (Fuentealba et al., 2007). Post-translational modification is another way to fine-tune the BMP response after BMP ligand activation.

During embryonic development, BMP signaling can be activated locally or over a long distance in order to specific certain cell fates. As illustrated above, there are multiple mechanisms to achieve spatial regulation of BMP signaling activity. To complete our understanding of how BMP ligands are spatially regulated, identifying BMP targets is also

essential. Despite the wealth of information on BMP signaling, a limited number of bona fide direct targets of BMP signaling have been identified: *Brinker*, *Vent2*, *Id1-4*, *Msx2* (Brugger et al., 2004), *Gata2* (Friedle et al., 2002). Therefore, to fully understand the transcriptional controls mediated by BMP signaling, we need to find all the key molecular players. Together with what we already know of how BMPs are spatially regulated and identifying more BMP targets genes, it is possible to uncover additional mechanisms regulating BMP activity during embryo development.

BMPs in Embryonic Stem Cells and Mouse Embryo:

Mouse ES cells were first isolated in 1981 from the inner cell mass (ICM) of mouse blastocysts (Evans et al., 1981, Martin, 1981). ES cells can differentiate into a wide variety of cell types, in particular, the three germ layers (ectoderm, mesoderm, and endoderm tissues). The maintenance of mouse ES cells in a self-renewal state requires the presence of cytokine leukemia inhibitory factor (LIF) and BMP4 (Ying et al., 2003). The binding of LIF to the Gp130/LIFReceptor-β activates the Jak-STAT3 pathway (Niwa et al., 1998). This subsequently leads to the activation expression of Klf4 (kruppel-like zinc finger transcription factor), which maintains Sry-box containing gene 2 (*Sox2*), a pluripotency transcription factor expressed in ES cells (Niwa et al., 2009). BMPs are normally present in ES culture media supplemented with fetal calf serum (FCS). Depletion of BMP ligands from the FCS serum causes spontaneously differentiation (Ying et al., 2003, Reversade et al., 2005). Some of the target genes regulated by BMP4 are the *Id* genes (Hollnagel et al., 1999; Ying et al., 2003). Importantly, ES cells overexpressing the *Id* genes can self-renew in the

absence of BMP4. This demonstrates that the induction of the *Id* genes is critical for BMP4/Smad-mediated pathway in ES cell self-renewal.

The interplay between extrinsic cues and intrinsic transcriptional responses determines ES cell self-renewal (Niwa et al., 2007; Chen et al., 2008) and lineage commitment. The intrinsic pluripotency core transcriptional factors in ES cells are Oct4 (Pou5f1), Sox2, and Nanog. Oct4 is essential for early embryonic development (Nichols et al., 1998; Boiani et al., 2005). Oct4 depletion in ES cells leads to trophoblast differentiation and a derepression of Caudal-related homeobox 2 (*Cdx2*) (Niwa et al., 2005). Additionally, Oct4 can interact with Sox2 to coregulate a number of genes (Boyer et al., 2005). A third well-studied transcription factor is Nanog, a homedomain protein, which can also sustain pluripotency in ES cells even in the absence of LIF (Mitsui et al., 2002; Chambers et al., 2003). *Oct4, Sox2*, and *Nanog* form a core regulatory network in ES cells self-renewal and pluripotency. An extrinsic cue such as BMP could be a part of this network regulation. However, little work has been done to examine how BMPs (an extrinsic factor) could be regulated by the *Oct4/Sox2/Nanog* network. Some evidence shows that *Oct4* depletion leads to a reduction of Smad1 and Stat3 binding based on chromatin imummoprecipitation (ChIP) (Chen et al., 2008). Additionally, *Oct4* binding to the *BMP4* promoter region has been observed (Sridharan et al., 2009). These data suggest that BMP4 may play a role in the establishment of ES cell self-renewal through a possible *Oct4* interaction.

The mechanisms of BMP signaling during early embryogenesis have been studied extensively in many different model organisms (*Drosophila*, *Xenopus*, and *zebrafish*). However, evidence of BMP activity in early preimplantation mouse embryos is limited. Coucouvanis and Martin (1999) showed by *in situ* hybridization that *BMP4* mRNA is

present exclusively in the ICM cells of the blastocyst. Additionally, BMP inhibition using a dominant negative *BMPRIB* cDNA blocked both cavitation and expression of *Hnf4a*, a marker of visceral endoderm, *in vitro* using cultured embryoid bodies (PSA1 embryonal) from carcinoma cells. This finding is the first to implicate BMPs in the cavitation of mouse embryos. Lastly, single-cell RNA-Seq of mouse embryos at the 2-cell stage to epiblast stage shows the presence of BMP ligands (Tang et al., 2011).

Dissertation Objectives/Goals:

This dissertation will first focus on characterizing the activity of an evolutionarily conserved BRE in mouse ES cells. I would like to test whether the *X.Id3* BRE-reporter is important for mediating BRE-dependent BMP signaling in mouse ES cells. This will provide further evidence that the *X.Id3* BRE identified our lab is highly conserved. Secondly, I would like to ascertain if the *X.Id3* BRE is responsible for regulating all BMP targets or if it only modulates a subset of BMP target genes using a high-throughput sequencing approach. Our study will address whether there is a universal BRE that can regulate BMP pathway target genes or if additional regulatory elements are required for BMP signaling. Third, I will investigate the role of BMPs in early mouse embryo development and whether active BMP signaling is present. This will demonstrate the importance of BMPs during early mouse embryogenesis. Lastly, I will examine how BMP can integrate into the *Oct4/Sox2/Nanog* regulatory network to regulate ES self-renewal. Our findings will shed more light into the molecular role of BMP signaling during early mouse development.

Chapter 1 References:

Andersson O, Reissmann E, Ibáñez CF. Growth differentiation factor 11 signals through the transforming growth factor-beta receptor ALK5 to regionalize the anterior-posterior axis. EMBO Rep. 2006 Aug;7:831-7.

Arnold SJ, Maretto S, Islam A, Bikoff EK, Robertson EJ. Dose-dependent Smad1, Smad5 and Smad8 signaling in the early mouse embryo. Dev Biol. 2006 Aug 1;296:104-18.

Attisano L, Cárcamo J, Ventura F, Weis FM, Massagué J, Wrana JL. Identification of human activin and TGF beta type I receptors that form heteromeric kinase complexes with type II receptors. Cell. 1993 Nov 19;75:671-80.

Bassing CH, Yingling JM, Howe DJ, Wang T, He WW, Gustafson ML, Shah P, Donahoe PK, Wang XF. A transforming growth factor beta type I receptor that signals to activate gene expression. Science. 1994 Jan 7;263:87-9.

Bharathy S, Xie W, Yingling JM, Reiss M. Cancer-associated transforming growth factor beta type II receptor gene mutant causes activation of bone morphogenic protein-Smads and invasive phenotype. Cancer Res. 2008 Mar 15;68:1656-66.

Blahna MT, Hata A. Smad-mediated regulation of microRNA biosynthesis. FEBS Lett. 2012 Jul 4;586:1906-12.

Boesen CC, Motyka SA, Patamawenu A, Sun PD. Crystallization and preliminary crystallographic studies of human TGF-beta type II receptor ligand-binding domain. Acta Crystallogr D Biol Crystallogr. 2002 Jul;58:1214-6.

Boiani M, Schöler HR. Regulatory networks in embryo-derived pluripotent stem cells. Nat Rev Mol Cell Biol. 2005 Nov;6:872-84.

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell. 2005 Sep 23;122:947-56.

Brugger SM, Merrill AE, Torres-Vazquez J, Wu N, Ting MC, Cho JY, Dobias SL, Yi SE, Lyons K, Bell JR, Arora K, Warrior R, Maxson R. A phylogenetically conserved cis-regulatory module in the Msx2 promoter is sufficient for BMP-dependent transcription in murine and *Drosophila* embryos. Development. 2004 Oct;131:5153-65.

Brummel TJ, Twombly V, Marqués G, Wrana JL, Newfeld SJ, Attisano L, Massagué J, O'Connor MB, Gelbart WM. Characterization and relationship of Dpp receptors encoded by the saxophone and thick veins genes in *Drosophila*. Cell. 1994 Jul 29;78:251-61.

Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell. 2003 May 30;113:643-55.

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK, Clarke ND, Wei CL, Ng HH. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell. 2008 Jun 13;133:1106-17.

Cheng LE, Reed RR. Zfp423/OAZ participates in a developmental switch during olfactory neurogenesis. Neuron. 2007 May 24;54:547-57.

Coucouvanis E, Martin GR. BMP signaling plays a role in visceral endoderm differentiation and cavitation in the early mouse embryo. Development. 1999 Feb;126(3):535-46.

Daly AC, Randall RA, Hill CS. Transforming growth factor beta-induced Smad1/5 phosphorylation in epithelial cells is mediated by novel receptor complexes and is essential for anchorage-independent growth. Mol Cell Biol. 2008 Nov;28:6889-902.

Dennler S, Itoh S, Vivien D, ten Dijke P, Huet S, Gauthier JM. Direct binding of Smad3 and Smad4 to critical TGF beta-inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. EMBO J. 1998 Jun 1;17:3091-100.

Derynck R, Feng XH. TGF-beta receptor signaling. Biochim Biophys Acta. 1997 Oct 24;1333:F105-50.

Derynck R, Zhang Y, Feng XH. Smads: transcriptional activators of TGF-beta responses. Cell. 1998 Dec 11;95:737-40.

Derynck R, Zhang YE. Smad-dependent and Smad-independent pathways in TGF-beta family signalling. Nature. 2003 Oct 9;425:577-84.

Descargues P, Sil AK, Sano Y, Korchynskyi O, Han G, Owens P, Wang XJ, Karin M. IKKalpha is a critical coregulator of a Smad4-independent TGFbeta-Smad2/3 signaling pathway that controls keratinocyte differentiation. Proc Natl Acad Sci U S A. 2008 Feb 19;105:2487-92.

Drissi MH, Li X, Sheu TJ, Zuscik MJ, Schwarz EM, Puzas JE, Rosier RN, O'Keefe RJ. Runx2/Cbfa1 stimulation by retinoic acid is potentiated by BMP2 signaling through interaction with Smad1 on the collagen X promoter in chondrocytes. J Cell Biochem. 2003 Dec 15;90:1287-98.

Ebner R, Chen RH, Lawler S, Zioncheck T, Derynck R. Determination of type I receptor specificity by the type II receptors for TGF-beta or activin. Science. 1993 Nov 5;262:900-2.

Entchev EV, Schwabedissen A, González-Gaitán M. Gradient formation of the TGF-beta homolog Dpp. Cell. 2000 Dec 8;103:981-91.

Evans MJ, Kaufman M. 1981. Establishment in culture of pluripotential cells from mouse embryos. Nature 292:154–56.

Feng XH, Derynck R. Specificity and versatility in tgf-beta signaling through Smads. Annu Rev Cell Dev Biol. 2005;21:659-93.

Franzén P, ten Dijke P, Ichijo H, Yamashita H, Schulz P, Heldin CH, Miyazono K. Cloning of a TGF beta type I receptor that forms a heteromeric complex with the TGF beta type II receptor. Cell. 1993 Nov 19;75:681-92.

Friedle H, Knöchel W. Cooperative interaction of Xvent-2 and GATA-2 in the activation of the ventral homeobox gene Xvent-1B. J Biol Chem. 2002 Jun 28;277:23872-81.

Fuentealba LC, Eivers E, Ikeda A, Hurtado C, Kuroda H, Pera EM, De Robertis EM. Integrating patterning signals: Wnt/GSK3 regulates the duration of the BMP/Smad1 signal. Cell. 2007 Nov 30;131:980-93.

Fürthauer M, Van Celst J, Thisse C, Thisse B. Fgf signalling controls the dorsoventral patterning of the zebrafish embryo. Development. 2004 Jun;131:2853-64.

Gao S, Steffen J, Laughon A. Dpp-responsive silencers are bound by a trimeric Mad-Medea complex. J Biol Chem. 2005 Oct 28;280:36158-64.

Gilchrist RB, Ritter LJ, Myllymaa S, Kaivo-Oja N, Dragovic RA, Hickey TE, Ritvos O, Mottershead DG. Molecular basis of oocyte-paracrine signalling that promotes granulosa cell proliferation. J Cell Sci. 2006 Sep 15;119:3811-21.

Goldman DC, Hackenmiller R, Nakayama T, Sopory S, Wong C, Kulessa H, Christian JL. Mutation of an upstream cleavage site in the BMP4 prodomain leads to tissue-specific loss of activity. Development. 2006 May;133:1933-42.

Goumans MJ, Valdimarsdottir G, Itoh S, Lebrin F, Larsson J, Mummery C, Karlsson S, ten Dijke P. Activin receptor-like kinase (ALK)1 is an antagonistic mediator of lateral TGFbeta/ALK5 signaling. Mol Cell. 2003 Oct;12:817-28.

Greenwald J, Groppe J, Gray P, Wiater E, Kwiatkowski W, Vale W, Choe S. The BMP7/ActRII extracellular domain complex provides new insights into the cooperative nature of receptor assembly. Mol Cell. 2003 Mar;11:605-17.

Greenwald J, Vega ME, Allendorph GP, Fischer WH, Vale W, Choe S. A flexible activin explains the membrane-dependent cooperative assembly of TGF-beta family receptors. Mol Cell. 2004 Aug 13;15:485-9.

Griffith DL, Keck PC, Sampath TK, Rueger DC, Carlson WD. Three-dimensional structure of recombinant human osteogenic protein 1: structural paradigm for the transforming growth factor beta superfamily. Proc Natl Acad Sci U S A. 1996 Jan 23;93:878-83.

Gunnell LM, Jonason JH, Loiselle AE, Kohn A, Schwarz EM, Hilton MJ, O'Keefe RJ. TAK1 regulates cartilage and joint development via the MAPK and BMP signaling pathways. J Bone Miner Res. 2010 Aug;25:1784-97.

Hart PJ, Deep S, Taylor AB, Shu Z, Hinck CS, Hinck AP. Crystal structure of the human TbetaR2 ectodomain--TGF-beta3 complex. Nat Struct Biol. 2002 Mar;9:203-8.

Hassel S, Schmitt S, Hartung A, Roth M, Nohe A, Petersen N, Ehrlich M, Henis YI, Sebald W, Knaus P. Initiation of Smad-dependent and Smad-independent signaling via distinct BMP-receptor complexes. J Bone Joint Surg Am. 2003;85-A Suppl 3:44-51.

Hata A, Seoane J, Lagna G, Montalvo E, Hemmati-Brivanlou A, Massagué J. OAZ uses distinct DNA- and protein-binding zinc fingers in separate BMP-Smad and Olf signaling pathways. Cell. 2000 Jan 21;100:229-40.

Hester M, Thompson JC, Mills J, Liu Y, El-Hodiri HM, Weinstein M. Smad1 and Smad8 function similarly in mammalian central nervous system development. Mol Cell Biol. 2005 Jun;25:4683-92.

Hollnagel A, Oehlmann V, Heymer J, Rüther U, Nordheim A. Id genes are direct targets of bone morphogenetic protein induction in embryonic stem cells. J Biol Chem. 1999 Jul 9;274:19838-45.

Hester M, Thompson JC, Mills J, Liu Y, El-Hodiri HM, Weinstein M.
Karaulanov E, Knöchel W, Niehrs C. Transcriptional regulation of BMP4 synexpression in transgenic *Xenopus*. EMBO J. 2004 Feb 25;23:844-56.

Huminiecki L, Goldovsky L, Freilich S, Moustakas A, Ouzounis C, Heldin CH. Emergence, development and diversification of the TGF-beta signalling pathway within the animal kingdom. BMC Evol Biol. 2009 Feb 3;9:28. doi: 10.1186/1471-2148-9-28.

Jadrich JL, O'Connor MB, Coucouvanis E. The TGF beta activated kinase TAK1 regulates vascular development in vivo. Development. 2006 Apr;133:1529-41.

Katagiri T, Imada M, Yanai T, Suda T, Takahashi N, Kamijo R. Identification of a BMP-responsive element in Id1, the gene for inhibition of myogenesis. Genes Cells. 2002 Sep;7:949-60.

Keller S, Nickel J, Zhang JL, Sebald W, Mueller TD. Molecular recognition of BMP-2 and BMP receptor IA. Nat Struct Mol Biol. 2004 May;11:481-8.

Kirsch T, Sebald W, Dreyer MK. Crystal structure of the BMP-2-BRIA ectodomain complex. Nat Struct Biol. 2000 Jun;7:492-6.

Koenig BB, Cook JS, Wolsing DH, Ting J, Tiesman JP, Correa PE, Olson CA, Pecquet AL, Ventura F, Grant RA, et al. Characterization and cloning of a receptor for BMP-2 and BMP-4 from NIH 3T3 cells. Mol Cell Biol. 1994 Sep;14:5961-74.

Korchynskyi O, ten Dijke P. Identification and functional characterization of distinct critically important bone morphogenetic protein-specific response elements in the Id1 promoter. J Biol Chem. 2002 Feb 15;277:4883-91.

Krattinger A, Gendre N, Ramaekers A, Grillenzoni N, Stocker RF. DmOAZ, the unique *Drosophila* melanogaster OAZ homologue is involved in posterior spiracle development. Dev Genes Evol. 2007 Mar;217:197-208.

Letsou A, Arora K, Wrana JL, Simin K, Twombly V, Jamal J, Staehling-Hampton K, Hoffmann FM, Gelbart WM, Massagué J, et al. *Drosophila* Dpp signaling is mediated by the punt gene product: a dual ligand-binding type II receptor of the TGF beta receptor family. Cell. 1995 Mar 24;80:899-908.

Leung T, Bischof J, Söll I, Niessing D, Zhang D, Ma J, Jäckle H, Driever W. Bozozok directly represses bmp2b transcription and mediates the earliest dorsoventral asymmetry of bmp2b expression in zebrafish. Development. 2003 Aug;130:3639-49.

Levy L, Hill CS. Smad4 dependency defines two classes of transforming growth factor {beta} (TGF-{beta}) target genes and distinguishes TGF-{beta}-induced epithelial-mesenchymal transition from its antiproliferative and migratory responses. Mol Cell Biol. 2005 Sep;25:8108-25.

Lim SK, Hoffmann FM. Smad4 cooperates with lymphoid enhancer-binding factor 1/T cell-specific factor to increase c-myc expression in the absence of TGF-beta signaling. Proc Natl Acad Sci U S A. 2006 Dec 5;103:18580-5.

Little SC, Mullins MC. Bone morphogenetic protein heterodimers assemble heteromeric type I receptor complexes to pattern the dorsoventral axis. Nat Cell Biol. 2009 May;11:637-43.

Liu F, Ventura F, Doody J, Massagué J. Human type II receptor for bone morphogenic proteins (BMPs): extension of the two-kinase receptor model to the BMPs. Mol Cell Biol. 1995 Jul;15:3479-86.

Liu Y, Festing M, Thompson JC, Hester M, Rankin S, El-Hodiri HM, Zorn AM, Weinstein M. Smad2 and Smad3 coordinately regulate craniofacial and endodermal development. Dev Biol. 2004 Jun 15;270:411-26.

López-Rovira T, Chalaux E, Massagué J, Rosa JL, Ventura F. Direct binding of Smad1 and Smad4 to two distinct motifs mediates bone morphogenetic protein-specific transcriptional activation of Id1 gene. J Biol Chem. 2002 Feb 1;277:3176-85.

Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A. 1981 Dec;78:7634-8.

Marty T, Müller B, Basler K, Affolter M. Schnurri mediates Dpp-dependent repression of brinker transcription. Nat Cell Biol. 2000 Oct;2:745-9.

Massagué J, Seoane J, Wotton D. Smad transcription factors. Genes Dev. 2005 Dec 1;19:2783-810.

McDonald NQ, Hendrickson WA. A structural superfamily of growth factors containing a cystine knot motif. Cell. 1993 May 7;73:421-4.

Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell. 2003 May 30;113:631-42.

Moriguchi T, Kuroyanagi N, Yamaguchi K, Gotoh Y, Irie K, Kano T, Shirakabe K, Muro Y, Shibuya H, Matsumoto K, Nishida E, Hagiwara M. A novel kinase cascade mediated by mitogen-activated protein kinase kinase 6 and MKK3. J Biol Chem. 1996 Jun 7;271:13675-9.

Morisato D, Anderson KV. Signaling pathways that establish the dorsal-ventral pattern of the *Drosophila* embryo. Annu Rev Genet. 1995;29:371-99.

Moustakas A, Heldin CH. Non-Smad TGF-beta signals. J Cell Sci. 2005 Aug 15;118(Pt 16):3573-84.

Nahmad M, Stathopoulos A. Dynamic interpretation of hedgehog signaling in the *Drosophila* wing disc. PLoS Biol. 2009 Sep;7:e1000202.

Natsume T, Tomita S, Iemura S, Kinto N, Yamaguchi A, Ueno N. Interaction between soluble type I receptor for bone morphogenetic protein and bone morphogenetic protein-4. J Biol Chem. 1997 Apr 25;272:11535-40.

Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Schöler H, Smith A. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell. 1998 Oct 30;95:379-91.

Niwa H, Burdon T, Chambers I, Smith A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev. 1998 Jul 1;12:2048-60.

Niwa H, Ogawa K, Shimosato D, Adachi K. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature. 2009 Jul 2;460:118-22.

Niwa H, Toyooka Y, Shimosato D, Strumpf D, Takahashi K, Yagi R, Rossant J. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell. 2005 Dec 2;123:917-29.

Niwa H. How is pluripotency determined and maintained? Development. 2007 Feb;134:635-46.

Nohno T, Ishikawa T, Saito T, Hosokawa K, Noji S, Wolsing DH, Rosenbaum JS. Identification of a human type II receptor for bone morphogenetic protein-4 that forms differential heteromeric complexes with bone morphogenetic protein type I receptors. J Biol Chem. 1995 Sep 22;270:22522-6.

Novel regulators of bone formation: molecular clones and activities. Science. 1988 Dec 16;242:1528-34.

O'Connor MB, Umulis D, Othmer HG, Blair SS. Shaping BMP morphogen gradients in the *Drosophila* embryo and pupal wing. Development. 2006 Jan;133:183-93.

Onichtchouk D, Chen YG, Dosch R, Gawantka V, Delius H, Massagué J, Niehrs C. Silencing of TGF-beta signalling by the pseudoreceptor BAMBI. Nature. 1999 Sep 30;401:480-5.

Pangas SA, Li X, Umans L, Zwijsen A, Huylebroeck D, Gutierrez C, Wang D, Martin JF, Jamin SP, Behringer RR, Robertson EJ, Matzuk MM. Conditional deletion of Smad1 and Smad5 in somatic cells of male and female gonads leads to metastatic tumor development in mice. Mol Cell Biol. 2008 Jan;28:248-57.

Penton A, Chen Y, Staehling-Hampton K, Wrana JL, Attisano L, Szidonya J, Cassill JA, Massagué J, Hoffmann FM. Identification of two bone morphogenetic protein type I receptors in *Drosophila* and evidence that Brk25D is a decapentaplegic receptor. Cell. 1994 Jul 29;78:239-50.

Peterson KJ, Cotton JA, Gehling JG, Pisani D. The Ediacaran emergence of bilaterians: congruence between the genetic and the geological fossil records. Philos Trans R Soc Lond B Biol Sci. 2008 Apr 27;363:1435-43.

Pyrowolakis G, Hartmann B, Müller B, Basler K, Affolter M. A simple molecular complex mediates widespread BMP-induced repression during *Drosophila* development. Dev Cell. 2004 Aug;7:229-40.

Qiao B, Padilla SR, Benya PD. Transforming growth factor (TGF)-beta-activated kinase 1 mimics and mediates TGF-beta-induced stimulation of type II collagen synthesis in chondrocytes independent of Col2a1 transcription and Smad3 signaling. J Biol Chem. 2005 Apr 29;280:17562-71.

Rebbapragada A, Benchabane H, Wrana JL, Celeste AJ, Attisano L. Myostatin signals through a transforming growth factor beta-like signaling pathway to block adipogenesis. Mol Cell Biol. 2003 Oct;23:7230-42.

Reilly GC, Golden EB, Grasso-Knight G, Leboy PS. Differential effects of ERK and p38 signaling in BMP-2 stimulated hypertrophy of cultured chick sternal chondrocytes. Cell Commun Signal. 2005 Feb 3;3:3.

Retting KN, Song B, Yoon BS, Lyons KM. BMP canonical Smad signaling through Smad1 and Smad5 is required for endochondral bone formation. Development. 2009 Apr;136:1093-104.

Reversade B, Kuroda H, Lee H, Mays A, De Robertis EM. Depletion of Bmp2, Bmp4, Bmp7 and Spemann organizer signals induces massive brain formation in *Xenopus* embryos. Development. 2005 Aug;132:3381-92.

Roelen BA, Goumans MJ, van Rooijen MA, Mummery CL. Differential expression of BMP receptors in early mouse development. Int J Dev Biol. 1997 Aug;41:541-9.

Rosenzweig BL, Imamura T, Okadome T, Cox GN, Yamashita H, ten Dijke P, Heldin CH, Miyazono K. Cloning and characterization of a human type II receptor for bone morphogenetic proteins. Proc Natl Acad Sci U S A. 1995 Aug 15;92:7632-6.

Saitoh M, Nishitoh H, Amagasa T, Miyazono K, Takagi M, Ichijo H. Identification of important regions in the cytoplasmic juxtamembrane domain of type I receptor that separate signaling pathways of transforming growth factor-beta. J Biol Chem. 1996 Feb 2;271:2769-75.

Sapkota G, Alarcón C, Spagnoli FM, Brivanlou AH, Massagué J. Balancing BMP signaling through integrated inputs into the Smad1 linker. Mol Cell. 2007 Feb 9;25:441-54.

Schmierer B, Hill CS. TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility. Nat Rev Mol Cell Biol. 2007 Dec;8:970-82.

Sebald W, Nickel J, Zhang JL, Mueller TD. Molecular recognition in bone morphogenetic protein (BMP)/receptor interaction. Biol Chem. 2004 Aug;385:697-710.

Shi X, Yang X, Chen D, Chang Z, Cao X. Smad1 interacts with homeobox DNA-binding proteins in bone morphogenetic protein signaling. J Biol Chem. 1999 May 7;274:13711-7.

Shi Y. Structural insights on Smad function in TGFbeta signaling. Bioessays. 2001 Mar;23:223-32.

Shi Y, Massagué J. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. Cell. 2003 Jun 13;113:685-700.

Shimmi O, Umulis D, Othmer H, O'Connor MB. Facilitated transport of a Dpp/Scw heterodimer by Sog/Tsg leads to robust patterning of the *Drosophila* blastoderm embryo. Cell. 2005 Mar 25;120:873-86.

Shirakabe K, Yamaguchi K, Shibuya H, Irie K, Matsuda S, Moriguchi T, Gotoh Y, Matsumoto K, Nishida E. TAK1 mediates the ceramide signaling to stress-activated protein kinase/c-Jun N-terminal kinase. J Biol Chem. 1997 Mar 28;272:8141-4.

Sopory S, Kwon S, Wehrli M, Christian JL. Regulation of Dpp activity by tissue-specific cleavage of an upstream site within the prodomain. Dev Biol. 2010 Oct 1;346:102-12.

Sorrentino A, Thakur N, Grimsby S, Marcusson A, von Bulow V, Schuster N, Zhang S, Heldin CH, Landström M. The type I TGF-beta receptor engages TRAF6 to activate TAK1 in a receptor kinase-independent manner. Nat Cell Biol. 2008 Oct;10:1199-207.

Sridharan R, Tchieu J, Mason MJ, Yachechko R, Kuoy E, Horvath S, Zhou Q, Plath K. Role of the murine reprogramming factors in the induction of pluripotency. Cell. 2009 Jan 23;136:364-77.

Stanton LA, Sabari S, Sampaio AV, Underhill TM, Beier F. p38 MAP kinase signalling is required for hypertrophic chondrocyte differentiation. Biochem J. 2004 Feb 15;378(Pt 1):53-62.

Tang F, Barbacioru C, Nordman E, Bao S, Lee C, Wang X, Tuch BB, Heard E, Lao K, Surani MA. Deterministic and stochastic allele specific gene expression in single mouse blastomeres. PLoS One. 2011;6:e21208.

Tanimoto H, Itoh S, ten Dijke P, Tabata T. Hedgehog creates a gradient of DPP activity in *Drosophila* wing imaginal discs. Mol Cell. 2000 Jan;5:59-71.

Teleman AA, Cohen SM. Dpp gradient formation in the *Drosophila* wing imaginal disc. Cell. 2000 Dec 8;103:971-80.

ten Dijke P, Yamashita H, Ichijo H, Franzén P, Laiho M, Miyazono K, Heldin CH. Characterization of type I receptors for transforming growth factor-beta and activin. Science. 1994 Apr 1;264:101-4.

Urist MR. Bone: formation by autoinduction. Science. 1965 Nov 12;150(3698):893-9.
van Grunsven LA, Taelman V, Michiels C, Verstappen G, Souopgui J, Nichane M, Moens E, Opdecamp K, Vanhomwegen J, Kricha S, Huylebroeck D, Bellefroid EJ. XSip1 neuralizing activity involves the co-repressor CtBP and occurs through BMP dependent and independent mechanisms. Dev Biol. 2007 Jun 1;306:34-49.

von Bubnoff A, Peiffer DA, Blitz IL, Hayata T, Ogata S, Zeng Q, Trunnell M, Cho KW. Phylogenetic footprinting and genome scanning identify vertebrate BMP response elements and new target genes. Dev Biol. 2005 May 15;281:210-26.

Wang Q, Wei X, Zhu T, Zhang M, Shen R, Xing L, O'Keefe RJ, Chen D. Bone morphogenetic protein 2 activates Smad6 gene transcription through bone-specific transcription factor Runx2. J Biol Chem. 2007 Apr 6;282:10742-8.

Wozney JM, Rosen V, Celeste AJ, Mitsock LM, Whitters MJ, Kriz RW, Hewick RM, Wang EA. Novel regulators of bone formation: molecular clones and activities. Science. 1988 Dec 16;242:1528-34.

Wrana JL, Attisano L, Cárcamo J, Zentella A, Doody J, Laiho M, Wang XF, Massagué J. TGF beta signals through a heteromeric protein kinase receptor complex. Cell. 1992 Dec 11;71:1003-14.

Wrana JL, Attisano L, Wieser R, Ventura F, Massagué J. Mechanism of activation of the TGF-beta receptor. Nature. 1994 Aug 4;370(6488):341-7.

Yao LC, Blitz IL, Peiffer DA, Phin S, Wang Y, Ogata S, Cho KW, Arora K, Warrior R. Schnurri transcription factors from *Drosophila* and vertebrates can mediate Bmp signaling through a phylogenetically conserved mechanism. Development. 2006 Oct;133:4025-34.

Yao LC, Phin S, Cho J, Rushlow C, Arora K, Warrior R. Multiple modular promoter elements drive graded brinker expression in response to the Dpp morphogen gradient. Development. 2008 Jun;135:2183-92.

Yamaguchi K, Shirakabe K, Shibuya H, Irie K, Oishi I, Ueno N, Taniguchi T, Nishida E, Matsumoto K. Identification of a member of the MAPKKK family as a potential mediator of TGF-beta signal transduction. Science. 1995 Dec 22;270:2008-11.

Yamashita H, ten Dijke P, Huylebroeck D, Sampath TK, Andries M, Smith JC, Heldin CH, Miyazono K. Osteogenic protein-1 binds to activin type II receptors and induces certain activin-like effects. J Cell Biol. 1995 Jul;130:217-26.

Yamashita M, Fatyol K, Jin C, Wang X, Liu Z, Zhang YE. TRAF6 mediates Smad-independent activation of JNK and p38 by TGF-beta. Mol Cell. 2008 Sep 26;31:918-24.

Ying QL, Nichols J, Chambers I, Smith A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell. 2003 Oct 31;115:281-92.

Zawel L, Dai JL, Buckhaults P, Zhou S, Kinzler KW, Vogelstein B, Kern SE. Human Smad3 and Smad4 are sequence-specific transcription activators. Mol Cell. 1998 Mar;1:611-7.

Zecca M, Basler K, Struhl G. Sequential organizing activities of engrailed, hedgehog and decapentaplegic in the *Drosophila* wing. Development. 1995 Aug;121:2265-78.

Zhang YE. Non-Smad pathways in TGF-beta signaling. Cell Res. 2009 Jan;19:128-39.

Zhang YW, Yasui N, Ito K, Huang G, Fujii M, Hanai J, Nogami H, Ochi T, Miyazono K, Ito Y. A RUNX2/PEBP2alpha A/CBFA1 mutation displaying impaired transactivation and Smad interaction in cleidocranial dysplasia. Proc Natl Acad Sci U S A. 2000 Sep 12;97:10549-54.

Chapter2

The BMP-Responsive Element (BRE) is an evolutionarily conserved module for BMP signaling in mouse embryonic stem cells.

Research in the Cho lab has demonstrated in *Xenopus* embryos that BMP4 can regulate the *Vent2* and *Id3* (Inhibitors of differentiation) genes via a recognized specific target sequence known as the BRE (BMP Responsive Element: 5'-GRCGNCNNNNNGTCTG-3') motif. BRE is an evolutionarily conserved regulatory module that can mediate BMP response in various organisms ranging from *Drosophila*, *Xenopus*, and *Zebrafish* (von Bubnoff et al., 2005; Yao et al., 2008; Alexander et al., 2011). The BMP response element identified in the *Xenopus* Id3 promoter by our lab was not the first BRE identified. The Korchynskyi, ten Dijke, and Mummery groups have previously identified a different BRE in the promoter of *Id1* gene (Korchynskyi et al., 2002; Monteiro et al., 2004; Monteiro et al., 2008). In their studies, they took two short sequences (-1105/-1080 and -1052/-1032) upstream of the *Id1* promoter that they considered to be the most important for BMP response and fused two copies of this "BRE" sequence in opposing orientation to an adenoviral minimal promoter construct. This BRE-luciferase reporter can activate Smad1/5/8 through the canonical BMP signaling pathway. Previously, published studies by these groups have previously generated BRE reporter mouse ES cell lines (BRE-luc, BRE-lac1, and BRE-lac2) and transgenic BRE reporter mice (BRE-lac1 and BRE-lac2) (Korchynskyi et al., 2002; Monteiro et al., 2004; Monteiro et al., 2008).

Here, I would like to investigate whether the BMP responsive element, an evolutionarily conserved module, can also mediate the BMP response in mouse embryonic

stem cells. Although, the BRE have been shown to function in *Drosophila* to mediate *Brinker* repression (Yao et al., 2006; Yao et al., 2008) and activation of *Id3* genes in *Xenopus* (von Bubnoff et al., 2005), the question as to whether the BRE can mediate BMP signaling in mouse ES cells remains untested.

Material and Methods:

Generation of the BRE-gal Reporter Mouse Embryonic Stem Cell Line:

The 7xBRE-gal reporter gene was created by digesting the 7xBRE-*XId3*-pCX GFP3 (von Bubnoff et al., 2005) construct with HindIII and SpeI. Once the 7xBRE concatamer and *X.Id3* (-201/+70) minimal promoter was isolated, this sequence was subcloned into the HindIII and XbaI digested pBAT-Gal (Maretto et al., 2003) vector, thus generating a 7xBRE-201-*XId3*-nlsLacZ/PGKNeo construct, also referred to as BRE-gal. Mouse ES cells were electrophorated with the linearized 7xBRE-201-*X.Id3*-nlsLacZ/PGKNeo plasmid. Neomycin-resistant ES cell colonies were subjected to Southern blot to confirm single-copy integration. From this, six colonies were analyzed for normal chromosome count, and two BRE-gal mouse ES cell lines were established. This was performed by Aixu Sun in the Cho lab.

Mouse ES Cell Culture Conditions:

E14Tg2a mouse ES cells were cultured on 0.1% gelatin-coated plates containing GMEM medium (Sigma Aldrich, #G5153) with 10% FBS (Hyclone, catalog #SH30071.03), 0.1mM NEAA, 1mM Sodium Pyruvate, 1000U/ml LIF (Millipore, catalog #ESG1107), and 0.1mM 2-

Mercaptoethanol. Cells were maintained at 37°C, 5% CO2, and split using 0.25% trypsin (Life Technologies, catalog #15400-054). Cells were split every other day and split using a 1:8 ratio. Frozen cell stocks were stored using 10% DMSO and 90%FBS.

ONPG/B-galactosidase Assay:

The levels of β-galactosidase expressed in cells were measured using an ONPG (ortho-nitrophenyl-beta-D-galactopyranoside) colorimetric assay. β-galactosidase catalyzes the hydrolysis of ortho-nitrophenyl-beta-D-galactopyranoside(ONPG), a lactose analog, into galactose and ortho-nitrophenol (ONP), which produces a yellow color that can be quantified using a spectrophometer.

BRE-gal mouse ES cells were treated with various secreted ligands (TGF-β (R&D), WNT3a (R&D), Activins (Sigma), GDF3, FGF4 (fibroblast growth factor 4), BMP2 (R&D), and BMP4 (R&D) at two different concentrations: 5ng/ml and 20ng/ml) for 12hrs in regular E14 culture media. Cells were harvested 12hrs after stimulation by washing cells once with 1xPBS and then lysing the cells in 400ul of lysed buffer (25 mM Tris-PO_4 (pH 7.8), 15% glycerol, 2% CHAPS {3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate}, 1% lecithin, 1% bovine serum albumin, 4 mM EGTA (pH 8.0), 8 mM $MgCl_2$, and 0.4 mM phenylmethylsulfonyl fluoride). Cells were spun down after lysis and the supernatant for each treatment was collected for ONPG assay. 100μl of lysate and 100 μl of ONPG solution was mixed and incubated at 37°C for 15 minutes and the reaction was stopped by 1M sodium bicarbonate. After stopping reaction, the intensity of ONPG reaction was measured at a wavelength of 420nm in a spectrophotometer to quantitate β-galactosidase expression upon various ligand treatments.

X-Gal Staining:

BRE-gal mouse ES cells were treated with different BMP ligands (BMP2, 4, &7) for 24hrs. The cells were washed once with 1xPBS (without calcium and magnesium chloride) and fixed with 0.05% glutaraldheyde for 15 minutes and washed consecutively with 1xPBS in 10min intervals for 3 times. Cells were then incubated in 1mg/ml X-gal solution (2mM MgCl2, 5mM K ferrocyanide, 5mM K ferricyanide in 1X PBS) contain 1mg/ml of X-gal for 4hrs at 37^{o}C and washed with 1XPBS several times before imaging.

Cycloheximide (CHX) Treatment of Cells:

BRE-gal mouse ES cells were pre-treated with 3-[2-(3,5-Dimethyl-2-oxocyclohexyl)-2-hydroxyethyl]glutarimide (cycloheximide, CHX) (Sigma: C4859) at a concentration of 20ug/ml for 2hrs. Cells were then replaced with fresh media containing CHX or CHX with BMP4 (10ng/ml) for another 4hrs. As control, cells were also treated with regular culture media (with serum) or BMP4 separately for 4hrs. Total RNA was then isolated from each treatment using 500ul of Trizol (Life Technologies, catalog #15596-026) following manufacture's instruction. Reverse transcription was performed using MMLV Reverse Transcriptase (Invitrogen, catalog #28025-013) according to manufacture's guideline. Quantitative PCR was performed using SYBR Green (Roche, catalog #04-707-516-001) and the Roche Lightcycler 480. qPCR was performed using the following LacZ primers:

Forward: 5'-GTCTTCTGACCCAGACTACCTTGTA-3'

Reverse: 5'- CAAAGTAGGACCCGTACTCATTCT-3'

Western Immunoblotting:

BRE-gal mouse ES cells treated with LDN193189 (Stemgent, catalog #04-0074, lot#1861) alone or with BMP4 were lysed in RIPA buffer (150mM NaCl, 50mM Tris-HCl pH 8.0, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate, 25mM β-

glycerophosphate, 100nM sodium fluoride, and 1mM sodium orthovanadate). Lysate containing 15μg of protein in a 40μl volume were electrophoresed through a 12% PAGE-SDS gel. Primary antibodies used were: anti-p-Smad1/5/8 (1:500, Millipore), anti-Smad1/5/8 (1:200, Santa Cruz Biotechnology), anti-HA (1:200, Santa Cruz Biotechnology), anti-β-tubulin (1:10,000, Sigma). ECL kit (GE Healthcare, catalog #RPN2232) was used for visualization.

Generation of the 7XBRE (Del-2)-Luciferase Reporter Construct:

The 7XBRE (Del-2) –Luciferase construct was generated using complementary oligos in which two nucleotides from the 5 spacer region of the BRE were deleted. Oligos:

A) 5'-GATCCTCTGGTCACAGGATAATAATCCTGACGCCAGAGTCTGGAGGTCA-3'

B) 5'-GATCTGACCTCCAGACTCTGGCGTCAGGATTATTATCCTGTGACCGAG-3'

Each oligo was phosphorylated using T4 polynucleotide kinase (PNK) according to manufacture's instruction (Invitrogen). The oligos were annealed together (30μl of 10μM concentration of both phosphorylated oligos) in a 50μl reaction of 1X Tris-EDTA buffer and incubated it at 95°C for 4 minutes and then slowly the reaction was cooled to room temperature. The annealed oligos (20μg total BRE(Del-2) DNA) were multimerized using T4 ligase (Invitrogen). The ligation product was ran on a 2% agarose gel to check for multimerization of oligos. The mulitmerized ligated product was cleaved with BamHI and BglII in a 250μL at 37°C for 3hrs. The confirmation of a ~350bp band indicating a multimerized BRE(Del-2) DNA fragment was followed by concentration of the digested reaction into a 50μL volume by sodium acetated precipitation. The concentrated reaction was then ran on a 1% agarose gel and gel extraction using a GeneClean kit was performed. Glycogen precipitation was done after gel extraction of the digested DNA.

After precipitation, the 7xBRE (Del-2) DNA fragment was cloned into the -201xId3-poLuc construct digested with BamHI. The resulting 7xBRE (Del-2)- 201xId3-poLuc construct was transformed into DH5α cells. Restriction digest test of colonies using BamHI and HindIII confirmed the presence of the multimerized mutant BRE. Each clone was sequenced and the orientation of oligonucleotides was determined.

Luciferase and β-Galactosidase Assays:

The BRE(7X)-luciferase wild type (WT) and BRE(7X)-luciferase mutant (MT) constructs were each co-transfected with the pCMV-β-galactosidase plasmid into wildtype E14 mouse ES cells using Lipofectamine 2000 (Invitrogen) according to manufacturer's direction. As a positive control, the pCMV-luc construct was also transfected into wildtype E14 mouse ES cells. Three biological replicates for each DNA transfection were prepared. . Cell media was changed 6–8 hrs post-transfection and cells were harvested 24 hours post-transfection in the presence and absence of BMP4 for 8hrs. Cells were lysed using a cell lysis buffer (25 mM Tris-PO_4 (pH 7.8), 15% glycerol, 2% CHAPS {3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate}, 1% lecithin, 1% bovine serum albumin, 4 mM EGTA (pH 8), 8 mM $MgCl_2$, and 0.4 mM phenylmethylsulfonyl fluoride) and the lysate was subject to luciferase assay. 50ul of cell lysate was mixed 100ul of luciferase assay reagent (20mM tricine. NaOH pH7.8, 1mM magnesium carbonate, 0.1mM EDTA, 5mM DTT, 2.7mM Coenzyme A, 0.5mM ATP, 0.5mM Na-Luciferin (Sigma)). Reaction mixture was measured using a luminometer (Analytical Luminescence Lab, Moonlight 2010). Glow luminescence was measured for 10sec to obtain RLU (Renilla luciferase unit) values. RLU values were normalized by performing a β-galactosidase/ONPG assay on each sample.

RNA Isolation & Quantitative Polymerase Chain Reaction (qPCR):

RNA was isolated after treatment of mouse ES cells with 10ng/ml of hBMP4 (R&D Systems, catalog #314-BP) using 500μL of Trizol (Life Technologies, catalog #15596-026). Reverse transcription was performed using MMLV Reverse Transcriptase (Invitrogen, catalog #28025-013) according to manufacture's guideline. Quantitative PCR was performed using SYBR Green (Roche, catalog #04-707-516-001). The following qPCR primers were used:

Id3:

Forward: 5'-CTGTCGGAACGTAGCCTGG-3'

Reverse: 5'-GTGGTTCATGTCGTCCAAGAG-3'

Results and Discussion:

Chapter 2.1: BREs mediate the BMP response in mouse embryonic stem cells.

Our lab generated a mouse ES cell line harboring a multimerized BRE (~46bp) upstream of the *Xenopus Id3* promoter. The advantages of generating a BRE-gal mouse ES line are two fold: a) it provides us an immediate transcriptional output to quantitate the transcriptional response mediated by BRE, and b) it abolishes the need for conventional biochemical and immunological techniques to detect BMP activity, which are time consuming and expensive. The BRE-gal reporter line consists of a stable mouse ES cell line carrying 7 copies of the 46bp BRE, which is linked to a minimal *X.Id3* promoter (-201/+70bp) linked to a LacZ/β-gal reporter (Figure 2.1).

The BRE-gal mouse ES reporter line was treated with various ligands, including BMPs, to demonstrate that the BRE-gal reporter gene in mouse ES cells were only responsive to BMP ligands. BRE-gal ES cells were treated with the following ligands: BMP2, BMP4, Activin, FGF4, TGF-β, WNT3A, and GDF3 for 12hrs and then subjected to an ONPG

(ortho-nitrophenyl-β-galactoside) assay following treatment. The ONPG assay is a quick colorimetric measurement assay for the detection of β-galactosidase activity (Fowler et al., 1977). As seen in Figure 2.2, BMP2 and BMP4 elicited a robust response (~4 fold induction), whereas other ligand treatments (Activin, FGF4, TGF-β, Wnt3A, and GDF3) did not change β-galactosidase activity. The BRE-gal mouse ES cell line responded to only BMPs and not to Activins, FGFs, Wnts, and TGF-β ligands. These results indicate that the evolutionarily conserved *Xenopus Id3* BRE module is responsive in mouse ES cells and responds specifically to only BMPs.

To further characterize this response, the transcript levels of BMP target genes were measured using quantitative RT-PCR to show that activated BMP targets were being induced in the presence of BMP4 (Figure 2.3). Inhibitors of differentiation genes know as the *Id* genes play key roles in mouse ES cell pluripotency and the *Id3* gene, in particular, is responsible for self-renewal in ES cells and a known BMP direct target (Ying et al., 2003). In BRE-gal cells treated with BMP4, the *Id3* gene is markedly increased compared to control (regular E14 media without ligand) (Figure 2.3). This clearly demonstrates that the BMP4 signaling pathway is activated in the BRE-gal mouse ES cell since a known BMP direct target gene (*Id3*) is upregulated.

Furthermore, we also demonstrated that the BRE-gal response is a direct response to BMP4 ligand treatment. We treated cells with BMP4 in the presence of a protein synthesis inhibitor, cycloheximide (McKeehan et al., 1969), and examined the transcriptional response of the BRE-gal ES cells. In the presence of cycloheximide, indirect target genes cannot be induce as the induction requires protein synthesis. Thus, only those genes upregulated in the presence of cycloheximide are direct targets. In Figure 2.4, we

observed that in the presence of cycloheximide and BMP4 there is an increase in *LacZ*/β-galactosidase mRNA levels, indicating that the *LacZ* induction is a direct response to BMP4 stimulation. However, we also noticed that *LacZ* transcripts are also induced in the presence of cycloheximide. This induction of the *LacZ* gene/transcript by cycloheximide in not a new phenomenon and have been previously reported (Hu et al., 1993). It can be hypothesized from this finding that a repressor may tightly regulate *Id3* expression. Upon cycloheximide treatment, repressor binding to the *X.Id3* BRE is relieved to allow for an upregulation of *LacZ* transcripts. Although the experiment does shows that cycloheximide treatment can induce *LacZ* transcription, the level of *LacZ* transcripts in the CHX+BMP4 treatment is much higher than CHX alone based on quantitative RT-PCR analysis. It can be concluded that *LacZ* induction in BRE-gal by BMP4 is direct and occurs in the absence of protein synthesis.

Chapter 2.2: BRE-gal reporter mouse ES cells responds to BMP4.

The 7xBRE-gal reporter mouse ES cell line is a good tool to monitor BMP signaling activity. Upon BMP (BMP2, 4, or 7) stimulation, the BRE-gal ES cells stained positive for X-gal (Figure 2.5). The minimal background X-gal staining observed in the untreated BRE-gal ES cells is primarily due to the low amount of BMPs known to be present in the serum of ES culture media. A brief comparison between our BRE reporter line and the published BRE (Korchynskyi et al., 2002; Monteiro et al., 2004; Monteiro et al., 2008) indicate some key differences. The first difference is that most of our BRE-gal ES cells appear respond to BMP ligand treatment. Our BRE-gal ES cells responded to BMPs even at concentrations as low as 5ng/mL of BMP4 (Figure 2.2), compared to the Korchynskyi BRE-lac1 mouse ES cells,

which showed a non-uniform reporter response using 20ng/mL concentration of BMP4 (Monteiro et al., 2004). It should be noted that these discrepancies could be attributed to different manufacturers of the ligands used by the two labs. Additionally, the site of integration of our BRE-gal transgene and the Korchynskyi BRE-lac1 transgene are different due to random integration of the clones. It would be useful to integrate both our BRE-gal transgene and the Korchynskyi BRE-lac1 transgene into the same *ROSA26* locus to gain a full understanding of the differences between these two BREs.

Our BRE-gal reporter was generated using seven copies of the BRE concatemerized together into our reporter construct, whereas the Korchynskyi "BRE" reporter construct only contained two copies of their "BRE". Using seven copies of our BRE into the BRE-gal construct may have help to amplify the BMP responses. Anna Javier, in our lab, demonstrated that our BRE-gal mice detected expression of BMP activity in developing embryos not previously reported by the Korchynskyi BRE-lac1 mouse reporter line (Monteiro et al., 2004; Monteiro et al., 2008). Embryos from our BRE-gal transgenic mouse reporter line revealed sites of BRE-mediated BMP activity in the chorion, otic vesicle, and somites (Javier et al., 2012), which were not previously reported.

Since most of the BRE-gal ES cells stained positive for X-gal, it is could be hypothesized that BMPs utilizes paracrine and/or autocrine signaling mechanisms for self-renewal (Figure 2.6). BMP2 has been shown in avian explant studies to be an essential paracrine signal for heart development (Yamagishi et al., 1999). Furthermore, in the Korchynskyi BRE-lac1 mouse, reporter activity was observed in not only the myocardium, but the endocardium and pericardium as well, which highly supports the idea that paracrine BMP signaling is taking place (Monteiro et al., 2004) during development. For

autocrine signaling, it is possible that there is a positive feedback loop controlling the transcriptional response of cells to either stay in a self-renewal or differentiate state upon BMP stimulation. Evidence has indicated that LIF-mediate controlled of mouse embryonic stem self-renewal (Davey et al., 2006; Davey et al., 2007) via the Jak/Stat3 pathway functions in an autoregulatory manner. It is therefore attractive to suggest that exogenous BMPs can induce a similar autoregulatory loop through Smad1/5/8 pathway to maintain a robust pathway responsiveness in ES cells as it undergo self-renewal.

It is interesting to note that varying degrees/intensities of X-gal staining (Figure 2.7) were observed in some BRE-gal ES cells. An explanation for the varying degrees in the X-gal intensity can be due to ligand-receptor interactions. Ligand-receptor interactions have been demonstrated to be important in self-renewal and cell fate determination. It is well known that signaling through ligand-receptor complexes are not static, suggesting cells that respond readily to BMP (hence higher β-galactosidase and more ligand-receptor complexes) may more likely undergo self-renewal versus ES cells with low number of ligand-receptor complexes and less β-galactosidase to differentiate (Viswanathan et al., 2002) (Figure 2.8).

Chapter 2.3: The BRE is a conserved module important for BMP signaling.

A great deal of progress has been made in identifying the components that mediate and modulate BMP signaling. However, the functions regulating BMP signaling in vertebrates is poorly understood especially in early mouse development. Published works have suggested that the BRE found in the promoters of BMP target genes can function to activate or repress target genes depending on the organism and context (see Chapter1).

Here we hypothesized that the *Xenopus Id3* BRE is a highly conserved element that can activate the BMP response in mouse ES cells. If this *Xenopus Id3* BRE is mutated, it can lead to the abolishment of BMP signaling in mouse ES cells.

To abolish BRE-mediated BMP4 activation, we made mutated version of 7X BRE-luciferase construct, containing a two-nucleotide deletion in the five base spacer region of each BRE (7X BRE-Del2-luciferase) (Figure 2.9). The five-spacer region between the Smad1 (GC binding site) and Smad4 binding site (SBE) was shown to be important for BRE-mediate BMP signaling. A simple two-nucleotide spacer deletion within the space region of BRE completely abolished BMP response in *Drosophila* (Yao et al., 2008). We therefore deleted the last two nucleotides of the five-nucleotide spacer, keeping the Smad1 and the Smad4 binding site intact (Figure 2.10A). The wild type 7XBRE-luc and mutant 7XBRE-luc constructs were transfected into E14 mouse ES cells and luciferase activity was measured in the presence and absence of BMP4 stimulation. Transfection of the wild type 7xBRE–luciferase construct did not induce luciferase activity in the absence of BMP4, but luciferase activity was markedly upregulated upon BMP4 stimulation (Figure 2.10B). Conversely, we observed an abolishment of luciferase activity in the mutant luciferase construct (7X BRE-Del2-luciferase) after BMP4 stimulation (Figure 2.10B).

Chapter 2.4: BRE-mediated BMP response is specific to only BMPs.

To further show BRE- dependent activation in mouse ES cells is BMP specific, we utilized Noggin, a BMP antagonist, and a small chemical inhibitor (LDN193189) to inhibit BMP signaling. Noggin has a high binding affinity (picomolar) for BMP2, 4, and 7 ligands (Zimmerman et al., 1996) and therefore sequesters any BMPs from binding to the BMP

receptors (Zimmerman et al., 1996) to activate the BMP signaling cascade. LDN193189 antagonizes BMP signaling by selectively targeting the BMP receptor istoypes ALK2 and ALK3 by inhibiting their kinase activity. LDN193189, a variant of dorsomorphin, has been shown to inhibit phoshorylation of Smad1,5,8 (Boergermann et al., 2010). The advantage of LDN193189 over dorsomorphin is that it exhibits a higher specificity for BMP receptors and can be used at a lower concentration to mitigate off target effects (Yu et al., 2008).

In Figure 2.11, Panel A shows that almost all cells are stained in blue after 10ng/ml BMP treatment for 12hrs, whereas in the presence of BMP4 and Noggin (Panel B) there is a substantial decrease in X-gal staining. However, Panel B does still show residual X-gal staining primarily due to some BMPs found in the culture media. Nonetheless, Noggin does inhibit BMP4 signaling. LDN193189 was efficient in inhibiting BMP4 signaling as seen in Figure 2.12. In the presence of LDN193189, BRE-gal expression was blocked in almost all the cells, indicating that LDN193189 is a potent inhibitor of BMP4 signalling cascade. Western blot analysis of cells stimulated with BMP4 together with LDN193189 showed a significant decrease of phosphorylated smads levels (Figure 2.13).

The advantage of LDN is that it can robustly inhibit BMP4 signaling and LDN is a relatively more inexpensive alternative to Noggin. Additionally, LDN193189 has been shown to have minimal off-target effects if concentration of LDN is used is below 1μM range (Boergermann et al., 2010) in a cell culture system. Our working concentration of 80nM LDN193189 in mouse embryonic stem cells is well below the concentration to have any influence in off-target effects. The use of a small molecule inhibitor such as LDN193189 to inhibit BMP signaling is a useful pharmacological tool to investigate BMP signaling during early development.

Conclusion:

In the present study, we characterized the specificity of the BRE in mouse ES cells. We demonstrated the conservation of the 46bp BRE from *Drosophila* to the mouse system. This conclusion is based on observations that showed 1) the mutated BRE failed to respond to BMP signaling, 2) the BMP antagonist, Noggin, and LDN193189 inhibited BRE-gal response to BMP signaling. Consistent with previous studies in *Xenopus* embryos, the BRE in mouse ES cells is responsible for activating the BMP transcriptional response and that the 5-bp spacer within the BRE is required for this activation.

Figure 2.1

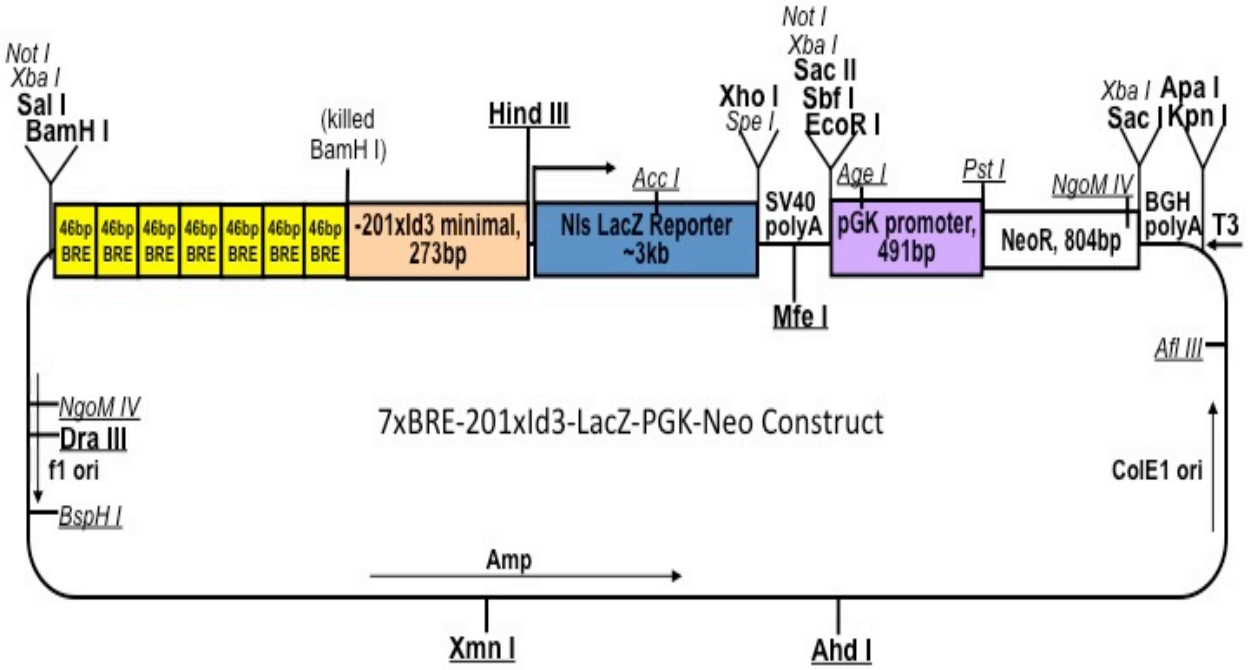


Figure 2.1: 7XBRE-201*X.Id3*-nlsLacZ-PGK-Neo Construct. The construct consists of 7 copies of multimerize BRE (273bp) linked to a minimal Id3 promoter (-201bp). Immediately following the minimal promoter is a LacZ reporter containing a nuclear localization signal. Construct can be linearized with NotI.

Figure 2.2

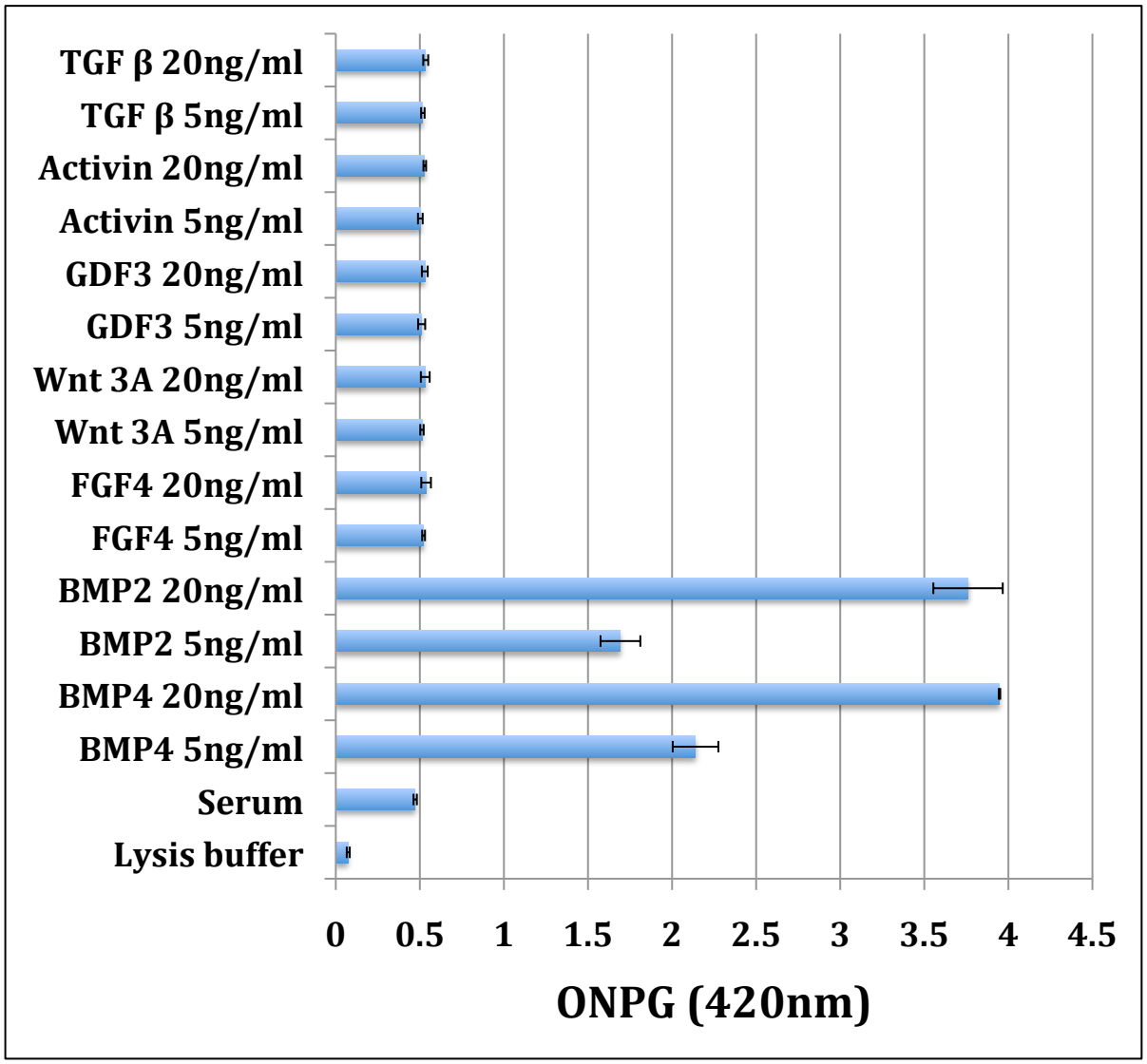


Figure 2.2: BRE-gal mouse ES cells respond specifically to BMP ligands. BRE-gal mouse ES cells were treated with indicated growth factors for 12hrs in regular E14 culture media. Cell lysates were isolated after stimulation with different ligands, and subjected to ONPG (*ortho*-Nitrophenyl-β-galactoside) assay to measure β-galactosidase expression levels. BMP4 and BMP2 homodimers activated LacZ reporter in BRE-gal mouse ES cells at a concentration as low as 5ng/ml. Non-BMP growth factors (e.g. Wnt, FGF, GDF, and TGF-β, and Activin ligands) were unable to elicit a transcriptional response in BRE-gal reporter cells.

Figure 2.3:

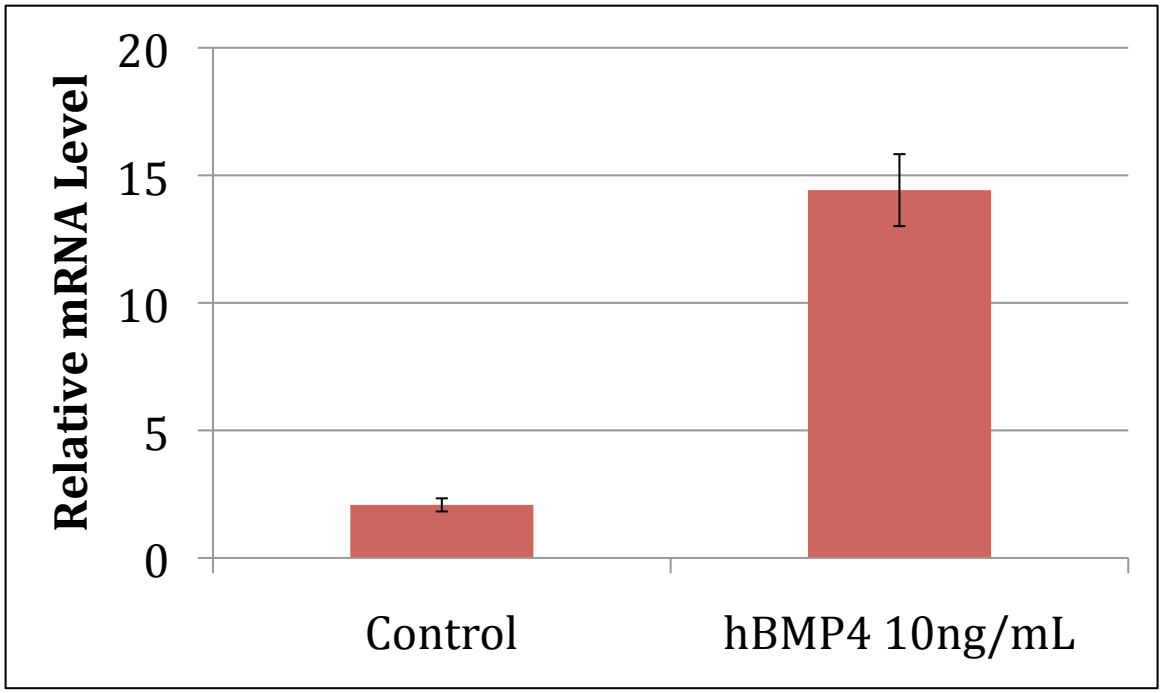


Figure 2.3: BMP induction of the endogenous *Id3* gene. Quantitative RT-PCR of *Id3* transcripts was performed in BRE-gal cells treated with BMP4 for 6hrs and regular culture media (Control).

Figure 2.4:

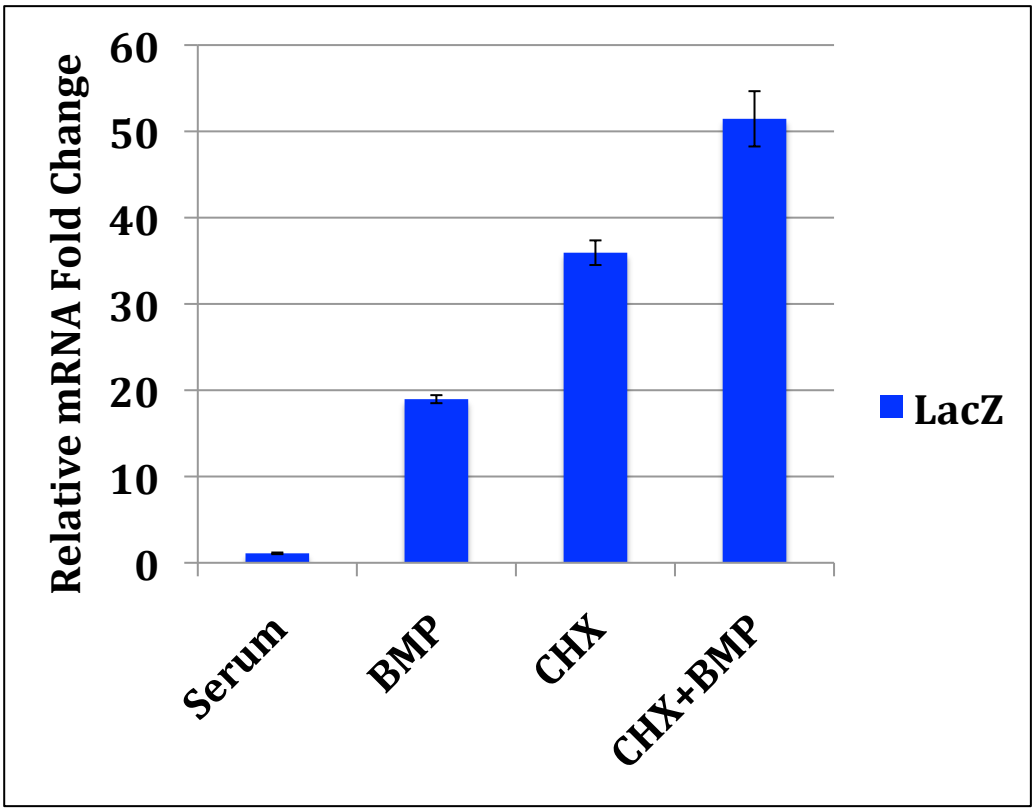
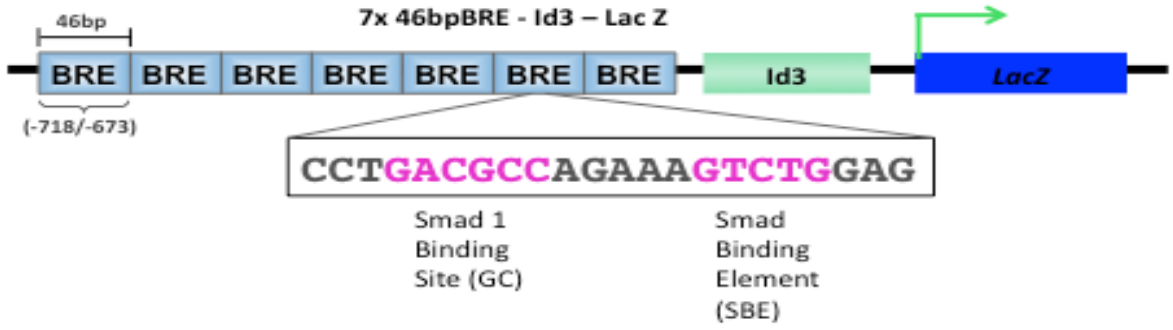


Figure 2.4: BMP4 directly activates BRE-LacZ expression in the presence of cycloheximide (CHX). LacZ is a direct target of BMP4 stimulation. BRE-gal mouse ES cells were induced with BMP4. For CHX treatment, cells were first pre-treated with CHX for 2hrs and then incubated together with either CHX or CHX+BMP4 for 4hrs. Quantitative RT-PCR was subsequently performed to measure LacZ transcript levels.

Figure 2.5:

A



B

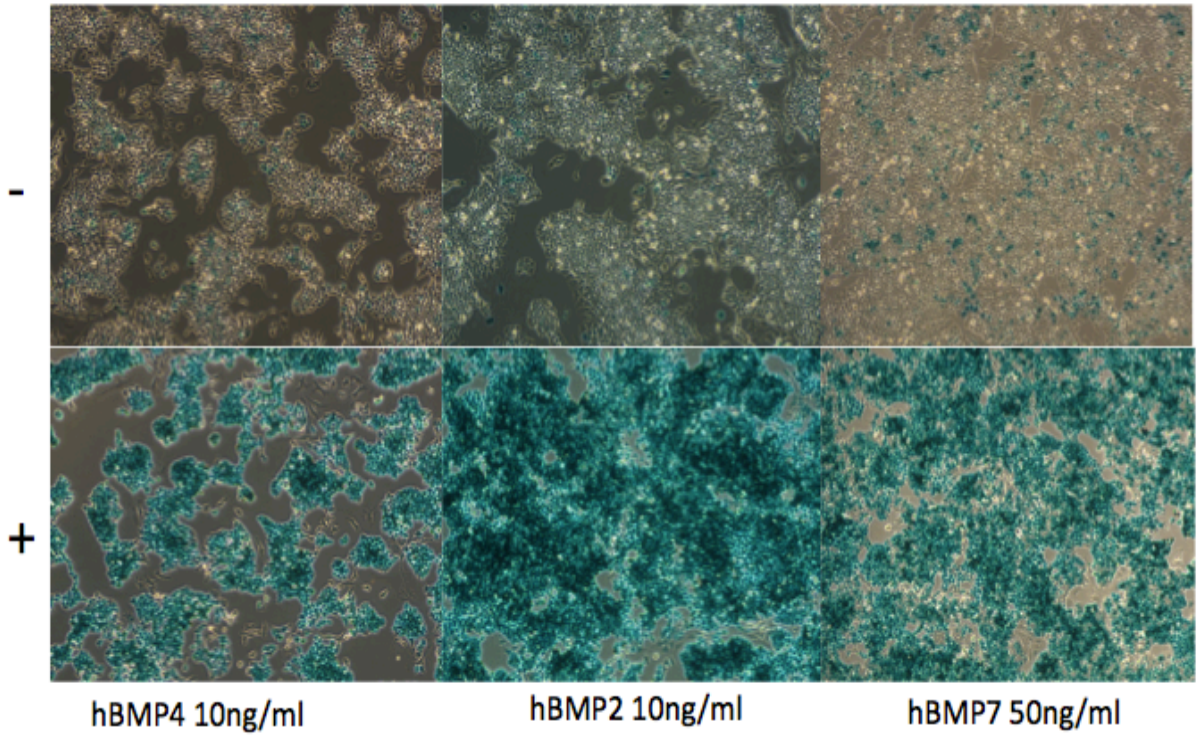


Figure 2.5: The BRE-gal Reporter mouse ES cell line responds to BMPs. A) A construct map of the BRE-gal Reporter mouse ES cell. The construct harbors 7 copies of the multimerized BRE in a minimal promoter Id3 linked to the LacZ/Galactosidase reporter gene. Each BRE contains a Smad1 Binding Site (GC-rich) and Smad Binding Element (Smad4) B) The BRE-gal ES cells respond to different BMP ligands (BMP4, BMP2, and BMP7). A concentration of 10ng/ml was able to elicit a response in BRE-gal mouse ES cells as shown by X-gal staining. A higher dose of BMP7 at 50ng/ml was needed for BRE-gal response but nevertheless, it responded to BMP7.

Figure 2.6:

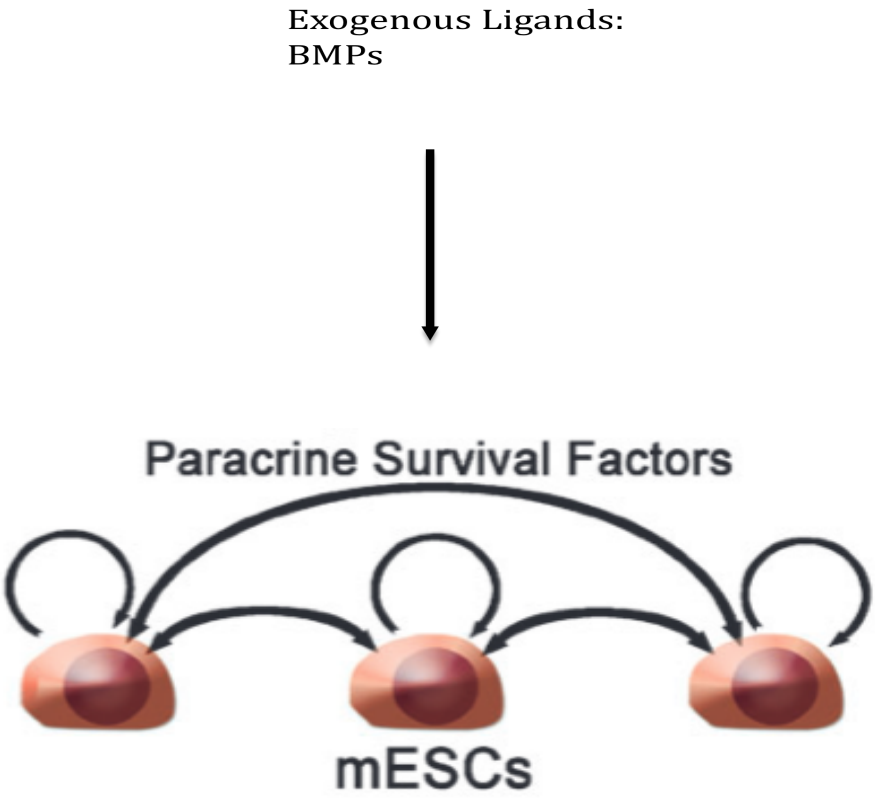


Figure 2.6: Model of paracrine and autocrine signaling in mouse ES cells. Exogenous ligands like BMP4 can potentially stimulate paracrine and autocrine survival/self-renewal factors in mouse ES cells (Yamagishi et al., 1999; Davey et al., 2007; Davey et al., 2006).

Figure 2.7

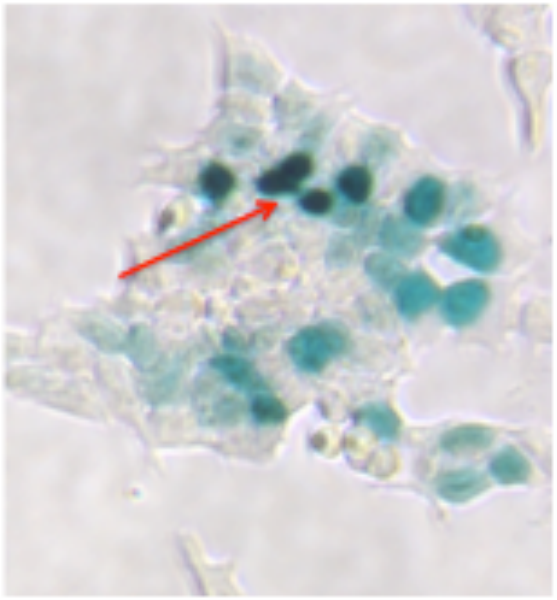


Figure 2.7: BRE-gal mouse ES cell responds to BMP4 but intensity of BRE-gal cells varies. After treatment of BMP4 for 24hrs and X-gal staining for 4hr, cells with darker X-gal staining was observed in some cells. This could be due to autocrine or paracrine signaling within individual cells.

Figure 2.8:

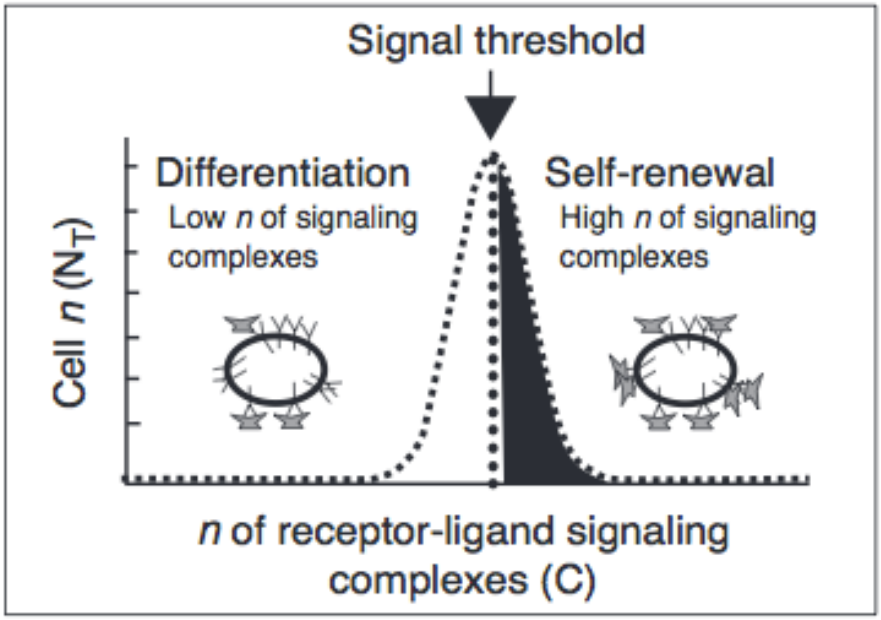


Figure 2.8: Ligand-receptor signaling threshold model for ES cell for self-renewal or differentiation. The number of receptor-ligand signaling complexes determine if a cell undergo differentiation or self-renewal based on the signaling threshold. Cell with a high numbers of signaling complexes can undergo self-renewal while cells with numbers of signaling complexes will undergo differentiation. This could explain why some cells stained with X-gal more intensely than others in the BRE-gal mouse ES cells. (Viswanathan et al., 2002)

Figure 2.9:

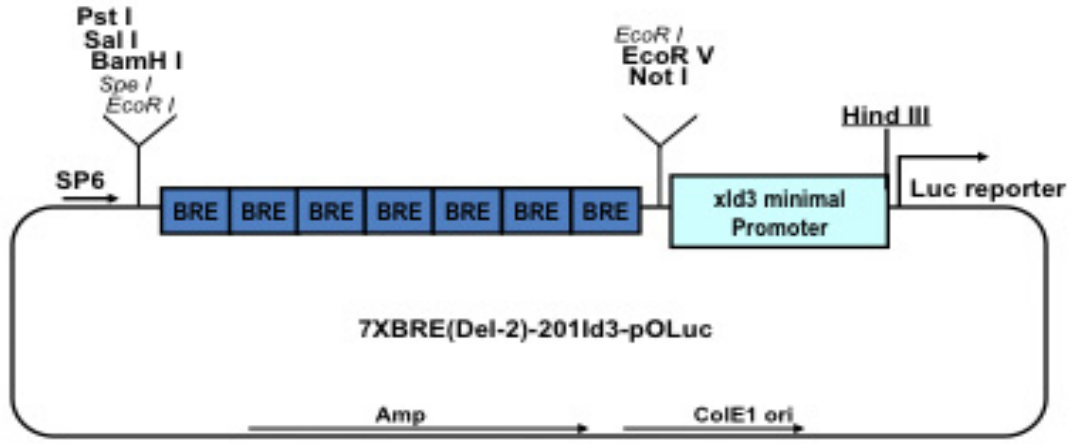


Figure 2.9: Generation of a Mutant 7XBRE(Del-2)-201-*XId3*-poLuc construct. The BRE(Del-2) DNA fragment was generated from annealing oligos containing a 2 nucleotide deletion of the 5 spacer region within the BRE. The annealed oligos were then multimerized and ligated into the poLuc plasmid containing the xId3 minimal promoter.

Figure 2.10:

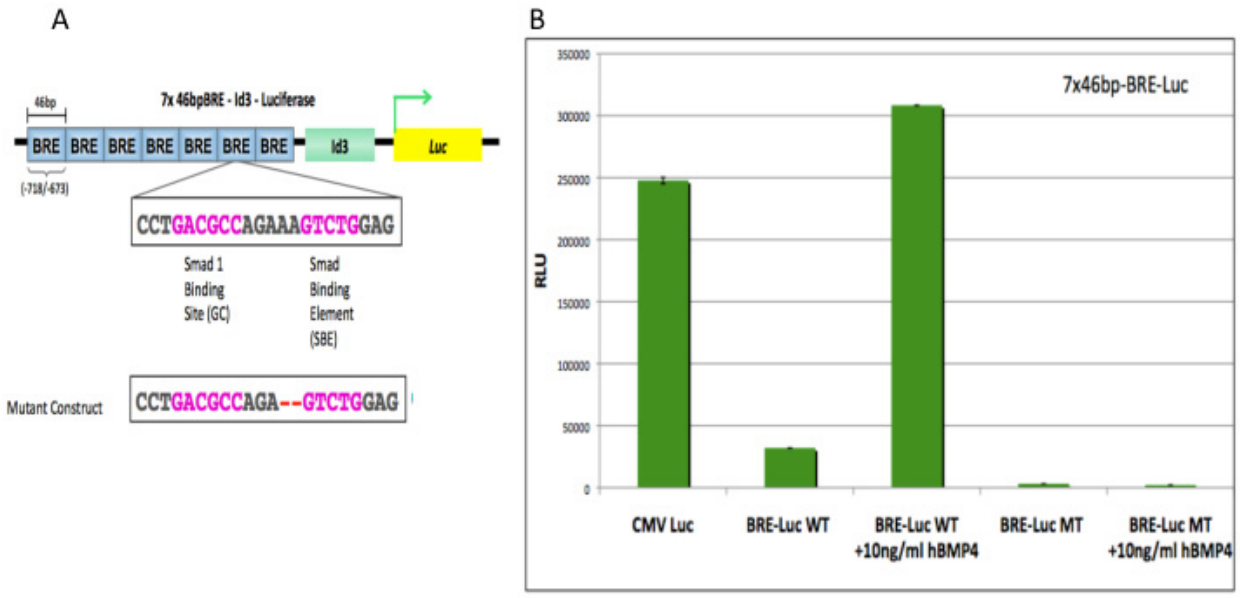


Figure 2.10: BRE is necessary to mediate the BMP response in mouse ES cells. A) The wild type BRE consists of a Smad1 binding site and a Smad4 binding site (SBE). The mutant BRE consists of a 2-nucleotide deletion in the 5-spacer region between the Smad1 and Smad 4 binding sites. B) Wild type and mutant BRE-luciferase (BRE-Luc) constructs were transfected into mouse ES cells for 24hrs and then treated with (-/+) BMP4 (10ng/ml) for 6hrs. Mutant BRE-Luc construct abolished luciferase activity. The wild type BRE-Luc construct in the presence of BMP4 showed a ~300 fold increase in luciferase activity.

Figure 2.11

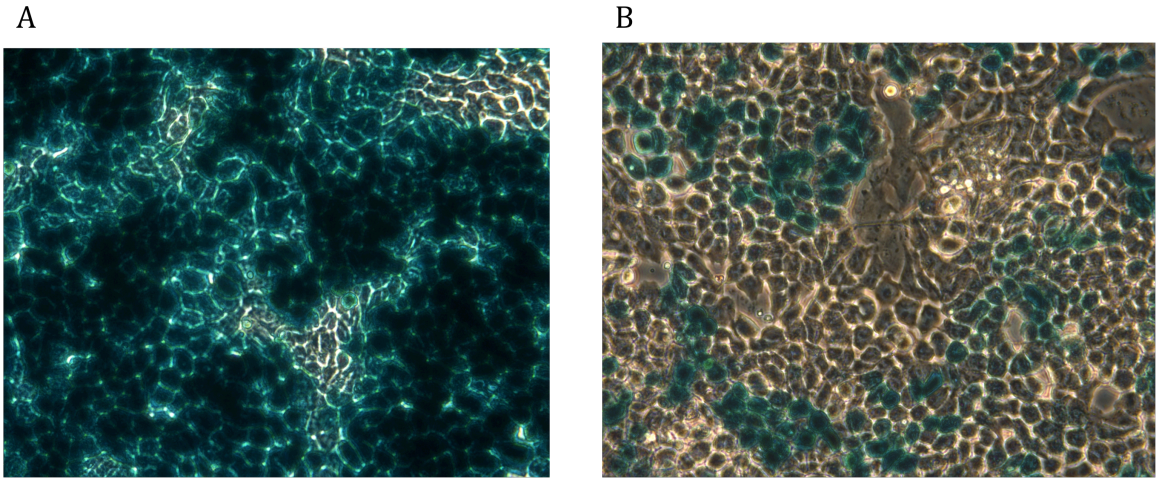


Figure 2.11: Inhibition of BMP4 signaling in BRE-gal mouse ES cells with Noggin. A) BMP4 Alone 10ng/ml for 24hrs B) BMP4 (10ng/ml) + Noggin (50ng/ml) for 24hrs following by X-gal staining. A reduction in X-gal staining is observed in the presence of Noggin.

Figure 2.12

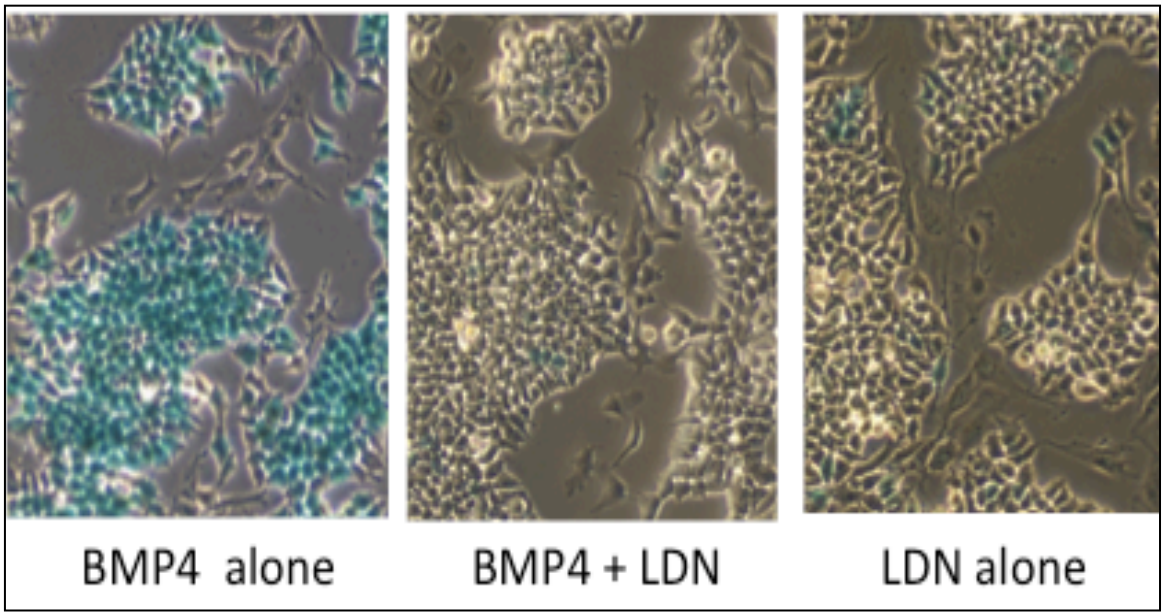


Figure 2.12: LDN193189, a small chemical inhibitor, inhibits BMP activity in BRE-gal mouse ES cells. LDN disrupts the Ser/Thr kinase activity of BMP Type I receptors. In the left panel, BMP4 alone stimulated LacZ expression in the BRE-gal ES cells. In the middle panel, BMP4 + LDN treatment inhibited BMP signaling in ES cells showing no observed X-gal staining. In the right panel, no X-gal staining was observed with LDN alone (control).

Figure 2.13:

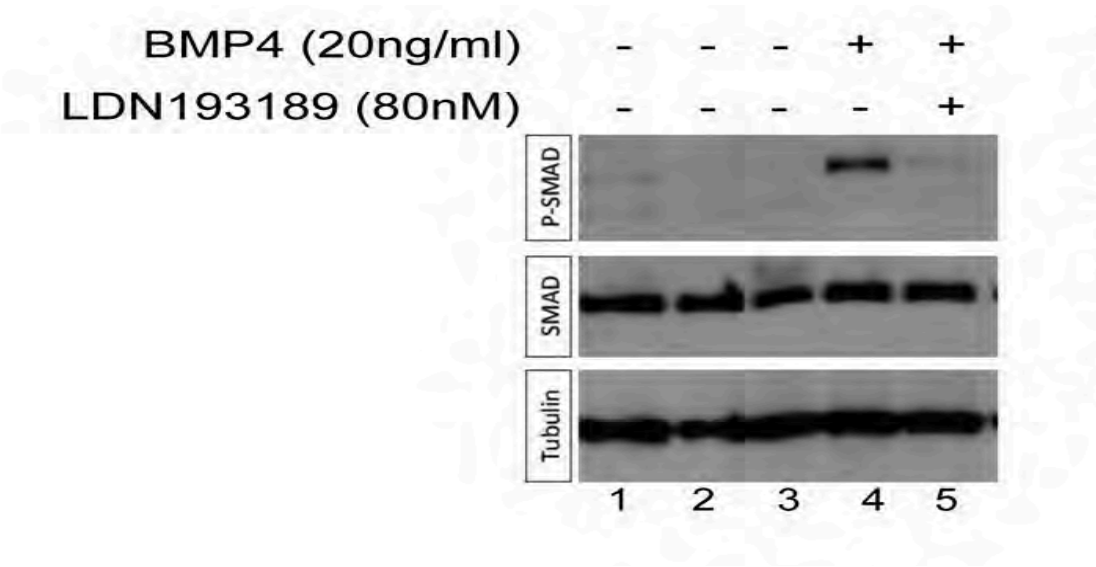


Figure 2.13: LDN193189 inhibits the phosphorylation of Smad1/5/8 in BRE-gal mouse ES cells. Western blot of phosphorylation of Smad1/5/8 increased upon BMP4 induction (lanes 4), which was blocked in the presence of LDN193189 (lane 5). Pan-Smad1/5/8 antibody was use to demonstrate that Smad levels were similar among all samples. Tubulin was used as our western blot loading control.

Chapter 2 References:

Alexander C, Zuniga E, Blitz IL, Wada N, Le Pabic P, Javidan Y, Zhang T, Cho KW, Crump JG, Schilling TF. Combinatorial roles for BMPs and Endothelin 1 in patterning the dorsal-ventral axis of the craniofacial skeleton. Development. 2011 Dec;138:5135-46.

Boergermann JH, Kopf J, Yu PB, Knaus P. Dorsomorphin and LDN-193189 inhibit BMP-mediated Smad, p38 and Akt signalling in C2C12 cells. Int J Biochem Cell Biol. 2010 Nov;42:1802-7.

Dani C, Chambers I, Johnstone S, Robertson M, Ebrahimi B, Saito M, Taga T, Li M, Burdon T, Nichols J, Smith A. Paracrine induction of stem cell renewal by LIF-deficient cells: a new ES cell regulatory pathway. Dev Biol. 1998 Nov 1;203:149-62.

Davey RE, Onishi K, Mahdavi A, Zandstra PW. LIF-mediated control of embryonic stem cell self-renewal emerges due to an autoregulatory loop. FASEB J. 2007 Jul;21:2020-32.

Davey RE, Zandstra PW. Spatial organization of embryonic stem cell responsiveness to autocrine gp130 ligands reveals an autoregulatory stem cell niche. Stem Cells. 2006 Nov;24:2538-48.

Fowler AV, Zabin I. The amino acid sequence of beta-galactosidase of Escherichia coli. Proc Natl Acad Sci U S A. 1977 Apr;74:1507-10.

Gao S, Steffen J, Laughon A. Dpp-responsive silencers are bound by a trimeric Mad-Medea complex. J Biol Chem. 2005 Oct 28;280:36158-64.

Hu ZW, Hoffman BB. Cycloheximide induces the alpha 1B adrenergic receptor gene by activation of transcription in DDT1 MF-2 smooth muscle cells. Mol Pharmacol. 1993 Dec;44:1105-12.

Javier AL, Doan LT, Luong M, Reyes de Mochel NS, Sun A, Monuki ES, Cho KW. Bmp indicator mice reveal dynamic regulation of transcriptional response. PLoS One. 2012;7:e42566.

Korchynskyi O, ten Dijke P. Identification and functional characterization of distinct critically important bone morphogenetic protein-specific response elements in the Id1 promoter. J Biol Chem. 2002 Feb 15;277:4883-91.

McKeehan W, Hardesty B. The mechanism of cycloheximide inhibition of protein synthesis in rabbit reticulocytes. Biochem Biophys Res Commun. 1969 Aug 15;36:625-30.

Monteiro RM, de Sousa Lopes SM, Bialecka M, de Boer S, Zwijsen A, Mummery CL. Real time monitoring of BMP Smads transcriptional activity during mouse development. Genesis. 2008 Jul;46:335-46.

Monteiro RM, de Sousa Lopes SM, Korchynskyi O, ten Dijke P, Mummery CL. Spatio-temporal activation of Smad1 and Smad5 in vivo: monitoring transcriptional activity of Smad proteins. J Cell Sci. 2004 Sep 15;117:4653-63.

Pyrowolakis G, Hartmann B, Müller B, Basler K, Affolter M. A simple molecular complex mediates widespread BMP-induced repression during *Drosophila* development. Dev Cell. 2004 Aug;7:229-40.

Sasano Y, Mizoguchi I, Takahashi I, Kagayama M, Saito T, Kuboki Y. BMPs induce endochondral ossification in rats when implanted ectopically within a carrier made of fibrous glass membrane. Anat Rec. 1997 Apr;247:472-8.

ten Dijke P, Yamashita H, Sampath TK, Reddi AH, Estevez M, Riddle DL, Ichijo H, Heldin CH, Miyazono K. Identification of type I receptors for osteogenic protein-1 and bone morphogenetic protein-4. J Biol Chem. 1994 Jun 24;269:16985-8.

Urist MR. Bone: formation by autoinduction. Science. 1965 Nov 12;150(3698):893-9.
Viswanathan S, Davey RE, Cheng D, Raghu RC, Lauffenburger DA, Zandstra PW. Clonal evolution of stem and differentiated cells can be predicted by integrating cell-intrinsic and -extrinsic parameters. Biotechnol Appl Biochem. 2005 Oct;42:119-31.

Viswanathan S, Benatar T, Rose-John S, Lauffenburger DA, Zandstra PW. Ligand/receptor signaling threshold (LIST) model accounts for gp130-mediated embryonic stem cell self-renewal responses to LIF and HIL-6. Stem Cells. 2002;20:119-38.

von Bubnoff A, Peiffer DA, Blitz IL, Hayata T, Ogata S, Zeng Q, Trunnell M, Cho KW. Phylogenetic footprinting and genome scanning identify vertebrate BMP response elements and new target genes. Dev Biol. 2005 May 15;281:210-26.

Yao LC, Blitz IL, Peiffer DA, Phin S, Wang Y, Ogata S, Cho KW, Arora K, Warrior R. Schnurri transcription factors from *Drosophila* and vertebrates can mediate Bmp signaling through a phylogenetically conserved mechanism. Development. 2006 Oct;133:4025-34.

Yao LC, Phin S, Cho J, Rushlow C, Arora K, Warrior R. Multiple modular promoter elements drive graded brinker expression in response to the Dpp morphogen gradient. Development. 2008 Jun;135:2183-92.

Ying QL, Nichols J, Chambers I, Smith A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell. 2003 Oct 31;115:281-92.

Ying QL, Stavridis M, Griffiths D, Li M, Smith A. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. Nat Biotechnol. 2003 Feb;21:183-6.

Yu PB, Deng DY, Lai CS, Hong CC, Cuny GD, Bouxsein ML, Hong DW, McManus PM, Katagiri T, Sachidanandan C, Kamiya N, Fukuda T, Mishina Y, Peterson RT, Bloch KD. BMP type I receptor inhibition reduces heterotopic [corrected] ossification. Nat Med. 2008 Dec;14:1363-9.

Zimmerman LB, De Jesús-Escobar JM, Harland RM. The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4. Cell. 1996 Aug 23;86:599-606.

Chapter 3

BREs are important for modulating a subset of BMP target gene.

Comprehensive gene expression profiling was first made possible with the introduction of microarrays, which allowed researchers to survey thousands of genes at once in many samples. However, microarrays had some shortcomings: a) The genes being studied were restricted to only known genes, and b) The sensitivity and dynamic range of microarrays were limited. In the last several years, massive parallel RNA-sequencing has become the choice for researchers to study gene expression profiling. It overcomes the problems attributed to microarrays by sequencing RNA directly (Ozsolak et al., 2009) or from cDNA after reverse-transcription (Cloonan et al., 2008; Mortazavi et al., 2008). Quantitation of gene expression levels is based on the numbers of reads (or "hits") counted for a particular gene among millions of reads generated. RNA-Seq has a dynamic range (10^6) and overcomes the sensitivity issues of microarrays.

To understand how BMP signaling affects gene expression in mouse ES cells and to identify new target genes that affect the property of ES cell, one must characterize the primary response genes. In this study, I took our BRE-gal mouse ES cells and identified direct target genes regulating BMP signaling activity. The advantage of using the BRE-gal mouse ES cells generated in our lab is that our line is responsive to BMPs and we can monitor the effectiveness of BMP treatment.

Utilizing RNA-Seq, I identified a set of differentially expressed (DE) target genes upon BMP stimulation. I confirmed that these DE genes are differentially expressed using quantitative RT-PCR. Next, I determined whether these targets are direct BMP targets using a protein synthesis inhibitor, cycloheximide. Through this approach, I was able to

identify bona fide BMP targets. Several labs have identified BMP targets in mouse using different approaches: 1) gene expression arrays in mouse ES cells treated with BMP4 (R1) (Li et al., 2012), and 2) Smad1/5/8 chromatin-immunoprecipitation sequencing (Fei et al., 2010).

To better understand the role of BMP signaling in ES cell self-renewal and cell fate determination, one must identify all the molecular players. In this present study, I report on a subset of BMP targets regulated by BRE and some that are not.

Materials and Methods

RNA Isolation, Quantitative RT-PCR, and RNA-Seq Library Generation:

RNA was isolated after treatment of moue ES or BRE-gal ES cells with 10ng/ml hBMP4 (R&D Systems, catalog #314-BP) using Trizol (Life Technologies, catalog #15596-026). 1μg of Total RNA was used for reverse transcription using MMLV Reverse Transcriptase (Invitrogen, catalog #28025-013). Quantitative PCR was performed using SYBR Green (Roche, catalog #04-707-516-001). Primers for PCR are listed in Table 3.1. For RNA-Seq, total RNA profiles were recorded using a Bioanalyzer 2100 (Agilent). RNA quality was determined by the RNA Integrity Number (RIN) value. RNA samples showing RIN values greater than 9.0 were used for library generation. RNA sequencing library was prepared using a Bioo Scientific Kit (Cat#5143-01). 1μg of Total RNA (1μg) was used to generate sequencing libraries. RNA-Seq transcriptome sequencing was performed by UCI Genomics High-Throughput Facility using the Illumina Hi-Seq 2000 platform.

Table 3.1: RNA-Seq Primers

Mouse Gene	Forward Primer	Reverse Primer
Skil	GTACATGTGTGACAAAGTGGTTGC	GACTGTTTGGAGGTGGACTTACTT
Micall2	GGAACTTAAGCCACCATCTACCT	GTGAGAGCTTTGCTAGTCACAGAAG
Hsp5a	GGATAAGAGAGAGGGAGAGAAGAAC	ACCCAGGTGAGTATCTCCATTAGT
Id1	GACATGAACGGCTGCTACTCAC	GACTTCAGACTCCGAGTTCAGC
Id2	GTGAGGTCCGTTAGGAAAAACAG	GTCGTTCATGTTGTAGAGCAGACT
Id3	CTGTCGGAACGTAGCCTGG	GTGGTTCATGTCGTCCAAGAG
Id4	AGTGCGATATGAACGACTGCTAC	AGGATGTAGTCGATAACGTGCTG
Cxcl3	ATCTCCTCAGGTGTCCACTACTTC	GGTTTATCTCCTCAGACCCTAGAGA
Pcdh8	CTTGCAGCTATTTAGTCTCTGCTG	GACACTTTCATGTGCAGGTCTTCT
Tchh	AGTTCGTAGTAACAGGCTCTTCACC	AGCTCTCTCTTCTCTTCCTCTTCGT
Sumo3	TAG TCA CCT GAC ACT GAC CTA CAT C	GTA AAC ACA ACA GAG CCA ACA GAC
Grem1	CCAGTGCAACTCCTTCTACATCC	GTAGCTCAGGACAGTTGAGTGTGAC
Msx2	CTT ACA GAG ACC CAG GTC AAA ATC	GAG TTG ATA GGG AAG GGC AGA CT
Coch	CTC AGA CAT TGG AGC CAA GAT AG	CAG CTA GGA CGT TCT CTT TGG TAT
CTGF	CAC TCT GCC AGT GGA GTT CAA AT	TGT CTC CGT ACA TCT TCC TGT AGT
Gata2	GTCAAGTACCAAGTGTCACTCTCC	CATAGGAGGGATAGGTGGGTATC
R3hdml	GTCAGAGGATCTCTGCCAAGTT	CAGTTAGAGCTGCTGCCTAGAAG

Evx1	CAGACTTCACTTGTGCCTCCAC	CACCTCCTCTCTCTGGTCCAAC
Sox7	TACTCACCGGAGTTCACACCTAGT	GACATCCAGAAACAGAGGACATC
Stard4	AGTGTTGAGTGGAGTGAGACGA	GACTCTGGCTTGGACTGTCTTT
Itgb1	GGAGAAAACTGTGATGCCGTAT	TTGTAGCTAAATGGGCTGGTG
Dusp5	ATTACGTCAAGCAGAGGAGGAG	AGGGCAGGATCTCAGATTCATA

RNA-Seq Computational Pipeline:

RNA-Seq data analysis was accomplished in two sequential steps: (1) map all raw reads to the mouse reference transcriptome (mm9/RefGene) as the gene model, and (2) count the read fragments mapped to each individual gene and quantify expression in terms of RPKM (Reads Per Kilobase per Million mapped reads). Raw RNA-seq reads were mapped using the Bowtie 0.12.8 software allowing a two-read mismatch in the alignment (-m 1) to the reference transcriptome. Cufflinks 2.0.2 was used for read quantitation and differential analysis (Trapnell et al 2012). Data were then clustered using Cluster3.0 (http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm) to generate a map of genes according to expression differences between two samples. Cluster3.0 parameter: Cluster "Genes", centered median, and centroid linkage. Cluster output (e.g. javatree.cdt) was input into Java Treeview, which allows visualization of the hierarchical clustering in a form of a heatmap diagram.

Gene Ontology Analysis:

Gene Ontology (GO) analysis was applied using the PANTHER database (http://www.pantherdb.org/) to interpret the molecular and biological processes that were enriched in the top 200 upregulated genes in our RNA-Seq analysis.

MotifMap Analysis:

MotifMap contains comprehensive maps of candidate regulatory elements encoded in the genome (http://motifmap.ics.uci.edu/) for various model organisms. The Motif Search tool in MotifMap allowed us to perform genome-wide search for putative transcription factor binding sites. The BRE motif (5'-GRNGNC-N5-GTCTG-3) was searched using the Motif Search tool. All putative binding sites were selected based on a normalized log-odd (NLOD) score of greater than 0.80. The NLOD is a probability score indicating the likelihood of BRE-like motifs to be present within the genome (Daily et al., 2011). I identified 1683 genes containing putative BRE sites located within 10kb to the closest gene.

Cycloheximide Experiment:

BRE-gal mouse ES cells were pre-treated with cycloheximide (C4859, Sigma) at a concentration of 20ug/ml for 2hrs. Cells were then treated with fresh CHX alone or CHX together with BMP4 (10ng/ml) for additional 4hrs. As control, cells were also treated with regular E14 media only. Total RNA was isolated and cDNA synthesized as described before. Quantitative PCR was performed using SYBR Green (Roche, catalog #04-707-516-001) and primer above in Table 3.1.

PhyloP Analysis:

The conservation of each nucleotide within each predicted BRE of Id1-4, Msx2, Gata2, and Grem1 was determined using the conservation track of the UCSC Genome Browser. The measurement of evolutionary conservation was done using PhyloP

(phylogenetic *P*-values). Multiple alignments using PhyloP30wayPlacental for 30 different vertebrates were used to determine nucleotide conservation. Conservation of each nucleotide of the BRE motif was displayed in a heatmap.

Mulan Analysis:

Multiple sequence alignment analysis is a powerful approach to localize conserved coding and noncoding genomic elements. Mulan (http://mulan.dcode.org/) is capable of comparing both closely and distantly related genomes to identify conserved elements over a broad range of evolutionary time. To identify potential conserved elements within the promoter, intron, and downstream regions of a given gene, genomic sequences 10kb upstream of ATG, the entire gene body including introns and exons, and 5kb downstream of a gene were aligned between the mouse (mm9) and human (hg18) locus.

***R3hdml* and *Dusp5* Promoter Cloning:**

10kb upstream ATG, the first intron, and +5kb downstream of the stop codon for the *R3hdml* and *Dusp5* genes were cloned into the unique BamHI and HindIII sites located within the multiple cloning site of Id3 (-201)-pOLuc. Primers to each cloned region were designed using Primer3 and their sequences are listed in Table 3.2. Upstream primers contained a BamHI and downstream primers contained a HindIII restriction site to assist the ligation of each PCR amplified fragment into the 201Id3-pOLuc plasmid. Genomic amplicons were generated using Touchdown PCR (Cycle1: 98^{o}C for 30secs, Cycle 2: 98^{o}C for 10secs, 63^{o}C for 30sec[Decrease temp after cycle 1 by 0.5^{o}C every 1 cycle], 72^{o}C for X min for 13 cycles; Cycle 3: 98^{o}C for 10secs, 55^{o}C for 30sec, 72^{o}C for X min for 32 cycles; Cycle 4: 72^{o}C for 10mins) and Pfx Platinum DNA polymerase (Life Technologies) or Phusion Polymerase (NEB) according to manufacture's guidelines. Amplicons (genomic

fragment insert) and the Id3 (-201) pOLuc plasmid (vector) were then digested with BamHI and HindIII. The vector was phosphatase using calf intestinal phosphatase (NEB) after restriction enzyme digestion. DNA fragments were gel extracted and ligated using T4 DNA ligase (Life Technologies). Ligated products were transformed into DH5α cells and positive colonies were grown for DNA isolation. Isolated DNA was subjected to DNA sequencing.

Table 3.2: Cloning Primers for *Dusp5* and *R3hdml*

R3hdml Upstream:

Cloning Region	Forward	Reverse
R3hdml Clone1 (-2560bp/0kb)	GCATCTAGA CCT TAT CTA GAG CAA AGA AGG CCT TAG	TCTGGATCC AGC TGT CTG GGA TGG TGT ACC TGT CTG
R3hdml Clone 2 (-5500bp/-3400bp)	GCATCTAGA ATA CTG CTT ACT AGC TTG CTC CTC ATG GCT	TCTGGATCC GTT TTC TGA CCT ACT GCT GAG TGA G
R3hdml Clone1b (-3400bp) For	GCATCTAGA CTC ACT CAG CAG TAG GTC AGA AAA C	
R3hdml Clone 3 (-7899bp/-5500bp)	CTA TAG GAT CCC AGG GGA TTC ATT TGA A	TCTGGATCC AAG CCA TGA GGA GCA AGC TAG TAA GCA GTA T
R3hdml Clone 4 (-10kb/-7899bp)	GCATCTAGA TTC CCA AAC ATC TCC CTA GCC TGG CCT	TTC AAA TGA ATC CCC TGG GAT CCT ATA G
R3hdml (-2560bp) Rev		TCTGGATCC TAA GGC TGA TCT ATT TCC TCG AGG CAG
R3hdml (-2920bp) Rev		TCTGGATCC AAG TGG AAC CAC AGG GCT GGA GAG A
R3hdml (-2920bp)	CAC TTT CTA GAT GAA GAA ACT GAG ACA CA	GCATCTAGA CAC TCC CAC TCT CAG ATA AAG CCT CCA T

Dusp5 Upstream:

Cloning Region	Forward	Reverse
Dusp5 Clone 1 (-2200bp/0bp)	GCATCTAGA T GTG TCT CCA TCA CAT CCT TGT GTC	TCTGGATCC GCC ACC CAG CGG CCC GAC GCT GCG
Dusp5 Clone 2 (-6940bp/-3000bp)	GAT GGG ATC CGA TTG TAA ATT	TCTGGATCC GAG ACA CAA GGA TGT GAT GGA GAC ACA
Dusp5 Clone 3 (-7599bp /-6940bp)	GCATCTAGA ATT AAA GTT GTG CGC CAC CAC TGC	AAT TTA CAA TCG GAT CCC ATC AAT
Dusp5 Clone (-10kb/7599bp)	GCATCTAGA AAG TAC ACT GTA GCT GTC CTC AGA CAC	TCTGGATCC GTG TCT GAG GAC AGC TAC AGT GTA CTT
Dusp5 Clone (-7239bp)	GCATCTAGA CTT CCC AAG TAC TAG GAT TAA AGG TGC ACG	TCTGGATCC CGT GCA CCT TTA ATC CTA GTA CTT GGG AAG
Dusp5 Clone 4 For	GCATCTAGA CCA GTC TAG GAA ATA AGA CAG GAG AGC CAA -	
Dusp5 Clone (-2200bp)	GCATCTAGA CAG GGT TAC TAG GCT GTG CCT CTA TGG TAC	TCTGGATCC GTA CCA TAG AGG CAC AGC CTA GTA ACC CTG
Dusp5 (-2610bp) Rev		GGA TGG AGG CTT TAT CTG AGA GTG GGA GTG
Dusp5 Clone (-9759bp) For	GCATCTAGA GTG AGC CTC AGT CTC TAA ATC TGT TAC AGT CTC TAA ATC TGT TAC	
Dusp5 (-9,220bp) For	TCA GCA TTC TAG AAC CTC CCT GCC CCC AC	

R3hdml Downstream:

Cloning Region	Forward	Reverse
R3hdml Clone (+0bp) For	CG GCC GC TCTAGA ACT GGC CCA ATG TTA TGT CAC CCC TAG C	
R3hdml Clone (+541bp)	CG GCC GC TCTAGA CTA AAG GTC ACC CTC AGC TAT TAG AG	C TTT CTA GGATCC CTC TAA TAG CTG AGG GTG ACC TTT AG
R3hdml Clone (+1501bp)	CG GCC GC TCTAGA ACT ACT CTA CTT GGT GGA AGA GGA AG	C TTT CTA GGATCC CTT CCT CTT CCA CCA AGT AGA GTA GT
R3hdml Clone (+2161bp)	CG GCC GC TCTAGA ACC TAC AGA AGC CAG AGA GGA TAT C	C TTT CTA GGATCC GAT ATC CTC TCT GGC TTC TGT AGG T
R3hdml Clone (+3301kb) Rev		C TTT CTA GGATCC GTG TGA GAG AAG GTT TAG ACC AGA C
R3hdml Clone (+ 4261bp)	CG GCC GC TCTAGA GGT TCC CTA AGT GAC TGG TTA CTG TT	C TTT CTA GGATCC AAC AGT AAC CAG TCA CTT AGG GAA CC
R3hdml Clone (+ 6181bp)	CG GCC GC TCTAGA GCT GGT ATC TAA CAT AGA GCT GCT G	C TTT CTA GGATCC CAG CAG CTC TAT GTT AGA TAC CAG C

Dusp5 Downstream:

Cloning Region	Forward	Reverse
Dusp5 Clone (+0kb) For	CG GCC GC TCTAGA GAC CGT GGT CGG TGG CTA CCA GCG AGT	
Dusp5 Clone (+200bp) For	CG GCC GC TCTAGA AGT ACG TGC TGA CTG CTG TAC TTC CGG	
Dusp5 Clone (+1681bp)	CG GCC GC TCTAGA GAG GAA AGG CTG ACT AGA GAT AGT TG	C TTT CTA GGATCC CAA CTA TCT CTA GTC AGC CTT TCC TC
Dusp5 Clone (+3kb) Rev		CG GCC GC TCTAGA GAT AGC TCT GGG ACT CTA ACT CCA TTTG
Dusp5 Clone (+2600bp)	CG GCC GC TCTAGA GAT AGC TCT GGG ACT CTA ACT CCA TTTG	C TTT CTA GGATCC CAA ATG GAG TTA GAG TCC CAG AGC TAT C
Dusp5 Clone (+3600bp) Rev		C TTT CTA GGATCC AGT ATA GCT AAA AGC CCA CCC CTA
Dusp5 Clone (+5900bp) For	CG GCC GC TCTAGA CCA CCT CTG TAC TCT GTG AGA ATG AAA G	
Dusp5 Clone (+6076bp)	CG GCC GC TCTAGA GTA GAA GAG AAG CGT TCC TAG ACC T	C TTT CTA GGATCC AGG TCT AGG AAC GCT TCT CTT CTA C -3
Dusp5 Clone (+8701bp) For	CG GCC GC TCTAGA CAG CCT CCT AAG TAG CTG GAA TTA AAG	
Dusp5 Clone (+9220bp) Rev		C TTT CTA GGATCC GTA CAA GAT TGA CTA CAG GGG AGA G

R3hdml Intron1:

Cloning Region	Forward	Reverse
R3hdml Intron (+0bp) For	CG GCC GC TCTAGA GTG AGT CCC AAA CTT TCC CTG TGC A	
R3hdml Intron (+181bp) For	CG GCC GC TCTAGA GGA ATA CTA TCT GTA GAG GGC CTT G	
R3hdml Intron (+1200bp)	CG GCC GC TCTAGA GTA TAG AAC ACT GGG ATG GTG TCT C	C TTT CTA GGATCC GAG ACA CCA TCC CAG TGT TCT ATA C
R3hdml Intron (+2341bp)	CG GCC GC TCTAGA GTC TAC ATA TCT GGG TGT ACA GGG A	C TTT CTA GGATCC TCC CTG TAC ACC CAG ATA TGT AGA C
R3hdml Intron (+3000bp) Rev		C TTT CTA GGATCC CTG ATG AAG GAC AGA TCA TGA ATT T

Dusp5 Intron1:

Cloning Region	Forward	Reverse
Dusp5 Intron (+0bp) For	CG GCC GC TCTAGA GTG AGT TCG TCC GGG AGC AGA CTG C	
Dusp5 Intron Clone (+4000bp) Rev		C TTT CTA GGATCC CTT CAG AAA TAA AGG GAT AAA CAA A
Dusp5 Intron Clone (+781bp)	CG GCC GC TCTAGA GAC TTA TTT CAC CTG CTG TCG AG ATA	C TTT CTA GGATCC TAT CTC GAC AGC AGG TGA AAT AAG TC
Dusp5 Clone (+1741bp)	CG GCC GC TCTAGA CTT ATA CTT CCT ACA GTT CCG TCT CC	C TTT CTA GGATCC GGA GAC GGA ACT GTA GGA AGT ATA AG
Dusp5 Clone (+2701bp)	CG GCC GC TCTAGA GTA ACA TGA CAG CAA GTG AGT CCT	C TTT CTA GGATCC AGG ACT CAC TTG CTG TCA TGT TAC

Luciferase Assay:

Various reporter gene constructs were transfected (Lipofectamine, Life Technologies) into wild type mouse ES cells for 24hrs and then stimulated for 8hrs with BMP4. After treatment, cells were lysed and subject to luciferase assay as follows. Cell lysate (50μl) was mixed with 100μl of luciferase assay reagent (20mM Tricine. NaOH pH7.8, 1mM magnesium carbonate, 0.1mM EDTA, 5mM DTT, 2.7mM Coenzyme A, 0.5mM ATP, 0.5mM Na-Luciferin). Reaction mixture was measured using a luminometer (Analytical Luminescence Lab, Moonlight 2010). Glow luminescence was measured for 10sec to obtain RLU (Renilla luciferase unit) values. RLU values were normalized by performing a β-galactosidase/ONPG assay on each sample.

Results and Discussion:

Chapter 3.1: Gene expression profile of BRE-gal ES cells after BMP4 stimulation.

To identify direct BMP4 target genes, RNA was isolated from BRE-gal cells from untreated cells and cells treated with BMP4 for 4hrs to prepare RNA-Seq libraries (Figure 3.1). Both samples were sequenced following library generation. The 4hrs time point was chosen because we know immediate early genes (primary response genes) such as the *Id3* gene are induced within 4hrs of BMP4 treatment.

The computational pipeline for RNA-Seq quantitation is shown in Figure 3.2. RNA samples isolated from each treatment time point were collected to determine its RNA integrity value (RIN). Both samples had a RIN value above 9.7. Sequencing of samples was done using 50bp single-end reads. For the untreated/control sample, we obtained 21 million reads and for the 4hrs BMP4 treatment, we obtained 15.6 million reads. Following

mapping of the sequencing reads to the mouse transcriptome (mm9), we obtained 9.8 (46.6%) million mapped reads for the untreated sample and 8.3 (53%) million mapped reads for the 4hrs BMP4 treated sample. RNA-Seq quantitation was done using Bowtie mapping and Cufflink software (Langbead et al., 2009; Trapnell et al., 2012). In Figure 3.3, the correlation of RPKM (Reads Per Kilobase of transcript per Million mapped reads, a measurement of transcript levels (Mortazavi et al., 2008; Trapnell et al., 2012)) values in the control (untreated ES) sample and BMP4 treated sample show genes both upregulated and downregulated in response to BMP4 treatment. From the results, 1538 genes were determined to be upregulated 2 fold or more, and 615 genes were downregulated 2 fold or more. A hierarchical cluster diagram (Figure 3.4) shows that a number of genes were upregulated or downregulated upon BMP4 treatment. Our analysis focused mainly on upregulated genes at this time. Of the 1538 genes, 22 genes were validated with quantitative qPCR to be upregulated from a list of the top 50 upregulated genes (Figure 3.5, 3.6 and 3.10). Some of the highly induced genes include bona-fide BMP target genes such as the inhibitor of differentiation genes (*Id1-4*) (Hollnagel A et al., 1999), *Msx2* (Brugger et al., 2004), and *Gata2* (Friedle et al., 2002), thus confirming that our RNA-Seq libraries were good.

To identify novel BMP targets, the top 200 induced genes were selected for the PANTHER (Protein Analysis Through Evolutionary Relationships) software (www.pantherdb.org) to determine whether any known molecular and biological functions can be assigned to these top 200 genes. PANTHER is a curated biological database of gene/protein families that can be used to identify gene function (Mi et al., 2013). This software is useful to classify genes and their functions from high-throughput sequencing

analysis. Figure 3.7A shows the molecular functions of the top 200 genes and most were be involved in "binding" and "nucleic acid binding transcription factor activities". This finding is not surprising since ES cells tend to utilize combinatorial binding of transcriptional factors to regulate the multitude of processes from self-renewal to priming cells for differentiation. Additionally, Figure 3.7B shows the biological functions of these top 200 genes. These genes were to be mainly involved in developmental processes as well as cellular and metabolic processes.

Chapter 3.2: A bioinformatics approach to identifying BREs.

From a series of molecular analysis combined together with RNA-Seq, I was able to identify a set of genes that were upregulated upon BMP4 stimulation and validated independently through quantitative qPCR analysis. Next, I investigated whether these BMP4 targets contain the BRE motif. We utilized a software program called MotifMap to search for the presence of the BRE (5'-GRNGNC-N5-GTCTG-3') motif near the genes. Motif-Map uses transcription factor databases such as Jaspar and Transfac (Daily et al., 2011) and apply their respective binding matrices to search for binding sites within a reference genome (Daily et al., 2011). Motif-Map produces a Normalized Log-Odds (NLOD) score derived from the position weight matrix of a transcription factor or a specified motif. Binding sites of each transcription factor motif during the scanning of the genome will generate a NLOD score to predict the likelihood a given transcription factor binding site is present. The higher the NLOD score, the more likely the transcription factor motif is present. In Figure 3.8, we observed that most of the known BMP targets such as the Id genes, *Gata2*, *Msx2* and *Grem1*, contained a consensus BRE motif. Interestingly, the BRE

motif is not present in *Sumo3*, *Dusp5*, and *R3hdml* promoters. The parameters of the search for BRE were 5kb upstream of the transcription start site (TSS) and +1kb downstream of the stop codon. A NLOD score of greater than 0.95 is considered significant (FDR <0.35). The lower NLOD cutoff in our analysis was 0.80, which was determined empirically by looking at the predicted BREs manually within the mouse genome. NLOD score lower than 0.80 did not show any predicted motifs resembling the BRE while most predicted BRE with a NLOD score greater than 0.95 matched the BRE motif. It should be noted that the bioinformatics identification of the BRE in these genes does not demonstrate that BMP responsiveness is mediated via BRE. Predicted BREs must be tested to determine experimentally if they are present and functional.

To extend our analysis, we searched the entire mouse reference genome (mm9) for BREs using the MotifMap software program. 1683 genes had a predicted BREs detected using the mm9 genome. The cutoff of the NLOD score was >0.80 and BREs were searched within 10kb upstream and downstream of the TSS and stop codon. However, if the NLOD cutoff score > 0.95, only 46 genes remained to have a BRE. These 1683 genes containing predicted BRE sites were then intersected with the 22 validated BMP target genes from our RNA-Seq analysis. Only 7 genes out of the 22 validated genes contained a predicted BRE (Figure 3.9). Using PhyloP conservation track from the UCSC Browser, a heatmap was generated to show the degree of conservation of the predicted BREs among different mammals (Mouse, Rat, Human, Orangutan, Dog, Horse, Opposum) in these 7 genes. The BRE motif (+ strand: 5'-GRCGNC-N5-GTCTG-3'; - strand: 5'- CAGAC-N5-GNCGYC -3') shows that most of the predicted BREs had a high degree of conservation among mammals (Figure 3.10). *Id1* and *Id4* have the highest degree of conservation not only in mammals but also to

the stickleback fish. Using MotifMap, we were able to demonstrate that validated known direct BMP targets (*Id1*, *Id2*, *Id3*, *Id4*, *Msx2*, *Gata2*, and *Grem1*) did contain BREs. The finding is consistent with the hypothesis that BRE motif is used to mediate a subset of BMP signaling. Since we found BRE only among 7 out of 22 BMP-regulated target genes, it supports the notion that only a subset of BMP target genes is transcriptionally induced via BRE. I propose that the other 15 BMP target genes are regulated through an alternate cis-regulatory element and that most likely do not require Schnurri. However, I cannot exclude the possibility that BREs may be present outside of the DNA sequences/region we examined.

Chapter 3.3: Identification of Novel BMP targets.

The identification of a set of BMP target genes from our RNA-Seq analysis allowed us to gather information on early genes that are upregulated in response to BMP4 induction. Before we address which cis-regulatory elements may be involved in regulating the non-BRE validated BMP target genes, we decided to first ascertain whether those non-BRE validated BMP target genes (~15 genes) are direct BMP targets. A cycloheximide experiment was performed to determine if the validated BMP target genes were direct BMP targets. Two genes, *R3hdml* and *Dusp5*, were shown to be BMP direct targets (Figure 3.11) based on cycloheximide block experiments. Scanning of the promoter -10kb upstream, +5kb downstream of the stop codon, and introns of the *Dusp5* and *R3hdml* genes showed no predicted BRE in both these target genes (Mochiview*:* www.johnsonlab.ucsf.edu/mochi.html*)*. The result indicates additional cis-regulatory elements other than the BRE mediating may be present to respond to BMP signaling.

Since both *R3hdml* and *Dusp5* appear to function independently of the BRE, I wished to determine which regulatory module could be used to mediate BMP signaling. To address this question, I cloned the various DNA fragments covering the promoter, UTRs, and the first intron of these genes. I used a local multiple alignment software, Mulan (http://mulan.dcode.org/), to identify transcription factor binding sites that are evolutionarily conserved across multiple species. The program allowed comparison of both closely and distantly related genomes to identify conserved elements over a broad range of evolutionary time (Loots et al., 2007). Sequences covering 10kb of the promoter upstream the transcription start site (TSS), the entire gene body, and 5kb downstream of the gene's stop codon were analyzed by Mulan. Evolutionarily conserved regions (ECRs) were identified by comparing the mouse and human genome. Results from Mulan for the *R3hdml* and *Dusp5* genes are shown in Figure 3.12. The promoters (-10kb), downstream regions (+5kb), and the first introns of these two genes were then cloned into a minimal *Id3* promoter-luciferase construct (201-*X.Id3*-pOLuc) and transfected into wild type E14 mouse ES cells. Transfection was followed by a luciferase assay. Cloned genomic fragments (Figure 3.13) were assayed to determine whether luciferase induction occurred in response to BMP4 stimulation. To date, none of the constructs cloned and tested have shown luciferase induction (Figure 3.14). Since my collection of subclones only encompassed ~60% of the proposed genomic region for *Dusp5* and *R3hdml*, it is entirely possible that I failed to clone the responsible cis-regulatory element. Alternatively, the readout of the limited regulatory region present in the cloned fragments might be too weak to provide a strong response. It may be necessary to transfect a constitutively active BMP receptor plasmid (Varley et al., 1998) in our assay to amplify the luciferase signal of our

cloned fragment. Constitutive active BMP Type I receptor activates the BMP-Smad pathway in a ligand-independent manner and can provide a stronger BMP signaling to induce luciferase signal in our clones.

If *R3hdml* and *Dusp5* are indeed involved in the BMP signaling cascade, it would be a novel finding linking *R3hdml* to BMP target signaling and also implicating BMP4-mediated *Dusp5* expression in mouse ES cell self-renewal. When we examined the expression of *Dusp5* in *Xenopus* (Xenbase: http://www.xenbase.org) and BMP4 ligand expression (Javier et al., 2012) in mouse, we noted that expression of *Dusp5* and BMP4 overlap. For example, Anna Javier showed that BMP4 expression is present in the neural tube, pharyngeal arches, forebrain, and otic vesicle (Javier et al., 2012). Similarly, *Dusp5* expression in *Xenopus* was observed in the same structures (http://www.xenbase.org/gene/showgene.do?method=displayGeneSummary&geneId=957571). Additionally, in zebrafish, *Dusp5* expression is present in the ventral mesoderm (www.zfin.org) and BMP4 is also expressed in the same region (Selever et al., 2004). However, the expression data for *R3hdml* was not in Xenbase and ZFin. Based on these observations, we propose that *Dusp5* is a highly conserved direct BMP4 target since I detected similar expression patterns among mouse, *Xenopus*, and zebrafish. It would be interesting to obeserve the knockdown or overexpression phenotypes of these genes.

Conclusion:

We have successfully identified direct BMP4 targets (*Dusp5* and *R3hdml)* using RNA-Seq. The dual-specificity protein phosphatase 5 (*Dusp5*, also known as HVH3 and B23) belongs to the Dusp family (also known as mitogen activated protein kinases (MAPK)

phosphatase or MKPs). Dusp typically dephosphorylates phosphothreonine and phosphotyrosine residues on MAPKs. *Dusp5* can be induced by both heat shock and serum (Ishibashi et. al., 1994; Kwak et al., 1995). Unlike other Dusps, *Dusp5* can inactivate extracelluar signal-regulated kinase 2 (Erk2) and is responsible for the nuclear translocation of Erk2 *in vivo* (Mandl et al., 2005) to negatively regulate Erk signaling. *Dusp5* transcripts increases from blastula stage 8.5-9 and functions in conjunction with *Dusp1* to negatively regulate FGF signaling (Branney et al., 2009). Additionally, *Dusp5* is important for angioblasts formation and is essential for vascular development in zebrafish embryos (Pramanik et al., 2008). In mouse embryonic stem cells, *Dusp5* has been shown to enhance self-renewal and *Dusp5* knockdown reduces the expression of key pluripotentcy factors Oct4 and Nanog (Chen et al., 2011). A recent study has implicated Dusp9, an Erk specific phosphatase, as a critical mediator of BMP4 signaling to control the ERK signaling pathway (Li et al., 2012). ERKs are intracellular mediators of MAPKs that are involved in growth and differentiation of various cell types (Kolch et al., 2005; Silva et al., 2008; Ying et al., 2008). Dusp9 overexpression was shown to inhibit ERK target genes (Li et al., 2012). Interestingly, Li et al., 2012 showed that *Dusp5* was not upregulated after BMP4 treatment. In our study, *Dusp5* is not only upregulated but is also a direct BMP target based on our cycloheximide block experiments using mouse ES cells. This discrepancy between our findings and published findings could be due to the differences in cell lines. Our study used the E14.Tg2a mouse ES cells whereas the Li et al group used the R1 mouse ES cell line. The R1 mouse ES cell line was established from from (129/Sv x129/Sv-CP) F1 3.5-day blastocyst whereas our E14.TG2a was derived from the 129/Ola mice strain.

The *R3hdml* (R3H domain containing-like) is a 253 amino acid cysteine-rich

secretory protein belonging to the CRISP (Cysteine-Rich Secretory Protein) family. It is considered a putative serine protease inhibitor (Krätzschmar et al., 1996). CRISP proteins have molecular weights of about 20-30 kDa and are characterized by the presence of sixteen conserved cysteine residues (Krätzschmar et al., 1996). Most of the cysteine residues are clustered in the C-terminal domain of the molecule. Not much is known about *R3hdml* function in development until now.

In the present study, our results demonstrate that only a subset of BMP4 target genes are regulated by the BRE and that a novel cis-regulatory module/regulatory element may be responsible for regulating other non-BRE BMP target genes. Notably, we found two genes, *R3hdml* and *Dusp5*, are direct BMP targets and do not contain a predicted BRE in its promoter, intron, and downstream regions. These two previously unidentified BMP target genes may play key roles in mouse ES self-renewal and physiology. It is interesting to note that although BMPs play such diverse roles in many biological processes, we did not identify a large number of direct BMP targets. This suggests that only a handful of genes are potential "master" regulators of BMP signaling. Taken together, the wider implications of this finding may explain why different developmental processes could be regulated by one particular growth factor, BMP4, during the various developmental stages.

Figure 3.1

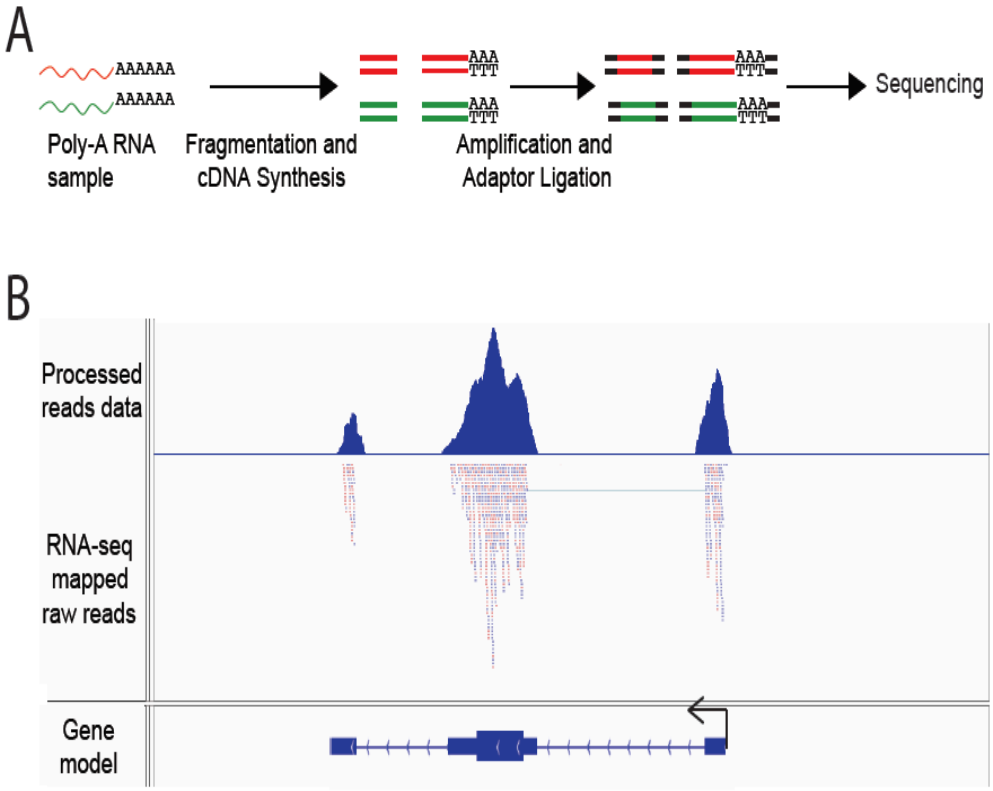


Figure 3.1: RNA-Seq Approach. A) This illustrate the typical workflow of how single stranded poly-A mRNA are converted to double stranded cDNA with adaptor on its ends to generate a sequencing library. B) After sequencing reads are computationally processed, reads are visualized on a genome browser. The genome browser consists of a gene model (right to left), mapped raw reads, and relative peaks (based on raw reads). Most of the reads are mapped to exon region of the gene.

Figure 3.2

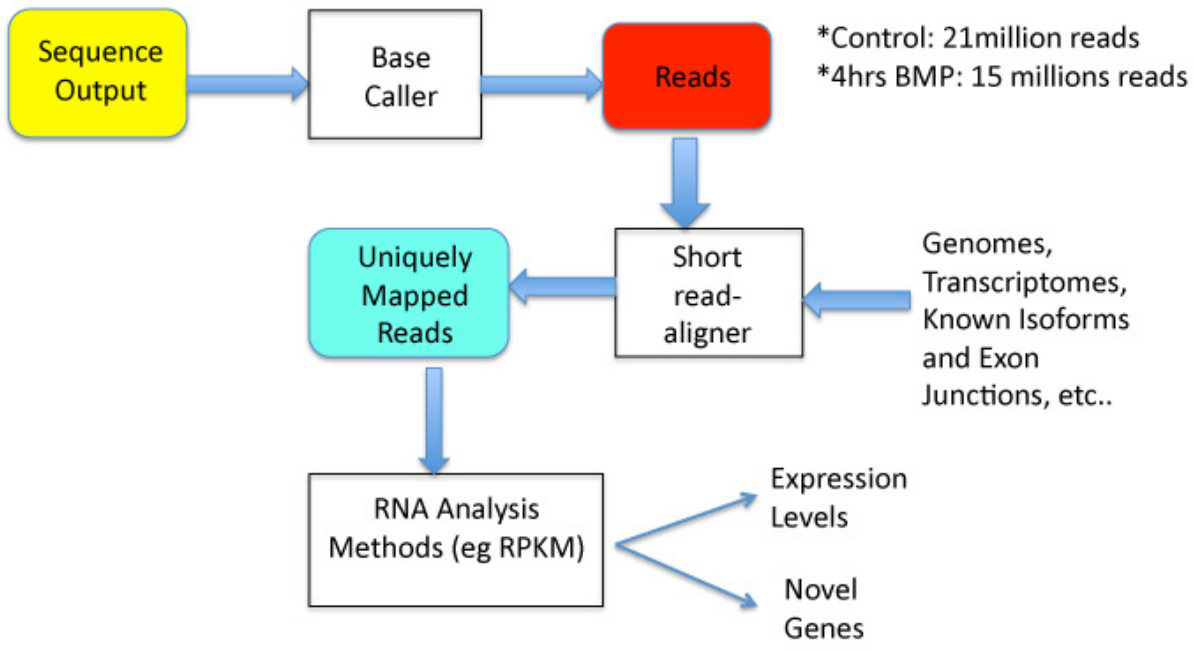


Figure 3.2: RNA-Seq Computational Pipeline: The Illumina Hi-Seq protocol generates a large sequence text file which is then input to a base caller to determine which base is present and to check for quality of each base read. The reads generated from the base caller is then aligned to a reference transcriptome. Only unique mapped reads to a gene are quantitated. Read quantitation are measured in RPKM (Reads Per Kilobase of transcript per Million mapped reads). The RPKM gives a measurement of transcript level.

Figure 3.3:

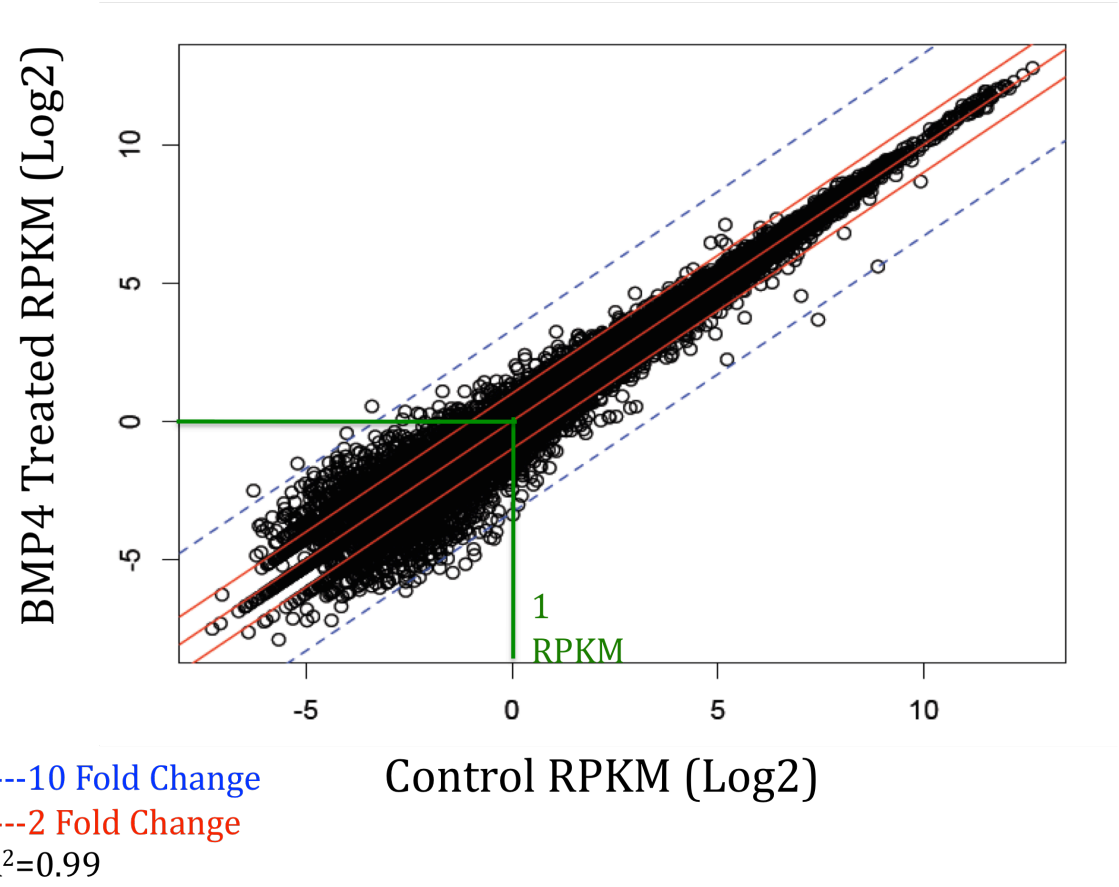


Figure 3.3: Correlation of Control Sample and BMP-Treated Sample. RPKM: Reads Per Kilobase of transcript per Million mapped reads, a measurement of transcript levels (Trapnell et al., 2012; Mortazavi et al., 2008)). After computationally processing RNA-Seq reads, RPKM values of all the genes in both treated and untreated (control) samples were graphed in scatterplot. Genes with a 10-fold change upregulate or downregulated are indicated in blue. Genes with a 2-fold change are indicated in red.

Figure 3.4:

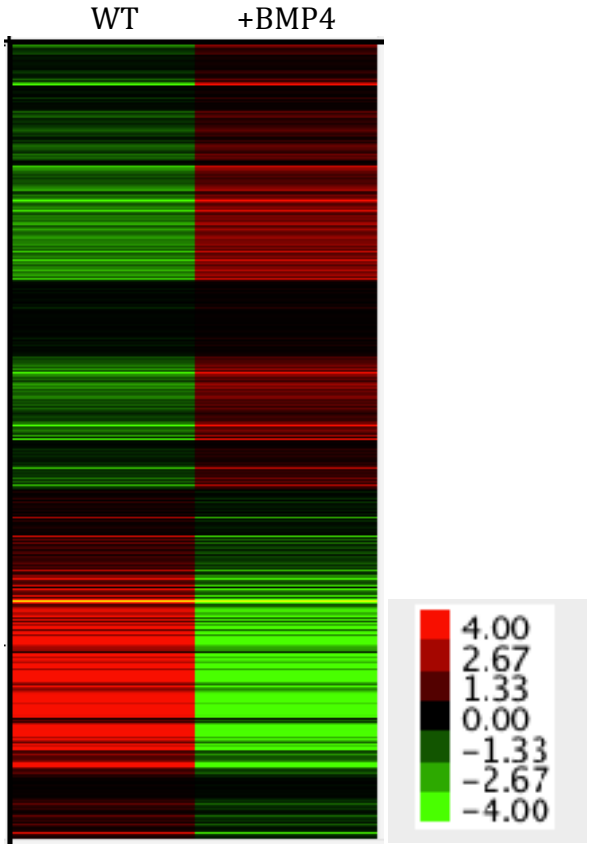


Figure 3.4: Cluster diagram of gene expression in wild type mouse ES cells and BMP4-treated mouse ES cells. Genes are organized by hierarchical clustering based on similarity in gene expression pattern. Red= >4 fold upregulation of gene expression; Green = represent > 4 fold downregulation of gene expression from the median. Black =No change in gene expression.

Figure 3.5:

RNA-Seq Genes Upregulated	Description
Sumo3	Small Ubiquitin-like small modifier (SUMO): posttranslationally modify numerous cellular proteins
Hsp5a	Heat shock protein
Msx2	Homeobox gene; known BMP4 target in murine and drosophila embryos
Coch	Identified in nerve fibers between the auditory ganglion and sensory epithelium
CTGF	Connective tissue growth factor (CTGF): a cysteine-rich, matrix-associated, heparin-binding protein
Grem1	Gremlin is an inhibitor in the TGF beta signaling pathway.
Gata2	BMP4 signaling has been implicated in activating expression of Gata2 during hematopoeitic development
Evx1	Even-skipped homeobox gene and expression is modulated by BMP-2 in developing mouse limb bud
Sox7	SRY-box containing gene 7 has been suggested to play a crucial role in parietal endoderm differentiation
Id, Id2, Id3, Id4	Inhibitor of differentiation and known to be expressed upon BMP4 stimulation in ES cells
Cxcl3	Small cytokine belonging to the CXC chemokine family and controls migration and adhesion of monocytes
Pcdh8	Belongs to the protocadherin gene family, a subfamily of the cadherin superfamily
Tchh	Trichohyalin confers mechanical strength to the hair follicle; BMP-4 facilitate Tchh transcription in mouse
Miccall2	The interaction of JRAB/MICAL-L2 coordinates the assembly of tight junctions and adherens junctions
Skil	nuclear protooncoprotein; Interacts with Mothers against decapentaplegic homolog 3 (aka SMAD3)
R3hdml	R3H domain containing-like
Dusp5	Dual-Specificity Phosphatase 9
Stard4	StAR-related lipid transfer protein 4
Itgb1	Integrin, Beta 1

Figure 3.5: A list of RNA-Seq genes that were upregulated and validated by quantitative RT-PCR. Genes highlighted in red are known BMP targets. Genes highlighted in green are novel direct BMP targets. A short description of the upregulated genes is seen to the right of the listed gene name.

Figure 3.6:

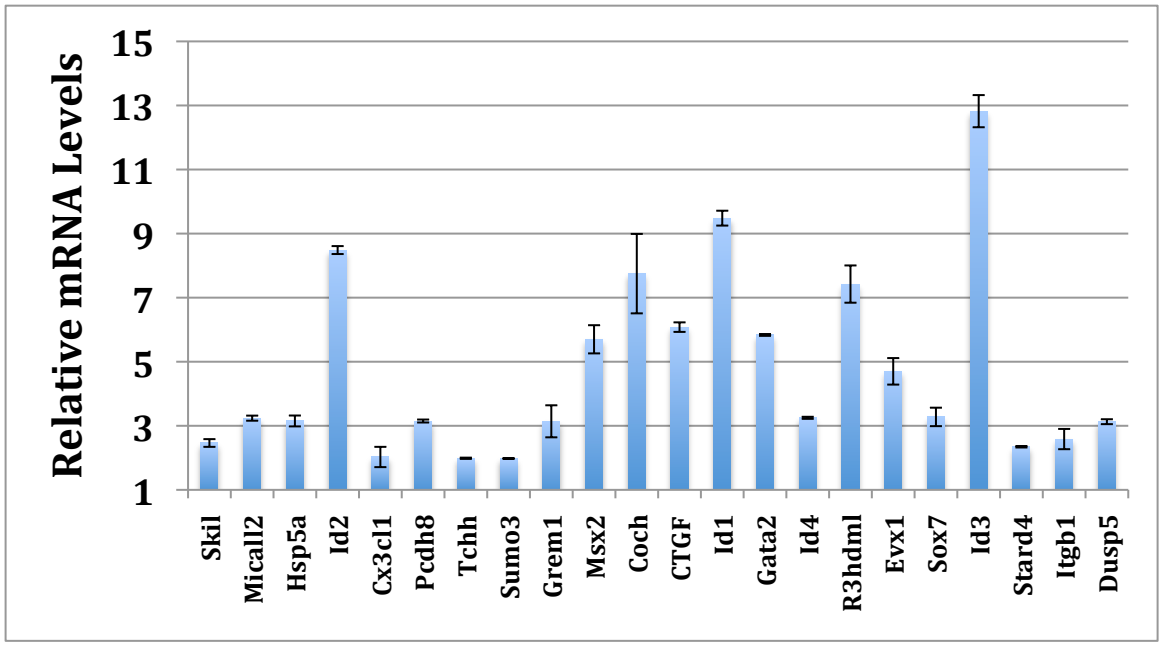
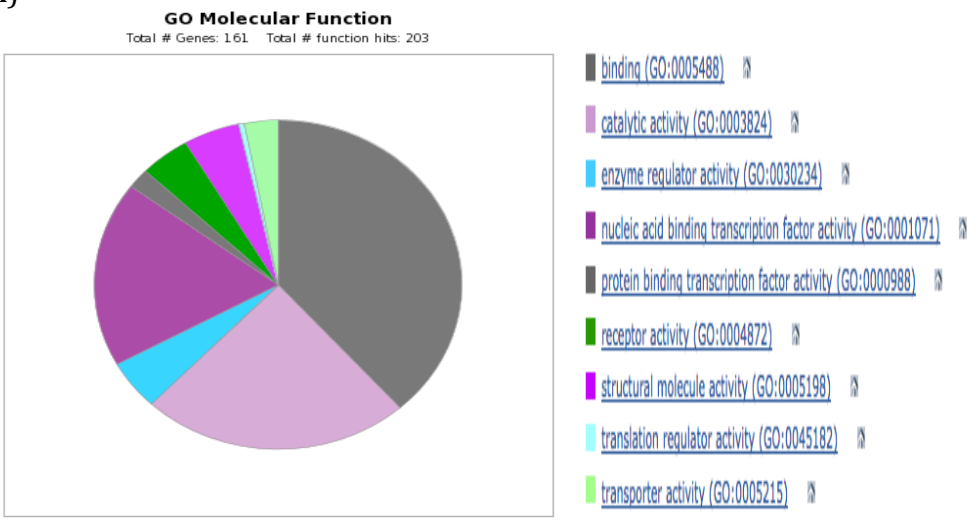


Figure 3.6: Quantitative RT-PCR validation of the upregulated RNA-Seq genes. Mouse ES cells were treated with and without BMP4 for 4hrs. RNA was isolated and cDNA was generated to perform quantitative real-time PCR for each gene. mRNA levels for each gene were normalized to the control (untreated) sample, which was at 1.

Figure 3.7:

A)



B)

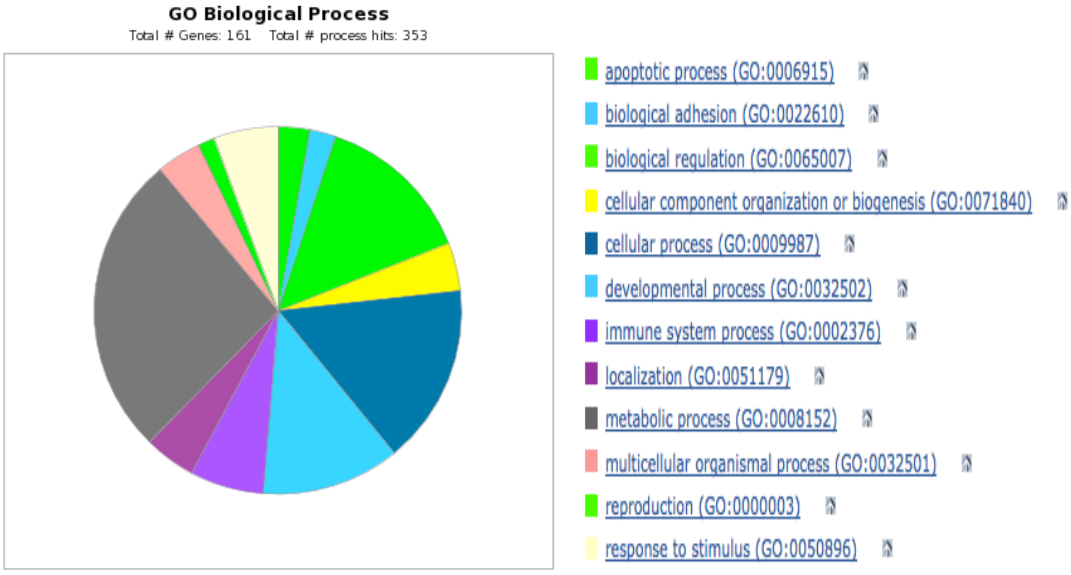
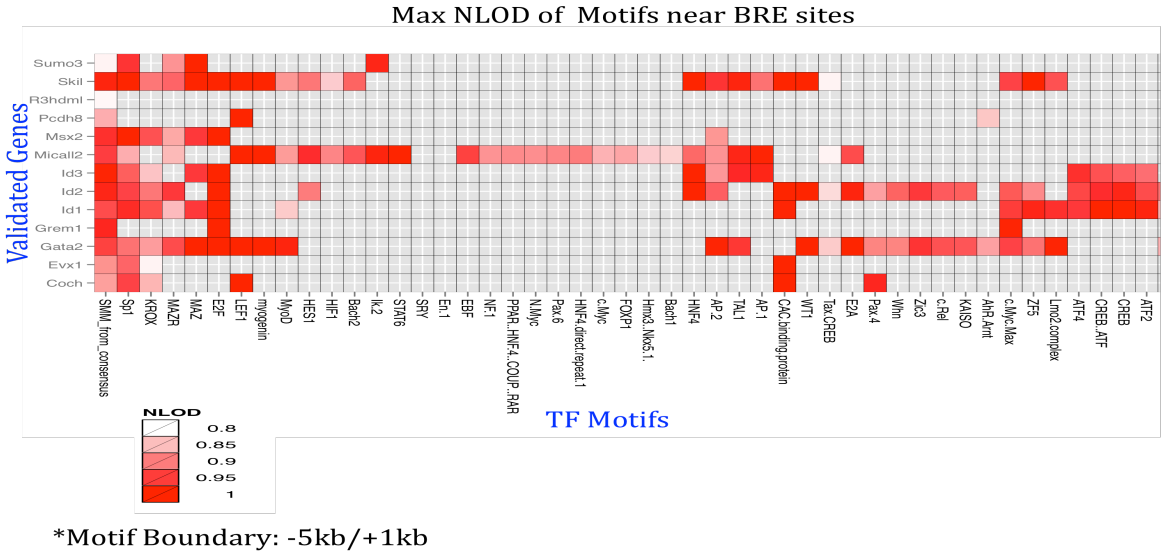


Figure 3.7: Gene Ontology (GO) of the top 200 upregulated genes by BMP4 as revealed by RNA-Seq. A) Gene Ontology based on Molecular Function. Many of the genes upregulated upon BMP4 stimulation in mouse ES cells are involved in binding, catalytic, and nucleic acid binding transcription factor activity. B) Gene Ontology Based on Biological Processes. A number of the genes upregulated upon BMP4 stimulation in mouse ES cells are involved in developmental process.

Figure 3.8:



*Motif Boundary: -5kb/+1kb

Figure 3.8: Many validated BMP inducible genes contain a BRE in their promoter regions. Transcription factor motifs were search 5kb upstream the transcription start site and +1kb downstream of the stop codon for each BMP4 validated gene using MotifMap. The X-axis contains a list of transcription motifs (TF). The y-axis contains a list of validated genes. The likelihood of containing a validated gene is based on the normalized log-odd (NLOD) score. The higher the NLOD score (darkest red) the more likely the BRE is present. The cuttoff of NLOD score was >0.80.

Figure 3.9:

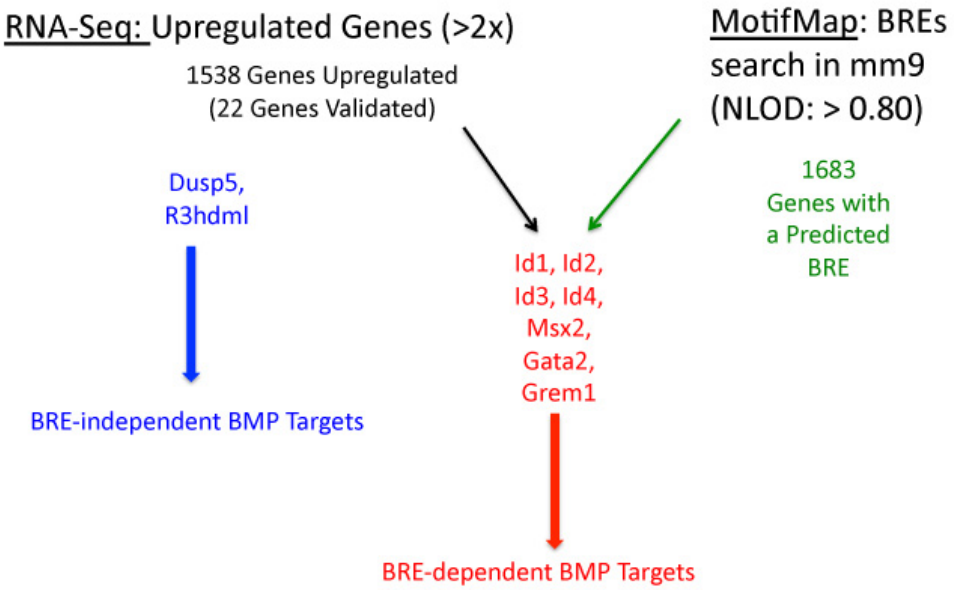
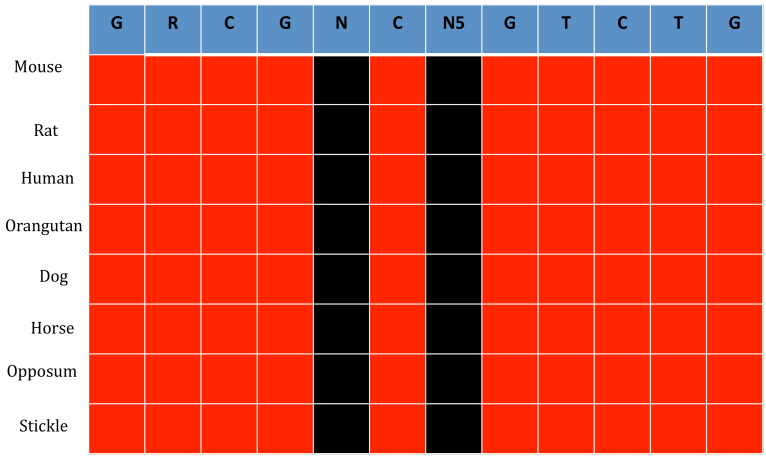


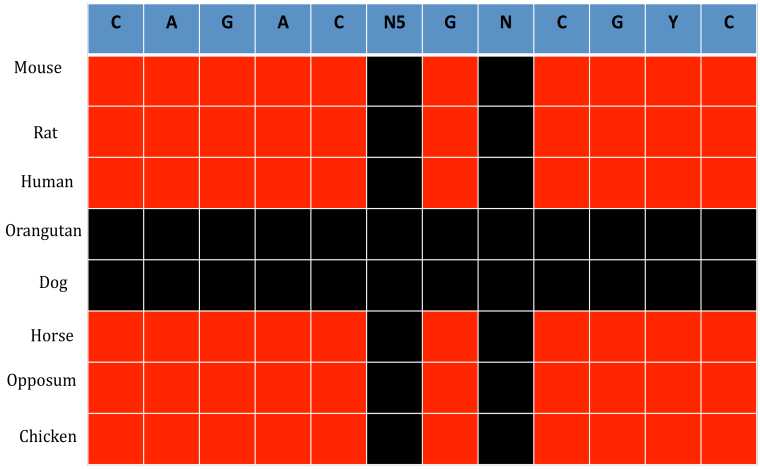
Figure 3.9: BREs are present in only a subset of BMP target genes. There were 1538 upregulated genes (> 2 Fold) in our RNA-Seq analysis. BREs were searched using MotifMap across the mm9 mouse genome using a normalized log-odd score (NLOD Cutoff: > 0.80). Approximately 1683 genes contained a predicted BRE within 10kb upstream and downstream. The predicted BREs were then intersected with the 22 validated genes (See Figure 3.6) upregulated from RNA-Seq anlaysis. Only 7 genes were known BMP targets and contained a predicted BRE. *R3hdml* and *Dusp5* are also direct BMP targets but do not contain a predicted BRE.

Figure 3.10:

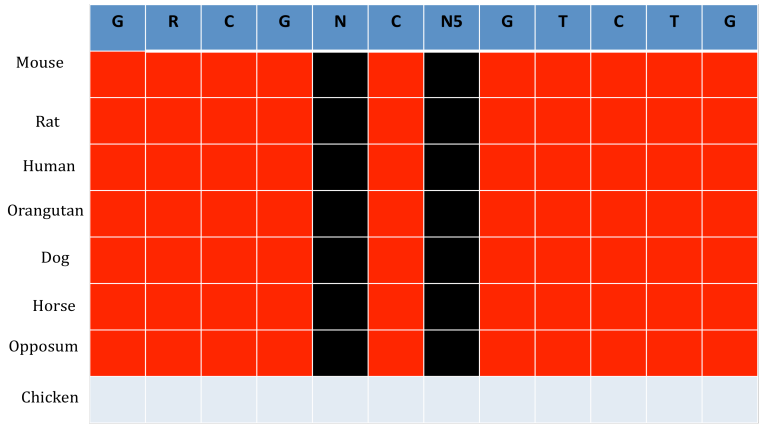
Id1: [chr2: 152565206-152565222] +



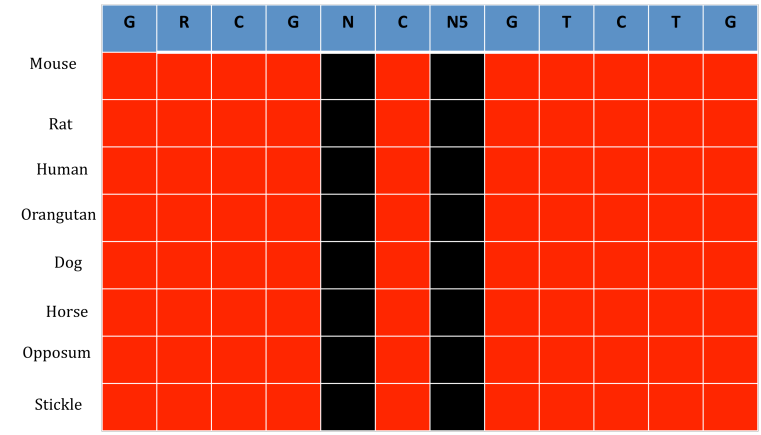
Id2: [chr12: 25783717-25783733] -



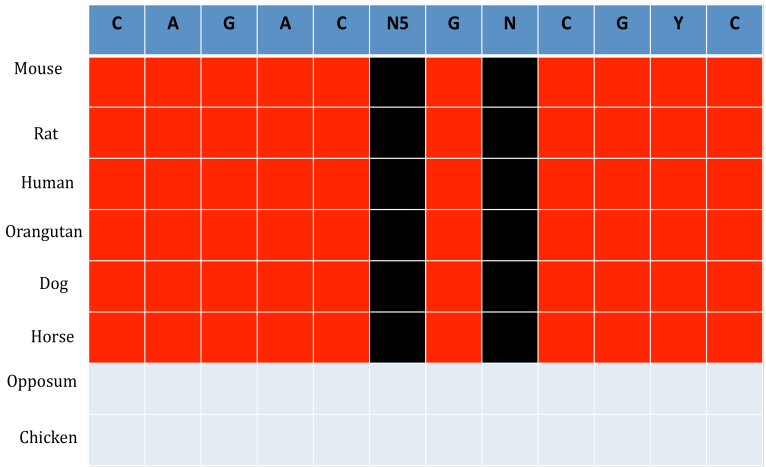
Id3: [chr4: 135696454-135696470] +



Id4: [chr13: 48358258 - 48358274] +



Msx2: [chr13:53571738 - 53571754] -



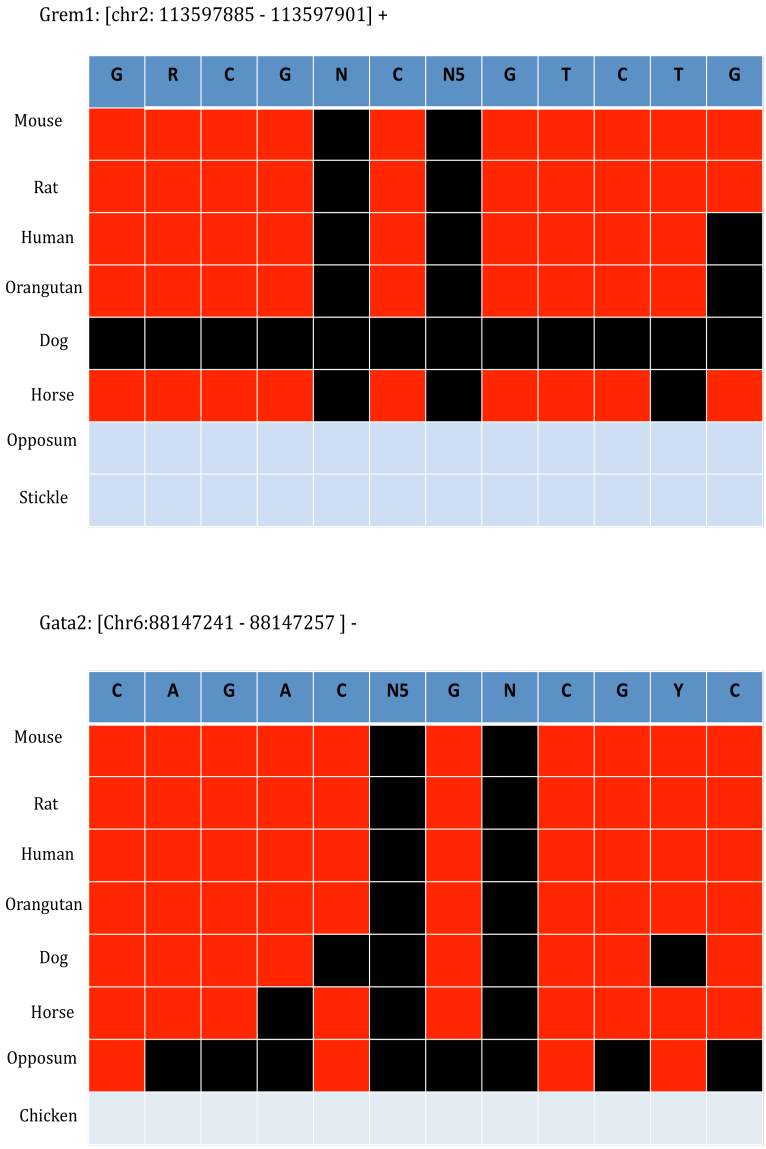


Figure 3.10: Heatmap of nucleotide conservation of predicted BRE among 7 BMP targets. The conservation of each nucleotide within the BRE was searched using PhyloP (UCSC Genome Browser). The Y-axis contains a list of organisms in which corresponding homologs were identified and the X-axis is the BRE motif (Positive strand: 5'-GRCGNC-N5-GTCTG-3'; Negative strand: 5'- CAGAC-N5-GNCGYC -3'). Red = Conservation; Black = No Conservation; Light Blue = No information

Figure 3.11:

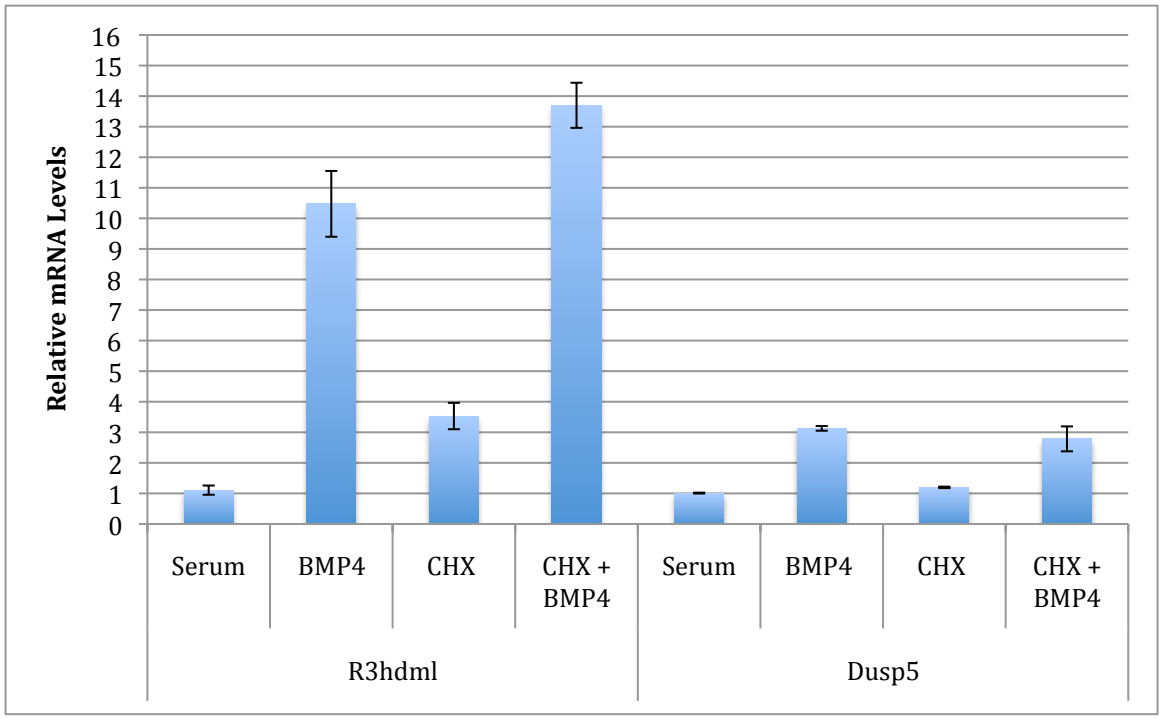
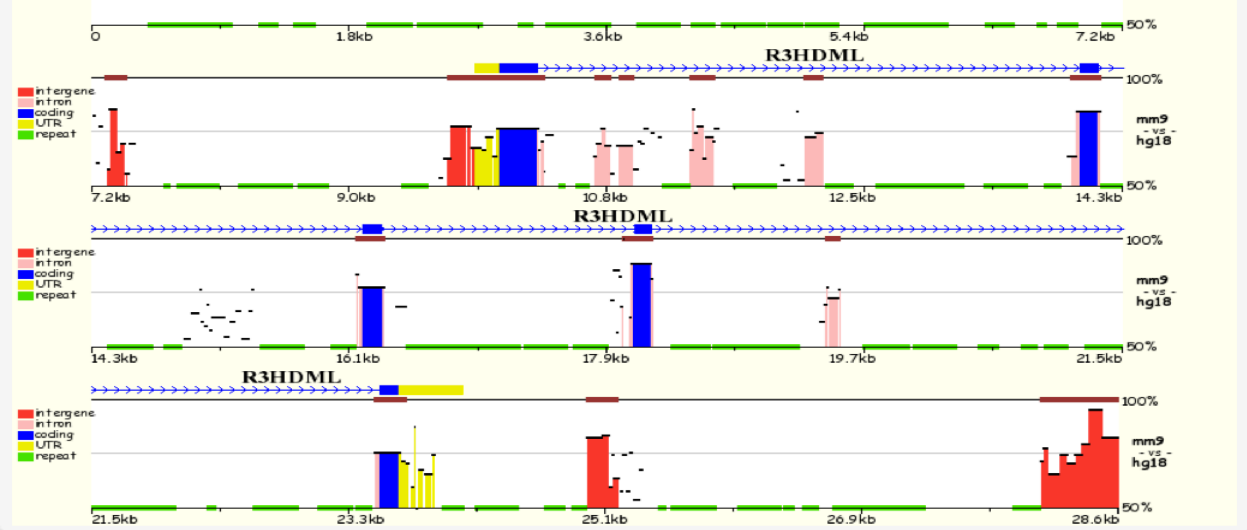


Figure 3.11: *R3hdml* and *Dusp5* are immediate response genes after BMP4 stimulation. Mouse ES cells were stimulated with treated with BMP4 alone, cycloheximide (CHX) alone, and CHX and BMP4 together for 4 hours. Quantitative RT-PCR was performed to check the expression levels of *R3hdml* and *Dusp5* transcripts. BMP induced the expression of *R3hdml* and *Dusp5*.

Figure 3.12:

A)



B)

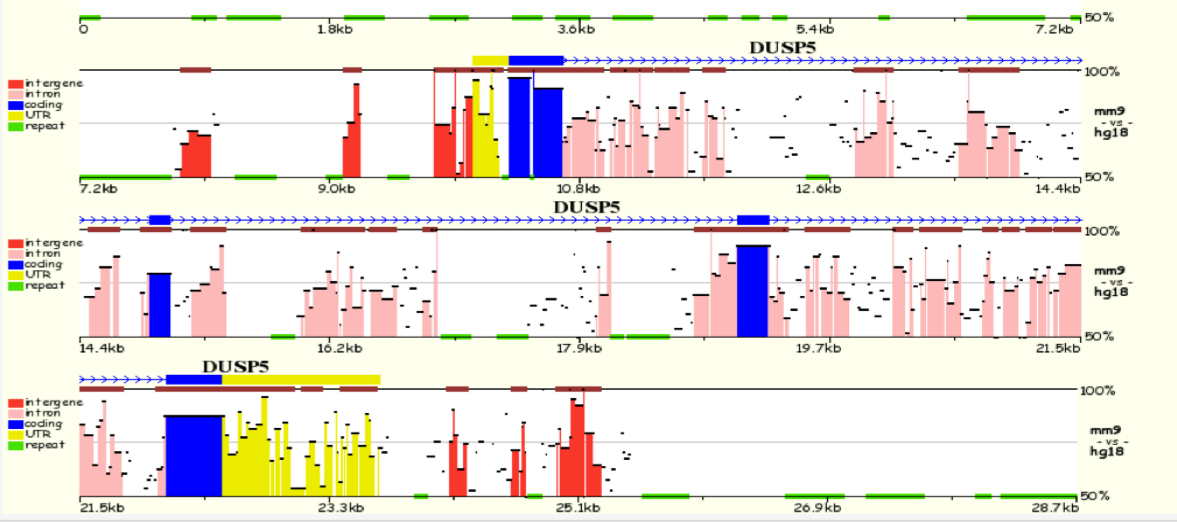


Figure 3.12: Evolutionarily conserved region (ECR) of the *R3hdml* and *Dusp5* genes predicted by Mulan (http://mulan.dcode.org/). ECR of the two genes: *R3hdml* (Panel A) and *Dusp5* (Panel B) (-10kb Upstream to +5kb downstream) between the mouse and human genome are indicated. Red = Intergenic region, Pink = Intron region, Blue = Coding region, Yellow = Untranslated region (UTR), Green = Repeat region, Burgundy = Evolutionarily conserved region.

Figure 3.13:

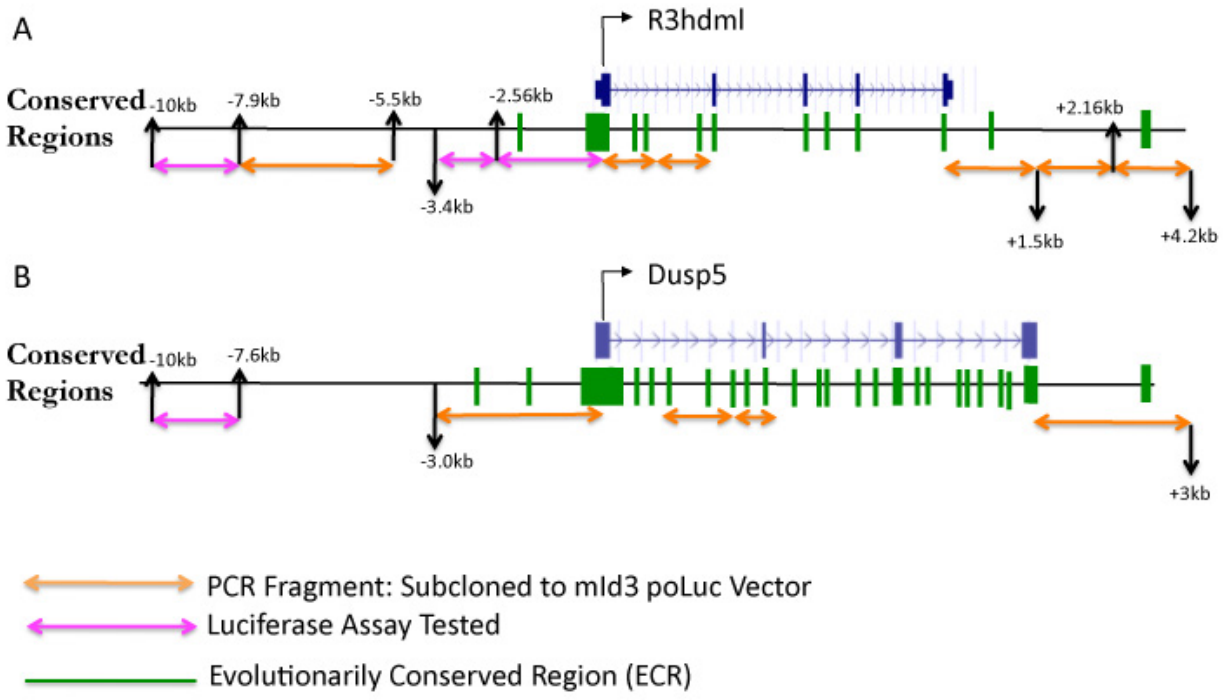


Figure 3.13: Diagram of cloning for *R3hdml* (Panel A) and *Dusp5* (Panel B). Magenta arrow shows tested region with luciferase assay. Orange arrow indicates PCR fragment has been successfully amplified and is in the process of being subcloned or have been cloned into the Id3 (-201) pOLuc vector. Green indicate evolutionarily conserved regions determined by Mulan.

Figure 3.14:

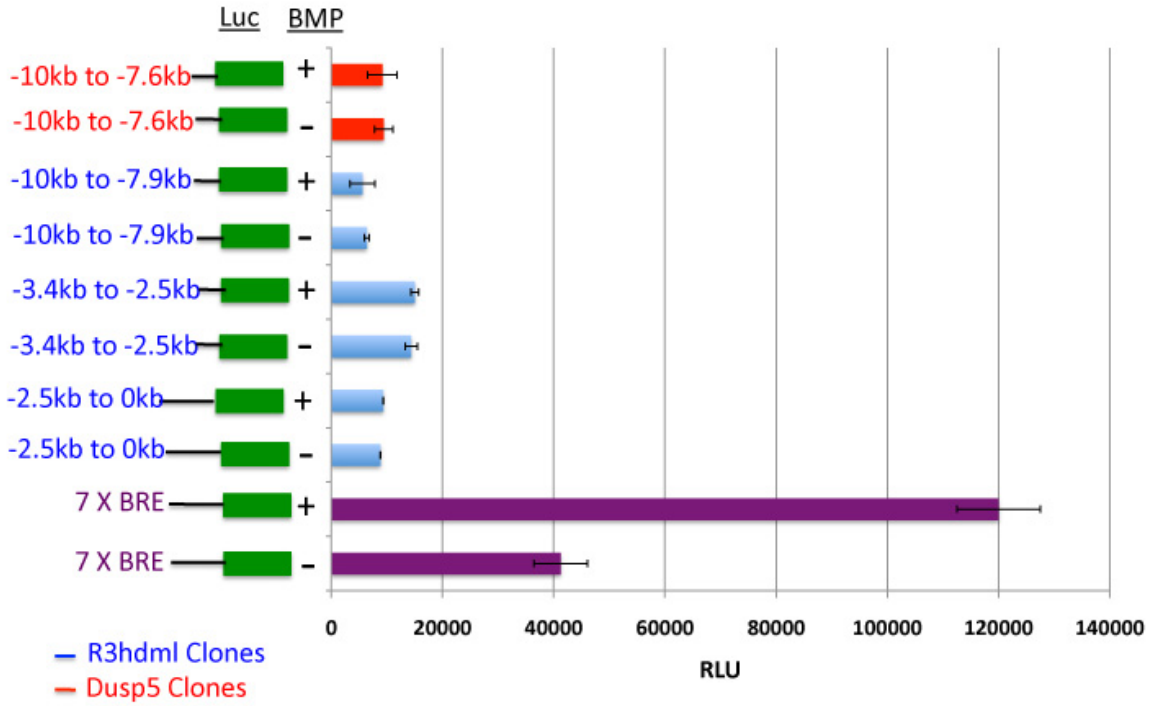


Figure 3.14: Luciferase assay of cloned promoter regions of *R3hdml* and *Dusp5* genes. Cloned promoters in the *Id3* (-201) pOLuc constructs were transfected into mouse ES cells and treated with and without BMPs. *R3hdml* clones are indicated in blue and *Dusp5* clones in red. A 7xBRE pOLuc construct was used as a positive control in the luciferase assay.

Chapter 3 References:

Branney PA, Faas L, Steane SE, Pownall ME, Isaacs HV. Characterisation of the fibroblast growth factor dependent transcriptome in early development. PLoS One. 2009;4:e4951.

Brugger SM, Merrill AE, Torres-Vazquez J, Wu N, Ting MC, Cho JY, Dobias SL, Yi SE, Lyons K, Bell JR, Arora K, Warrior R, Maxson R. A phylogenetically conserved cis-regulatory module in the Msx2 promoter is sufficient for BMP-dependent transcription in murine and *Drosophila* embryos. Development. 2004 Oct;131:5153-65.

Chen Q, Zhou Y, Zhao X, Zhang M. Effect of dual-specificity protein phosphatase 5 on pluripotency maintenance and differentiation of mouse embryonic stem cells. J Cell Biochem. 2011 Nov;112:3185-93.

Cloonan N, Forrest AR, Kolle G, Gardiner BB, Faulkner GJ, Brown MK, Taylor DF, Steptoe AL, Wani S, Bethel G, Robertson AJ, Perkins AC, Bruce SJ, Lee CC, Ranade SS, Peckham HE, Manning JM, McKernan KJ, Grimmond SM. Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nat Methods. 2008 Jul;5:613-9.

Daily K, Patel VR, Rigor P, Xie X, Baldi P. Motifmap: MotifMap: integrative genome-wide maps of regulatory motif sites for model species. BMC Bioinformatics. 2011 Dec 30;12:495.
Friedle H, Knöchel W. Cooperative interaction of Xvent-2 and GATA-2 in the activation of the ventral homeobox gene Xvent-1B. J Biol Chem. 2002 Jun 28;277:23872-81.

Hollnagel A, Oehlmann V, Heymer J, Rüther U, Nordheim A. Id genes are direct targets of bone morphogenetic protein induction in embryonic stem cells. J Biol Chem. 1999 Jul 9;274:19838-45.

Ishibashi T, Bottaro DP, Michieli P, Kelley CA, Aaronson SA. A novel dual specificity phosphatase induced by serum stimulation and heat shock. J Biol Chem. 1994 Nov 25;269:29897-902.

Javier AL, Doan LT, Luong M, Reyes de Mochel NS, Sun A, Monuki ES, Cho KW. Bmp indicator mice reveal dynamic regulation of transcriptional response. PLoS One. 2012;7:e42566.

Kolch W. Coordinating ERK/MAPK signalling through scaffolds and inhibitors. Nat Rev Mol Cell Biol. 2005 Nov;6:827-37.

Krätzschmar J, Haendler B, Eberspaecher U, Roosterman D, Donner P, Schleuning WD. The human cysteine-rich secretory protein (CRISP) family. Primary structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3. Eur J Biochem. 1996 Mar 15;236:827-36.

Kwak SP, Dixon JE. Multiple dual specificity protein tyrosine phosphatases are expressed and regulated differentially in liver cell lines. J Biol Chem. 1995 Jan 20;270(3):1156-60.

Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25

Li Z, Fei T, Zhang J, Zhu G, Wang L, Lu D, Chi X, Teng Y, Hou N, Yang X, Zhang H, Han JD, Chen YG. BMP4 Signaling Acts via dual-specificity phosphatase 9 to control ERK activity in mouse embryonic stem cells. Cell Stem Cell. 2012 Feb 3;10:171-82.

Loots GG, Ovcharenko I. Mulan: multiple-sequence alignment to predict functional elements in genomic sequences. Methods Mol Biol. 2007;395:237-54.

Mandl M, Slack DN, Keyse SM. Specific inactivation and nuclear anchoring of extracellular signal-regulated kinase 2 by the inducible dual-specificity protein phosphatase *DUSP5*. Mol Cell Biol. 2005 Mar;25:1830-45.

Mi H, Muruganujan A, Thomas PD. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res. 2013 Jan;41:D377-86.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008 Jul;5:621-8.

Ozsolak F, Platt AR, Jones DR, Reifenberger JG, Sass LE, McInerney P, Thompson JF, Bowers J, Jarosz M, Milos PM. Direct RNA Sequencing. Nature. 2009 Oct 8;461:814-8.

Pramanik K, Chun CZ, Garnaas MK, Samant GV, Li K, Horswill MA, North PE, Ramchandran R. Dusp-5 and Snrk-1 coordinately function during vascular development and disease. Blood. 2009 Jan 29;113:1184-91.

Selever J, Liu W, Lu MF, Behringer RR, Martin JF. Bmp4 in limb bud mesoderm regulates digit pattern by controlling AER development. Dev Biol. 2004 Dec 15;276:268-79.

Silva J, Barrandon O, Nichols J, Kawaguchi J, Theunissen TW, Smith A. Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol. 2008 Oct 21;6:e253.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012 Mar 1;7:562-78.

Ying QL, Nichols J, Chambers I, Smith A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell. 2003 Oct 31;115:281-92.

Chapter 4

BMP signaling in early mouse embryonic development

During early embryonic development, the fertilized mammalian egg undergoes multiple cell divisions to form a morula that will eventually give rise to a blastocyst consisting of two cell populations: the trophectoderm (TE) and the inner cell mass (ICM). Each will give rise to two distinct cell lineages. The TE will give rise to the placenta and extra-embryonic tissues, whereas the ICM cells will give rise to the epiblast and the primitive endoderm (PE) (Beddington et al., 1999; Rossant et al., 2009). The earliest cell fate decision during mammalian embryonic development after fertilization is the specification of the ICM and TE (Figure 4.1).

The BMP signaling pathway plays a critical role in many developmental processes and loss of BMP signaling components are embryonic lethal. In mice lacking BMP4, BMP Type II receptor (BMPRII), or BMP Type I receptor (BMPRI), defects include improper mesoderm differentiation and failure to initiate gastrulation (Mishina et al., 1995; Winnier et al., 1995; Beppu et al., 2000). In addition, BMP4 is important for allocation of primordial germ cells (PGCs) during development (Lawson et al., 1995). Furthermore, deletion of *Smad1* in mice can cause defects in extraembryonic tissues, decrease in PGC numbers, and lethality at embryonic day 10.5 (E10.5) (Zhao et al., 2003; Miyazono et al., 2010).

Although BMP4 has been implicated in various developmental processes, the molecular mechanisms underlying preimplantation blastocyst stage development are still unclear. Data supporting BMP activity in early embryogenesis has been limited due to difficulty in manipulating mouse embryos. Experiments performed by Coucouvanis and Martin (Coucouvanis et al., 1995) demonstrated that *BMP4* is exclusively expressed in the

ICM as revealed by *in situ* hybridization. Furthermore, inhibiting BMP4 can lead to disruption in cavitation in embryoid bodies (Coucouvanis et al., 1995). Additionally, BMP transducers, such as Smad1 and Smad5, are also expressed in the epiblast of the early embryo (Chang et al., 1999; Chang et al., 2001; Tremblay et al., 2001). Moreover, *Id2*, a BMP target gene, has been shown to be upregulated in 16-cell (E3.0) and 32-cell stage embryo (E3.25-E3.5) along with Sox2, a pluripotency transcription factor (Guo et al., 2010). The Id proteins were first identified in 1990 (Benezra et al., 1990). Currently, four members, *Id1* through *Id4*, have been identified in mammals (Benezra et al., 1990; Christy et al., 1991; Sun et al., 1991; Riechmann et al., 1994). During development, all members of the Id family are expressed in many organs and tissues, showing overlapping but distinct expression patterns in particular developmental stages and localized regions (Evans et al., 1993; Riechmann et al., 1995; Zhu et al., 1995; Jen et al., 1996, 1997).

It can therefore be hypothesized that BMP signaling may be active in the ICM and that BMP target genes are important for regulating Id expression during the first lineage segregation of the ICM and TE during early mouse embryogenesis.

Materials and Methods:

Single-Cell RNA-Seq Analysis:

Single-Cell RNA Seq published dataset was obtained from GEO database (GSE22182) to the mouse reference transcriptome (mm9). Details for computational protocols for SOLid sequencing platform and analysis can be found in: Tang F, Barbacioru C, Nordman E, Bao S, Lee C, Wang X, Tuch BB, Heard E, Lao K, Surani MA. Deterministic and stochastic allele specific gene expression in single mouse blastomeres. PLoS One. 2011;6(6):e21208.

Mouse ES Cell Culture:

E14Tg2a mouse ES cells were cultured on 0.1% gelatin-coated plates containing GMEM medium (Sigma Aldrich, #G5153) with 10% FBS (Hyclone, catalog #SH30071.03), 0.1mM NEAA, 1mM Sodium Pyruvate, 1000U/ml LIF (Millipore, catalog #ESG1107), and 0.1mM 2-Mercaptoethanol. Cells were maintained at 37°C, 5% CO_2, and split using 0.25% trypsin (Life Technologies, catalog #15400-054).

Quantitative Polymerase Chain Reaction (qPCR):

E14Tg2a mouse ES cells were treated with 20ng/ml hBMP4 (R&D Systems), 80nM LDN193189 (Stemgent), or both for 8hrs. RNA was isolated after the following treatment using Trizol (Life Technologies). Reverse transcription was performed using MMLV Reverse Transcriptase. Quantitative PCR reactions were performed using Sybr Green (Roche).

qPCR primer sequences for Id genes:

mId1

Forward: 5'- GACATGAACGGCTGCTACTCAC -3'

Reverse: 5'- GACTTCAGACTCCGAGTTCAGC -3'

mId2

Forward: 5'- GTGAGGTCCGTTAGGAAAAACAG -3'

Reverse: 5'- GTCGTTCATGTTGTAGAGCAGACT -3'

mId3

Forward: 5'- CTGTCGGAACGTAGCCTGG -3'

Reverse: 5'- GTGGTTCATGTCGTCCAAGAG -3'

mId4

Forward: 5'- AGTGCGATATGAACGACTGCTAC -3'

Reverse: 5'- AGGATGTAGTCGATAACGTGCTG -3'

Mouse Embryo Acquisition:

Three-week old CD1 females (Charles River) were superovulated using pregnant mare serum (PMSG, Sigma Aldrich, catalog #G-4877, 2000IU/vial) followed ~ 45 hours later with human chorionic gonadotropin (hCG, Sigma Aldrich, catalog #C-1063, 2500IU/vial). All females received 5 IU doses of each hormone intraperitoneally. After hCG (Sigma Aldrich, catalog #C-1063, 2500IU/vial) injections, mating were set up. Females were in checked the following morning for mating plugs, at which they were considered e0.5 if a vaginal plug was present. Embryos were collected at desired embryonic day by flushing uterine horns or oviducts with Dulbecco's Modified Eagle Medium (DMEM, Life Technologies, catalog #11965-092) with Hepes (10mM) media (Life Technologies, catalog #15630-106).

Immunofluorescence Staining:

Preimplantation stage mouse embryos were rinsed in Acid Tyrode's solution, (Sigma Aldrich), followed by a 30 minutes fixation with 3.7% formaldehyde. Embryos were then incubated with primary anti-Id2 antibody (1:100, Santa Cruz Biotechnology). Additionally, embryos were stained with Hoechst (2µg/ml, Sigma Aldrich) for nuclear staining. Embryos were placed on a glass slide coated with a 1% agarose pad and compressed to a 3:1 aspect ratio. All confocal images were acquired using 63X 1.4 NA 22 objective, on an Axioobserver Z1 Zeiss 780 confocal microscope with Zen2009 software. Z-

stack images were acquired at 0.3μm intervals.

Results and Discussion:

Chapter 4.1: BMP Ligand Expression in Early Mouse Development as Revealed by Single-Cell RNA-Seq

Single-cell transcriptomic analysis technique is a powerful tool to profile global gene expression. Most RNA-Seq analyses performed currently are at "population level," in which transcriptome profiles are an average from millions of cells. A single-cell RNA-Seq approach has the following advantages: a) It can detect subtle differences between seemingly identical cells to extract biologically important information, b) Cell lineages can be traced and heterogeneity within a cell population can be detected, c) Rare cells can be now be identified and characterized from this approach. Although single-cell RNA-Seq has many advantages, it should be noted that single-cell RNA-Seq has a higher technical noise due to the fact the starting material is low for this type of sequencing (Brennecke et al., 2013).

To determine if BMP components are expressed in preimplantation embryos, I analyzed a published dataset from Tang et al. This paper published single-cell RNA-Seq from early mouse embryos: 2-cell, 4-cell, 8-cell, TE, ICM, ES cells (ESC), and epiblast (E4.5) cells. I observed that BMP2 is expressed at low levels (~1RPKM) throughout the different developmental stages. BMP6 and BMP7 are expressed early on (2-cell stage) but their expression declined as development progressed (4-cell stage to epiblast) (Figure 4.2). BMPRIA (ALK3) and BMPRIB (ALK6) are differentially expressed zygotically (2-cell to epiblast). The Type II receptor, BMPRII, is expressed highly (~38 RPKM) at the 2-cell stage

but gradually declines as it reaches E4.5. BMP4 expression remains relatively low (~1RPKM-11RPKM) during the earlier mouse stages (2-8cell stages). However, during the first lineage segregation between the ICM and TE, BMP4 is highly expressed (~284 RPKM) in the ICM but not in the TE (Figure 4.2). This result is consistent with the observation of Coucouvanis et al. where they showed the expression of *BMP4* in the ICM. While RNA-Seq data reveals dynamic spatiotemporal expression patterns of BMP ligands and receptors, they do not reveal if BMPs are functional. Research in the Cho lab has shown that BMP signaling is active as early as the 4-cell stage (Reyes de Mochel et al., 2014).

Chapter 4.2: BMP Targets in Early Mouse Embryos: Inhibitors of Differentiation Genes

The inhibitors of differentiation (Id) proteins belong to the helix-loop-helix (HLH) family of transcription factors. They are inhibitors of basic HLH (bHLH) transcription factors. bHLH factors functions as active transcription factors to promote cell differentiation by activating tissue-specific gene expression (Murre et al., 1989). bHLH transcription factors contain a basic DNA-binding domain and a helix-loop-helix region that mediate protein–protein interactions. Unlike bHLH transcription factors, the Id proteins lack a basic DNA-binding domain. However, the Id proteins can still heterodimerize with other bHLH transcription factors via its HLH (Benezra et al., 1990; Jen et al., 1992; Norton et al., 1998), and function as a dominant-negative transcriptional regulator.

In mouse ES cells, all four *Id* genes are upregulated upon BMP stimulation (Figure 4.3). However, in the presence of a BMP inhibitor, LDN193189, the *Id* genes are downregulated (Figure 4.4). In mouse embryogenesis, both *Id1* and *Id3* are expressed in

the tissues derived from the inner cell mass, whereas *Id2* is expressed in tissues derived from trophectoderm (Jen et al., 1997). *Id4* expression is absent during this period of development (Jen et al., 1997). This differential expression pattern within the *Id* genes suggests a non-redundant role for these genes in antagonizing the activity of bHLH transcription factors during early mouse development.

To obtain spatial-temporal expression patterns of the *Id* genes in preimplantation early mouse embryonic development, we utilized the single-cell RNA Seq data (Tang et al., 2011). Our analysis of the RNA-Seq data shows that *Id2* and *Id3* are differentially expressed during the ICM and TE segregation (Figure 4.5). *Id2* is expressed highly in the trophectoderm whereas *Id3* is highly expressed in inner cell mass. Interestingly, in ES cells, *Id2* and *Id3* expression are at similar levels, whereas in the ICM of the embryo, *Id3* (~424 RPKM) expression is much higher than *Id2* (~6RPKM). This difference may be attributed to that fact that ES cells are cultured *in vitro*, which sometimes may not always precisely reflect *in vivo* conditions. Additionally, Id2 immunostaining of E3.5 mouse embryos shows Id2 expression primarily in the trophectoderm but not the ICM (Figure 4.6). A possible explanation for this differential expression of the Id genes is microRNAs. MicroRNAs bind to complementary sites within the 3′UTRs (3'Untranslated regions) of cognate mRNAs, leading to translational repression or cleavage (Lim et al., 2005) of the targeted messages. It is known that different clusters of microRNAs are present within the trophectoderm and ICM (Viswanathan et al., 2009; Tang et al., 2010) to potentially downregulate specific genes in the ICM or TE. Additionally, studies have shown that Id proteins (Id1-3) are subject to K48-linked polyubiquitination and subsequent degradation by the 26S proteasome (Bounpheng et al., 1999). Conversely, the deubiquitinating enzyme USP1 promotes Id

protein stability. The polyubiquitination and deubiquitination of Id proteins can also explain the differential expression of the Id2 protein in the ICM and TE.

It should be noted in our immunostaining that Id2 expression is mostly localized in the cytoplasm. The small size the of the Id proteins (13-18 kDa) suggests that Id proteins enter and exit the nucleus by passive diffusion but several studies have indicated that other pathways may regulate their subcellular localization (Deed et al., 1996; Wang et al., 2001; Tux et al., 2003; Samanta et al., 2004). Research has shown Id2 and Id4 can sequester nuclear bHLH proteins in the cytoplasm to promote differentiation of cultured neural progenitors (Samanta et al., 2004). Furthermore, Kurooka et al. suggested that Id2 is exported from the nucleus and has an inhibitory effect on repression of E-box-mediated transcription. Additionally, Kurooka et al. also identified a nuclear export signal (NES) in mouse *Id2* responsible for exporting Id2 from the nucleus via the nuclear receptor, chromosome region maintenance 1 (CRM1). Currently, it is unclear which mechanism (passive diffusion or Id2 nuclear export via CRM1) is utilized for Id2 cytoplasmic localization in the TE.

Although we have not performed Id3 immunostaining in E3.5 embryos, we predict Id3 expression to be strongly expressed in the ICM. These findings show that this restricted expression of specific Id proteins occur earlier than previously reported (Jen et al., 1996; Jen et al., 1997).

The striking difference in *Id* gene expression during mouse blastocyst development suggests possible roles for *Id* genes during the first lineage segregation between the ICM and TE. The role of *Id* genes in lineage commitment is not novel and can be seen in the establishment of the spinal cord during embryogenesis (Jen et al., 1997). Although, *Id*

genes have been implicated in stem cell maintenance (Ying et al., 2003), *Id* genes also play a role during neural crest formation (Kee et al., 2005; Nichane et al., 2008). During neurogenesis, all members of the *Id* genes are expressed along the dorsal ventral axis of the neural tube (Jen et al., 1997). In the early stages of spinal cord development, *Id1* and *Id3* expression are found in mitotically active and less-differentiated neuroblasts located in the ventricular zone (Jen et al., 1997). On the other hand, *Id2* and *Id4* are expressed in the presumptive postmitotic neurons (Jen et al., 1997). Previously reported findings (Jen et al., 1997) along with our observation of differential *Id* gene expression in the mouse blastocyst support the possibility that the *Id* genes are just not simply inhibitors of transcription factors but may play a broader role in cell fate determination as well. The shared function among all the *Id* genes is to sequester other bHLH transcription factors and prevent them from binding to DNA. It is well known that *Id1* and *Id3* share overlapping expression in almost all tissues at the same developmental stages (Ellmeier et al., 1995), thus suggesting a functional redundancy between these two *Id* genes. *Id2* and *Id4* do not share such overlapping expression. The differences in dimerization partners of the *Id* genes and the cellular context may contribute to the varying activities/expression of each *Id* gene. Therefore, it is likely that the *Id1* and *Id3* genes are important for ICM identity and *Id2* for TE identity.

Conclusions:

At present, data indicating BMP activity in pre-implantation embryos has been scarce. Utilizing published single-cell RNA-Seq dataset, we were able to visualize the spatial distribution of BMP ligands and its targets at a specific developmental window.

Here we present data showing that BMP activity is localized in the ICM via RNA-Seq. This data is also supported by work done by Soledad Mochel showing strong BMP expression in the ICM and weak BMP expression in the TE of BRE-gal blastocysts (Reyes de Mochel et al., 2014). It is possible that BMP signaling regulating ICM and TE identity may be mediated via utilization of different BMP ligands and receptors, and thus different transcriptional machineries.

We also provided preliminary evidence that the *Id* genes are differentially expressed in the mouse blastocyst. However, more experiments need to be done to determine how *Id* genes and their binding partners regulate ICM and trophectoderm identity. Moreover, it is equally important to better define the spatiotemporal expression patterns of BMP signaling components within the embryo since different BMP ligands and receptors can activate different transcriptional machineries by selecting various transcription factor partners. In conclusion, identifying the role of BMP signaling in controlling cell fate decisions and cellular behaviors in early mouse embryogenesis offers many intriguing insights in this critical developmental window.

Figure 4.1

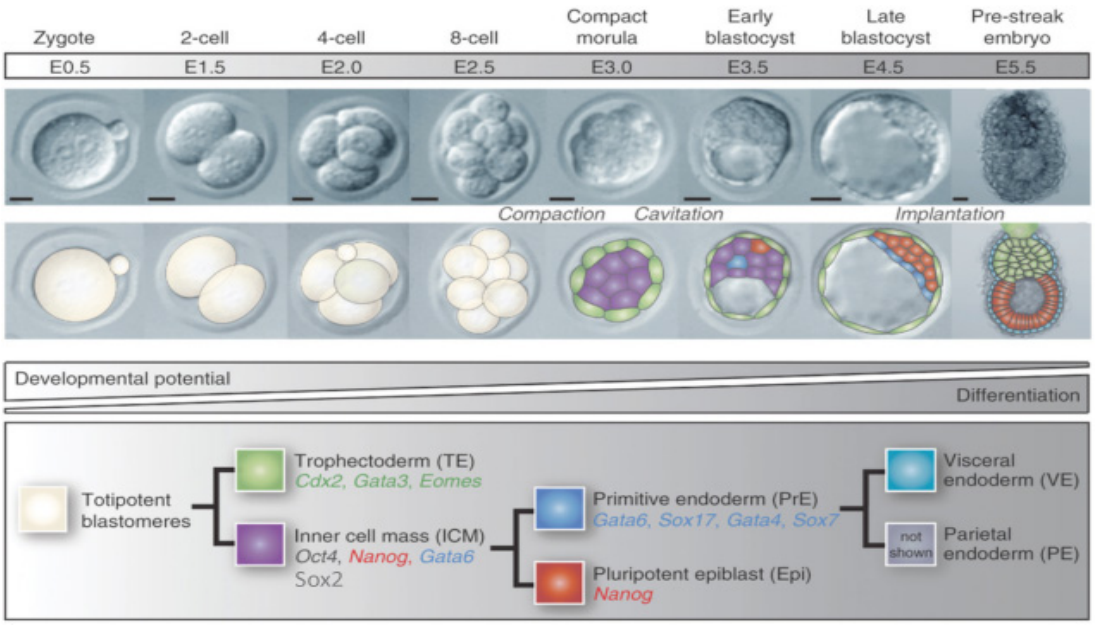


Figure 4.1: Early Mouse Embryonic Development. Stages of mouse preimplantation stages. Fertilization of sperm with oocyte forms a zygote. The zygote undergo successive cell divisions to form a blastocyst at day 3.5 (E3.5). The early blastocyst contains the inner cell mass (ICM) and trophectoderm (TE). The ICM will give rise to the epiblast (EPI) and the primitive endoderm (PE). The epiblast will give rise to the 3 primary germ layers of the embryo proper: endoderm, mesoderm, ectoderm. The primitive endoderm give rise to the parietal and visceral endoderm (Modified from Niakan et al., 2013).

Figure 4.2

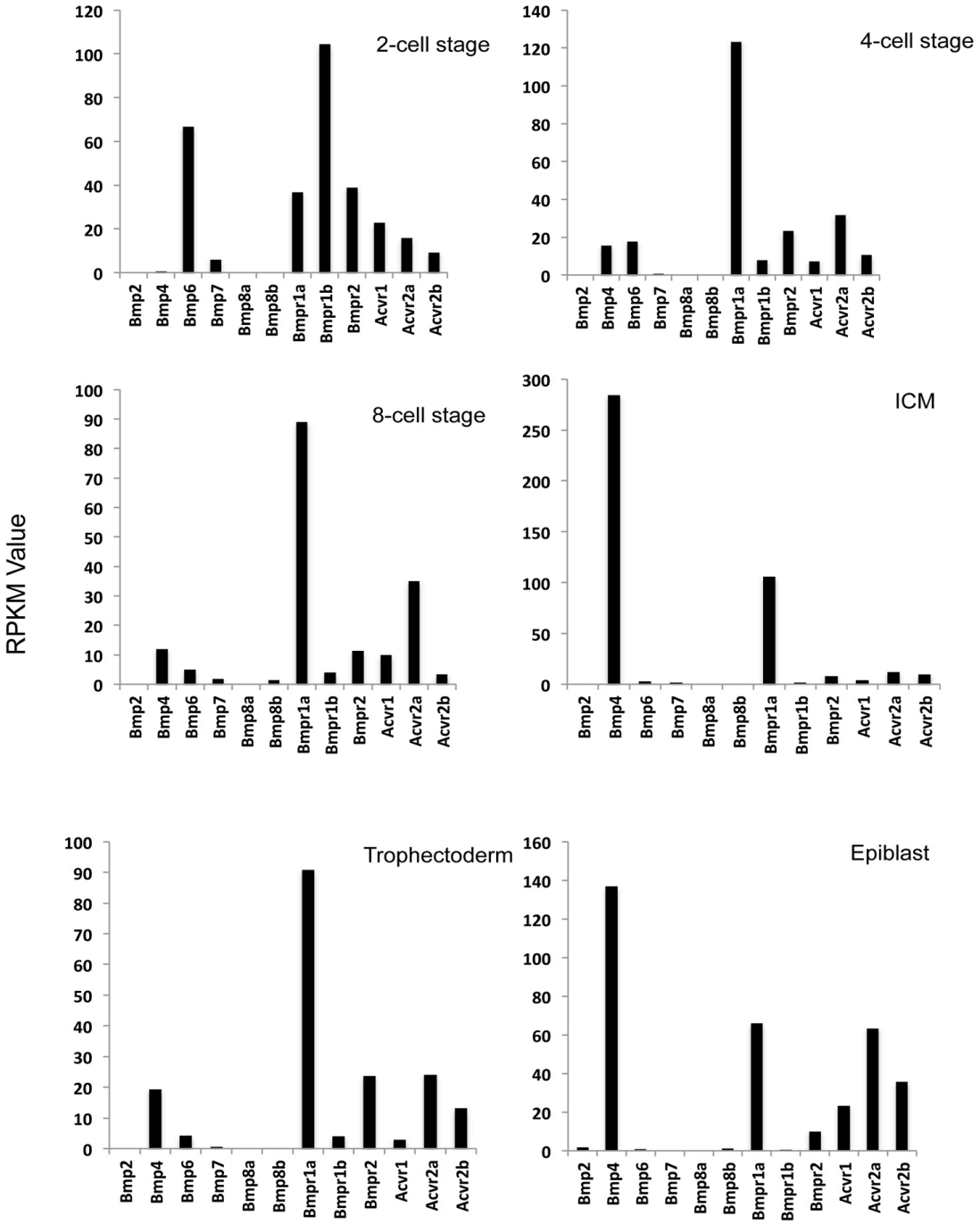


Figure 4.2: Expression profiles of TGF–β-related ligands and receptors from single-cell RNA-Seq. RPKM levels were obtained for BMP and Activin/TGF–β ligands and receptors at different developmental stages. Inner Cell Mass (ICM), Trophectoderm (TE), and Epiblast. BMP4 is enriched in the ICM stage. Published data was obtained from Tang et al., 2011.

Figure 4.3:

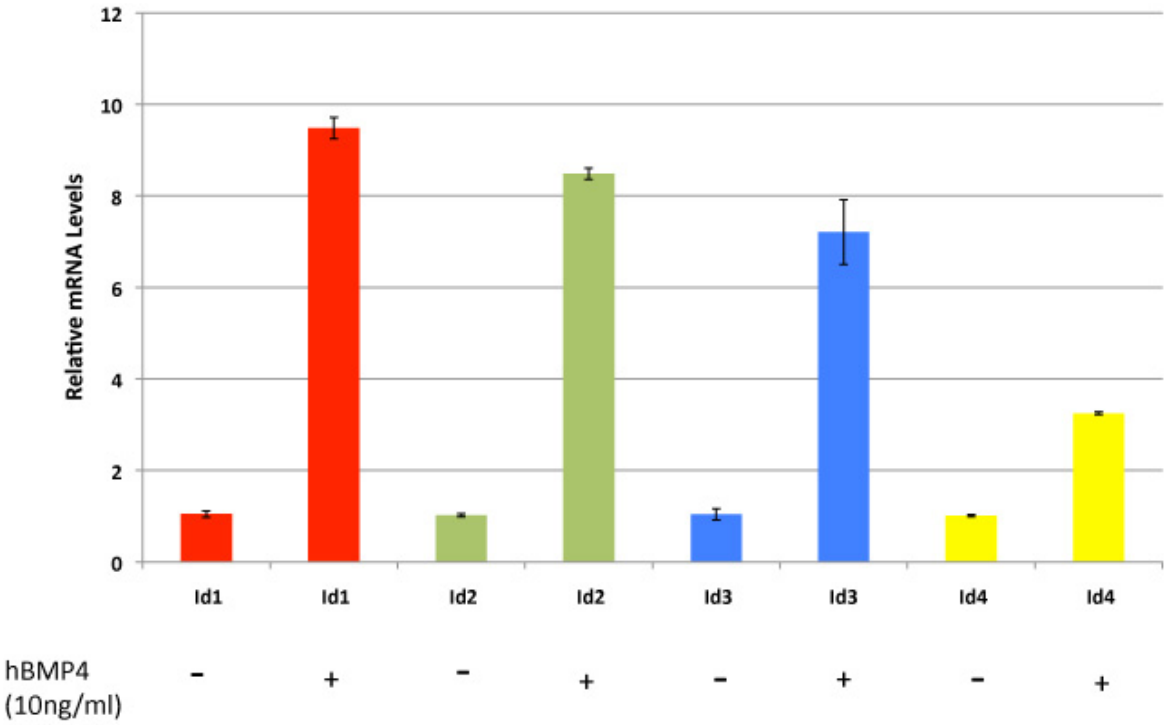


Figure 4.3: Induction of *Id* genes by BMP4 in mouse ES cells. Mouse ES cells were cultured in the presence and absence of BMP4. RNA was isolated and quantitative real-time PCR was performed for *Id1-4* genes. BMP4 induced expression of all Id genes. In ES cells, *Id2* and *Id3* expression are at similar levels, whereas in the ICM of the embryo, *Id3* (~424 RPKM) expression is much higher than *Id2* (~6RPKM).

Figure 4.4:

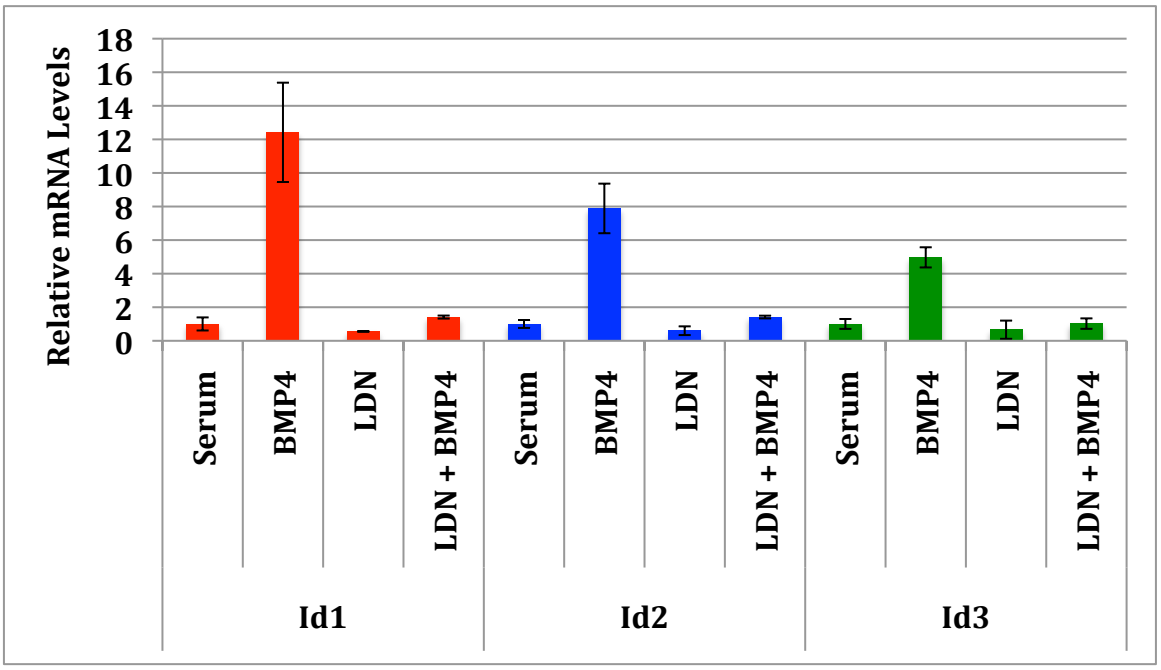


Figure 4.4: LDN193189 specifically inhibit BMP signaling in mouse ES cells. Mouse ES cells were treated in the presence of BMP4 (20ng/ml) alone, LDN193189 (80nM) alone, or both. qPCR analyses of *Id* genes were performed. For all *Id* genes, their transcript levels decreased after LDN193189 treatment.

Figure 4.5

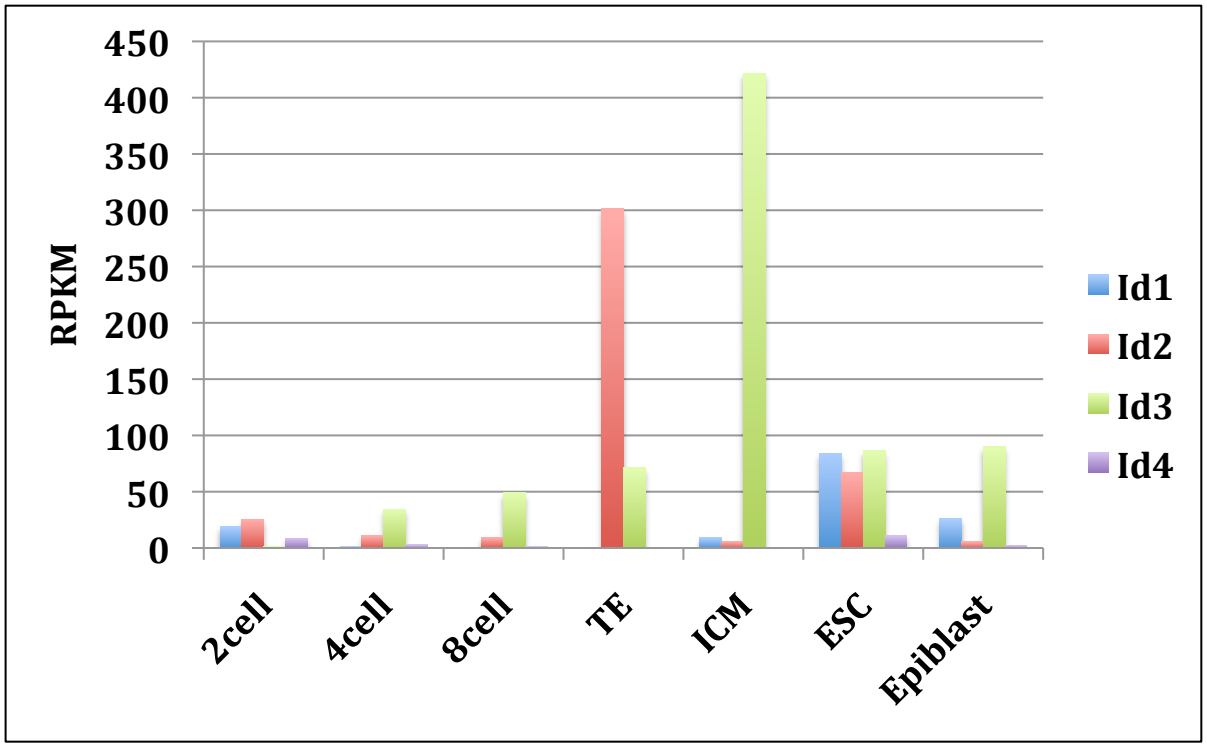


Figure 4.5: Expression of *Id* genes during early mouse development from single-cell RNA-Seq. RPKM levels were obtained for BMP ligands/receptors and Activin receptors at different developmental stages: 2-cell stage, 4-cell stage, 8-cell stage, Inner Cell Mass (ICM), Trophectoderm, and Epiblast. *Id2* is enriched in trophectoderm whereas *Id3* is enriched in ICM. Published data was obtained from Tang et al., 2011.

Figure 4.6

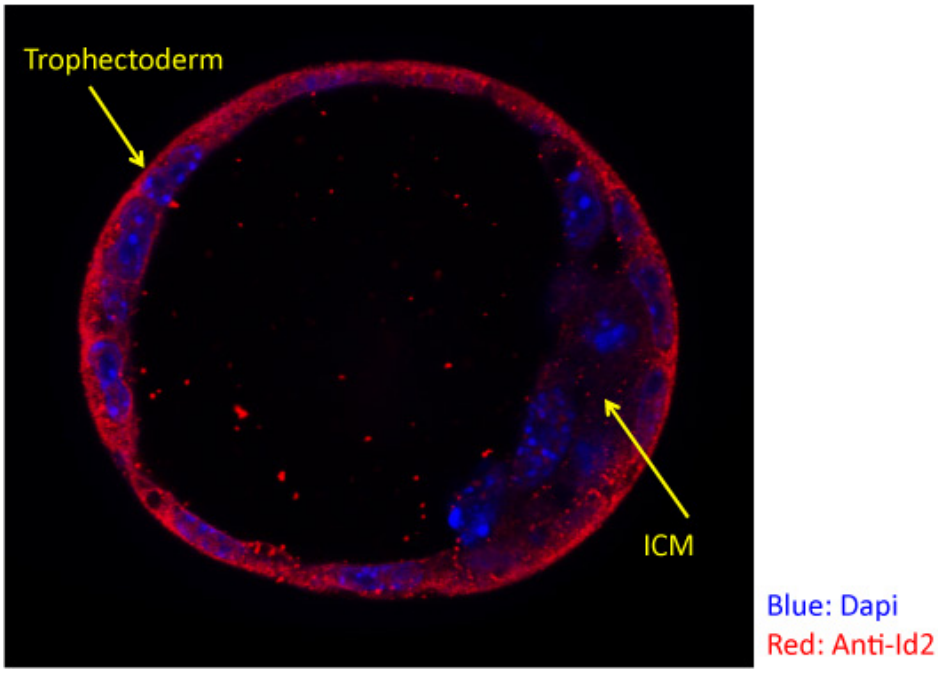


Figure 4.6: Id2 immunostaining expression in trophectoderm of E3.5 mouse blastocyst. Anti-Id2 immunostaining shows Id2 expression primarily in the trophectoderm and not the ICM. Id2 is localized strongly in the cytoplasm.

Chapter 4 References:

Beddington RS, Robertson EJ. Axis development and early asymmetry in mammals. Cell. 1999 Jan 22;96:195-209.

Benezra R, Davis RL, Lassar A, Tapscott S, Thayer M, Lockshon D, Weintraub H. Id: a negative regulator of helix-loop-helix DNA binding proteins. Control of terminal myogenic differentiation. Ann N Y Acad Sci. 1990;599:1-11.

Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell. 1990 Apr 6;61:49-59.

Beppu H, Kawabata M, Hamamoto T, Chytil A, Minowa O, Noda T, Miyazono K. BMP type II receptor is required for gastrulation and early development of mouse embryos. Dev Biol. 2000 May 1;221:249-58.

Biggs J, Murphy EV, Israel MA. A human Id-like helix-loop-helix protein expressed during early development. Proc Natl Acad Sci U S A. 1992 Feb 15;89:1512-6.

Bounpheng MA, Dimas JJ, Dodds SG, Christy BA. Degradation of Id proteins by the ubiquitin-proteasome pathway. FASEB J. 1999 Dec;13:2257-64.

Brennecke P, Anders S, Kim JK, Kołodziejczyk AA, Zhang X, Proserpio V, Baying B, Benes V, Teichmann SA, Marioni JC, Heisler MG. Accounting for technical noise in single-cell RNA-seq experiments. Nat Methods. 2013 Nov;10:1093-5.

Chang H, Huylebroeck D, Verschueren K, Guo Q, Matzuk MM, Zwijsen A. Smad5 knockout mice die at mid-gestation due to multiple embryonic and extraembryonic defects. Development. 1999 Apr;126:1631-42.

Chang H, Matzuk MM. Smad5 is required for mouse primordial germ cell development. Mech Dev. 2001 Jun;104:61-7.

Christy BA, Sanders LK, Lau LF, Copeland NG, Jenkins NA, Nathans D. An Id-related helix-loop-helix protein encoded by a growth factor-inducible gene. Proc Natl Acad Sci U S A. 1991 Mar 1;88:1815-9.

Coucouvanis E, Martin GR. BMP signaling plays a role in visceral endoderm differentiation and cavitation in the early mouse embryo. Development. 1999 Feb;126:535-46.

Deed RW, Bianchi SM, Atherton GT, Johnston D, Santibanez-Koref M, Murphy JJ, Norton JD. An immediate early human gene encodes an Id-like helix-loop-helix protein and is regulated by protein kinase C activation in diverse cell types. Oncogene. 1993 Mar;8:599-607.

Deed RW, Armitage S, Norton JD. Nuclear localization and regulation of Id protein through an E protein-mediated chaperone mechanism. J Biol Chem. 1996 Sep 27;271:23603-6.

Doan LT, Javier AL, Furr NM, Nguyen KL, Cho KW, Monuki ES. A Bmp reporter with ultrasensitive characteristics reveals that high Bmp signaling is not required for cortical hem fate. PLoS One. 2012;7:e44009.

Ellmeier W, Weith A. Expression of the helix-loop-helix gene Id3 during murine embryonic development. Dev Dyn. 1995 Jun;203:163-73.

Evans SM, O'Brien TX. Expression of the helix-loop-helix factor Id during mouse embryonic development. Dev Biol. 1993 Oct;159:485-99.

Guo G, Huss M, Tong GQ, Wang C, Li Sun L, Clarke ND, Robson P. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. Dev Cell. 2010 Apr 20;18:675-85.

Javier AL, Doan LT, Luong M, Reyes de Mochel NS, Sun A, Monuki ES, Cho KW. Bmp indicator mice reveal dynamic regulation of transcriptional response. PLoS One. 2012;7:e42566.

Jen Y, Manova K, Benezra R. Each member of the Id gene family exhibits a unique expression pattern in mouse gastrulation and neurogenesis. Dev Dyn. 1997 Jan;208:92-106.

Jen Y, Manova K, Benezra R. Expression patterns of Id1, Id2, and Id3 are highly related but distinct from that of Id4 during mouse embryogenesis. Dev Dyn. 1996 Nov;207:235-52.

Jen Y, Weintraub H, Benezra R. Overexpression of Id protein inhibits the muscle differentiation program: in vivo association of Id with E2A proteins. Genes Dev. 1992 Aug;6:1466-79.

Kee Y, Bronner-Fraser M. To proliferate or to die: role of Id3 in cell cycle progression and survival of neural crest progenitors. Genes Dev. 2005 Mar 15;19:744-55.

Kurooka H, Yokota Y. Nucleo-cytoplasmic shuttling of Id2, a negative regulator of basic helix-loop-helix transcription factors. J Biol Chem. 2005 Feb 11;280:4313-20.

Lawson KA, Dunn NR, Roelen BA, Zeinstra LM, Davis AM, Wright CV, Korving JP, Hogan BL. Bmp4 is required for the generation of primordial germ cells in the mouse embryo. Genes Dev. 1999 Feb 15;13:424-36.

Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature. 2005 Feb 17;433:769-73.

Mazerbourg S, Sangkuhl K, Luo CW, Sudo S, Klein C, Hsueh AJ. Identification of receptors and signaling pathways for orphan bone morphogenetic protein/growth differentiation factor ligands based on genomic analyses. J Biol Chem. 2005 Sep 16;280:32122-32.

Mishina Y, Suzuki A, Ueno N, Behringer RR. Bmpr encodes a type I bone morphogenetic protein receptor that is essential for gastrulation during mouse embryogenesis. Genes Dev. 1995 Dec 15;9:3027-37.

Miyazono K, Kamiya Y, Morikawa M. Bone morphogenetic protein receptors and signal transduction. J Biochem. 2010 Jan;147:35-51.

Murre C, McCaw PS, Baltimore D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell. 1989 Mar 10;56:777-83.

Niakan KK, Schrode N, Cho LT, Hadjantonakis AK. Derivation of extraembryonic endoderm stem (XEN) cells from mouse embryos and embryonic stem cells. Nat Protoc. 2013 Jun;8:1028-41.

Nichane M, de Crozé N, Ren X, Souopgui J, Monsoro-Burq AH, Bellefroid EJ. Hairy2-Id3 interactions play an essential role in *Xenopus* neural crest progenitor specification. Dev Biol. 2008 Oct 15;322:355-67.

Norton JD, Deed RW, Craggs G, Sablitzky F. Id helix-loop-helix proteins in cell growth and differentiation. Trends Cell Biol. 1998 Feb;8:58-65.

Reyes de Mochel NS, Luong M, Chiang M, Javier AL, Luu E, Toshihiko F, MacGregor GR, Cinquin O, Cho KW. BMP signaling is required for cell cleavage in preimplantation-mouse embryos. Dev Biol. 2015 Jan 1;397:45-55.

Riechmann V, Sablitzky F. Mutually exclusive expression of two dominant-negative helix-loop-helix (dnHLH) genes, Id4 and Id3, in the developing brain of the mouse suggests distinct regulatory roles of these dnHLH proteins during cellular proliferation and differentiation of the nervous system. Cell Growth Differ. 1995 Jul;6:837-43.

Riechmann V, van Crüchten I, Sablitzky F. The expression pattern of Id4, a novel dominant negative helix-loop-helix protein, is distinct from Id1, Id2 and Id3. Nucleic Acids Res. 1994 Mar 11.

Roelen BA, Goumans MJ, van Rooijen MA, Mummery CL. Differential expression of BMP receptors in early mouse development. Differential expression of BMP receptors in early mouse development. Int J Dev Biol. 1997 Aug;41:541-9.

Rossant J, Tam PP. Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. Development. 2009 March;136:701–713.

Samanta J, Kessler JA. Interactions between ID and OLIG proteins mediate the inhibitory effects of BMP4 on oligodendroglial differentiation. Development. 2004 Sep;131:4131-42.

Sun XH, Copeland NG, Jenkins NA, Baltimore D. Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. Mol Cell Biol. 1991 Nov;11:5603-11.

Tang F, Barbacioru C, Bao S, Lee C, Nordman E, Wang X, Lao K, Surani MA. Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. Cell Stem Cell. 2010 May 7;6:468-78.

Tang F, Barbacioru C, Nordman E, Bao S, Lee C, Wang X, Tuch BB, Heard E, Lao K, Surani MA. Deterministic and stochastic allele specific gene expression in single mouse blastomeres. PLoS One. 2011;6:e21208.

Tremblay KD, Dunn NR, Robertson EJ. Mouse embryos lacking Smad1 signals display defects in extra-embryonic tissues and germ cell formation. Development. 2001 Sep;128:3609-21.

Tu X, Baffa R, Luke S, Prisco M, Baserga R. Intracellular redistribution of nuclear and nucleolar proteins during differentiation of 32D murine hemopoietic cells. Exp Cell Res. 2003 Aug 1;288:119-30.

Viswanathan SR, Mermel CH, Lu J, Lu CW, Golub TR, Daley GQ. microRNA expression during trophectoderm specification. PLoS One. 2009 Jul 3;4:e6143.

Wang S, Sdrulla A, Johnson JE, Yokota Y, Barres BA. A role for the helix-loop-helix protein Id2 in the control of oligodendrocyte development. Neuron. 2001 Mar;29:603-14.

Winnier G, Blessing M, Labosky PA, Hogan BL. Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. Genes Dev. 1995 Sep 1;9:2105-16.

Ying QL, Nichols J, Chambers I, Smith A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell. 2003 Oct 31;115:281-92.

Zhao GQ. Consequences of knocking out BMP signaling in the mouse. Genesis. 2003 Jan;35:43-56.

Zhu W, Dahmen J, Bulfone A, Rigolet M, Hernandez MC, Kuo WL, Puelles L, Rubenstein JL, Israel MA. Id gene expression during development and molecular cloning of the human Id-1 gene. Brain Res Mol Brain Res. 1995 Jun;30:312-26.

Chapter 5

The Role of BMP Signaling in Oct4 Depleted Mouse ES cells.

The first cell differentiation event in mouse development is the establishment of two distinct cell lineages: the trophectoderm (TE) and the inner cell mass (ICM). Transcription factors are critical regulators of cell fate determination. Oct4 (Pou5f1), a key transcription factor, in early mouse development was first identified nuclear extracts from different embryonic stages (Lenardo et al., 1989; Scholer et al., 1989; Okamoto et al., 1990). Maternal *Oct4* mRNA can be found in unfertilized oocyte. After fertilization, a decrease in maternal *Oct4* occurs between the 2- and 4-cell stages of development. Following zygotic genome activation, upregulation of *Oct4* expression is observed at the 2.5-dpc (days post coitum) morula stage and 3.5-dpc blastocyst. Rapid downregulation of Oct4 protein levels begins in cells differentiating into the trophectoderm lineage, while Oct4 remains strongly expressed in the ICM (Okamoto et al., 1990; Rosner et al., 1990; Scholer et al., 1990). *Oct4* null mutant embryos fail to properly segregate into the ICM and TE and readily differentiate into trophoblast giant cells (derived from TE). Downregulation of *Oct4* in mouse ES cells induces trophectoderm formation (Niwa et al., 2000). Maintenance of *Oct4* expression is critical for sustaining ES self-renewal and pluripotency.

Another well-studied transcription factor important in early mouse development is Sox2. *Sox2* belongs to the *Sry* gene family and contains a DNA-binding domain referred to as a high-mobility group (HMG) domain. Like Oct4, Sox2 is a maternal factor that is specifically expressed in the ICM and primitive ectoderm (Keramari et al., 2010). *Sox2* expression is detected in both the ICM and TE, but its expression eventually becomes restricted to the ICM (Avilion et al., 2003, Sarkar et al., 2013). Maternal and zygotic

knockdown of *Sox2* results in an incomplete trophectoderm in fertilized embryos, which failed to progress beyond the morula stage (Keramari et al., 2010). This suggests that *Sox2* is involved in the segregation of ICM and TE in preimplantation embryos. Furthermore, *Oct4* and *Sox2* work cooperatively to regulate self-renewal in ESC (Yuan et al., 1995; Boyer et al., 2005).

Similar to Oct4 and Sox2, the Nanog transcription factor is also a master regulator of ES cell pluriotentcy and self-renewal. Nanog, Oct4, and Sox2 transcription factors have been shown to repress the expression of developmental genes while modulating their own expression levels by binding to each other's promoter regions (Mitsui et al., 2003; Boyer et al., 2005). Mouse embryos lacking *Nanog* fail to develop beyond the blastocyst stage due to the absence of epiblasts. *Nanog*-/- ES cells can be derived from the *Nanog*-/- blastocyst, but display slow differentiation to extra embryonic endoderm *in vitro* (Mitsui et al., 2003). When *Nanog* is overexpressed, mouse ES cells also do not need LIF to maintain its pluripotency. Overall, these finding suggests *Nanog* as a critical player in sustaining ES stem cell properties by preventing it from giving rise to the primitive endoderm lineage (Chambers et al., 2003; Mitsui et al., 2003; Kuroda et al., 2005; Rodda et al 2005).

During trophectoderm specification, the caudal-like homeodomain transcription factor, Cdx2, is a critical player. *Cdx2* expression can be observed at the 8-cell compact stage in a stochastic manner and later restricted to the TE (Dietrich et al., 2007; Ralston et al., 2008). *Cdx2*-null embryos show a collapsed blastocoel, retain positive *Oct4* and *Nanog* expression, and fail to out grow into trophoblast cells (Strumpf et al., 2005). *Oct4* and *Cdx2* mutually suppress the expression of one another in the ICM and TE (Niwa et al., 2005). Similar mutual suppression is also observed between *Cdx2* and *Nanog* (Chen et al., 2009).

Mouse preimplantation embryo is a very attractive model to study regulatory networks that are involved in controlling cell fate decisions, in particular, the ICM and trophectoderm lineage segregation. The molecular crosstalks of pluripotent transcription factors such as Oct4, Nanog, Sox2 within the ICM and Cdx2 in the trophectoderm, have been studied extensively in reference to early mouse development. *Oct4*, *Sox2*, and *Nanog* are important in the formation and maintenance of the ICM and *Cdx2* in the TE. Null embryos for these transcription factors are incapable of deriving ES cells from the ICM. In *Oct4* and *Nanog*-null blastocysts, they lack epiblast and primitive endoderm characteristics and can only give rise to trophoblast derivatives. Similarly, Sox2 null embryos lack the ability to sustain ICM identity (Guo et al., 2010). To fully understand the molecular mechanisms and key players during the formation of these two distinct lineages, an investigation into how intrinsic and extrinsic factors may influence ICM and TE formation is needed. It has been suggested that BMP, an extrinsic factor, may be a player in cross talks between these pluripotent transcription factors (Chen et al., 2008; Sridharan et al., 2009). Evidence suggest that *Bmp4* may also be a bona fide target of Oct4, Sox2, and Klf4 grouped transcription factors in induced pluripotent cells (iPS) and that *Oct4* may contain a Smad1 binding site (Sridharan et al., 2009) in its promoter region. Additionally, there is also proof suggesting that Nanog and Smad1 can physically interact through the conserved C-terminal MH2 domain of Smad1 (Suzuki et al., 2006). We hypothesize that BMPs are possible key players in early mouse development and can integrate into the *Oct4*/*Sox2*/*Nanog* regulatory network to sustain ES properties.

Material and Methods:

Mouse ES Cell Culture Conditions:

Zhbtc4 mouse ES cells were cultured on 0.1% gelatin-coated plates containing GMEM medium (Sigma Aldrich, #G5153) with 10% FBS (Hyclone, catalog #SH30071.03), 0.1mM NEAA, 1mM Sodium Pyruvate, 1000U/ml LIF (Millipore, catalog #ESG1107), and 0.1mM 2-Mercaptoethanol. Cells were maintained at 37°C, 5% CO_2, and split using 0.25% trypsin (Life Technologies, catalog #15400-054). Tetracycline (MP Biomedical) was added to the media to differentiate the cells as needed.

Western Immunoblotting:

Zhbtc4 mouse ES cells were collected in RIPA buffer (150mM NaCl, 50mM Tris-HCl pH8.0, 1% NP40, 0.5% sodium deoxycholate, 0.1% SDS, 25mM β-glycerophosphate, 100nM NaF, and 1mM Na_3VO_4) and 20µg of protein lysates were electrophoresed through a 10% PAGE-SDS gel. Following transfer of gel to PVDF membrane, it was probed with the following antibodies: anti-Oct4 (1:200; SC-5279, Santa Cruz Biotechnology) and anti-β-tubulin (1:10,000; T5168, Sigma). ECL kit (GE Healthcare, catalog #RPN2232) was used for visualization.

Quantitative Real-time PCR:

RNA was isolated after treatment of Zhbtc4 cells with tetracycline using Trizol (Life Technologies, catalog #15596-026). Reverse transcription was performed using MMLV Reverse Transcriptase (Invitrogen, catalog #28025-013). Quantitative PCR was performed using SYBR Green (Roche, catalog #04-707-516-001).

Table 5.1: Primers for qPCR in Zhbtc4 ES cells

Gene	Forward	Reverse
mPouf1	AGAGGATCACCTTGGGGTACA	CGAAGCGACAGATGGTGGTC
mCdx2	GCAGTCCCTAGGAAGCCAAGTGA	CTCTCGGAGAGCCCGAGTGTG
mBMP2	CACCAGGTTAGTGAATCAGAACAC	AGACACCTGGGTTCTCCTCTAAAT
mBMP4	GTGCAGACCCTAGTCAACTCTGTTA	CCTCTACCACCATCTCCTGATAAT
mBMP7	ACCTCCAGGGAAAGCATAATTC	GGTACTGAAGACGGCCTTGTAG
mId1	GACATGAACGGCTGCTACTCAC	GACTTCAGACTCCGAGTTCAGC
mId2	GTGAGGTCCGTTAGGAAAAACAG	GTCGTTCATGTTGTAGAGCAGACT
mId3	CTGTCGGAACGTAGCCTGG	GTGGTTCATGTCGTCCAAGAG

Zhbtc4 RNA-Seq Library Generation:

RNA was extracted from Zhbtc4 ES cells post tetracycline treatment: 0hrs, 24hrs, 48hrs, 72hrs using Trizol (Life Technologies, catalog #15596-026). Total RNA profiles were recorded using a Bioanalyzer 2100 (Agilent). RNA quality was determined by the RNA Integrity Number (RIN) value. RIN values >9.0 were used for library generation. Following RNA quality check, RNA sequencing libraries were prepared using the Bioo Scientific Kit (Cat# 5143-01). 1ug of total RNA was used to generate the sequencing libraries. RNA-Seq transcriptome sequencing was performed by UCI Genomics High-Throughput Facility using the Illumina Hi-Seq 2000 platform.

RNA-Seq Computational Pipeline:

RNA-Seq data analysis was accomplished in two sequential steps: (1) map all raw reads to the mouse reference transcriptome (mm9/RefGene) as the gene model, and (2) count the read fragments mapped to each individual gene and quantify expression in terms of RPKM (Reads Per Kilobase per Million mapped reads). Raw RNA-seq reads were mapped using the Bowtie 0.12.8 software allowing a two-read mismatch in the alignment (-m 1) parameters to the reference transcriptome. Output of mappings were then used in Express (http://bio.math.berkeley.edu/eXpress/overview.html) for read quantiation and Edge R for differential analysis (http://www.bioconductor.org/packages/release/bioc/html/edgeR.html). Data were then clustered using Cluster3.0 (http://bonsai.hgc.jp/~mdehoon/software/cluster/software.html) to generate a map of genes that were constant (no change) or changing (upregulated or downregulated). Cluster3.0 parameter: Cluster "Genes", centered median, and centroid linkage. Cluster output (e.g. javatree.cdt) was input into Java Treeview. Treeview allows visualization of the hierarchical clustering in a form of a heatmap diagram.

Nuclei Isolation and DNase I Isolation:

Zhbtc4 ES cells (~30 million) were collected and resuspended in 1xPBS media. Cells were then spun at 500xg at 4^{o}C for 5mins to pellet cells. The supernatant was then removed and the cells were washed with 1xPBS. The cells were pelleted again. The cells nuclei were later resusupended in 6ml of ice cold Buffer A (15mM Tris-HCl, pH 8.0; 15mM NaCl, 60mM KCl, 1mM EDTA, pH 8.0, 0.5mM EGTA, pH 8.0, 0.5mM Spermidine). Next, equal volume of 2x IGEPAL (Buffer A with IGEPAL CA-630 at a final concentration of 0.06%) was added to the resupended cell mixture and placed on ice for 5-6mins. We then pelleted the

nuclei for 5 minutes at 500xg at 4°C and then washed nuclei with cold Buffer A. After washing, the nuclei were pelleted for 5 minutes at 500xg at 4°C. The supernatant was removed and the nuclei was resuspended with 1X DNASE I Digestion Buffer (6mM CaCl2, 75mM NaCl) containing DNASE I enzyme in 15mL tube. The reaction mixture was incubated for 3minutes at 37°C in a water bath. Immediately after the incubation, STOP Buffer (50mM Tris-HCl, pH 8.0, 100mM NaCl, 0.10% SDS, 100mM EDTA, pH 8.0, 50μg/ml Proteinase K) was added to the mixture. The mixture was then transferred to a 55°C water bath and incubated overnight. The next day, RNase A was added to the mixture and incubated for 1 hr. Phenol-chloroform extraction was performed twice. DNase I-digested DNA fragments were precipitated using 1/10 volume of 3M sodium acetate and 1 volume of isopropanol and icubated overnight at -20°C. The precipiation mixture was spun at 3,000xg for 20minutes to pellet the DNA fragments. The DNA fragments were finally resupsended in 600μl of TE.

Gel Extraction for Size Selection:

The DNase I-digested DNA fragments were ran on a 1.5% low-melt agarose gel for size selection. We gel extracted DNA fragments between 100bp to 500bp. A Qiagen gel extraction kit was used to retrieve the DNA fragments. DNA concentrations were obtained using the Invitrogen Qubit machine.

Quantitative PCR Validation of DNase I-Treated Samples:

Quantitative PCR was used to verify the efficiency of the DNase I digestion by identification of known DNase hypersensitive sites (DHSs). We selected six known promoter regions that would capture DHSs and designed PCR primers to those regions. Each primer pair was designed to generate a 70-100bp product using Primer3.

Quantitative PCR was performed using SYBR Green.

Table 5.2: Primers for DNase-Seq Analysis

Promoter	Forward	Reverse
mRpl41	CTGCCTTACTAGAGTGCGTGTTATC	AGTCCCATACTGGAAACAGGTG
mOct4	AGCAACTGGTTTGTGAGGTGTC	GAGGCCTTCATTTTCAACCTTC
mSox2	GAGTTGTCAAGGCAGAGAAGAGAGT	CCTAGTCTTAAAGAGGCAGCAAAG
mNanog	GGGATGTCTTTAGATCAGAGGATG	AGAAGCTGTAAGGTGACCCAGACT
mCdx2	GTACTGCGGAGGACTGACAAAGT	ACAAGGACGTGAGCATGTATCCTA
mFGFR1	CAAGTTAGACCAGACTGTAGAGCAG	GGGCCTTGTAGTTCTCAAGCTG
mEomes	CTAAAGCATGCAGTTGGGAGAG	AGCACTCTCCAGCGAGTAGAAGT
mIvl	CAACCCCTCCTAACTGAGATCAT	GTAGGATACCTTTCTCACTCCCTTC
mTtn	GGTGTAAGATTAGTACAGGGTGTGC	CCCTAAACACTTCCCATTAGTGAC
mKrt86	ATGGAGGAGGTAGAAGGTGACTAAC	ATAGTTCTTCTCCGTGCCCCTTA

DNase I Library Preparation:

DNase sequencing libraries were prepared using the Bioo Scientific Kit (Cat#5143-01) according to manufacture's direction. 50ng of DNA was used to generate the sequencing libraries. DNase sequencing was performed by UCI Genomics High-Throughput Facility using the Illumina Hi-Seq 2000 platform.

DNase-Seq Analysis:

DNase sequencing reads were mapped to Bowtie 0.12.8 software allowing a two-read mismatch in the alignment (-m 1) parameters to the reference genome (mm9). Prior

to read mapping, reads were trimmed at the 3' end by 14bp to ensure better alignment. Most likely, DNase I digestion will occur at the 5' end of the reads so trimming the reads would not affect the mapping results significantly. The mapped output from Bowtie was then input into Hotspot V4 (http://www.uwencode.org/proj/hotspot/) for peakcalling. Peaks were then visualized using the UCSC genome browser. The associated genes of the called peaks were annotated using the Homer software (http://homer.salk.edu/homer).

Results and Discussion:

Chapter 5.1: Repression of Oct4 in mouse ES cells leads to differentiation

Oct4 is an essential transcription factor in mouse embryogenesis and deletion of *Oct4* results in lost of pluriopotency and promotion of ICM/ES cells into extraembryonic lineages (Nichols et al., 1998; Niwa et al., 2000). Transcription factors such as Oct4 often act as molecular switch to activate or repress gene expression akin to a binary on/off switch. To investigate the expression and repression of Oct4, we obtained a conditional *Oct4* mouse ES cell (Zhbtc4) line from Dr. Hitoshi Niwa (RIKEN Institute, Japan). The Zhbtc4 ES cells have had both endogenous alleles of *Oct4* inactivated by gene targeting and contain a regulatable tetracycline (Tc)-dependent transactivator *Oct4* transgene. These cells can be propagated as ES cells in the absence of Tc, in which the *Oct4* transgene is active. However, in the presence of Tc, the *Oct4* transgene is repressed (Niwa et al., 2000). The Zhbtc4 ES cell line is indistinguishable from wild type mouse ES cells in the absence of Tc (Niwa et al., 2000). The advantage of using the Zhbtc4 ES cells is two fold: A) The cell line can induce Oct4 knockdown when required, and B) The knockdown of Oct4 is homogeneous, which means there is a complete knockdown of Oct4 in all cells. The Zhbtc4

ES cells are morphologically compact and undifferentiated (Figure 5.1A) but upon Tc treatment for 48hrs, differentiation of cells can be observed by its flattened and spreadout morphology and enlarged nuclei (Figure 5.1A). The morphologically differentiated appearance of ES cells upon Tc treatment is similar to LIF (Ying et al., 2003) removal in wild type ES cells. We then examined when Oct4 knockdown occurs by performing a time course Tc treatment on the Zhbtc4 ES cells. Oct4 protein levels were assayed by western blot. The Tc treatment time course experiment was performed at 3hrs, 6hrs, 12hrs, and 18hrs. Western blot using an anti-Oct3/4 antibody showed that Oct4 protein levels gradually declined over time in the presence of Tc with virtually no Oct4 expression at 18hrs (Figure 5.1B). We next investigated the expression of marker genes to evaluate how they are affected upon Tc treatment. Quantitative real-time PCR was performed to measure *Oct4* and *Cdx2* transcript levels. During Tc treatment over 72hrs, *Oct4* transcript level was completely gone whereas the trophectoderm marker, *Cdx2*, transcript level was substantially increased after 24hrs with the highest expression observed at 48hrs (Figure 5.1C). The expression of *Cdx2* and cellular morphoglogy of Tc-treated Zhbtc4 ES cells demonstrated that Zhbtc4 differentiated cells are similar to trophectoderm cells. Both the loss and gain of function of markers of gene expression in the presence of Tc are consistent with observed phenotypic differences in the Zhbtc4 ES cells. These findings demonstrate that ES cells can be respecified to a different fate upon Oct4 depletion and that the Zhbtc4 cell line could be utilized as an *in vitro* model for studying early trophectoderm differentiation. We conclude that *Oct4* functions as a gatekeeper to prevent ES cell differentiation into the trophectoderm lineage.

Chapter 5.2: Oct4 depletion leads to downregulation of BMP4 and its target genes

In mouse, loss of *Oct4* in the embryo results in developmental arrest before the blastocyst stage (Nichols et al., 1998). Functional loss of *Oct4* produces severe gastrulation and patterning abnormalities and results in embryonic lethality. During early embryogenesis, several signaling pathways are activated to establish a genetic network to control cell fate decisions. Soledad Reyes de Mochel, in the Cho lab, has implicated BMP activity in early embryo development (Reyes de Mochel et al., 2014). Additionally, Soledad has also shown that BMP4 is active during the same time when the *Oct4/Sox2/Nanog* network becomes established. It can therefore be hypothesized that extrinsic factors, such as BMPs, may possibly feed into the *Oct4/Sox2/Nanog* transcriptional network to shed light into the molecular mechanisms that coordinate control of ES cell pluripotency and cellular differentiation.

Utilizing the Zhbtc4 ES cell line, we examined whether Oct4 depletion affected *BMP* transcript levels. We measured transcript levels for *BMP2*, *BMP4*, and *BMP7* and results showed that *BMP4* is significantly downregulated upon Oct4 knockdown whereas *BMP2* and *BMP7* show very moderate increases (Figure 5.2A). Since BMPs are essential for ES cell mainteance and knockdown of Oct4 leads to a downregulation of BMP4 ligand levels, these findings support the hypothesis that *Oct4* and *BMP4* are part of the same regulatory network. However, whether Oct4 directly or indirectly regulate *BMP4* expression has yet to be determined.

We also examined whether BMP target genes were affected after Oct4 depletion. The *Id1, Id2*, and *Id3* transcript levels were measured after tetracycline treatment in the Zhbtc4 ES cells. We observed that only *Id1* and *Id3* are downregulated during Oct4

knockdown whereas *Id2* is upregulated (Figure 5.2B). In Chapter 4, we indicated that *Id3* is mainly expressed in the ICM and the *Id2* in the trophectoderm. The upregulation of *Id2* is consistent with our previous findings (see Chapter 4) that *Id2* is important for maintaining trophectoderm identity. Since Oct4 depletion leads to ES cell differentiation to trophectoderm-like cells, the upregulation of *Id2*, confirms its involvement in the trophectoderm lineage. In human embryonic stem cells, Oct4 was demonstrated to negatively regulate *Id2* (Babaie et al., 2007) in differentiating ES cells and our findings suggest that a similar mechanism could be at play in mouse ES cells. In the Zhbtc4 ES cells, upon Oct4 knockdown, *Id2* repression is relieved and an upregulation of *Id2* expression is observed. Confirmation of whether *Id2* is directly regulated by Oct4 in mouse ES cells can be addressed by performing Oct4-ChIP qPCR. It would also be interesting to determine if Oct4 can also positively or negatively regulate additional BMP pathway target genes.

Chapter 5.3: Time Course RNA-Seq of Oct4 Depleted Mouse ES Cells

In vivo, Oct4 proteins are abundantly expressed in ES cells and are quickly downregulated as ES cells differentiate into the trophectoderm lineage (Niwa et al., 2000). Oct4 activates transcription of specific genes to maintain ES self-renewal and proliferation (Boyer et al., 2005). To define the temporal transcriptional signature during this differentiation process, we performed RNA-Seq on Zhbtc4 ES cell line to study the transcriptional dynamics during this critical developmental window.

Zhbtc4 ES cells were treated with tetracycline for a period of 3 days and total RNA was collected every 24hrs at O, 24, 48, and 72hrs. Biological replicates of RNA-seq libraries were prepared for each time point. After performing RNA-Seq analysis, 1480 differentially

expressed genes were observed (1% FDR cut-off) across the time course. As a quality control test, we wanted toconfirm whether mapped RNA-Seq reads decreased in the region covering the *Oct4* gene exons. Indeed mapped reads in the *Oct4* exon dropped precipitously after 24hrs of Tc treatment (Figure 5.2C). A heatmap of the differentially expressed genes are shown in Figure 5.3A. The heatmap shows hierarchical clustering of genes that are upregulated and downregulated upon differentiation. Interestingly, untreated Zhbtc4 (WT) and 24hrs tetracycline-treated Zhbtc4 cells (Tc_24hrs) clustered closely together indicating that their transcriptome profiles are still moderately similar even after Oct4 knockdown. However, after 48hrs of treatment (Tc_48hrs), the transcriptome profiles of the differentiating ES cells were dramatically different. At this point, genes that were previously downregulated (color:blue) are now upregulated and vice versa (color: red). As differentiation progressed, the transcriptome profile of ES cells at 72hrs tetracycline treatment was significantly different from the untreated Zhbtc4 ES cells. Also, it is apparent from our principal component analysis (PCA) that the 72hrs-differentiated cells are very different from the untreated and 24hrs Tc-treated ES cells as they cluster far apart from one another (Figure 5.3B).

It is well known that cellular differentiation is a process of sequential induction of regulatory genes that in turn initiates the expression of lineage specific target genes. Each developmental process requires the activation of a specific transcriptional program. Our data showed global gene expression changes that take place upon differentiation. These genes represent novel candidates that are likely to be important in driving mammalian preimplantation development. Together, this data set provides a valuable resource to

search for not only novel trophectoderm marker genes but also allows us to dissect gene regulatory mechanisms underlying progressive development of early mammalian embryos.

Chapter 5.4: Time Course DNase-Seq of Oct4 Depleted Mouse ES Cells

Identification of active gene regulatory elements is important to understanding the transcriptional control governing cell fate decisions during early embryogenesis. The presence of DNase I hypersensitive sites (DHS) has served as useful evidence for identifying active cis-regulatory elements including enhancers and promoters (Galas et al., 1978). DNase hypersensitivity assay takes advantage of the fact that DNase I preferentially cleaves DNA at sites of "open/accessible" chromatin. Most DNA is compacted into chromatin consisting of DNA tightly wound around nucleosomes and is inaccessible to DNase I treatment. Open chromatin regions are more susceptible to DNase I cleavage and most often these open chromatin regions contain promoter and enhancer regions. DNase I-sequencing adapts from traditional DNase I footprinting assay and leverages modern DNA sequencing to identify regions of the genome where regulatory factors interact with DNA to modify chromatin structure and gene transcription (Galas et al., 1978; Boyle et al., 2011; Pique-Regi et al., 2011; Song et al., 2011, Neph et al., 2012). Using this technique (Figure 5.4A), we identified cell and lineage specific regulators during Zhbtc4 differentiation.

To determine the optimal DNase I enzyme concentration condition to generate high-quality and reproducible DNase-Seq libraries, I first titrated the amount of DNase I enzyme required for optimal DNase I cleavage. I used concentrations between 20-80units to digest nuclei from untreated Zhbtc4 ES cells. After DNase I digestion, I performed qPCR to quantify open promoter regions of genes (*Rpl41, Oct4, Sox2, Nanog*) that DNase I is known

to target for cleavage and closed promoter regions of lowly expressed genes (*Ivl, Ttn, Krt86*). The open and closed chromatin regions were chosen based on their RNA-Seq levels and known regions of transcription factor binding. At 20units of DNase I treatment, I did not observe cleaved DNA fragments but at 40units of DNase I, I observed a marked increase in the recovery of DNA fragments corresponding to the open chromatin regions of the *Rpl41, Oct4, Sox2*, and *Nanog* genes and very little amount of DNA fragments in the closed regions of *Ivl, Ttn, Krt86* (Figure 5.4B) genes. As I continued to increase DNase I concentration, we see not only DNA fragments increased in the open chromatin regions but also in the closed chromatin regions. The increase in DNA fragments in the closed chromatin regions using higher DNase I treatments is undesirable since it would be hard to determine from the DNase-Seq analysis whether active regulatory regions were actually present or if they were a result of DNase I overdigestion. I decided to be conservative in my approach and used a 40units DNase I concentration for all our DNase-Seq libraries. We generated a DNase-Seq library on untreated Zhbtc4 cells as well as cells treated with Tc at 24hrs, 48hrs, and 72hrs using a similar titration protocol. In Tc-treated samples, we used different qPCR primers to assay the open chromatin regions (*Cdx2, Eomes, FGFR1, and Rpl41*) since chromatin accessibility changes as ES cells differentiate.

DNase-Seq libraries were generated for each tetracycline-treated sample and mapped to the mouse genome (mm9). Approximately, 20millions reads were mapped to the mouse genome with the exception of the Untreated-1 sample, which had 150 million reads. Peakcalling of DNase peaks was done using Hotspot Version 4 software. Taking total peaks called by Hotspot for each sample and dividing it by the totaled mapped reads, we obtained the mapping efficiencies for each sample. Our mapping efficiency was around

4% for almost all samples except for the Untreated-1 sample (32% efficiency) (Figure 5.4C). The low mapping efficiencies in our sample could be due to many factors. One such factor is the size selection of DNA fragments after DNase I treatment. I size-selected DNA fragments from 100-500bp to generate the DNase-Seq libraries but a recent report suggested that the longer DNA fragments (200-300bp) can lower mapping efficiency and require greater amount of sequencing (He et al., 2014). Additionally longer DNA fragments most likely span the coverage of nucleosomes (He et al., 2014) and could cofound the DNase-Seq analysis. The second factor could be that the 40units of DNase I was not optimal and refining our DNaseI concentration could increase mapping efficiency. However, it is most likely that our size selection must be revised to generate better DNase-seq libraries.

Although the coverage of my DNase-Seq libraries were not as good as those published by John Stamatoyannopoulos's group (Figure 5.4D), I still nonetheless believe that our data is still good enough to do preliminary analyses to view chromatin accessibility during the differentiation of the Zhbtc4 ES cells. When I viewed the promoter region of the *Oct4* gene in the UCSC genome browser, I noticed that in the untreated sample, DNase reads piled up in its promoter region suggesting that it is open and accessible by transcription factors to regulate *Oct4* transcription (Figure 5.5A). However, as differentiation progressed, the *Oct4* promoter became less open and therefore I observed less DNase I digested DNA fragment reads at the 24hrs, 48hrs, and 72hrs Tc-treated samples (Figure 5.5A). I observed a similar pattern within the *BMP4* and *Id3* promoters as well (Figure 5.5B, 5.5C). In contrast, the promoters of known trophectoderm genes (*Id2, Tead4*) were not DNase I sensitive in untreated Zhbtc4 ES cells, but upon Tc treatment, those regions became more accessible/open (Figure 5.5D, 5.5E). Although the promoter of the *Cdx2* gene

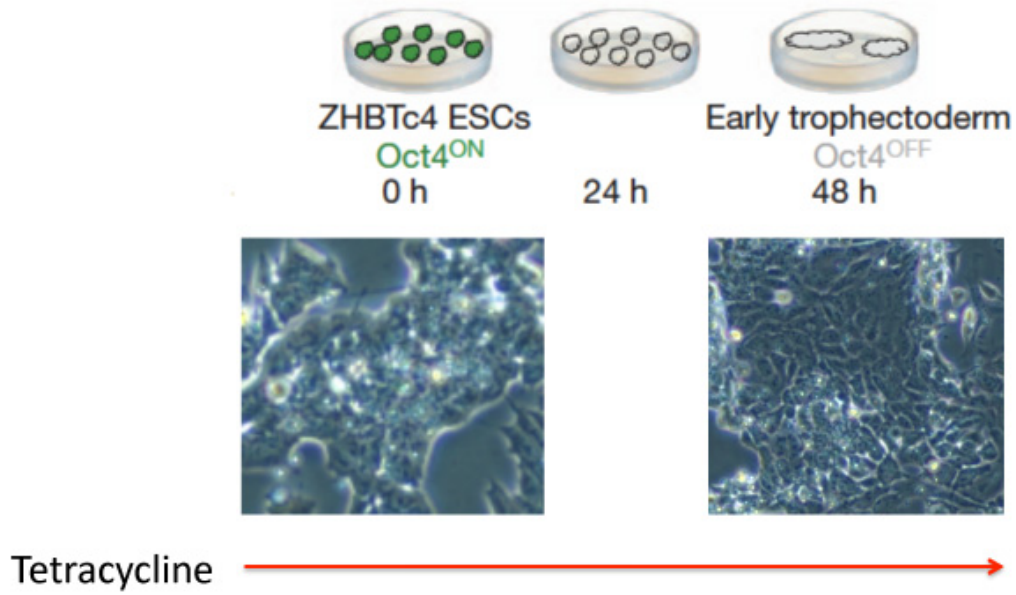
did not show changes in its chromatin state, I however did notice that hypersensitive sites decreased in the first intron of the *Cdx2* gene, where Oct4 is known to be bound (Figure 5.5F). It should be noted that DHS distribution in my samples could reflect variations in how genes interact with trans-activation factor complexes, and differences in the protein complexes in different chromosome territories (He et al., 2014).

Conclusions:

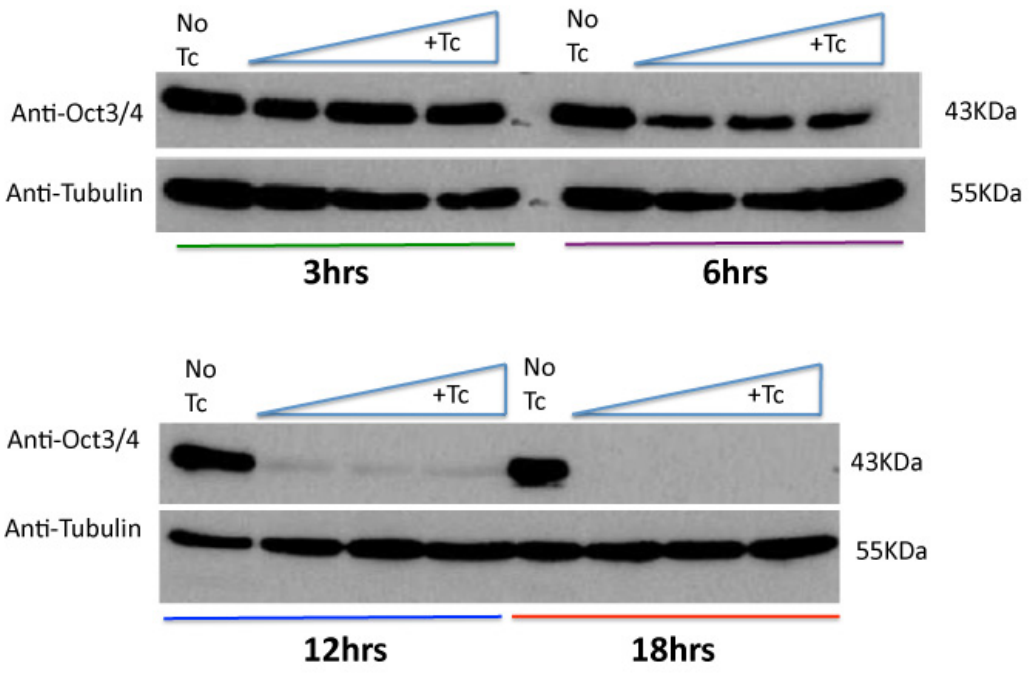
In summary, our findings extend our knowledge of the transcription programs occuring during the first lineage segregation of ICM and TE in mouse embryogenesis. Using the Zhbtc4 ES cell line as an *in vitro* model, I obtained a global transcriptome profile after *Oct4* knockdown in ES cells and subsequent differentiation of ES cells into trophectoderm-like cells. From our RNA-Seq time course dataset, I was able to reveal cluster of genes that were changing (either upregulated or downregulated) as a result of Oct4 depletion (Figure 5.3A). At each time point, I observed a unique transcriptional signature, suggesting a dynamic process is underway as ES cells differentiate. I also demonstrated that the BMP4 ligand and its target *Id* genes (*Id2* and *Id3*) are affected upon Oct4 depletion, suggesting BMPs may be involved in the *Oct4/Sox2/Nanog* network, directly or indirectly. The molecular mechanism(s) of how BMP signaling feeds into this particular network will need to be studied further. Also in this present study, I demonstrated that chromatin accessibility is affected as ES cells differentiate. Chromatin accessibility changes as different transcriptional requirements are needed during cellular differentiation. The results of my global study provide relevant insights into the molecular mechanisms during differentiation ES cells for future studies.

Figure 5.1

A)


ZHBTc4 ESCs
Oct4ON
0 h
24 h
Early trophectoderm
Oct4OFF
48 h
Tetracycline

B)


No Tc
+Tc
No Tc
+Tc
Anti-Oct3/4
43KDa
Anti-Tubulin
55KDa
3hrs
6hrs
No Tc
+Tc
No Tc
+Tc
Anti-Oct3/4
43KDa
Anti-Tubulin
55KDa
12hrs
18hrs

C)

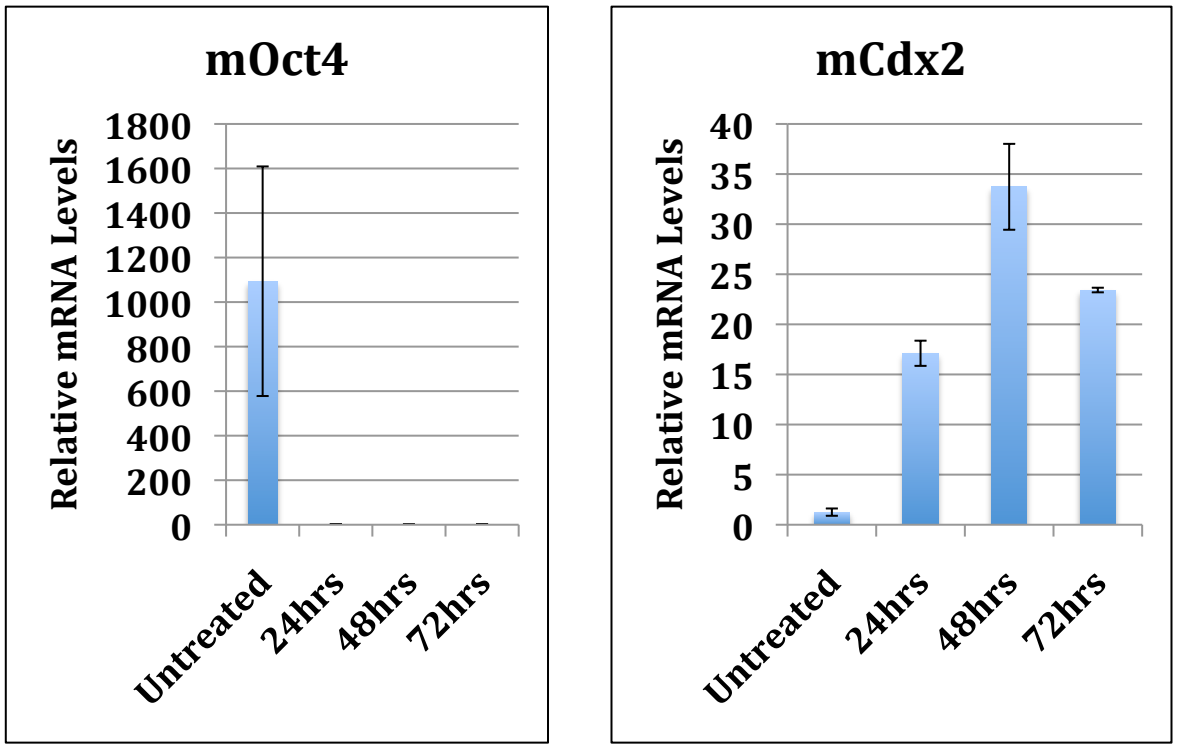
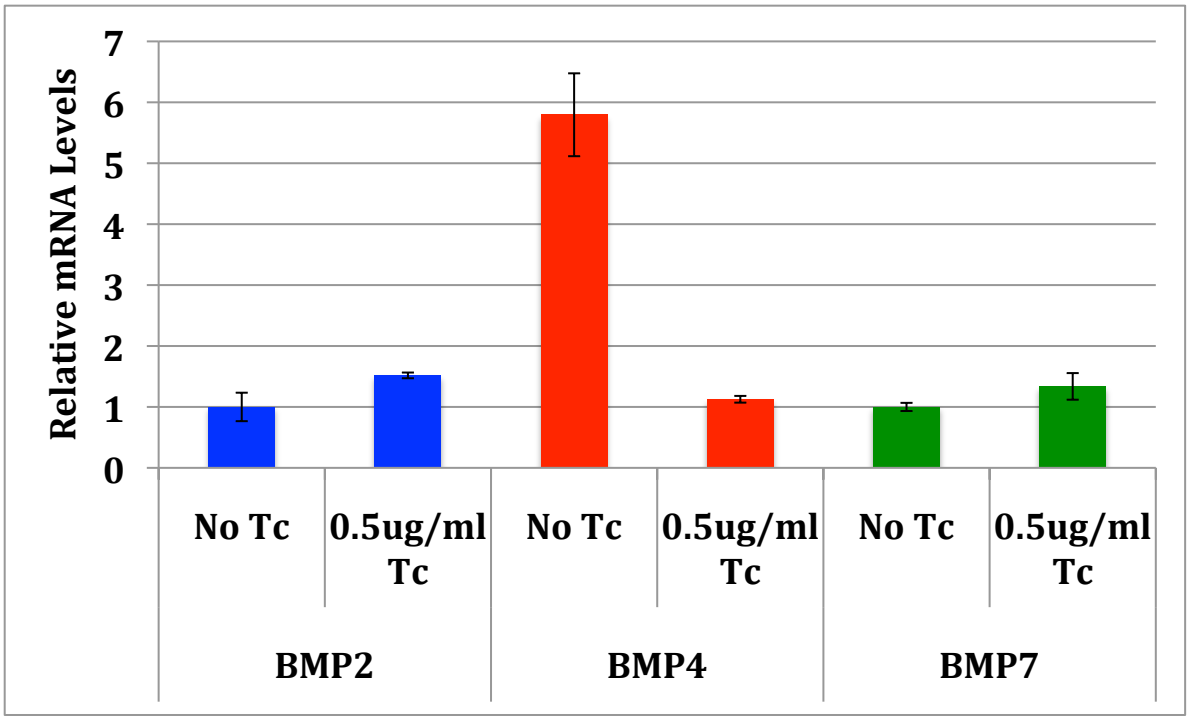


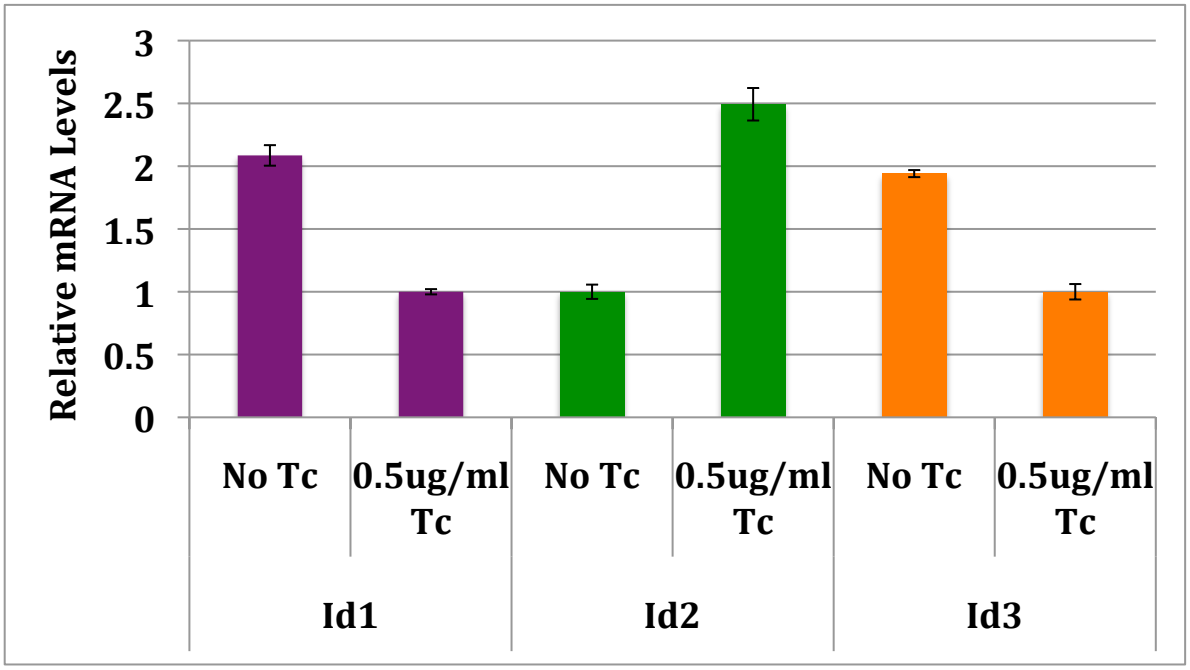
Figure 5.1: Characterization of Zhbtc4 mouse ES cell line. A) Schematic of Zhbtc4 ES cell line, an inducible Tetracycline-Off Oct4 knockdown ES cell line. Both endogenous alleles of Oct4 were inactivated by gene targeting and contained a regulatable tetracycline(Tc)-off Oct4 transgene. Upon Tc treatment, Zhbtc4 cells differentiate into trophectoderm-like cells after 48hrs. Photomicrographs of ES cells were at 20X magnification. B) Western blot of Oct4 knockdown upon tetracycline treatment. Time-course of tetracycline treatment in Zhbtc4 mouse ES cells: 3hrs, 6hrs, 12hrs, and 18hrs at different Tc concentrations (0.5μg/ml, 1μg/ml, or 2μg/ml) C) QPCR analyses of Zhbtc4 ES cells upon tetracycline treatment. Oct4 (Pou5f1) and Cdx2 transcript levels were measured at the following tetracycline treatment time points: 0hrs (untreated), 24hrs, 48hrs, and 72hrs. The Tc concentration was 0.5μg/ml.

Figure 5.2:

A)


Relative mRNA Levels
7
6
5
4
3
2
1
0
No Tc
0.5ug/ml Tc
No Tc
0.5ug/ml Tc
No Tc
0.5ug/ml Tc
BMP2
BMP4
BMP7

B)


Relative mRNA Levels
3
2.5
2
1.5
1
0.5
0
No Tc
0.5ug/ml Tc
No Tc
0.5ug/ml Tc
No Tc
0.5ug/ml Tc
Id1
Id2
Id3

C)

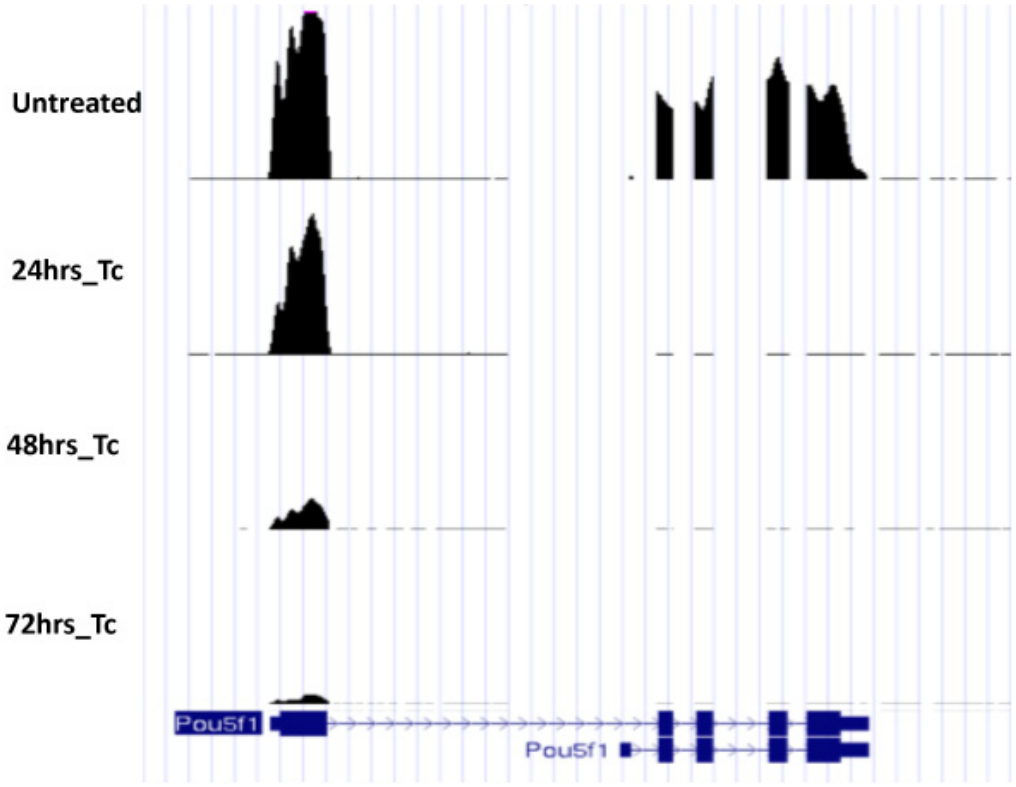
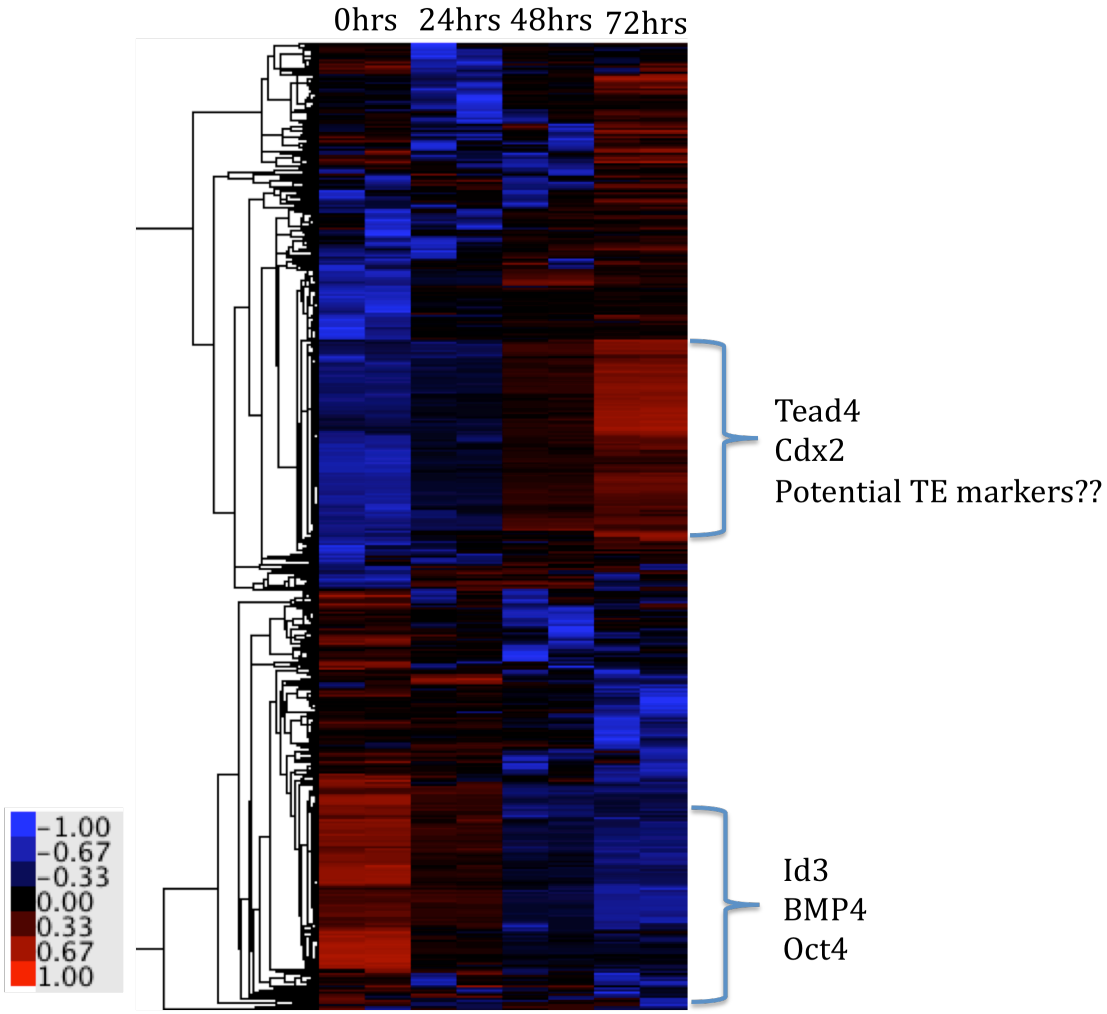


Figure 5.2: BMP and its targets are downregulated upon Oct4 knockdown in Zhbtc4 ES cells. QPCR analyses of A) BMP2, 4, and 7 transcript levels measured after tetracycline (Tc) treatment (0.5μg/ml) for 48hrs and B) BMP targets, *Id1*, *Id2*, and *Id3* transcript levels after Oct4 depletion (48hrs). *Id2* is upregulated during Zhbtc4 ES cell differentiation into trophectoderm-like cells. C) UCSC genome browser of RNA-Seq reads covering the Oct4 exons. Mapped reads decreased as tetracycline treatment time increased (24hrs, 48hrs, 72hrs).

Figure 5.3

A)


0hrs 24hrs 48hrs 72hrs
Tead4
Cdx2
Potential TE markers??
Id3
BMP4
Oct4
-1.00
-0.67
-0.33
0.00
0.33
0.67
1.00

B)

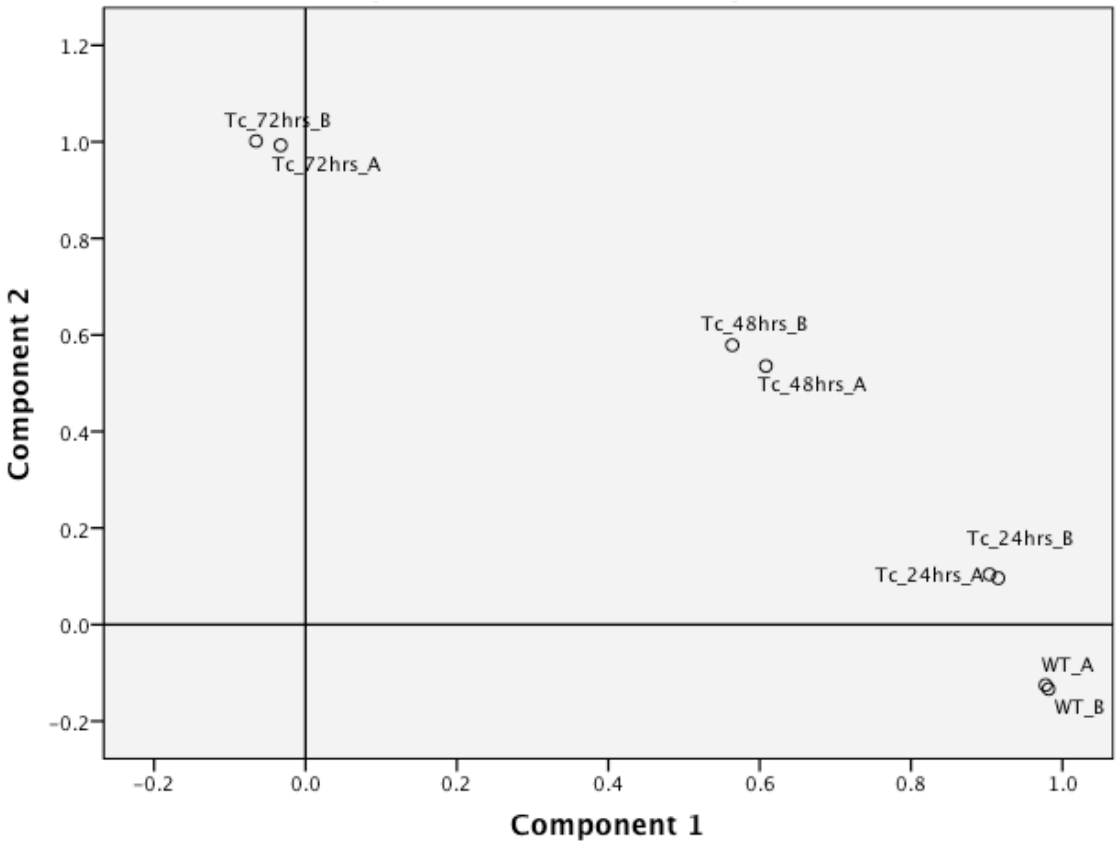
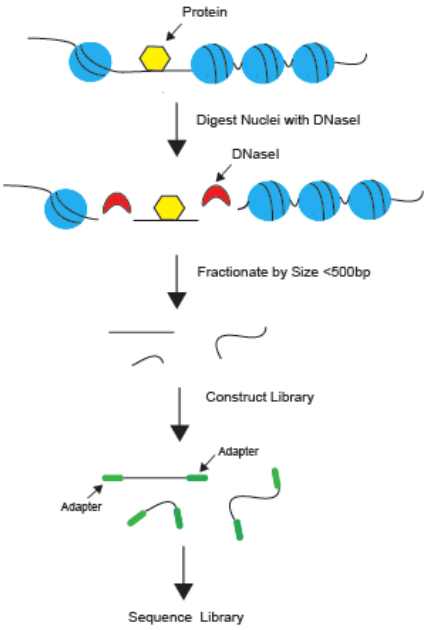


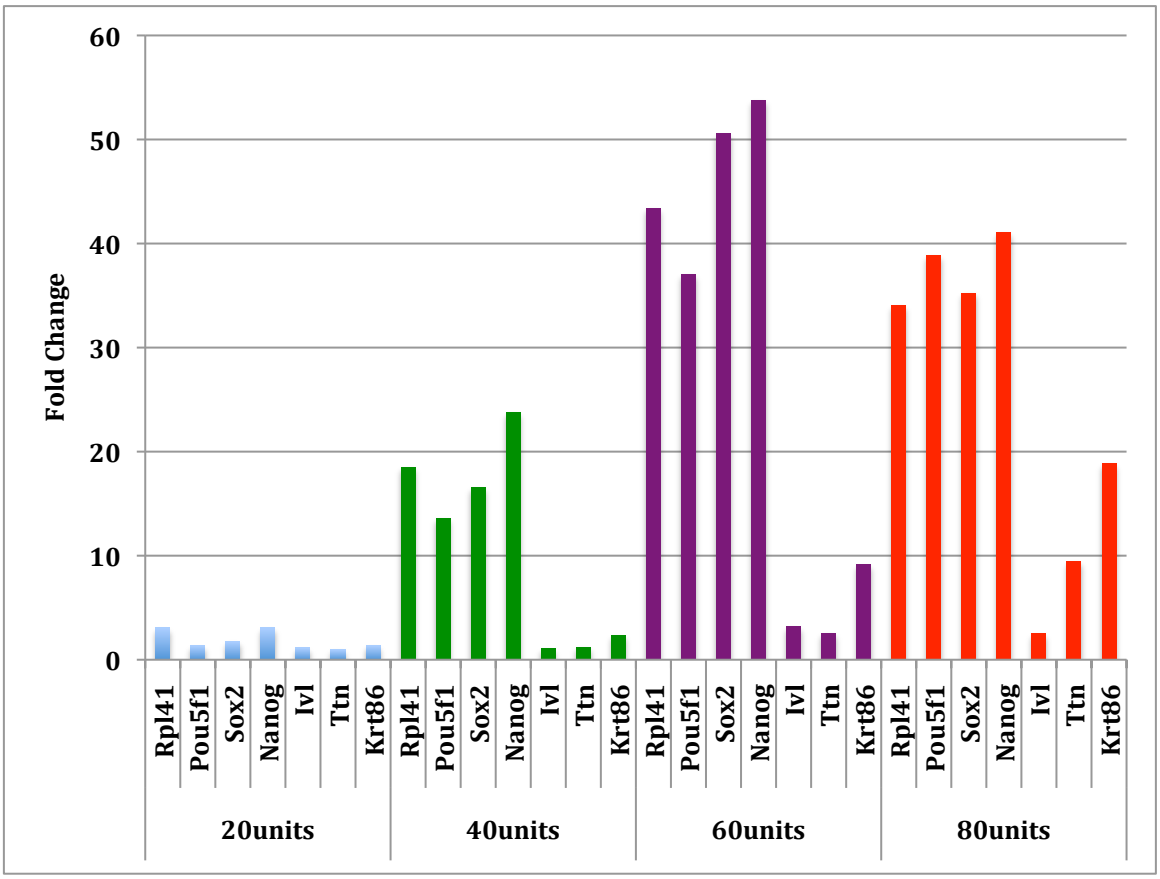
Figure 5.3: RNA-Seq time course analyses. A) Cluster diagram of gene expression in Oct4 depleted Zhbtc4 ES cells. RNA-Seq was performed on Zhbtc4 ES cells treated with Tc at the following time point: 0hrs, 24hrs, 48hrs, and 72hrs. Genes were organized by hierarchical clustering based on similarity in expression pattern. B) Principal component analysis (PCA) plot of untreated Zhbtc4 ES cells and Zhbtc4 cells treated with tetracycline for 24hrs, 48hrs, and 72hrs.

Figure 5.4:

A)


Protein
Digest Nuclei with DNaseI
DNaseI
Fractionate by Size <500bp
Construct Library
Adapter
Adapter
Sequence Library

B)


Fold Change
60
50
40
30
20
10
0
Rpl41
Pou5f1
Sox2
Nanog
Ivl
Ttn
Krt86
20units
40units
60units
80units

C)

Sample	Total Processed Reads	Unique Mapped Reads (Trim 14bp)	Samtools RMDUP	Coverage Read (Hotspot_V4: 1%FDR):RMDUP	Efficiency: (Coverage Reads/ Mapped Reads) x100
Tc -24hrs	27,373,907	19,496,389	15,519,701	725,591	4.67%
Tc-24hrs	29,239,314	21,080,396	19,464,836	904,203	4.64%
Tc -48hrs	29,650,571	21,227,880	19,791,664	915,166	4.62%
Tc -48hrs	28,351,658	20,531,620	18,900,946	931,321	4.92%
Tc -72hrs	29,969,833	21,407,594	19,533,080	911,774	4.66%
Tc-72hrs	25,233,936	18,302,782	16,890,417	946,110	5.60%
Untreated-1	275,423,783	155,456,357	52,658,847	17,220,000	32.70%
Untreated-2	36,854,838	26,418,440	24,086,956	1, 319,000	5.47%

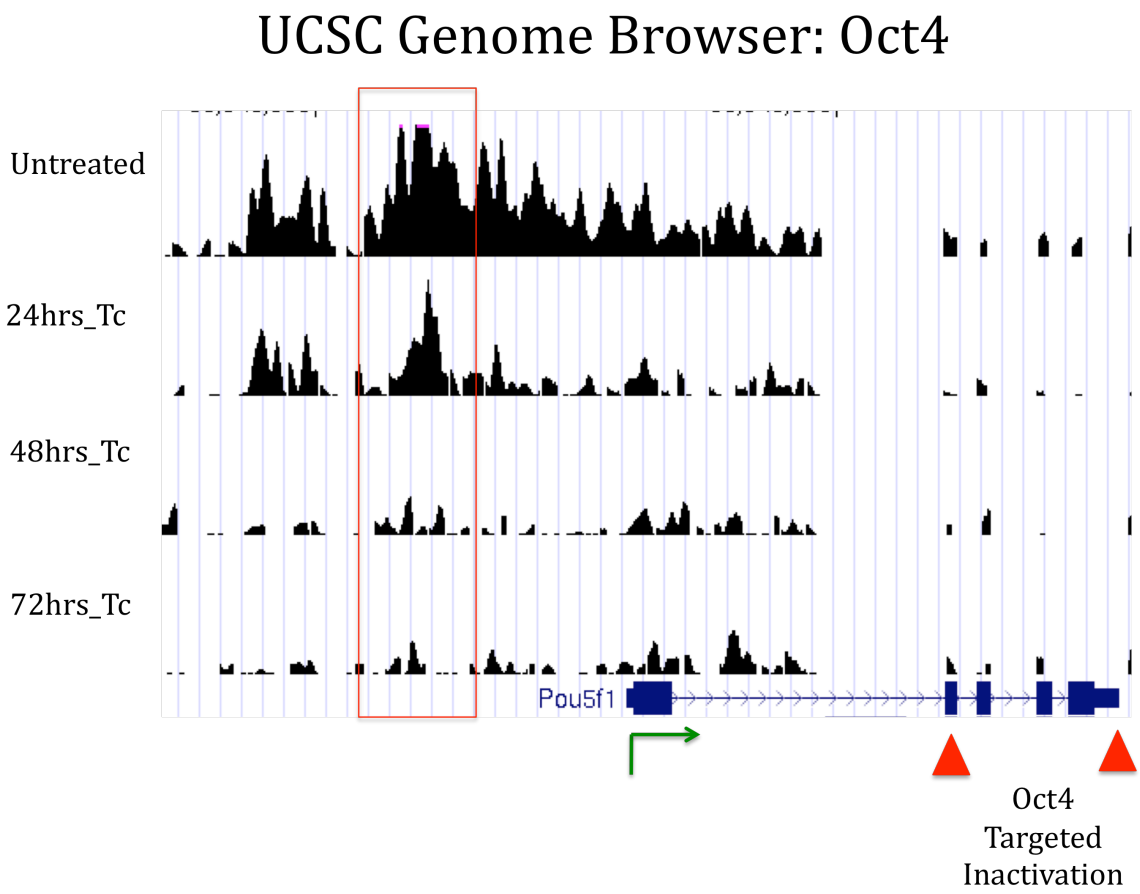
D)

Sample	Total Processed Reads	Unique Mapped Reads (Trim 14bp)	Samtools RMDUP	Coverage Read (Hotspot_V4:1% FDR): RMDUP	Efficiency: (Coverage Reads/ Mapped Reads) x100
Encode_0hrs_Rep1V2	41,291,067	30,152,405	22,463, 130	5,115,000	22.77%
Encode_0hrs_Rep2	44,798,543	32,065,589	26,450,121	4,297,000	16.24%
Encode_24hrs_Rep1	43,816,918	31,833,034	24,877,475	4,722,000	18.98%
Encode_24hrs_Rep2	38,264,936	27,686,875	23,487,299	3,825,000	16.28%

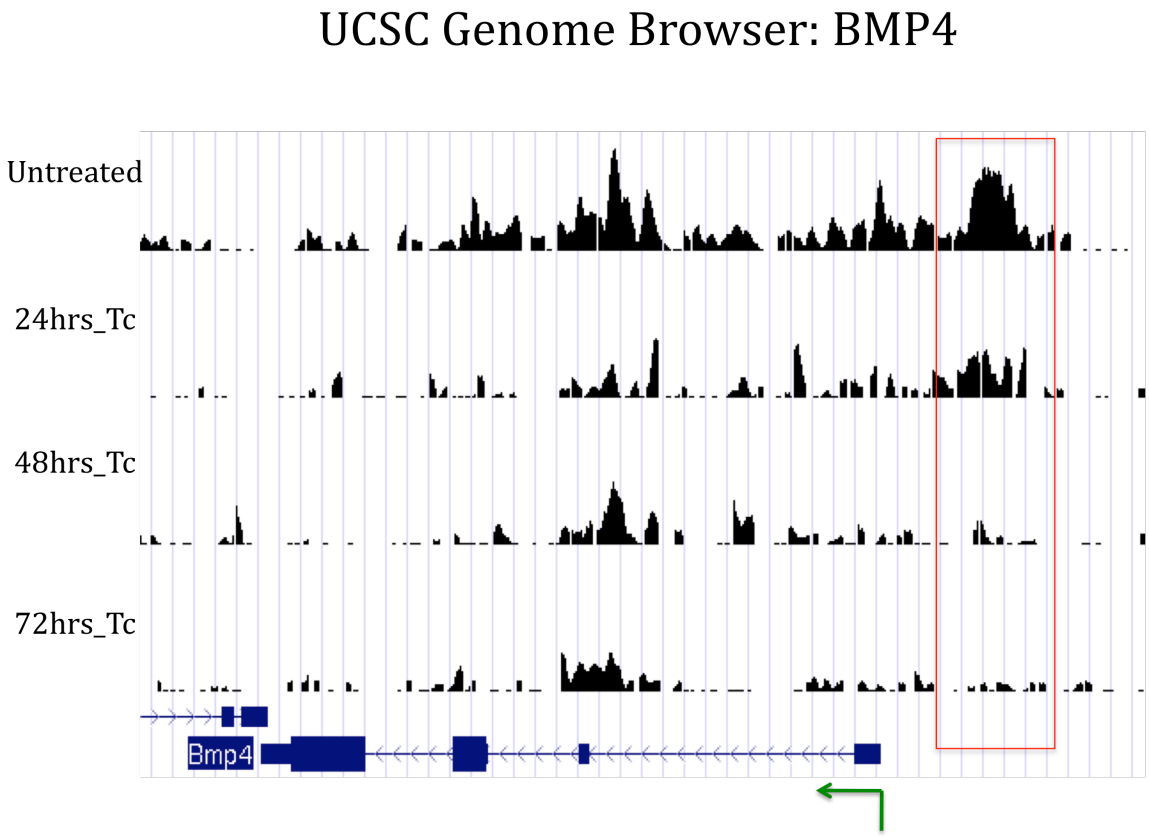
Figure 5.4: DNase-Seq approach and analyses A) Overview of DNase-Seq workflow. The top of this illustration shows a double-stranded DNA molecule wrapped around four nucleosomes (blue) and a relatively open segment of chromatin containing a cis-regulatory region with a bound protein (yellow). DNAse (red) cleavage occurs preferentially in the open chromatin region releasing it from the bulk DNA. After size fractionation, the DNA is adaptor ligated and amplified to create a library that is ready for high throughput sequencing. Since DNA from open chromatin regions is released in greater amounts than from closed chromatin, open chromatin regions are over-represented in the library. Therefore, sequence reads tend to pile up in genomic intervals containing open chromatin. B) Titration of DNase I for proper nuclei digestion. Quantititative PCR was performed on open and closed chromatin regions after DNase I treatment in untreated Zhbtc4 nuclei. C) DNase-Seq mapping efficiencies of Zhbtc4 ES cells at each tetracycline treatment time point. Analysis of each sample: 1) number of reads obtained from sequencing, 2) total mapped reads after trimming each read by 14bps at the 3' end, 3) removal of duplicate reads (Samtools RMDUP), 4) total number of reads under all peaks called by the Hotspot peakcalling software, and 5) the mapping efficiency of the sequencing library (#4/#2). D) Published Zhbtc4 DNase-Seq data from John Stamatoyannopoulos's group for untreated and 24hrs tetracycline-treated Zhbtc4 ES cells. DNase-seq analysis was performed as above in C.

Figure 5.5

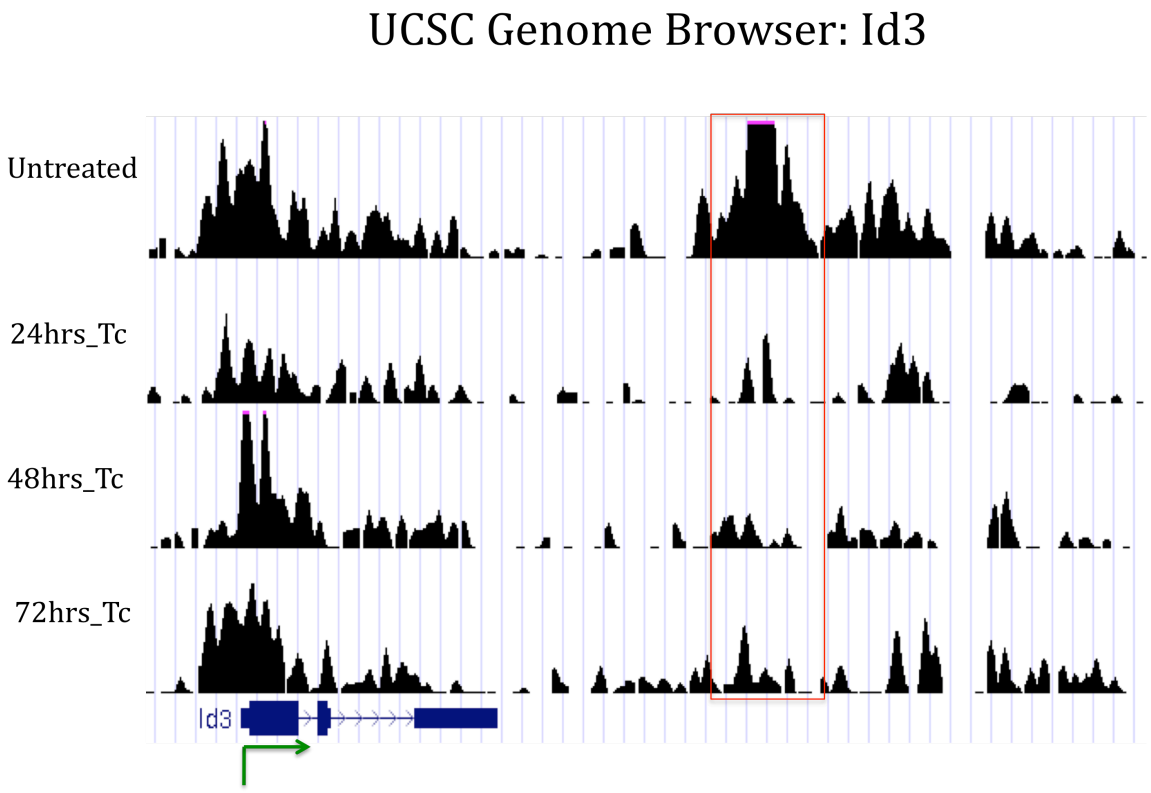
A)


UCSC Genome Browser: Oct4
Untreated
24hrs_Tc
48hrs_Tc
72hrs_Tc
Pou5f1
Oct4
Targeted
Inactivation

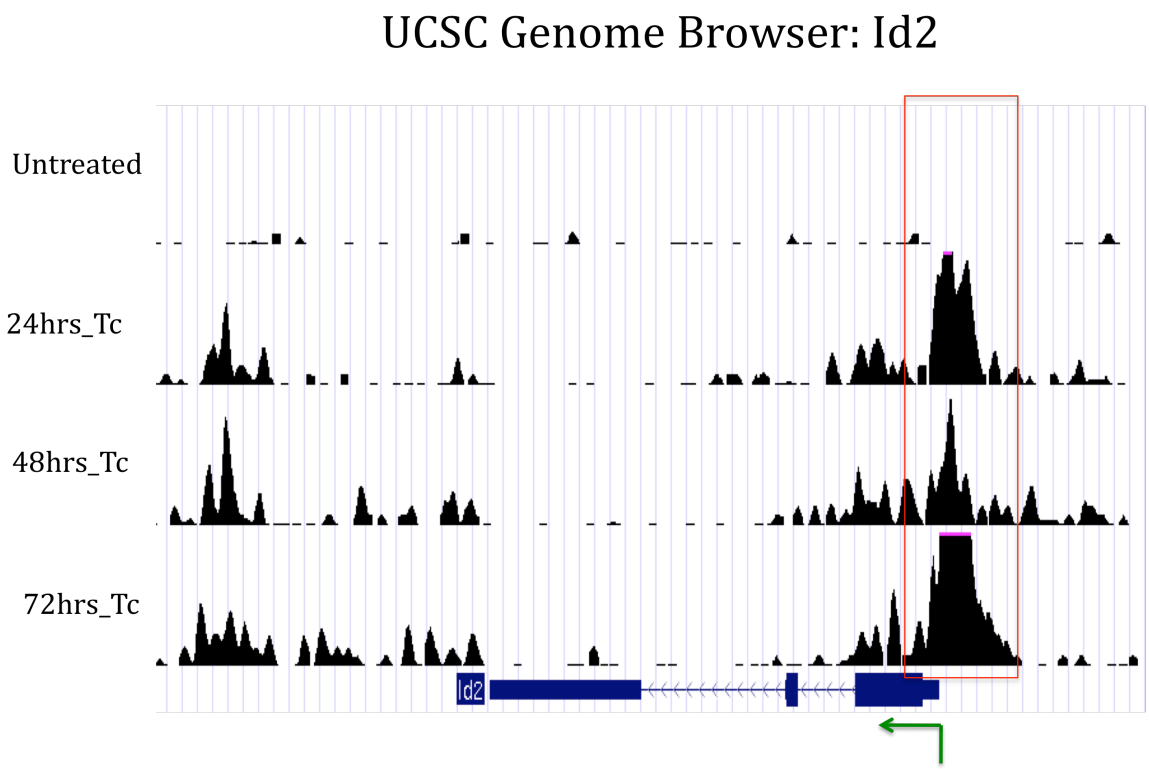
B)


UCSC Genome Browser: BMP4
Untreated
24hrs_Tc
48hrs_Tc
72hrs_Tc
Bmp4

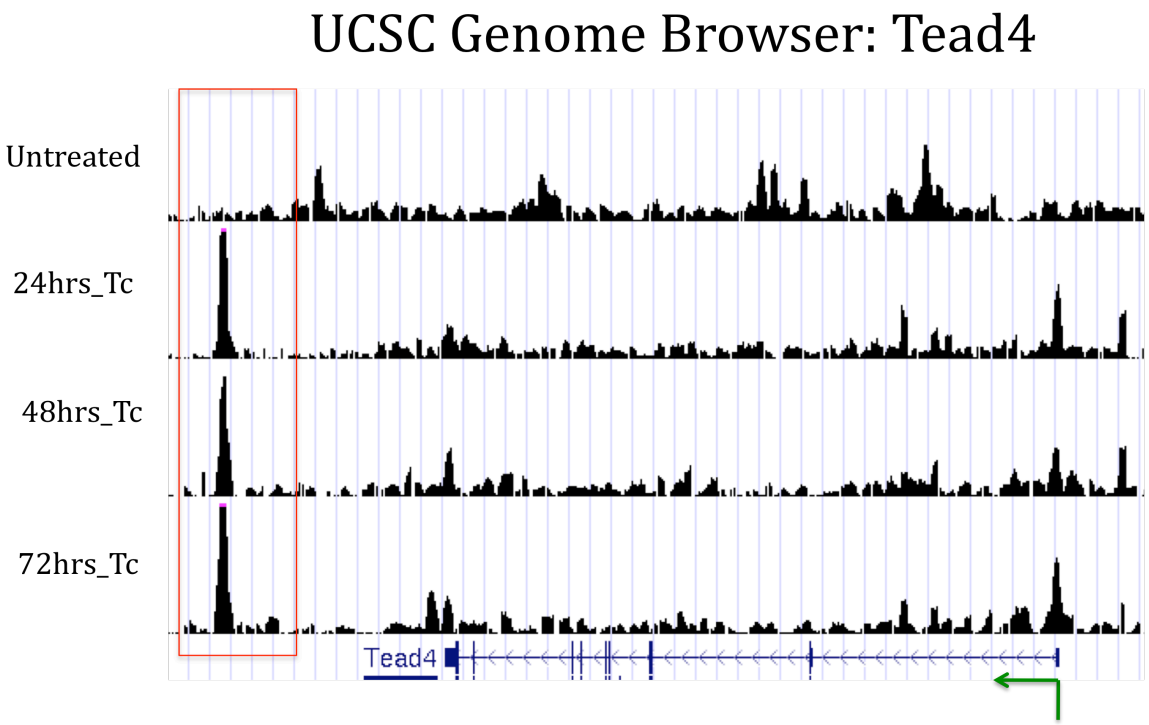
C)


UCSC Genome Browser: Id3
Untreated
24hrs_Tc
48hrs_Tc
72hrs_Tc
Id3

D)


UCSC Genome Browser: Id2
Untreated
24hrs_Tc
48hrs_Tc
72hrs_Tc
Id2

E)


UCSC Genome Browser: Tead4
Untreated
24hrs_Tc
48hrs_Tc
72hrs_Tc
Tead4

F)

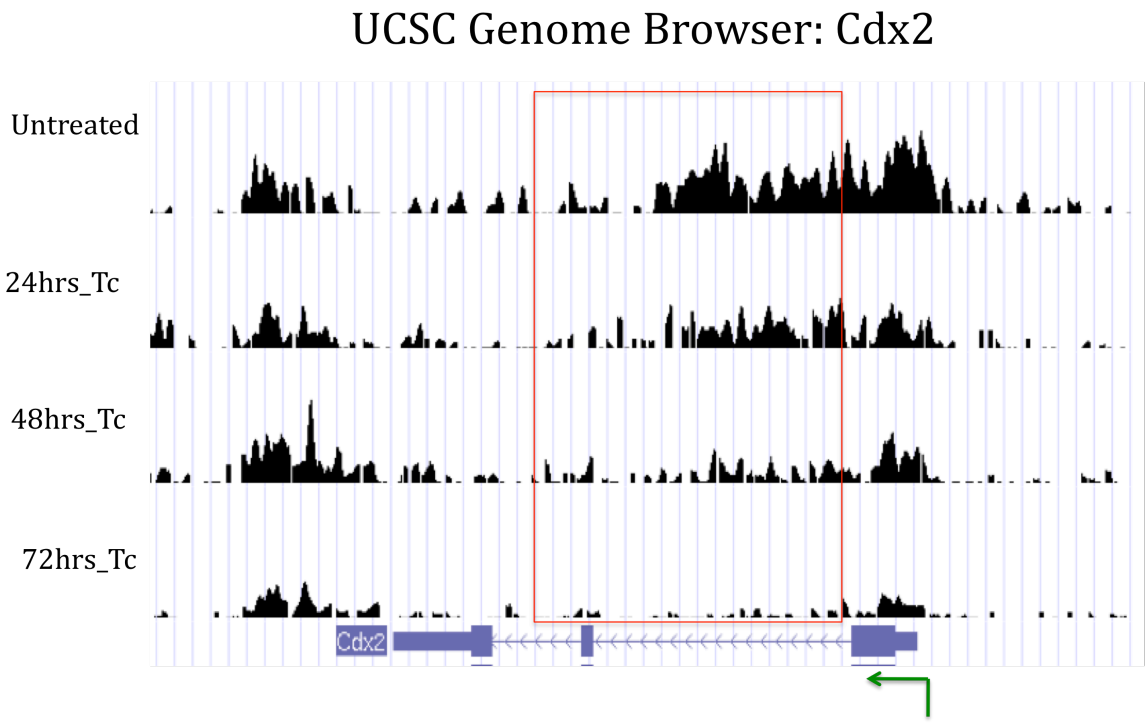


Figure 5.5: Visualization of Zhbtc4 DNase-Seq peaks on the UCSC genome browser at different time points of Zhbtc4 cells after tetracycline treatment: 0hrs (Untreated), 24hrs (24hrs_Tc), 48hrs (48_Tc), and 72hrs (72hrs_Tc). DNase-Seq reads for the following genes were viewed: A) *Oct4* (Pou5f1), B) *BMP4*, C) *Id3*, D) *Id2*, E) *Tead4* F) *Cdx2*

Chapter 5 References:

Babaie Y, Herwig R, Greber B, Brink TC, Wruck W, Groth D, Lehrach H, Burdon T, Adjaye J. Analysis of Oct4-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells. Stem Cells. 2007 Feb;25:500-10.

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell. 2005 Sep 23;122:947-56.

Boyle AP, Song L, Lee BK, London D, Keefe D, Birney E, Iyer VR, Crawford GE, Furey TS. High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells. Genome Res. 2011 Mar;21:456-64.

Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell. 2003 May 30;113:643-55.

Chen L, Yabuuchi A, Eminli S, Takeuchi A, Lu CW, Hochedlinger K, Daley GQ. Cross-regulation of the Nanog and Cdx2 promoters. Cell Res. 2009 Sep;19:1052-61.

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK, Clarke ND, Wei CL, Ng HH. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell. 2008 Jun 13;133:1106-17.

Dietrich JE, Hiiragi T. Stochastic patterning in the mouse pre-implantation embryo. Development. 2007 Dec;134:4219-31.

Galas DJ, Schmitz A. DNAse footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res. 1978 Sep;5:3157-70.

Guo G, Huss M, Tong GQ, Wang C, Li Sun L, Clarke ND, Robson P. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. Dev Cell. 2010 Apr 20;18:675-85.

He HH, Meyer CA, Hu SS, Chen MW, Zang C, Liu Y, Rao PK, Fei T, Xu H, Long H, Liu XS, Brown M. Refined DNase-seq protocol and data analysis reveals intrinsic bias in transcription factor footprint identification. Nat Methods. 2014 Jan;11:73-8.

Keramari M, Razavi J, Ingman KA, Patsch C, Edenhofer F, Ward CM, Kimber SJ. Sox2 is essential for formation of trophectoderm in the preimplantation embryo. PLoS One. 2010 Nov 12;5:e13952.

Kuroda T, Tada M, Kubota H, Kimura H, Hatano SY, Suemori H, Nakatsuji N, Tada T. Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. Mol Cell Biol. 2005 Mar;25:2475-85.

Lenardo MJ, Staudt L, Robbins P, Kuang A, Mulligan RC, Baltimore D. Repression of the IgH enhancer in teratocarcinoma cells associated with a novel octamer factor. Science. 1989 Jan 27;243:544-6.

Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell. 2003 May 30;113:631-42.

Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Schöler H, Smith A. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell. 1998 Oct 30;95:379-91.

Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet. 2000 Apr;24:372-6.

Niwa H, Toyooka Y, Shimosato D, Strumpf D, Takahashi K, Yagi R, Rossant J. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell. 2005 Dec 2;123:917-29.

Okamoto K, Okazawa H, Okuda A, Sakai M, Muramatsu M, Hamada H. A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. Cell. 1990 Feb 9;60:461-72.

Pique-Regi R, Degner JF, Pai AA, Gaffney DJ, Gilad Y, Pritchard JK. Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. Genome Res. 2011 Mar;21:447-55.

Ralston A, Rossant J. Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. Dev Biol. 2008 Jan 15;313:614-29.

Reyes de Mochel NS, Luong M, Chiang M, Javier AL, Luu E, Toshihiko F, MacGregor GR, Cinquin O, Cho KW. BMP signaling is required for cell cleavage in preimplantation-mouse embryos. Dev Biol. 2015 Jan 1;397:45-55.

Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng HH, Robson P. Transcriptional regulation of nanog by OCT4 and SOX2. J Biol Chem. 2005 Jul 1;280:24731-7.

Rosner MH, Vigano MA, Ozato K, Timmons PM, Poirier F, Rigby PW, Staudt LM. A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. Nature. 1990 Jun 21;345:686-92.

Sarkar A, Hochedlinger K. The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. Cell Stem Cell. 2013;12:15–30.

Schöler HR, Ruppert S, Suzuki N, Chowdhury K, Gruss P. New type of POU domain in germ line-specific protein Oct-4. Nature. 1990 Mar 29;344:435-9.

Song L, Zhang Z, Grasfeder LL, Boyle AP, Giresi PG, Lee BK, Sheffield NC, Gräf S, Huss M, Keefe D, Liu Z, London D, McDaniell RM, Shibata Y, Showers KA, Simon JM, Vales T, Wang T, Winter D, Zhang Z, Clarke ND, Birney E, Iyer VR, Crawford GE, Lieb JD, Furey TS. Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. Genome Res. 2011 Oct;21:1757-67.

Sridharan R, Tchieu J, Mason MJ, Yachechko R, Kuoy E, Horvath S, Zhou Q, Plath K. Role of the murine reprogramming factors in the induction of pluripotency. Cell. 2009 Jan 23;136:364-77.

Strumpf D, Mao CA, Yamanaka Y, Ralston A, Chawengsaksophak K, Beck F, Rossant J. Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development. 2005 May;132:2093-102.

Suzuki A, Raya A, Kawakami Y, Morita M, Matsui T, Nakashima K, Gage FH, Rodríguez-Esteban C, Izpisúa Belmonte JC. Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. Proc Natl Acad Sci U S A. 2006 Jul 5;103:10294-9.

Ying QL, Nichols J, Chambers I, Smith A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell. 2003 Oct 31;115:281-92.

Yuan H, Corbi N, Basilico C, Dailey L. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. Genes Dev. 1995 Nov 1;9:2635-45.

Chapter 6:

Conclusion

BMP signaling is important in early mouse development and the generation of a *X.Id3* BRE-gal reporter line in the Cho lab has been useful in studying BRE-mediated BMP signaling in mouse ES cells and early mouse embryogenesis. We demonstrated that our *X.Id3* BRE-gal ES cell reporter line responded directly to BMP ligand treatment and is responsible for activating the BMP transcriptional response in mouse ES cells. Additionally, our *X.Id3* BRE-gal mice detected expression of BMP activity in developing embryos not previously reported by the Korchynskyi BRE-lac1 mouse reporter line (Korchynskyi et al., 2002; Monteiro et al., 2004; Javier et al., 2012). The Korchynskyi BRE-lac1 was derived from the mouse Id1 promoter. Both BREs are similar to each other except for one nucleotide difference in the Smad1 binding site. However, we observed that our BRE-gal mouse ES cells responded to BMP4 concentrations as low as 5ng/ml where as the Korchynskyi BRE ES cells responded at 20ng/ml. This could be due to the 7 copies of BRE in our *X.Id3* BRE-gal line while the Korchynskyi BRE-lac1 line contained only two copies of BRE. Currently, we have not yet examined the threshold response to our BRE-gal reporter using different copy numbers of BRE. Andrea von Bufnoff, a former Cho lab member, found that the threshold sensitivity and amplitude of the BMP response improved significantly with increased BRE copy numbers in *Xenopus* (unpublished data). This could also be true in mouse ES cells but have not been tested. If so, we can potentially develop a series of BMP reporters (containing different BRE copy numbers) that can respond to different thresholds of BMP ligand concentrations. They can useful reagents for looking at the

threshold responses of target gene transcription and monitoring BMP gradient activity in living embryos.

It should be noted that our *X.Id3* BRE-gal reporter was randomly integrated into the mouse genome. Random integration may alter cell phenotype due to insertional mutagenesis in undefined gene regions. Also, the reliability of data obtained may be compromised by differences in transgene integration sites when comparing multiple transgenic cell lines (e.g. Korchynskyi BRE-lac1 and our *X.Id3* BRE-gal ES cell lines). The random integration of the BRE-gal transgene may explain why we observed different X-gal staining patterns in both ES cell lines and also the differential gene expression patterns between the Korchynskyi mouse *Id1* BRE and our *X.Id3* BRE transgenic mice (Monteiro et al., 2004; Javier et al., 2012). We believe that our current BRE-gal reporter is still reliable since Anna Javier observed BMP activity (X-gal) in many developmental structures (heart, neural tube, and pharyngeal arches) that are established domains of BMP activity. However, targeting the BRE-gal transgene and Korchynskyi BRE to the *ROSA26* locus would allow for a better comparative analysis. The *ROSA26* can be targeted without silencing or disrupting normal gene function. A drawback to random integration of our BRE-gal transgene is that BRE-gal expression can be either be artificially enhanced or dampened by neighboring genes depending on the integration site. Targeting our transgene to the *ROSA26* can remove artifacts in our studies and reveal true BMP activity. Additionally, the generation of the BRE-gal (Del-2) construct from our lab could be used to generate a reporter line deficient in BRE-mediated BMP signaling. The BRE-gal (Del-2) construct could similarly be targeted to the *ROSA26* locus for analysis of ES cells and mouse embryos that are deficient in BRE-mediated BMP signaling. We can perform comparative analyses

with our wild type BRE-gal ES cell and mouse lines with the mutant BRE-gal (Del-2) lines to observe how the nature of such a deletion affects BMP transcriptional responses and developmental programs.

From this study, it is clear that our BRE-gal reporter does not capture all BMP signaling activities. Given the diverse modes of Smad interactions with gene promoters (Feng et al., 2005; Miyazono, et al., 2005) it is unlikely that a single cis-regulatory element can modulate the different Smad binding combinations with various transcription factors. Adding another layer of complexity to the BMP response are transcription cofactors and how they interact with activated Smads to regulate BMP target genes. Furthermore, although BMPs typically transduce their signal through Smads 1/5/8, known non-canonical BMP pathways have been identified to be independent of the Smad proteins (Moustakas et al., 2005) and these non-canonical pathways do not utilize Smad binding sites.

From my RNA-Seq data of mouse ES treated with BMP4, two direct BMP targets, *R3hdml* and *Dusp5*, were identified that did not contain a predicted BRE based in our motif search of the genome. This highly suggested that a BRE-independent cis-regulatory element could also mediate BMP signaling. We know that in Activin/Nodal signaling pathway, there are four cis-regulatory modules that have been identified to date: Node specific enhancer (NDE), Assymmetric Enhancer (ASE), Posterior Epiblast Element (PPE), and Asymmetric Initiatior Element/Left Side Specific Enhancer (AIE/LSE) that are involved in many developmental processes with each activated at different times of development (Adachi et al., 1999; Norris et al., 1999; Krebs et al., 2003; Raya et al., 2003; Vincent et al., 2004; Saijoh et al., 2005; Ben-Haim et al., 2006). Similarly, this could also be true for the BMP pathway. It is unlikely that one universal BRE can regulate all BMP signaling activity.

Therefore, I cloned regions upstream and downstream along with the first intronic regions of the two direct BMP target genes (*Dusp5* and *R3hdml)* into a luciferase construct. Although my preliminary tests of a few of my clones did not yield the localized region responsible for BMP responsiveness, I have yet to test all my clones and predict that the responsible element may be in my untested clones. Once the responsible cis-regulatory element is isolated, we can test whether this element is active during embryonic development by investigating whether it can drive reporter gene expression (β-galactosidase, luciferase, or GFP) in ES cell *in vitro* and embryos *in vivo*. Furthermore, to assess the relevance of this regulatory element, we can generate mutations within the element and observe whether BMP responsiveness is affected. However, I cannot eliminate the possibility that the *Dusp5* and *R3hdml* genes may be regulated by non-canonical BMP signaling. If so, we would need to identify these mediators in regulating BRE-independent BMP signaling.

The role BMP signaling in early mouse development has been briefly characterized in this thesis as well. Based on our analysis of published single-cell RNA-Seq of early stages of mouse development, I was able to show BMP ligands and its target genes expression during the very early stages of mouse embryogenesis. Our findings indicate that the BMP4 ligand and the *Id3* gene are expressed primarily in inner cell mass whereas *Id2* is expressed in the trophectoderm. However, in ES cells, this differential expression is not present. This could be due to the fact that ES cells are cultured *in vitro* and may not always reflect *in vivo* conditions. It is known that the association of activated Smad complexes with specific DNA-binding cofactors (e.g. Schnurri) determines which genes are activated or repressed. It is possible that cofactors present in the ICM can bind tightly to the activated Smad

complex in the promoter of the *Id3* gene and regulate its expression while different Smad binding cofactors inhibit or repress *Id2* expression. Conversely, this could be occurring in the TE as well, where *Id2* is expressed while *Id3* inhibited. For future studies, I propose that the identification of these cell-type-specific cofactors to study differential regulation of the *Id* genes. Screening the proteome of these two different cell types for putative cofactors and combining it with published genomic dataset can lead to identification of ICM/TE-specific cofactors regulating BMP target genes. Alternatively, the differential regulation of *Id2* and *Id3* in ICM and TE could be explained by microRNA targeting. It has been shown that different microRNA clusters exist within the trophectoderm and ICM (Viswanathan et al., 2009; Tang et al., 2010) to potentially downregulate specific genes within the ICM and TE. If this is occurring, we can examine the interactions between the ICM or TE-specific microRNAs and its targeting site(s) within the *Id2* and *Id3* 3'UTRs using a luciferase reporter assay. Preliminary immunostaining showed Id2 expression in the trophectoderm and not the ICM provides early evidence for BMPs in ICM/TE specification. Immunostaining of Id3 will need to be carried out in the immediate future to confirm its presence in the ICM to support our hypothesis that the *Id* genes have non-redundant functions and play a role in cell fate determination and lineage commitment during early mouse embryogenesis.

The robust expression of BMP4 ligand in the ICM from our single-RNA Seq analysis is consistent with work done by Soledad Reyes de Mochel in the Cho lab and Coucouvanis et al. Through her experiments, Soledad experimentally showed that BMP activity was present in the ICM and to a lesser extent, the TE (Reyes de Mochel et al., 2014). This difference could be attributed to ligand and/or BMP receptor stability. Since the TE had

very weak BMP staining consistent with our single-cell RNA Seq data, it is possible that either BMP ligands and/or receptors are downregulated in the TE. Therefore, dynamic regulation of BMP ligands and their receptors as well as Smad1/4 binding partners could be contributing to the differential expression of BMPs in the blastocyst.

Oct4, Sox2, and *Nanog* are core regulators of ES cell pluripotency in mouse embryogenesis. Previous research has shown that aside from the *Oct4/Sox2/Nanog* network, BMPs are also important for maintaining ES self-renewal and pluripotency (Ying et al 2003). We hypothesized that BMP may be part of the network regulated by *Oct4, Sox2, and Nanog*. We demonstrated that Oct4 depletion downregulates BMP4 but not BMP2 and BMP7. Additionally, our findings demonstrated that *Id3* is downregulated as a result of Oct4 depletion. Conversely, *Id2* is upregulated upon Oct4 knockdown as the Zhbtc4 ES cells are differentiating into trophectoderm-like ES cells suggesting a novel role for *Id2* in TE identity. The molecular mechanism(s) of how Oct4 is directly of indirectly regulating *BMP4*, *Id3*, and *Id2* in the Zhbtc4 ES cells can be first ascertained by performing Oct4 ChIP-qPCR in Zhbtc4 ES cells.

The recent advent of massive parallel sequencing has provided increasing information on the cellular transcriptome and chromatin architecture. In my thesis, I performed RNA-Seq and DNase-Seq in Oct4 depleted ES cells and their subsequent differentiation. The RNA-Seq time course data set revealed a number of genes upregulated upon Zhbtc4 ES cell differentiation. These genes could be potential trophectoderm markers. We can therefore perform siRNA on these potential trophectoderm marker genes to see if they have arrested or delayed trophectoderm differentiation in the Zhbtc4 ES cells upon Oct4 depletion. Additionally, we can also determine if loss of these potential

trophectoderm markers lead to loss of known trophectoderm marker genes such as *Cdx2*, *Tead4*, and *Eomes*. We also employed DNase-Seq to identify cis-regulatory elements across the genome (Song et al., 2011). Utilizing this technique, we were able to demonstrate that the chromatin landscape changed as ES cells differentiated. In particular, I noticed that chromatin accessibility of *Oct4* and BMP pathway target genes (*Id3* and *BMP4)* were initially in an open state but switched into a closed state following differentiation. In contrast, trophectoderm markers (*Cdx2* and *Tead4*) were closed but became open after differentiation. In this present study, I have obtained a global profile of DHSs during ES cell differentiation. It would be interesting to determine how these individual DHSs are changing and what genes are associated with these changing DHSs. From both sequencing data sets, we can tease out candidate regulators important for trophectoderm differentiation by intersecting our RNA-Seq and DNase-Seq data sets. Future work would also include collecting transcription factor binding data by performing ChIP-Seq and adding these binding data sets into our RNA/DNase-Seq intersected analysis. I therefore propose collecting Oct4 ChIP-Seq and Cdx2 ChIP-Seq binding dataset during Zhbtc4 ES differentiation.

Overall, my thesis has provided new insights into the regulation of BMP signaling in mouse ES cells and early mouse embryogenesis.

Chapter 6 References:

Adachi H, Saijoh Y, Mochida K, Ohishi S, Hashiguchi H, Hirao A, Hamada H. Determination of left/right asymmetric expression of nodal by a left side-specific enhancer with sequence similarity to a lefty-2 enhancer. Genes Dev. 1999 Jun 15;13:1589-600.

Ben-Haim N, Lu C, Guzman-Ayala M, Pescatore L, Mesnard D, Bischofberger M, Naef F, Robertson EJ, Constam DB. The nodal precursor acting via activin receptors induces mesoderm by maintaining a source of its convertases and BMP4. Dev Cell. 2006 Sep;11:313-23.

Feng XH, Derynck R. Specificity and versatility in tgf-beta signaling through Smads. Annu Rev Cell Dev Biol. 2005;21:659-93.

Javier AL, Doan LT, Luong M, Reyes de Mochel NS, Sun A, Monuki ES, Cho KW. Bmp indicator mice reveal dynamic regulation of transcriptional response. PLoS One. 2012;7:e42566.

Korchynskyi O, ten Dijke P. Identification and functional characterization of distinct critically important bone morphogenetic protein-specific response elements in the Id1 promoter. J Biol Chem. 2002 Feb 15;277:4883-91.

Krebs LT, Iwai N, Nonaka S, Welsh IC, Lan Y, Jiang R, Saijoh Y, O'Brien TP, Hamada H, Gridley T. Notch signaling regulates left-right asymmetry determination by inducing Nodal expression. Genes Dev. 2003 May 15;17:1207-12.

Monteiro RM, de Sousa Lopes SM, Bialecka M, de Boer S, Zwijsen A, Mummery CL. Real time monitoring of BMP Smads transcriptional activity during mouse development. Genesis. 2008 Jul;46:335-46.

Monteiro RM, de Sousa Lopes SM, Korchynskyi O, ten Dijke P, Mummery CL. Spatio-temporal activation of Smad1 and Smad5 in vivo: monitoring transcriptional activity of Smad proteins. J Cell Sci. 2004 Sep 15;117:4653-63.

Moustakas A, Heldin CH. Non-Smad TGF-beta signals. J Cell Sci. 2005 Aug 15;118:3573-84.

Norris DP, Robertson EJ. Asymmetric and node-specific nodal expression patterns are controlled by two distinct cis-acting regulatory elements. Genes Dev. 1999 Jun 15;13:1575-88.

Raya A, Kawakami Y, Rodriguez-Esteban C, Buscher D, Koth CM, Itoh T, Morita M, Raya RM, Dubova I, Bessa JG, de la Pompa JL, Izpisua Belmonte JC. Notch activity induces Nodal expression and mediates the establishment of left-right asymmetry in vertebrate embryos. Genes Dev. 2003 May 15;17:1213-8.

Reyes de Mochel NS, Luong M, Chiang M, Javier AL, Luu E, Toshihiko F, MacGregor GR, Cinquin O, Cho KW. BMP signaling is required for cell cleavage in preimplantation-mouse embryos. Dev Biol. 2015 Jan 1;397:45-55.

Saijoh Y, Oki S, Tanaka C, Nakamura T, Adachi H, Yan YT, Shen MM, Hamada H. Two nodal-responsive enhancers control left-right asymmetric expression of Nodal. Dev Dyn. 2005 Apr;232:1031-6.

Song L, Zhang Z, Grasfeder LL, Boyle AP, Giresi PG, Lee BK, Sheffield NC, Gräf S, Huss M, Keefe D, Liu Z, London D, McDaniell RM, Shibata Y, Showers KA, Simon JM, Vales T, Wang T, Winter D, Zhang Z, Clarke ND, Birney E, Iyer VR, Crawford GE, Lieb JD, Furey TS. Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. Genome Res. 2011 Oct;21:1757-67.

Tang F, Barbacioru C, Bao S, Lee C, Nordman E, Wang X, Lao K, Surani MA. Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. Cell Stem Cell. 2010 May 7;6:468-78.

Vincent SD, Norris DP, Le Good JA, Constam DB, Robertson EJ. Asymmetric Nodal expression in the mouse is governed by the combinatorial activities of two distinct regulatory elements. Mech Dev. 2004 Nov;121:1403-15.

Viswanathan SR, Mermel CH, Lu J, Lu CW, Golub TR, Daley GQ. microRNA expression during trophectoderm specification. PLoS One. 2009 Jul 3;4:e6143.

Ying QL, Nichols J, Chambers I, Smith A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell. 2003 Oct 31;115:281-92.