UCSF
UC San Francisco Electronic Theses and Dissertations

Title

The Role of Polycomp-like 3 and Ofd1 in Embryonic Stem Cell Maintenance and Differentiation

Permalink

https://escholarship.org/uc/item/3488m43m

Author

Hunkapiller, Julie

Publication Date

2011

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

The Role of Polycomb-like 3 and Ofd1 in Embryonic Stem Cell Maintenance and Differentiation

by

Julie Hunkapiller

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

Graduate Division

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Acknowledgments

When I reflect back on my time in graduate school, one saying comes to mind, “It takes a village.” My journey has been inundated with joyous celebration and bitter disappointments, which I now understand are necessary components in the life of a scientist. Testing novel ideas requires many failures, yet outstanding results can induce embarrassing outbursts of cheering and clapping. I am extremely fortunate I have surrounded myself with people who wholeheartedly support and encourage me, so these are acknowledgements to my “village”.

First, I would like to thank all past and current members of the Reiter lab for being such a collaborative and fun group. Whether I have needed help performing mouse procedures or a drinking buddy at happy hour, I never had to go far. There are several people to whom I am particularly grateful. Kevin was instrumental in training me how to execute and think about science and was my go-to person for the first three years. When I first joined the lab, Veena was the only other person working on ES cells. She taught me how to be a better scientist and helped me navigate grad school. Since then, Nicole and Monika have joined Troupe ES. I am incredibly indebted to all of these girls for being there to bounce ideas around, cheer me up, or give me a day off on the weekend. You girls are awesome. I also want to thank Sunny for having my back (figuratively and literally). We joined the lab at the same time, and he has sat behind me and put up with me for the past five years. I can always count on him for a corny joke, an interesting discussion, or a weekend chat.

While everyone in lab contributed to my scientific growth and success, no one more so than Jeremy. Jeremy possesses all the brilliance of a great scientist, but it’s his easy going and optimistic attitude that makes him a great mentor. At times when I felt defeated by science, Jeremy was there to provide perspective, dust me off, and help me figure out how to move forward. I will always be grateful to Jeremy for allowing me to

pursue my project, even though it was completely outside the scientific focus of the lab. He took it upon himself to become familiar with the field I was studying and connect me to useful collaborators. This is an indication of his confidence and courage in science, as well as his commitment to his lab members. Jeremy's training and mentorship were invaluable to my success and will continue to be so.

During my time at UCSF, I have also received support from Barbara Panning and my thesis committee members, Miguel Ramalho-Santos and Geeta Narlikar, who have been very generous with their time and advice.

Although these ladies did not provide scientific guidance per se, they are my lifeline to sanity. I have known them for over a decade, and they have supported me through all life's twists and turns. Kayla, Lisa, Anna, Kristen, Jen, Emily, and Christine, you are the best friends anyone could ever ask for.

There is no way I could have made it this far without my family. All of my success and happiness is truly a testament to their unyielding love and encouragement. My parents are my biggest cheerleaders. Even though they don't always understand the science I'm doing, they always find the words I need to hear. I hope they know how much I love and appreciate them. My brother, Mark, has always been an example of what is kind and sincere. Even though he is younger than me, I look up to him.

Through some wonderful partnerships, I have gained more family members over the past few years. Mark's wife, Rachael, brings out the best in all of us, and is an amazing addition to the Gauldens. From the very beginning, my husband's family has been warm and welcoming. They have made me one of their own and have been exceptionally generous with their love and support. I am delighted to have them so close and to be a part of their family.

Finally, I want to thank the most important person in my life, my husband Nathan. Being a graduate student as well, he understood and supported me better than anyone

could have during this journey. Nathan is my rock, but besides that, he's the person that makes me smile the most. I am lucky and grateful to have found a true partner in life.

The Role of Polycomb-like 3 and Ofd1 in Embryonic Stem Cell Maintenance and Differentiation

By Julie Hunkapiller

Graduate Advisor: Jeremy F. Reiter M.D., Ph.D.

Embryonic stem cells (ESCs) hold tremendous promise for cell based therapeutics and as a biological tool. As differentiation of ESCs mimics early development, ESCs also serve as an excellent tool to study the underlying mechanisms of many developmental processes and cancer. The transition from ESC to somatic cell requires proper signaling as well as many intricate layers of gene regulation. Clarifying the interplay between these different processes is important for understanding the many transitions that happen during development as well as those that occur upon cancer metastasis.

, To elucidate some of the molecular mechanisms of ESC differentiation, we explored different areas of regulation within ESCs including cell signaling as well as epigenetic and transcription gene regulation. To best investigate these areas, we developed a recombinant gene technology called the Floxin system, in which we specifically and efficiently tagged genes within their endogenous loci in ESCs. The gene trap and tagged ESC lines were used to identify novel binding partners and investigate disease alleles.

Through the Floxin system, we demonstrate that Ofd1, a protein required for ciliogenesis, restrains neural differentiation in embryoid bodies (EBs). We also show that Ofd1 is essential for proper Hh and Wnt signaling and that disease alleles of Ofd1 cannot rescue proper neural differentiation. Secondly, we investigate Polycomb-like 3 (Pcl3), a protein involved in histone modifications and gene repression. We found that Pcl3 incorporates into Polycomb repressive complex 2 (PRC2) and promotes complex function. Diminishment of Pcl3 decreased PRC2 binding as well as ESC self-renewal,

indicating that Pcl3 enhances PRC2 targeting and ESC maintenance. Finally, we identify novel binding partners of Sall4, a transcription factor important for ESC pluripotency and reprogramming.

Notably, Ofd1, Pcl3, and Sall4 are all associated with human conditions and cancer. Thus besides understanding how these proteins function in ESC maintenance and differentiation, we have gained insight into the mechanisms by which disruption of these proteins contributes to human disease.

Table of Contents

Page

Chapter 1: Introduction……..1

Embryonic Stem Cell Maintenance and Pluripotency……1

Primary Cilia and Cell Signaling..……..3

Cilia and Development……5

Epigenetic Regulation of Embryonic Stem Cells: DNA Methylation, Chromatin Remodeling, and Histone Modifications……7

DNA Methylation……..7

Chromatin Remodeling……...8

Histones and Histone Modifications……9

Function and Regulation of Polycomb Repressive Complex 2……...10

Polycomb-like Proteins……..12

ESC Transcription Network and Spalt Family Transciption Factors.……14

Chapter 2: The ciliogenic protein Ofd1 regulates the neuronal differentiation of embryonic stem cells……..24

Abstract……24

Introduction……..25

Results……26

Discussion……34

Methods……..44

Chapter 3: Polycomb-like 3 promotes Polycomb repressive complex 2 binding to CpG islands and embryonic stem cell self-renewal…..……..49

Abstract……49

Introduction……..50

Results……52

Pcl3 is a component of PRC2……52

Pcl3 promotes ESC self-renewal……53

Pcl3 promotes trimethylation of H3K27……..…55

Pcl3 promotes PRC2 binding and function at a subset of PRC2 targets……..57

Pcl3 co-localizes with Suz12 at many PRC2 targets to promote H3K27me3……..59

Pcl3 regulates gene expression at a subset of PRC2 targets……..60

Pcl3 regulates Suz12, but not complex stability……..……61

The Tudor domain of Pcl3 is critical for its function……62

Pcl3 and Pcl2 are part of distinct PRC2 complexes that have overlapping targets……..…63

Pcl2 and Pcl3 co-localize with Suz12 at CpG islands……65

Discussion……66

Methods……103

Chapter 4: Pluripotency Factor Sall4 Associates with the Mi-2/Nurd Complex in Embryonic Stem Cells……..120

Introduction……120

Results……...121

Discussion……123

Methods……128

Chapter 5: Conclusions and Future Directions……..131

Bibliography……....138

Appendix: Neural induction and differentiation during embryonic development and in embryonic stem cells……157

Abstract……..157

Introduction……157

Neural induction through BMP antagonism……158

Beyond the Default Model……..……160

ES cells: “All cases are unique, and very similar to others”……..162

Conclusions and Perspectives……..……166

List of Tables and Figures

Tables **Page**

Table 1. *Gene expression changes upon Pcl3 shRNA ESCs by microarray*……..…..90

Table 2. *Gene expression changes in pre-plated Pcl3 shRNA ESCs by microarray*…...97

Figures **Page**

Figure 1. *Primary cilium components and structure*……..18

Figure 2. *Hedgehog signaling requires cilia for proper patterning and development*…...19

Figure 3. *Neural patterning and specification requires Shh, Wnt, and BMP signaling*….20

Figure 4. *Histones organize DNA into open or closed states*……...21

Figure 5. *Histone tail modifications are abundant and modulate gene expression*...…...22

Figure 6. *PRC2 trimethylates H3K27 to induce gene repression and binds accessory proteins which modulate function*……..23

Figure 7. *Ofd1^{Gt} EBs lack cilia and display altered Hh and Wnt signaling*…….38

Figure 8. *Ciliation in wild type, Ofd1^{Gt}, Ofd1^{Ofd1myc} EBs*……39

Figure 9. *Ofd1^{Gt} EBs exhibit increased neural differentiation*……...40

Figure 10. *Proliferation rates and BMP signaling are unaltered in Ofd1^{Gt} EBs*……..41

Figure 11. *Neural induction in mouse embryos lacking cilia and Ofd1 mutant missense EBs*……42

Figure 12. *Neural specification and patterning in wild type and Ofd1^{Gt} EBs and Kif3a mutant embryos*……...43

Figure 13. Pcl3 is a component of PRC2……71

Figure 14. Suz12-TAP and Pcl3-V5 localize to the nucleus……..……..72

Figure 15. Pcl3 promotes ESC self-renewal……..……...73

Figure 16. Depletion of Pcl3 does not affect differentiation to all three germ layers…....75

Figure 17. Pcl3 promotes PRC2 function……..……..77

Figure 18. Depletion of Suz12 and Pcl3……..……..……..78

Figure 19. Pcl3 affects Suz12 binding at a subset of PRC2 targets……..……79

Figure 20. Pcl3 knockdown causes the most significant depletion of Suz12 and H3K27me3 on chromosome 11.……....80

Figure 21. Pcl3 localizes to PRC2 targets……..……82

Figure 22. Pcl3 co-localizes with Suz12……83

Figure 23. Pcl3 regulates Suz12, but not complex stability……..……84

Figure 24. Pcl3 misregulates a subset of genes……85

Figure 25. Pcl3 requires Tudor domain residues for function……..……86

Figure 26. Pcl3 and Pcl2 participate in separate PRC2 complexes but overlap at many PRC2 binding sites……..87

Figure 27. Pcl2 and Pcl3 localize to CpG islands……..89

Figure 28. *Floxin mediated manipulation of the Sall4 locus*……...126

Figure 29. *Sall4-TAP associates with all members of the NuRD complex in ESC and differentiated cells*……..127

Figure 30. *In vivo neural induction*……..168

Figure 31. *In vitro neural induction*……..169

Chapter 1: Introduction

Normal mammalian development requires hundreds of cell types to be specified in an extremely controlled and precise manner. Each cell type formed is a result of extrinsic and intrinsic regulatory cues. Disrupting these cues can lead to many human conditions and even lethality. Therefore, exploring the mechanisms that induce specific and efficient transition from one cell type to another is fundamental to understanding development and disease.

Embryonic stem cells (ESCs) also have the capacity to form all cell types of the developing and adult organism. Furthermore, ESC differentiation utilizes many of the same regulatory cues as those used by the early embryo. Thus, ESCs are an excellent model system to understand the molecular mechanisms that regulate cellular transitions. Here, I will describe ESCs more extensively and outline the known extrinsic and instrinsic mechanisms that regulate ESC maintenance and self-renewal. I will also discuss several cues that control development and ESC differentiation including cell signaling, chromatin remodeling, and gene regulation.

Embryonic Stem Cell Maintenance and Pluripotency

Mouse embryonic stem cells (ESCs) are derived from the inner cell mass of blastocysts and can be maintained in culture indefinitely [1-3]. ESCs are grown on mouse embryonic fibroblast feeders (MEFs) or, in some cases, have been transitioned to feeder-free culture. If ESCs are cultured feeder-free, the medium must be supplemented with Leukemia Inhibitory Factor (LIF), which promotes ESC self-renewal by activating the JAK/STAT and MAPK signaling pathways [4,5]. ESCs also require

BMP/Smad signaling for maintenance, which is supplied through serum supplementation within the media [4,5].

Like the cells of the inner cell mass, ESCs are pluripotent, meaning they have the potential to form all cell types of the developing and mature organism [1,2,6]. Upon removal of LIF, ESCs spontaneously differentiate and can be grown in a monolayer or in low-adherent plates as embryoid bodies (EBs) [7,8]. Embryoid bodies form by ESC aggregation, and due to their three dimensional structure can mimic many features of the early embryo [6]. Like embryos, EBs specify all three germ layers and cavitate, which mimics early embryo invagination [9,10]. Moreover, specific neural subtypes can be produced in EBs by addition of Hh, BMP, or Wnt ligands or repressors (See Appendix) [11,12]. Due to the absence of neural tube structure, however, spontaneous differentiation often results in considerable variability between EBs.

Significant progress has been made recently on methods to differentiate ESCs to specific cell types including neurons, pancreas, heart, neural crest, and endothelial cells [13-19]. As a consequence, the desire to use ESCs as a model system for studying development and disease has expanded. In order to easily manipulate the mouse genome in ESCs, we developed the Floxin system to introduce exogenous DNA into endogenous loci [20]. A two step process, the Floxin system uses gene trap ESC lines in combination with Cre-lox technology to recombine DNA with high efficiency [20]. Specialized shuttle vectors are recombined within the endogenous locus to create tagged proteins or reporter constructs expressed from the endogenous promoter [20]. This technology can be used to study mechanisms of ESC pluripotency and differentiation as well as model human development and disease.

Understanding how ESCs differentiate to specific cell types requires insight into the molecular character of ESCs and somatic cells. This complicated transition involves cell signaling as well as transcriptional and epigenetic changes [21].

Primary Cilia and Cell Signaling

There are a number of signaling pathways that are important for ESC differentiation and normal development, two of which include the Hedgehog (Hh) and Wnt pathways. Work in recent years has revealed that these two pathways are uniquely regulated through primary cilia [22,23]. Primary cilia are a specialized type of cilia that can be found in most cell types of the developing embryo and adult [24,25]. Cells only generate a single primary cilium, yet all cilia are structured similarly and play important roles in development.

Cilia are microtubule based structures that protrude from the cell membrane and are found in organisms ranging from unicellular plants to humans (Fig. 1A) [24]. Until recently, mammalian cilia were thought to be exclusively involved in movement of liquids in the lungs, brain, and kidney. However, work in the past decade has revealed that their role in embryonic patterning and organ function is quite broad. For example, cilia are necessary for normal development of the brain, heart, kidney, limbs, and skeleton as well as for proper function of the eyes, ears, and nose [24-27].

Using the mother centriole as its base, cilia are built using kinesins, anterograde microtubule motors that carry and deposit cargo (Fig. 1B-C) [28]. Two of these kinesins, Kif3a and Kif3b, heterodimerize and bind carriers called intraflagellar transport (IFT) proteins, which bring cargo into the cilium (Fig. 1C). Once the cargo is deposited, IFT proteins bind dynein retrograde motors to remove proteins from the cilium (Fig. 1C) [28]. Thus, perturbations of intraflagellar transport through loss of Kif3a, Kif3b, or Ift88 cause defects in ciliogenesis [24,28]. In contrast, loss of dynein transporters results in accumulation of debris within the cilium. In either case, cilia function is disrupted [24,28].

Cilia have many functions including sensing the extracellular space, creating flow, and mediating cell signaling [24]. Loss of cilia perturbs all of these processes. For example, sensory cilia in the retina and olfactory bulb are essential for proper sight and smell [24]. Nodal cilia create flow that ultimately forms the left-right axis in the embryo [29]. Furthermore, mammalian Hh signaling cannot occur without cilia, and Wnt and Pdgf signaling are both regulated by cilia [22,23,30]. Thus, cilia are responsible for relaying signals in many facets of development and tissue hometostasis.

The reason that cilia are required for Hh signaling is because several Hh pathway components entail primary cilium localization for their function, [22,31,32]. In the absence of pathway induction, Patched (Ptch), a twelve-pass transmembrane protein, prevents the translocation of Smoothened (Smo), a seven-pass transmembrane protein, into the primary cilium (Fig. 2A). The presence of Sonic Hh, a Hh pathway ligand, allows Smo to be released from Ptch and move into the primary cilium (Fig. 2B) [22]. Gli proteins, the effectors of the Hh pathway, localize to the primary cilium where they are processed to either activator or repressor forms, depending of the status of Smo localization [31,33,34]. In short, Smo translocation to the cilium causes increased Gli1 processing to activator forms and decreased processing of repressor forms (Fig. 2B) [32]. Processed Gli proteins are then shuttled to the nucleus to induce or repress transcription of Hh target genes [35,36].

There are three Gli paralogs in mammals, Gli1, Gli2, and Gli3. Gli1 does not require processing and typically acts as an activator of Hh signaling [37]. In contrast, Gli2 and Gli3 both require cilia localization for processing (Fig. 2A-B) [31,33,34]. Gli2 has a role in both active and repressive transcription, while Gli3 typically acts as repressor [37]. In the absence of Hh ligand, Gli3 is processed to a truncated repressor form in the cilium. Upon pathway activation, Gli3 maintains its full length form and

cannot inhibit the pathway (Fig. 2A) [35,36,38,39]. This ratio of full length to truncated Gli3 protein influences whether Hh target genes will be activated or repressed.

Primary cilia also play a role in regulating Wnt signaling, which is important for promoting cell proliferation and self-renewal. During canonical pathway activation, Wnt ligand binds two co-receptors, Frizzled (Fz) and LRP5/6, inducing phosphorylation of LRP5/6 tails and Axin2 binding [40,41]. This event stabilizes β-catenin, which then enters the nucleus, binds transcriptional co-activators (TCF/LEF), and initiates transcription [40,41]. Several Wnt pathway components are found at the primary cilia, but how this localization contributes to Wnt pathway transduction or repression is unclear [42]. In contrast to Hh signaling, the cilium restrains Wnt signaling as disruption of cilia causes enhanced downstream Wnt activity [23,42]. Indeed, mutation of *Kif3a* or *Ift88* causes stabilization and increased nuclear localization of β-catenin as well as over-activation of a Wnt-inducible reporter [23]. Thus, loss of cilia causes both abrogated and enhanced signaling depending on the pathway.

Cilia and Development

Because cilia are so important in regulating cell signaling and sensing the extracellular space, it makes sense that cilia are also critical for normal development. Loss of cilia causes a number of human diseases, termed ciliopathies, some of which include Bardet-Biedl, Meckel-Gruber, Joubert, and Oral-facial-digital Syndromes [25-27]. Although phenotypes observed in these syndromes vary between patients and between diseases, there are some common features including cystic kidneys, mental retardation, retinal degeneration, obesity, and hand and facial malformations [43-45]. Severe cilia perturbations can be embryonic lethal in humans, and mouse embryos lacking cilia typically die between 10-12 days post-fertilization [27,46-48]. These embryos often

present with hydrocephalus, situs inversus, pericardial swelling, and neural and limb patterning defects [29,46-49].

Some of the phenotypes observed in ciliopathies and in mouse mutants lacking cilia are a result of disrupted Hh signaling. For example, the neural tube and limb bud are patterned in part through Hh signaling gradients (Fig. 2C, 3) [37]. Shh, initially produced in the notochord and eventually the floor plate, specifies the ventral portion of the neural tube (Fig. 3) [50,51]. The most ventral neural subtypes are the V3 interneurons and motor neurons, which require the highest levels of Shh for specification (Fig. 2C, 3) [50-52]. Loss of cilia or Hh signaling causes dorsalization of the neural tube, sometimes resulting in a complete loss of these most ventral neural subtypes (Fig. 2D) [37,52]. Embryos lacking cilia also present with extra digits due to Hh patterning defects in the developing limb bud (Fig. 2D) [31,33,37,38,47]. In contrast, over-induction of Hh signaling causes ventralization of the neural tube and loss of digits [37].

Extra digits or polydactyly is seen in multiple ciliopathies including Oral-facial-digital type 1 Syndrome (Ofd1 syndrome). This rare ciliopathy is X-linked disorder that affects the face, oral cavity, and digits and sometimes causes cystic kidneys and mental retardation [53]. Ofd1 syndrome is caused by deletions or missense mutations in the *Ofd1* gene [53]. This disease is seen almost exclusively in females, who are heterozygous for the mutant *Ofd1* allele, while males with mutant *Ofd1* die prenatally [53]. *Ofd1* encodes a distal centrosomal protein that acts like a cap to restrict the size of centrioles [54]. Thus, cells lacking Ofd1 have long centrioles and show mislocalization of some distal centriolar proteins [54]. These defects in centriolar structure and composition prevent ciliogenesis *in vitro* and *in vivo* [47,53,54].

Ofd1 mutant mice have many of the same characteristics seen in human Ofd1 patients [47,55]. Like human cases, male mice mutant for *Ofd1* die prenatally and exhibit polydactyly, deformities of the face and oral cavity, bone abnormalities, as well as

mispatterned neural tubes and limbs [47]. Female mice heterozygous for *Ofd1* die at birth unlike human patients with Ofd1 syndrome [47]. Differences in Ofd1 syndrome severity between female mice and humans are attributed to murine specific X-inactivation at this particular allele [47,55]. In mice, *Ofd1* is randomly X-inactivated so that some cells express wild-type *Ofd1* while other cells lack Ofd1 [55]. In humans, however, *Ofd1* is not X-inactivated, allowing both alleles to be expressed [55]. As a result, all human cells contain some wild type Ofd1 protein, resulting in less severe phenotypes than in the mice [47].

In short, cilia and cell signaling are incredibly important in mammalian development and participate in the early steps of instructing cells to become specified. However, the steps following, which include epigenetic and transcriptional gene regulation, ultimately determine whether a gene is activated or repressed. These intrinsic factors also play significant roles in development as well as ESC maintenance and differentiation.

Epigenetic Regulation of Embryonic Stem Cells: DNA Methylation, Chromatin Remodeling, and Histone Modifications

DNA Methylation

Epigenetic gene regulation consists of modifications and changes to chromatin that is independent of DNA sequence. There are several forms of epigenetic gene regulation, which include DNA methylation, chromatin remodeling, and histone modification [56-60]. DNA methylation is perhaps the best understood mechanism of gene regulation, which includes addition or removal of a methyl group to a cytosine or adenine. In general, highly methylated DNA is associated with repressed genes, while

hypomethylated DNA is associated with active transcription [56,61,62]. Hypomethylated cytosine-guanine dinucleotides (CpGs) are an exception to this generalization. Although 70-80% of CpGs in mammals are methylated, 60% of human promoters contain CpG dense regions that are hypomethylated [56,62]. These regions termed, CpG islands, remain unmethylated in every tissue during all stages of development, indicating that these regions are protected from methylation [60,63]. Situations where CpG islands become methylated are typically associated with X-inactivation or imprinting [56,60,62]. CpG islands are often bound by a myriad of transcription factors and chromatin modifying complexes. Thus, it is possible that proteins are recruited to these regions to regulate gene expression.

Chromatin remodeling

The two other forms of epigenetic gene regulation, chromatin remodeling and histone modification, do not involve modifications to the DNA but rather revolve around changes in chromatin structure. In eukaryotes, DNA is not just free floating but highly organized into chromatin [64]. Chromatin consists of DNA and the proteins which spool the DNA called histones (Figure 4A-B) [58]. There are six types of histones, H1, H2A, H2B, H3, H4, and H5 [64]. The core histones, H2A, H2B, H3, and H4, form dimers, which then associate to form an octamer [65]. Approximately 146 base pairs or nearly 2 rounds of DNA encircle the octamer to create a nucleosome (Figure 4D) [64]. Histone H1 or H5 act as cinchers at the entry and exit site to lock the DNA in place around the nucleosome [65].

In open chromatin or euchromatin, DNA is loosely bound around nucleosomes and appears as “beads on a string” (Figure 4A) [66]. This semi-structured form allows

binding of transcriptions factors, chromatin remodelers, and RNA polymerase II, which results in active or poised gene expression [58,64].

Another form of chromatin found during interphase is heterochromatin, which can be separated into constitutive heterochromatin and facultative heterochromatin (Figure 4B) [59]. Constitutive heterochromatin is characterized by tightly compacted and highly repetitive regions of DNA that are never expressed such as telomeres and centromeres. Facultative heterochromatin also contains condensed and silenced DNA but can decompact upon induction of specific developmental or environmental cues [58,59]. Chromatin remodeling proteins can mediate a heterochromatin to euchromatin transition, whereby silenced genes become active [58,59]. In contrast, signaling can also cause euchromatin to condense and become silenced facultative heterochromatin [58]. Hence, changes in chromatin structure at a macro level influence gene regulation by determining whether activating proteins can access genes. Yet chromatin can be altered molecularly as well, and these histone modifications allow for rapid and efficient gene activation or inhibition.

Histones and Histone Modifications

The core histones of the octamer have amino and carboxy terminal tails that protrude from the nucleosome (Figure 4D) [58,67]. These “tails” can be acetylated, methylated, phosphorylated, or ubiquitinated by different histone modifying complexes (Figure 4C and 5A) [67]. Each particular mark confers information and influences gene regulation [58]. Thus far, 37 modifications have been identified at 35 sites, and more likely exist (Figure 5A) [67]. Some modifications are associated with gene expression, like trimethylation of lysine 4 and lysine 36 on histone H3 (H3K4me3, H3K36me3) and acetylation of lysine 27 on histone H3 (H3K27ac) [58,67]. Others, however, are

associated with gene repression like trimethylation of lysine 9 and lysine 27 on histone H3 (H3K9me3, H3K27me3) and ubiquitination on lysine 119 on histone H2A (H2AK119ubiq) [58,67]. Despite knowing that numerous modifications exist, it is unclear how many of these marks are transferred and what effects they have on gene expression.

Relationships between different modifications show that some marks are mutually exclusive while the combination of others confers unique function [68]. For example, H3K27me3, a repressive mark, and H3K27ac, an active mark, are often mutually exclusive within specific regions [69]. Moreover loss of the H3K27me3, causes an increase in H3K27ac, suggesting that the antagonistic switch between these two marks acts as a regulator of gene expression [69].

ESCs have a unique chromatin signature, called bivalency, whereby approximately two thousand gene promoters display both active (H3K4me3) and repressive (H3K27me3) marks (Figure 5B) [70]. Among bivalent genes are many developmental regulators that are repressed in ES cells, but activated upon ESC differentiation in a cell-type specific manner [21]. This suggests that the bivalent signature poises genes for activation (Figure 5B) [70]. The rapid responsiveness at these genes is due to stalled RNA Polymerase II, which is released after removal of H3K27me3 (Figure 5B) [58,71]. This suggests that H3K27me3 and the Polycomb complex that makes this mark are the final obstacles to active transcription [71-74].

Function and Regulation of PRC2

Polycomb proteins are responsible for catalyzing several histone marks, including H3K27me3. Polycomb complexes exist in a wide-range of organisms including multicellular plants, flies, and mammals [57,62,75,76]. Polycomb proteins were first

identified in Drosophila for their role in segment patterning through Hox genes [77,78]. Since then, two Polycomb complexes have been identified in mammals, Polycomb repressive complex 1 (PRC1) and Polycomb repressive complex 2 (PRC2) [75]. PRC2 is composed of three core components Eed, Suz12, and Ezh2, a methyltransferase responsible for catalyzing di- and trimethylation of H3K27 (Figure 6A)[75]. Following PRC2 trimethylation of H3K27, PRC1 is recruited through an unknown mechanism and catalyzes H2A119ub [67]. Both of these marks are involved in initiating gene repression and subsequent chromatin condensation.

Abolishment of H3K27me3 causes many normally repressed genes to be expressed, suggesting that removal of this mark may be the final release to gene activation [79-82]. The existence of bivalent genes, which have both active (H3K4me3) and repressive (H3K27me3), supports this model [70]. Current data suggests that removal of H3K27me3 and subsequent release of RNA Polymerase II allows expression of bivalent genes upon ESC differentiation (Figure 5B) [58,71]. It is possible, however, that intermediate steps exist but have not been explored.

In Drosophila, PRC2 is recruited to defined binding sites called Polycomb repressive elements (PREs) where it induces gene repression [83-85]. While few mammalian PREs have been identified, PRC2 is known to associate with CpG islands [86,87]. PRC2 was thought to be specifically recruited to these regions, however a recent report suggests that PRC2 non-specifically binds CpG rich regions that are unoccupied by other chromatin binding complexes [60,88]. They show that PRC2 is recruited to CpG islands lacking activating sequences as well as GC-rich elements from *E.coli*. Thus, PRC2 recruitment may not require specific GC sequences as much as chromatin availability within CpG islands [88,89].

PRC2 components contain DNA, RNA, and histone binding domains, but how these proteins interact as a larger complex is unclear. Crystallization and binding

studies of the WD40-repeats of Eed demonstrate that it recognizes trimethylated H3K27 and serves to bind and propagate H3K27me3 during cell division (Figure 6B) [90]. Moreover, Ezh2, like other methyltransferases, contains a SET domain, which is essential for catalyzing the transfer of methyl groups from S-Adenosyl methionine to H3K27 (Figure 6B) [3,75,91,92].

Multiple accessory proteins regulate PRC2 function (Figure 6A-B) [75]. For example, retinoblastoma binding protein 4 and 7 (Rbbp4, Rbbp7) and adipocyte-enhancer binding protein 2 (Aebp2) bind PRC2 and enhance catalysis of H3K27me3 [75,93,94]. Despite conflicting reports about whether Jarid2 promotes or antagonizes PRC2, it interacts with PRC2 and is important for recruitment of PRC2 to target genes [72,95-98]. It is possible that the effects of Jarid2 on PRC2 function are cell type-dependent.

Polycomb-like proteins

Polycomb-like family proteins also regulate PRC2. Polycomb-like was originally identified in Drosophila and was named based on the resemblance of polycomb-like mutants to polycomb mutants [99,100]. Mammals have three homologues of Polycomb-like which share approximately 40% homology. All three Polycomb-like proteins contain an amino terminal Tudor domain closely linked to two carboxy PHD zinc fingers (Figure 6B) [75,101,102]. Both Tudor domains and PHD zinc fingers bind methylated or unmethylated histone arginines and lysines [103,104].

Polycomb-like 1 (Pcl1), also called PHD finger protein 1 (Phf1), is expressed in adult tissues and in the testis during development [101,105]. Pcl1 binds Ezh2 and promotes high levels of H3K27me3 in HeLA cells, 3T3 cells, and GC1spg, a male germ

cell line [105-107]. In addition, Pcl1 has PRC2-independent roles as it promotes non-homologous end-joining at sites of DNA double strand breaks [108].

Another paralog, Polycomb-like 2 (Pcl2) or Metal-regulatory transcription factor 2 (Mtf2), is a PRC2 regulator in ESCs and MEFs [101,109]. Interestingly, Pcl2 has both active and inhibitory roles in PRC2 function. In ESCs, knockdown of *Pcl2* causes reduced H3K27me3 at specific targets, like *Hox* genes and pluripotency transcription factors, yet global H3K27me3 levels increase compared to control [101,109,110]. In contrast, *Pcl2* mutant MEFS show increased H3K27me3 levels at the *Cdkn2a* locus, suggesting that Pcl2 can restrain PRC2 in certain contexts [109].

Polycomb-like 3, also called PHD finger protein 19, is alternatively spliced in humans to produce two isoforms: full length Pcl3 (hPcl3L) and a short isoform lacking the second PHD domain (hPcl3S) [111,112]. While both isoforms associate with PRC2, they do not exist within the same complex, and only the long isoform can homodimerize [111]. Sequence alignment of Drosophila, mouse, and human Pcl3 indicates that certain residues within the Tudor domain that may be important for histone binding [102]. In addition, Pcl3 is overexpressed in many cancers including liver, colon, lung, cervical, skin, rectal, and uterus cancers [112].

Crystallization of several PHD fingers revealed that the presence or absence of certain residues can be used to predict whether PHD fingers will bind methylated or unmethylated histones [113,114]. Based on sequence analysis of mouse Polycomb-like proteins, the amino PHD finger of Pcl1, Pcl2, and Pcl3 is predicted to bind unmethylated histones whereas the carboxy PHD finger is predicted to be methylated histone binder. Thus, it is possible that the Pcl3 Tudor domain and PHD zinc fingers cooperate to bind a specific histone signature, which promotes PRC2 recruitment or activity.

During development, Pcl proteins are important for tissue-specific gene regulation. Xenopus contain two Drosophila Polycomb-like homologues, which regulate

neural development [115,116]. Knockdown of Xenopus *polycomb-like 1* (*XPcl1*) causes increased fore- and midbrain tissue whereas overexpression represses multiple anterior neural target genes [115]. Xenopus polycomb-like 2 (XPcl2) also regulates anterior neural genes and helps pattern the anterior-posterior neural tube [116]. Notably, perturbation of other Polycomb proteins in Xenopus causes similar phenotypes [115,117].

In chicks, Pcl2 associates with Ezh2 and represses Shh expression specifically on the right side of Hensen's node, establishing left-right asymmetry [112]. Repression of *Pcl2* in chicks causes ectopic expression of Shh on the right side of the node along with randomized heart looping [112]. Overexpression of chick *Pcl2* in the mouse limb bud results in a near loss of Shh expression in the limb bud [112]. Interestingly, *Pcl2* mutant mice have normal left-right asymmetry yet show developmental abnormalities in axial skeleton formation [118]. These inconsistencies may be attributed to functional redundancy from Pcl1 or Pcl3 in mice.

Thus, PRC2 and its associated proteins are incredibly important for gene repression. The interplay between repressive complexes and activating transcription factors creates an environment where genes are rapidly and precisely regulated. This is particularly important in ESCs, where cells are induced to become numerous cells types.

ESC Transcription Network and Spalt Family Transcription Factors

ESCs are maintained through a complex network of factors that promote self-renewal and preserve pluripotency. The JAK/STAT and BMP/Smad signaling pathways are essential for ESC maintenance because they are responsible for inducing expression of transcription factors known to maintain pluripotency like Oct4, Nanog, and Sox2 [3-5,15]. Loss or overexpression of any one of these factors causes ESC

differentiation, indicating that levels of these factors must be very precisely monitored and regulated [3-5]. Besides Oct4, Nanog, and Sox2, several other transcription factors, C-myc, Sall4, and Klf4, promote ESC pluripotency and self-renewal [119,120]. Sall4 is particularly intriguing because it can function as an activator and repressor in gene regulation.

Sall4 is one of four SAL family proteins, Sall1-4, which are homologues of the Drosophila gene *spalt*. Spalt is necessary for proper embryonic development of the anterior head and posterior tail segments as well as for morphogenesis and organogenesis later in development [121,122]. Similarly, SAL family members are involved in early embryonic development as well as organogenesis and maintenance of stem cell pools [123]. Mutations in *Sall1* or *Sall4* can cause Townes-Brocks Syndrome, branchio-oto-renal syndrome, Duane Radial Ray Syndrome, Ivic Syndrome, and Holt-Oram Syndrome [121,123,124]. Most of these syndromes are autosomal dominant and cause defects in multiple organs and tissues including kidney, heart, eyes, bones, ears, and digits.

Sall4 expression initiates at the two cell stage and increases until blastocyst formation [125]. *Sall1* and *Sall3* are also expressed very early and loss of any of these proteins causes peri-implantation lethality [126-128]. Although *Sall4* mutant ES cells can be derived, it occurs at extremely low efficiency [129]. In addition, these cells have an increased propensity to spontaneously differentiate due to spurious expression of trophectoderm differentiation genes [129].

In addition to their role in development, Sall1 and Sall4 are implicated in ESC pluripotency and somatic cell reprogramming [119,130,131]. *Sall1* and *Sall4* are both highly expressed in ESCs and bind several core ESC transcription factors. For example, Sall1 binds Nanog and Sox2 to promote *Nanog* expression through autoregulation [130]. However, Sall1 can also independently repress mesodermal and ectodermal genes

[130]. Similar to Sall1, Sall4 associates with Nanog and Oct4 to create both positive and negative regulatory loops [131-134]. For example, Sall4-Nanog and Sall4-Oct4 complexes positively auto-regulate themselves, yet Sall4 competes with Oct4 to repress other SAL family members like Sall1 and Sall3 [131,132,134]. In addition, Sall4 acts as an enhancer of cell reprogramming upon ESC-somatic cell fusion [119]. Thus, Sall1 and Sall4 promote pluripotency twofold, through activation of pluripotency genes and repression of differentiation genes.

The role of Sall4 in positive and negative gene regulation is not limited to ESCs. In kidney cells and cancer cell lines, Sall4 localizes to heterochromatin and cooperates with CyclinD1, a cell cycle regulator, to inhibit gene expression [135]. During embryonic development, however, Sall4 interacts with Tbx5 to positively promote genes essential for heart and limb patterning [136]. Hence, different Sall4 binding partners may confer distinct functions that result in either active or repressive properties.

SAL family proteins have several conserved domains including multiple zinc-fingers, which can bind protein, DNA, and chromatin [123]. Besides containing several sumoylation sites, Sall1 also contains a conserved twelve amino acid motif that is sufficient to recruit the nucleosome remodeling and deacetylase (NuRD) complex (Mi-2 complex) [137]. As the NuRD complex primarily functions in gene repression, mutation of this motif prevents association with the NuRD complex and causes derepression of target genes such as *Gbx2* [138,139].

Although Oct4, Nanog, Sox2, C-myc, Sall4, and Klf4 make up the core group of ESC regulators, they connect to several dozen more transcription factors [15]. It is because of this very intricate and powerful network that ESCs have such unqiue characteristics. No other *in vitro* cell type can form all cell types of the adult organism and self-renew. Thus, besides being a useful tool, ESCs are a remarkable cell type to study and hold much promise for understanding and treating human disease.

In an effort to better understand the mechanisms that regulate the transition from ESCs to somatic cells, we investigated cues from all regulatory steps. We explored how cilia and cell signaling affect ESC differentiation to different cell types. We also assess the role of Polycomb-like 3 in PRC2 function and ESC self-renewal. Finally, we investigate binding partners of Sall4 to try to elucidate the activating and repressive roles of Sall4 in ESCs. These studies have provided insight into the extrinsic and intrinsic regulatory cues which allow ESCs to self-renew and differentiate.

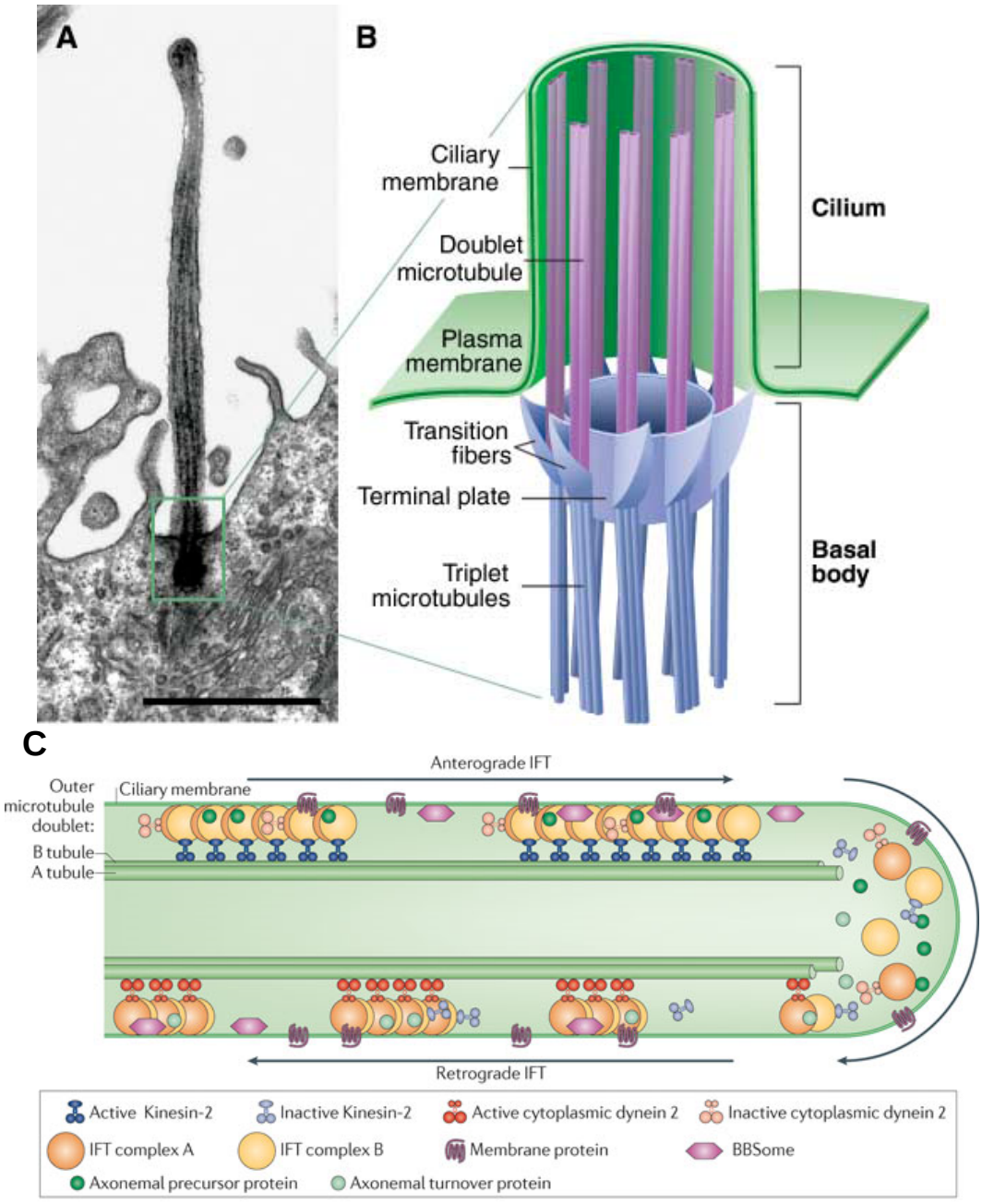


Figure 1. (Adapted from Singla et al. 2006 and Ishikawa and Marshall 2011) *Primary cilium components and structure [24,28].*

(A) Primary cilia extend out from the cell membrane to sense fluid flow and signaling gradients. **(B)** Primary cilia are composed of nine microtubular doublets, which extend from the mother centriole. **(C)** Cartoon illustrating intraflagellar transport. Anterograde transport uses Kinesin 2 motors, Kif3a and Kif3b, and intraflagellar transport proteins, like Ift88, to bring cargo into the cilium like membrane proteins or cilium components. Retrograde transport uses dynein motors to remove proteins from the cilium.

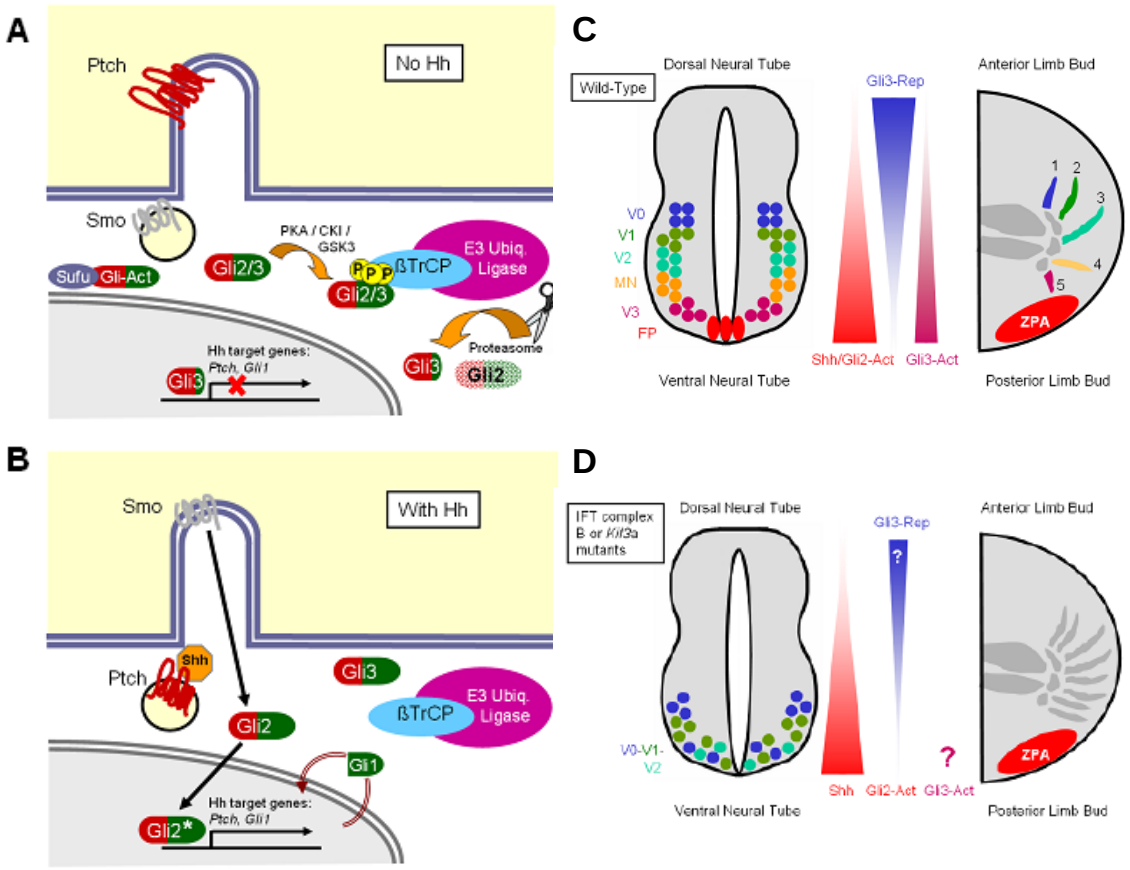


Figure 2. (Adapted from Wong and Reiter 2008) *Hedgehog signaling requires cilia for proper patterning and development [37].*

(A) In the absence of Hh signaling, Ptch prevents Smo localization to the cilium resulting in cilium-dependent Gli3 processing to Gli3 repressor and inhibition of Hh targets genes, *Ptch* and *Gli1*. In addition, Gli1 and Gli2 are inhibited and targeted for protein degradation respectively. **(B)** Upon Hh binding, Ptch releases Smo, which then localizes to the cilium and induces processing of Gli2 to its active form. Gli1 is relieved of its inhibition and both active Gli proteins enter the nucleus to activate Hh targets genes. **(C)** Both the ventral neural tube and the limb bud are patterned by Hedgehog gradients. In wild type embryos, Hh ligand produced in the notochord and floor plate cause six subtypes of neurons to be specified in the ventral neural tube. Wild-type limb buds are patterned both by active and repressive Gli proteins, which results in the formation of five digits. **(D)** Loss of primary cilia through mutation of Ift complex B or Kinesin 2 motor components causes decreased Hh signaling and dorsalization of the neural tube. In addition, loss of cilia in the limb bud eliminates both Gli3 repressor processing and active Hh signaling. Thus, the limb bud loses the vast majority of its location identity, resulting in the formation of additional digits.

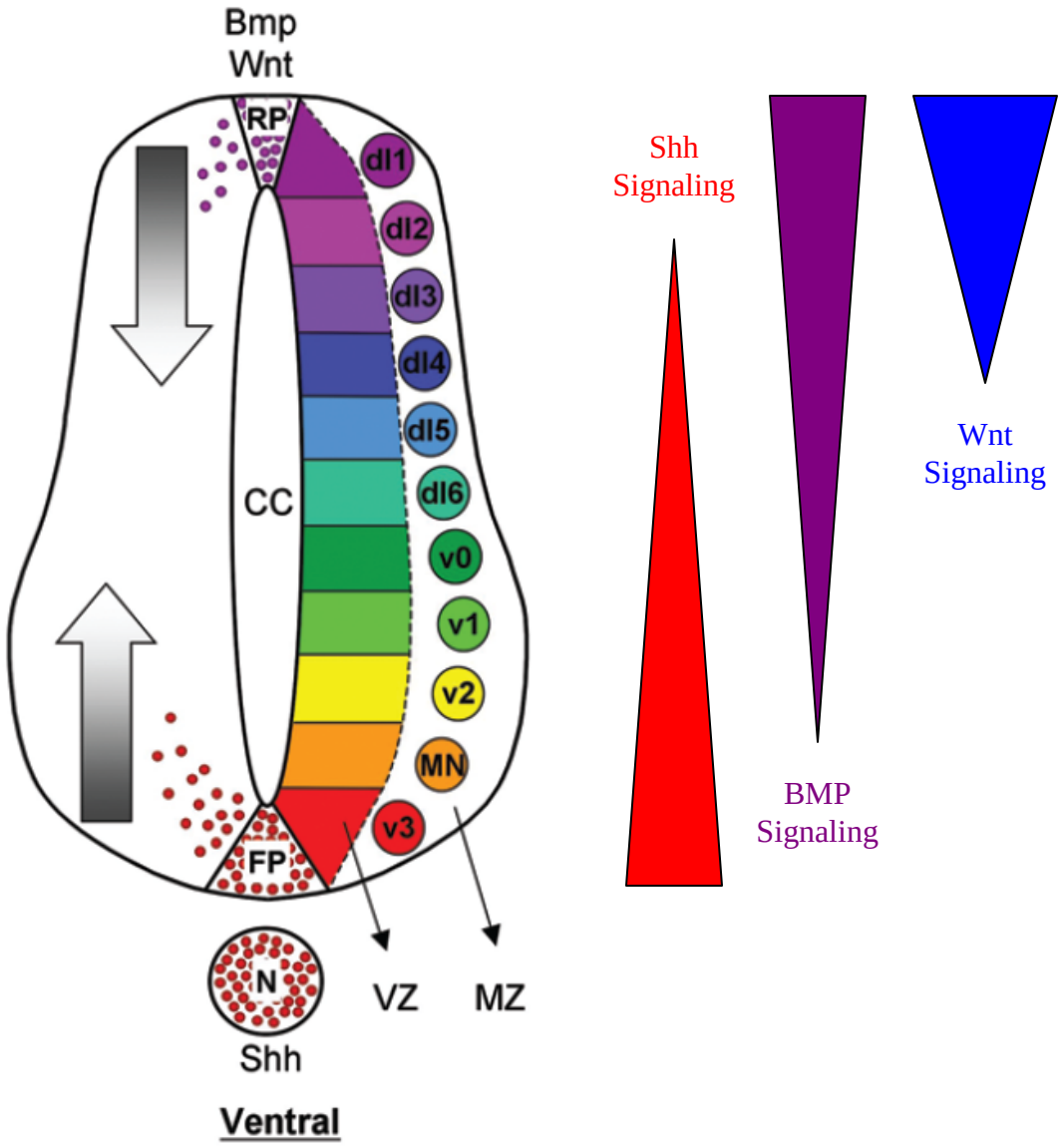


Figure 3. (Adapted from Ulloa and Briscoe 2007) *Neural patterning and specification requires Shh, Wnt, and BMP signaling [140].*

Shh is initially produced in the notochord (N) and subsequently the floor plate (FP) while BMP and Wnt signaling emanates from the roof plate. Levels of each of these signaling pathways provide positional identity through the expression of particular transcription factors.

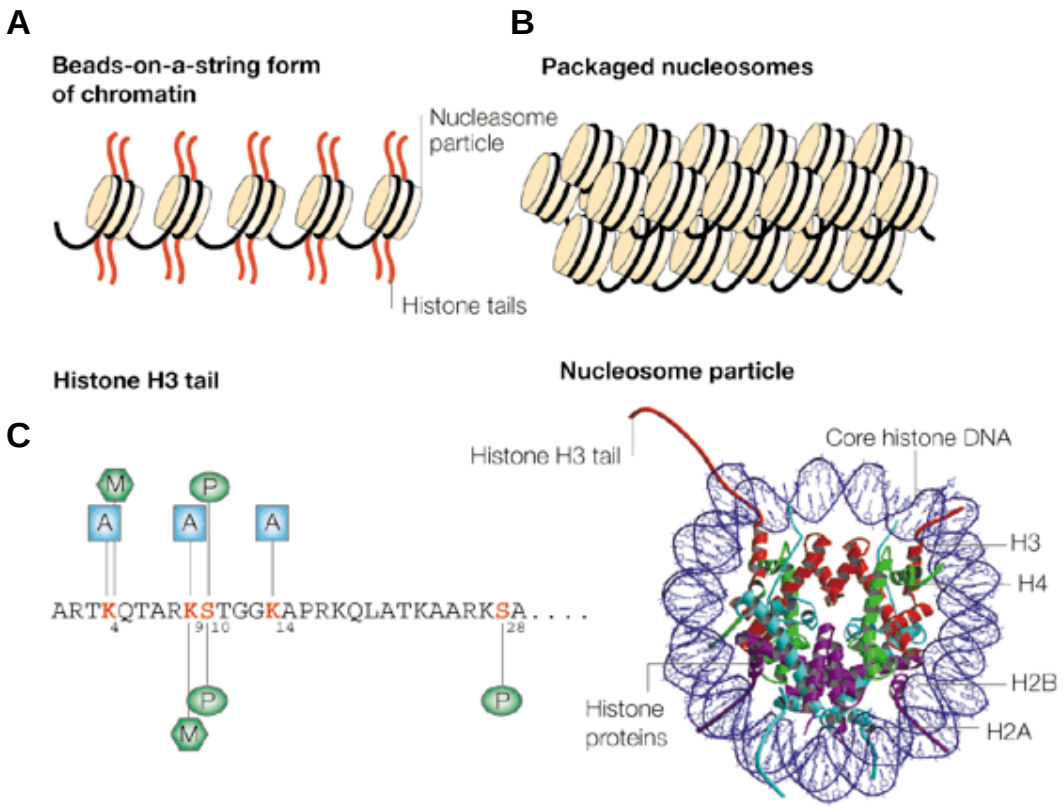


Figure 4. (Adapted from Marmorstein 2001) *Histones organize DNA into open or closed states* [66].

(A) In an open chromatin state, DNA is loosely wrapped around histone octamers, which allows binding of transcription factors, chromatin remodelers, and histone modifying complexes. Typically seen in euchromatin, this 11nm fiber organization represents beads on a string and lacks cinching histones H1 and H5 [64]. Red lines represent the histone tails, which protrude from the nucleosome to regulate recruitment and allow for chromatin modification [66]. **(B)** In the closed state, histone H1 and H5 cinch nucleosomes to create a condensed 34nm fiber. Unlike transiently repressed genes, which can maintain an open chromatin state, permanently repressed genes are highly compacted and prevent binding of polymerases or DNA repair machinery [64]. **(C)** Histones H3, H4, H2A, and H2B dimerize and then interact to form a nucleosome octamer. The tails of these histones can be modified at certain residues, which is illustrated for histone H3.

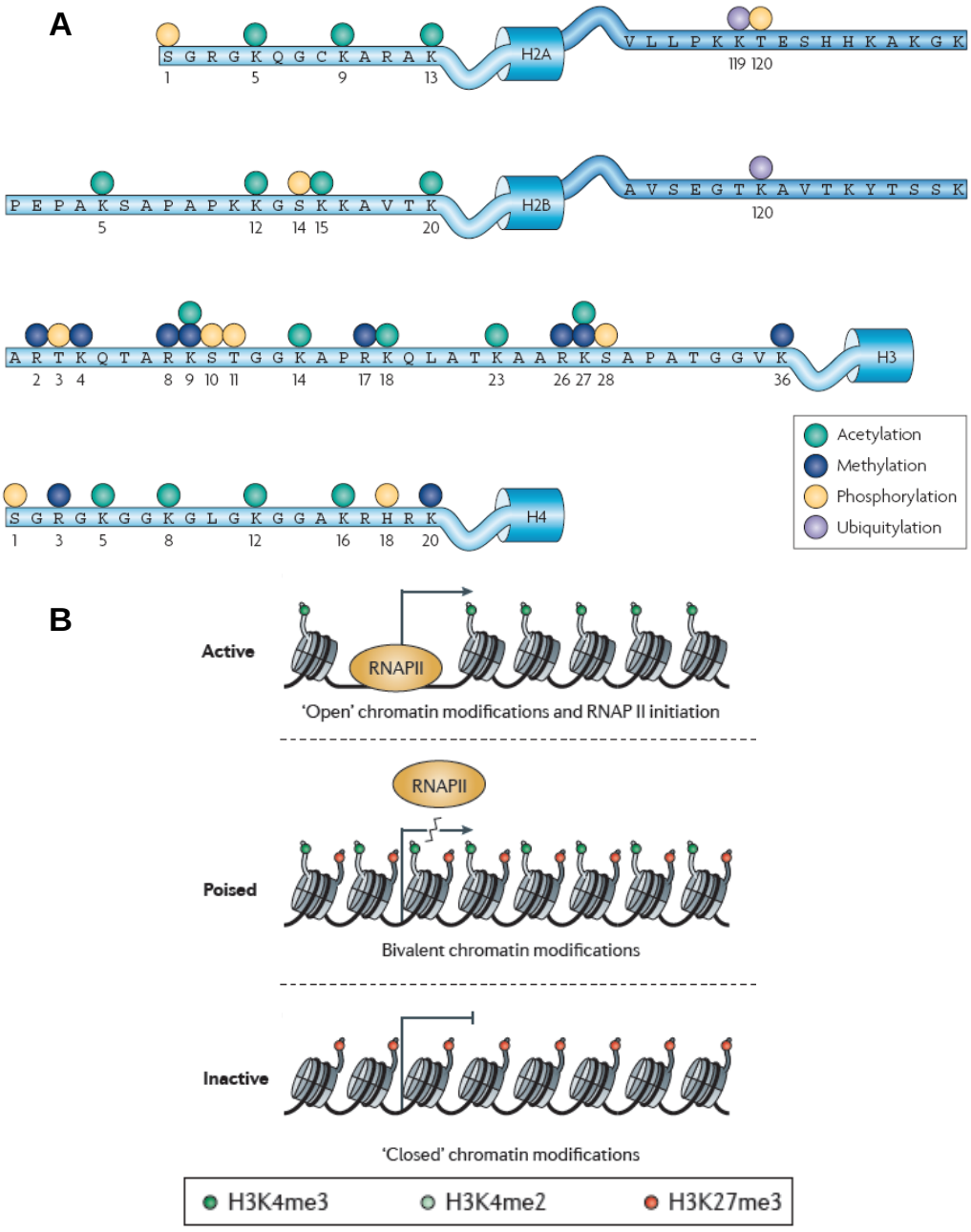


Figure 5. (Adapted from Spivakov and Fisher 2007 and Zhou et al. 2011) *Histone tail modifications are abundant and modulate gene expression* [58,67].

(A) Histone tails can be acetylated, methylated, phosphorylated, and ubiquitinated on many residues. These modifications can occur concurrently, or for some are mutually exclusive. Trimethylation of lysine 4 or lysine 36 on histone H3 is associated with gene activation, while trimethylation of lysine 27 and lysine 9 on histone H3 is associated with gene repression. **(B)** ESCs have a unique signature where some genes have both H3K4me3 and H3K27me3 called bivalent domains. This bivalent state causes RNA Polymerase II to stall, which keeps the genes repressed but poised for activation upon differentiation [70].

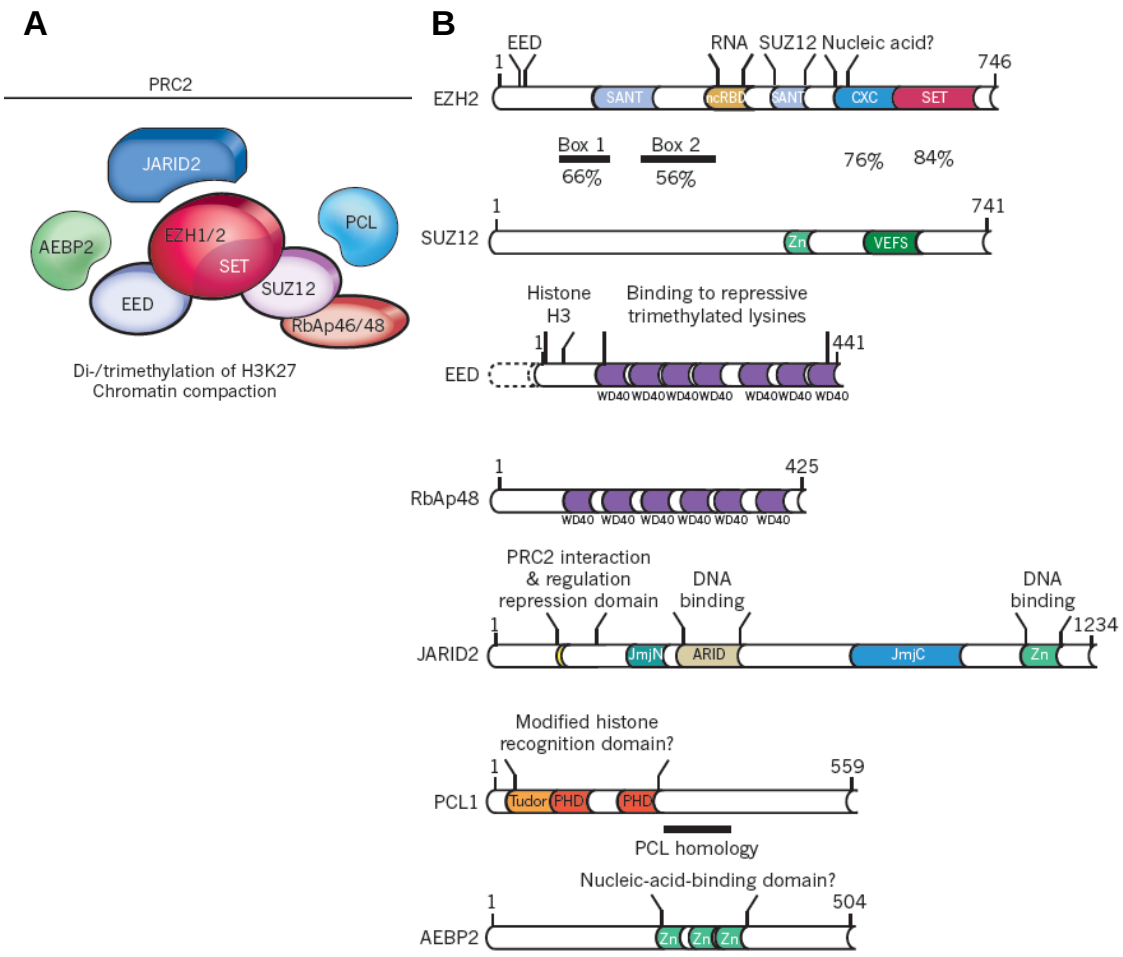


Figure 6. (Adapted from Margueron and Reinberg 2011) *PRC2 trimethylates H3K27 to induce gene repression and binds accessory proteins which modulate function [75]*.

(A) Polycomb repressive complex 2 (PRC2) contains three core components, Suz12, Eed, and methyltransferase Ezh2 and is responsible for di- and tri- methylating lysine 27 on histone H3. This mark, in addition to PRC1-mediated ubquitination of lysine 119 on H2A, are early events in the process of gene repression. Other PRC2 interacting proteins have been identified that regulate complex function including Rbbp4 (RbAp46), Rbbp7 (RbAp48), Aebp2, Jarid2, and Polycomb-like proteins. **(B)** Many diverse domains exist within PRC2 and interacting proteins. The Ezh2 SET domain is required for methyltransferase function, while the WD40 domain repeats within Eed create a three-dimensional structure that binds H3K27me3 to propagate the mark. Polycomb-like proteins contain a Tudor domain and two PHD finger domains which are predicted to bind unmodified and modified histones. Several other domains have been implicated in DNA or histone binding.

Chapter 2: The ciliogenic protein Ofd1 regulates the neuronal differentiation of embryonic stem cells

Julie Hunkapiller, Veena Singla, Allen Seol, and Jeremy F. Reiter. (2010). The ciliogenic protein Ofd1 regulates the neuronal differentiation of embryonic stem cells. *Stem Cells and Development*, May 20(5): 831-41. 2011.

ABSTRACT

Oral-Facial-Digital 1 (OFD1) Syndrome is an X-linked developmental disorder caused by mutations in the gene *Oral-Facial-Digital Syndrome 1* (*Ofd1*). OFD1 syndrome involves malformation of the face, oral cavity, and digits and may be characterized by cystic kidneys and mental retardation. Deletion or missense mutations in *Ofd1* also result in loss of primary cilia, a microtubule-based cellular projection that mediates multiple signaling pathways. *Ofd1* mutant mice display pleiotropic developmental phenotypes including neural, skeletal, and cardiac defects. To address how loss of Ofd1 and loss of primary cilia affect early differentiation decisions, we analyzed embryoid bodies (EBs) derived from *Ofd1* mutant embryonic stem (ES) cells. *Ofd1* mutant EBs do not form primary cilia and display defects in Hedgehog (Hh) and Wnt signaling. Additionally, we show that ES cells lacking Ofd1 display an increased capacity to differentiate into neurons. Nevertheless, neurons derived from *Ofd1* mutant ES cells fail to differentiate into V3 interneurons, a cell type dependent on ciliary function and Hh signaling. Thus, loss of Ofd1 affects ES cell interpretation of developmental cues and reveals that EBs model some aspects of ciliopathies, providing insights into the developmental origins of OFD1 syndrome and functions of cilia.

INTRODUCTION

Cilia participate in a broad range of developmental events and organ functions [24-27]. For example, cilia are necessary for normal development of the brain, heart, kidney, limbs, and skeleton as well as for sight, hearing and smell [24,26,141]. Furthermore, genes involved in primary cilia formation have been found to participate in multiple signaling pathways, such as those that transduce Hedgehog (Hh), Wnt, and PDGF signals [24,25,27,30].

Hh signal transduction is mediated through the primary cilium, and localization of several Hh pathway components to primary cilia is necessary for their function. Smoothened (Smo), a seven-pass transmembrane protein, moves into the primary cilium in the presence of Sonic Hedgehog (Shh) [22]. Gli proteins, the effectors of the Hh pathway, also localize to the primary cilium, and this localization is essential for formation of both activator and repressor forms [31]. Smo translocation to the cilium triggers the switch from formation of Gli repressors to Gli activators [31-34]. Gli3, for example, is processed to a truncated repressor form in the absence of Hh ligand [35,36,38,39]. This processing is inhibited by pathway activation in a cilium-dependent manner [31-33]. Gli proteins are presumed to shuttle from the cilium to the nucleus to control transcription of Hh target genes. Thus, the cilium coordinates multiple steps of the Hh pathway to regulate Hh pathway transcriptional activity.

As cilia play diverse roles in development and signaling, ciliary dysfunction manifests as human genetic syndromes known as ciliopathies, which include Meckel, Joubert, Senior-Loken, Bardet-Biedl, and OFD1 syndromes [25-27,37,43-45,53]. OFD1 syndrome is characterized by polydactyly and deformity of the oral cavity and face and is caused by mutations in the gene *Ofd1* [53,55]. *Ofd1* encodes a protein that localizes to the distal end of centrioles where it functions as a cap to regulate centriole length

[47,54]. As *OFD1* is X-linked, males lacking OFD1 do not form cilia resulting in prenatal lethality [47]. The developmental phenotypes displayed in *Ofd1* mutant mice resemble those seen in humans with OFD1. Likewise *Ofd1* mutant mice share many developmental abnormalities with other mouse mutants lacking cilia [29,46,48,49]. We recently described a mouse ES cell line that contains a gene trap insertion into the gene *Ofd1* (*Ofd1*^{Gt}) [20,54]. The *Ofd1*^{Gt} ES cell line is male and thus lacks both Ofd1 and primary cilia.

Embryonic stem (ES) cells are derived from the inner cell mass of the blastocyst and can differentiate into all cell types of the body [1,2,6]. In addition to representing a potential for stem-cell based therapies, ES cells are a tool for investigating cell fate decisions and the mechanisms of development [6,142]. Mouse ES cells are able to maintain their pluripotency in culture by activation of the JAK/STAT and BMP pathways [5]. Upon differentiation in suspension culture, ES cells form aggregates called embryoid bodies (EBs) [7,8]. EBs form three layers comprised of endoderm, mesoderm, and ectoderm and have the potential to form nearly all cell types of the embryo. [6-8].

Here, we used the *Ofd1*^{Gt} ES cell line to address the role of Ofd1 and primary cilia in ES cell differentiation. We found that *Ofd1*^{Gt} EBs have Hh signaling defects and exaggerated β-catenin-dependent pathway activation in response to Wnt3a. Furthermore, differentiated *Ofd1*^{Gt} ES cells displayed increased neural differentiation. Examination of mouse mutant embryos lacking cilia demonstrated that cilia do not limit neural differentiation *in vivo*. Nevertheless, *Ofd1*^{Gt} EBs do not form V3 interneurons similarly to mouse mutants with abrogated ciliogenesis. V3 interneurons require high levels of Hh signaling in the ventral neural tube for induction, thus indicating that EB differentiation recapitulates the role of cilia in Hh mediated neuronal patterning.

RESULTS

To understand how Ofd1 and primary cilia contribute to cell fate decisions and to establish whether ES cells may be an appropriate model system to address ciliary function, we examined whether differentiated ES cells possess primary cilia. Wild type ES cells were grown in non-adherent suspension culture in the absence of LIF to form EBs. When probed for Arl13b, a ciliary component, and γ-tubulin, a centrosomal component, wild type EBs displayed abundant cilia associated with centrosomes (Figures 7A, 8A). While approximately 15% of asynchronously dividing ES cells have cilia, quantification of EB cilia showed that 28% of EB cells possess cilia at day seven of differentiation (data not shown) [54]. Cilia were especially evident at the periphery of the EBs (Figure 8A).

Given that wild type EBs possess cilia, we addressed whether *Ofd1*^{Gt} EBs, which do not make Ofd1 protein, lack primary cilia like their ES cell counterparts (Figure 7B) [54]. Indeed, primary cilia were completely absent in the *Ofd1*^{Gt} EBs, although centrosomes were present (Figures 7A, 8B).

Previous studies have found that mutations that disrupt primary cilia abrogate vertebrate Hh signaling and can alter Wnt signaling [23,33,49]. To ascertain whether *Ofd1*^{Gt} EBs also display signaling abnormalities, we analyzed Gli processing and gene induction in the presence of either SAG, a small molecule activator of the Hh pathway, or Wnt3a [143]. Efficiency of Gli3 processing was assessed by measuring levels of truncated Gli3 compared to full-length Gli3. SAG induction in wild type EBs caused increased full-length Gli3 and reciprocally decreased truncated Gli3, indicating that pathway activation inhibits Gli3 processing to the repressor form in EBs (Figures 7C, 8C). In contrast, addition of SAG failed to change the amount of full-length or truncated Gli3 in *Ofd1*^{Gt} EBs (Figures 7C, 8C). Interestingly, *Ofd1*^{Gt} EBs displayed higher levels of

full-length Gli3 and Gli2 protein compared to wild type EBs (Figures 7C, 8C). Together, these results indicate that *Ofd1*^{Gt} EBs display defects in Gli processing and degradation.

To determine whether *Ofd1*^{Gt} EBs generate a transcriptional response to Hh pathway stimulation, we assayed downstream target genes, *Gli1* and *Patched1 (Ptch1)*, in the presence of Hh ligand. Wild type EBs exhibited a nearly threefold increase in *Gli1* and *Ptch* expression upon induction with Shh whereas *Gli1* and *Ptch* levels in the *Ofd1*^{Gt} EBs remained the same (Figure 7D). Notably, the increased full length Gli3 and Gli2 protein levels observed in *Ofd1*^{Gt} EBs did not correlate with increased activity, indicating that this stabilized protein is unable to activate the Hh transcriptional program. These data indicate that *Ofd1*^{Gt} EBs, like mouse mutants lacking cilia, cannot activate genes in response to Shh.

At least some genes essential for cilium formation are required to restrain canonical Wnt signaling in mice and mouse embryonic fibroblasts (MEFs) [23,144]. Therefore, we assessed whether *Ofd1*^{Gt} EBs display increased responsiveness to Wnt3a. We exposed wild type and *Ofd1*^{Gt} EBs to recombinant Wnt3a and assayed for transcriptional activation of Wnt target genes, *C-myc*, *CyclinD1*, *Axin2*, and *Follistatin*. Addition of Wnt3a in wild type EBs induced a response in each gene, but stimulation in *Ofd1*^{Gt} EBs with equal concentrations of Wnt3a induced a greater increase in gene expression (Figure 7E). Thus, *Ofd1*^{Gt} EBs overactivate canonical Wnt target genes in the presence of Wnt3a, similar to *Kif3a* mutant cells [23].

As Ofd1, cilia, and Hh and Wnt signaling are all critical for mammalian development, we examined how Ofd1 influences ES cell differentiation [40,145,146]. We analyzed the expression of a panel of differentiation markers in wild type and *Ofd1*^{Gt} EBs. Whereas markers of epithelium (*K18*), trophectoderm (*Eomes*), and endoderm (*Hnf4*) were expressed at similar levels in wild type and *Ofd1*^{Gt} EBs, *Ofd1*^{Gt} EBs expressed less *T-Brachyury*, an early mesodermal marker, than wild type EBs (Figure

9A). The most pronounced difference, however, was the increased expression of *Sox1*, a marker of neural precursors, in *Ofd1*^{Gt} EBs compared to wild type EBs (Figure 9A). Expression analysis of several other neural markers, including *Sox3*, *Nestin*, *Tuj1*, and *Ngn2*, also revealed a dramatic increase in *Ofd1*^{Gt} EBs, revealing a 5 to 35 fold change between *Ofd1*^{Gt} and wild type EBs (Figure 9B). These data suggest that in the absence of Ofd1, ES cells have an increased capacity to differentiate towards the neural lineage.

To test whether increased neural induction was associated with loss of Ofd1 protein or whether it was due to Ofd1-independent differences in the mutant cell line, we made use of a carboxy-terminal Myc tagged Ofd1 allele reintroduced into the endogenous locus (*Ofd1*^{Ofd1myc}) [20,54]. The *Ofd1*^{Ofd1myc} ES cells display lower levels of Ofd1 protein compared to wild type and approximately half the number of cilia (Figures 7B, 8B). Correlatively, *Ofd1*^{Ofd1myc} EBs express levels of neural markers intermediate to wild type and *Ofd1*^{Gt} EBs (Figure 9A-B). Thus, by re-expressing *Ofd1* at lower levels than wild type, we were able to partially rescue the phenotypes seen in the *Ofd1*^{Gt} EBs, suggesting that the increased neural differentiation observed in *Ofd1*^{Gt} EBs is due to loss of Ofd1 protein and primary cilia.

To distinguish whether increased neural marker expression observed in *Ofd1*^{Gt} EBs is a result of the presence of more neural cells or higher expression of neural genes within a normal number of neural cells, we analyzed wild type and *Ofd1*^{Gt} EBs by flow cytometry and immunofluorescence. We found that *Ofd1*^{Gt} EBs contained a higher percentage of neural cells as assayed by FACS for Sox1, Tuj1, and Nestin expressing cells (Figure 9C). Likewise, we observed expanded immunofluorescent staining of Nestin, Tuj1, and Ngn2 in *Ofd1*^{Gt} EBs compared to wild type EBs (Figure 9D-F). Of note, Ngn2 staining in wild type EBs at day twelve of differentiation was dramatically lower than *Ofd1*^{Gt} EBs, while at later time points, wild type EBs showed Ngn2 staining more comparable to, although still less than, *Ofd1*^{Gt} EBs (Figures 8D, 9F). This

suggests that Ngn2 may be induced at an earlier time point in *Ofd1*^{Gt} EBs than in wild type EBs, indicating that loss of Ofd1 changes the timing of neural differentiation.

There are several ways in which increased neural induction could occur in *Ofd1*^{Gt} EBs, including increased proliferation rates, decreased levels of apoptosis, or misregulation of cell fate decisions. Although, the population doubling rate of wild type and *Ofd1*^{Gt} ES cells is indistinguishable, Ofd1 may impinge upon proliferation rates of differentiated cells [54]. Consequently, we tested proliferation in wild type and *Ofd1*^{Gt} EBs by BrdU incorporation and phospho-histone H3 staining during differentiation. Both assays revealed that wild type and *Ofd1*^{Gt} EBs have similar rates of cell division throughout differentiation (Figure 10A-C). In addition, we tested levels of apoptosis by TUNEL staining, which was comparable between wild type and *Ofd1*^{Gt} EBs (data not shown). Based on these data, it seems most likely that *Ofd1*^{Gt} EBs form neurons at the expense of other cell types.

Mammalian neural induction involves the interplay of several signaling pathways, including the Bone Morphogenic Protein (BMP), Wnt, Nodal, Fibroblast Growth Factor (FGF), and Hh pathways [11,147]. In particular, inhibition of the BMP pathway is a key determinant of neural fate in many vertebrates [11,148-150]. Defects in BMP pathway signaling can lead to early and increased neural induction, and conversely, increased BMP signaling results in loss of forebrain development [151,152]. Thus, we tested BMP activity and the effects of BMP activation and inhibition on neural induction in wild type and *Ofd1*^{Gt} EBs. The BMP pathway is mediated through phosphorylation of Smad proteins 1/5/8, which then bind to Smad4, relocate to the nucleus, and initiate transcription of target genes [153]. To determine whether pathway activity is altered in *Ofd1*^{Gt} EBs, levels of phosphorylated Smad1/5 were measured at time points throughout differentiation. Phosphorylated Smad levels were indistinguishable between wild type and *Ofd1*^{Gt} EBs, suggesting that BMP activity is unaffected by loss of Ofd1 (Figure 10D).

To further test whether increased neural induction in *Ofd1*^{Gt} EBs is due to altered responses to BMPs or BMP antagonists, we added either BMP4 or Noggin, a BMP inhibitor, to differentiating EBs and assessed neural induction. In response to BMP4, both wild type and *Ofd1*^{Gt} EBs displayed decreased *Sox1* expression (Figure 10E). Conversely, addition of Noggin resulted in increased Sox1 expression in wild type and *Ofd1*^{Gt} EBs (Figure 10E).

These results were substantiated by analyzing neural induction following interruption of the BMP pathway downstream of ligand interaction. Smad1 is a mediator of the BMP pathway as it is phosphorylated upon BMP pathway activation and translocates to the nucleus with Smad4 to activate downstream BMP target genes. Wild type and *Ofd1*^{Gt} ES cells were transduced with a *Smad1* RNAi lentivirus and subsequently differentiated. Both wild type and *Ofd1*^{Gt} transduced EBs exhibited greater than 80% knockdown of *Smad1* RNA levels and increased *Sox1* and *Nestin* expression compared to control EBs (Figure 10F-G). Based on these data, BMP activity and response do not appear to depend upon Ofd1 as wild type and *Ofd1*^{Gt} EBs react to pathway activation and inhibition similarly.

Ofd1^{Gt} EBs have defects in Hh and Wnt signaling that are similar to mouse mutants lacking Kif3a, a kinesin essential for ciliary formation [23,48]. Thus, we wanted to determine if there was a general role for cilia in restriction of neurogenesis. While Kif3a has functions beyond cilia formation, many phenotypes in *Ift88*, *Kif3a*, and *Ofd1* null embryos are attributed to loss of cilia [23,46-48,144]. Therefore, we analyzed *Ift88* and *Kif3a* mutant embryos to determine whether neural induction is increased. Neural marker expression was assessed in E9.5 *Kif3a* and *Ift88* mutants and compared to their wild type littermates. Overall there was a 30-50% decrease in *Ngn2*, *Tuj1*, and *Nestin* expression in the *Kif3a* and *Ift88* mutants (Figure 11A). In situs of E7.5 and E8.5 embryos probed for brain markers *Sox2*, *Krox20*, and *Engrailed* indicated that there was

no appreciable difference in neural specification between wild type and mutant embryos (Figure 12A-C). To examine neural differentiation in greater detail, E9.5 wild type and *Ift88* mutant embryo neural tubes were examined for Nestin and Tuj1 (Figure 11B). Stage-matched mutant embryonic neural tubes were comparable to wild type, indicating that neural induction is unperturbed in mice lacking cilia. Thus, it appears that either cilia do not serve as a constraint of neurogenesis *in vivo* as they do in EBs or that increased neural differentiation is specific to loss of Ofd1.

Loss of cilia causes decreased processing of Gli3 into Gli3 repressor in both *Ift88* and *Kif3a* mutant embryos, consistent with what is seen in the $Ofd1^{Gt}$ EBs [33,154]. To determine whether reduced levels of Gli3 repressor affect neural induction, we examined neural markers in $Gli3^{xt/xt}$ mutant embryos. Similar to the *Ift88* and *Kif3a* mutants, a 30-50% decrease in neural marker expression was observed in $Gli3^{xt/xt}$ mutants as compared to their wild type littermates at E9.5 (Figure 11A). These results exclude Gli3 from mediating Ofd1-dependent effects on neural differentiation, and further indicate that the stabilized full-length Gli3 seen in the $Ofd1^{Gt}$ EBs does not function in neurogenesis.

To further dissect the differentiation potential of $Ofd1^{Gt}$ EBs, we tested whether $Ofd1^{Gt}$ EBs could form ventral neural subtypes. The ventral neural tube is patterned by Shh produced in the notochord and floor plate [155]. Mice lacking primary cilia have dorsalized neural tubes, including a loss of the most ventral neurons, V3 interneurons, and a decrease in motor neuron formation [32,37]. To determine whether this same phenotype occurs *in vitro*, wild type and $Ofd1^{Gt}$ EBs were assayed for Nkx2.2, a marker of V3 interneurons. Whereas wild type EBs showed widespread Nkx2.2 expression, $Ofd1^{Gt}$ EBs had very little or low Nkx2.2 expression (Figure 11C). In contrast, $Ofd1^{Gt}$ EBs form motor neurons at levels similar to wild type EBs as assessed by Islet1/2 levels (Figure 12D). Thus, neurons that require the highest levels of Hh signaling, V3

interneurons, do not form in the absence of cilia, whereas motor neurons, which require intermediate levels, are formed in *Ofd1*^{Gt} EBs.

One reason for the disparity between the roles that cilia play in neural induction in EBs and embryos may be differences in the prevalence of cilia in EBs and mouse embryos during early development. ES cells can be ciliated and become more highly ciliated upon differentiation, whereas blastocysts, from which ES cells are derived, have not been reported to possess cilia. The first evidence of ciliation in the mouse embryo is in the embryonic node at E7.5 [156]. As differences in the prevalence of cilia could affect how ES cells and inner cell mass cells respond to signaling pathways before and during neural induction, we stained wild type blastocysts to see if cilia were present. Most blastocysts displayed acetylated tubulin staining of mid-bodies and mitotic spindles (Figure 11D). Moreover, we observed several blastocysts that possessed long, acetylated tubulin-containing projections, but these projections did not extend from the basal body as indicated by co-staining with Rootletin, a marker of the ciliary rootlet (Figure 11D). As primary cilia derive from basal bodies, our interpretation is that these acetylated tubulin-positive projections are not cilia, but a distinct type of microtubular cellular structure.

Finally, to assess the effects of human disease-associated OFD1 mutations on neurogenesis, we examined whether four ES cell lines containing distinct OFD1-associated missense mutations in highly conserved residues or domains display defects in neural induction [54]. While *Ofd1* missense lines, G139S and KDD359-361FSY, retain some Ofd1 protein that localizes correctly to the centrioles, all missense mutant lines show a complete or partial loss of cilia [54]. Upon differentiation, all *Ofd1* missense mutant EBs displayed increased expression levels of neural markers *Sox1*, *Nestin*, *Sox3*, and *Tuj1* compared to wild type EBs (Figure 11E). ES cells with mutations G139S and KDD359-361FSY are ciliated at approximately 20% and 35% of wild type levels,

while mutations S75F and A80T prohibit Ofd1 centriolar localization and cilia formation [54]. Despite differences in ciliation frequency, neural induction was relatively similar amongst all four missense mutant lines (Figure 11E). Centrioles in the *Ofd1* missense mutant cells are structurally abnormal and lack distal appendages [54]. Thus, it is possible that the centriolar role of Ofd1 may contribute independently to its ciliogenic functions and its ability to restrain neural induction.

DISCUSSION

Primary cilia are essential for the development of diverse tissues and organs. We have found that Ofd1 is essential for EB ciliogenesis, restrains EB neurogenesis, and is essential for V3 interneuron differentiation. These phenotypes may be attributable to the demonstrated misregulation of Hh and Wnt signaling in *Ofd1*^{Gt} EBs. However, mouse embryos lacking Ift88 or Kif3a, other proteins essential for ciliogenesis, do not show increased neural induction, raising the possibility that the neural induction defect is due to altered centriolar structure.

To help elucidate the signaling mechanisms causing increased neural induction, we investigated the involvement of the BMP, Hh, and Wnt pathways, known regulators of neural induction and differentiation. Inhibition of the BMP pathway has been shown to be important for neural induction in diverse vertebrates, and in some species, BMP inhibition is sufficient to induce neural tissue [11,148,151,152,157]. Mammals require BMP signaling for proper neural induction, but also require inputs from additional pathways. When induced with BMP agonists or antagonists, wild type and *Ofd1*^{Gt} EBs displayed comparable changes in neural induction, suggesting that the cilium is not essential for interpretation of BMP signals. Likewise, levels of phosphorylated Smads, the downstream mediators of BMP signaling, were similar in wild type and *Ofd1*^{Gt} EBs,

indicating that BMP pathway activity is not dependent on Ofd1 or the primary cilium. Furthermore, disruption of the intracellular BMP signal transduction pathway using *Smad* RNAi resulted in comparable increases in neural marker expression. These experiments suggest that loss of cilia has no effect on BMP signaling, and misregulation of other pathways likely results in increased neurogenesis in *Ofd1*^{Gt} EBs.

Both Hh and Wnt signaling pathways are important for neural differentiation and specification [11,94,155]. Previous studies have shown that cells and mouse embryos lacking primary cilia have defective Hh signaling and can display overactive Wnt responsiveness [23]. Consistent with this data, we found that *Ofd1*^{Gt} EBs also have an overactive canonical response to Wnt3a. Additionally, we show that *Ofd1*^{Gt} EBs cannot activate downstream Hh target genes upon treatment with Shh and show abrogated processing of Gli3 to the truncated repressor form. Thus, *Ofd1*^{Gt} EBs recapitulate the known biochemical and transcriptional Hh signaling defects displayed by mouse embryos lacking cilia.

Despite similarities in signaling profiles between *Ofd1*^{Gt} EBs and embryos lacking cilia, we observed increased neural induction in *Ofd1*^{Gt} EBs, but not in *Ift88* or *Kif3a* mutants. One possible reason for the divergence in neural induction observed in embryos without cilia and unciliated EBs may be differences in the temporal regulation of ciliogenesis. Approximately 15% of undifferentiated ES cells are ciliated, whereas blastocysts do not appear to be. Early discrepancies in ciliation frequency could influence how signaling is mediated and consequently how cell fates are determined. Another fundamental difference between EB differentiation and embryo development is the spatial disorganization of EB tissues. During development, embryonic architecture limits the exposure of prospective neural tissue to other tissues, signals, and cell-cell contacts. These barriers are less present in the chaotic EB environment.

Variation of growth conditions could also account for neurogenesis differences between *Ofd1*^{Gt} EBs and *Ift88* and *Kif3a* mutant embryos. EBs are cultured *in vitro* in the presence of signaling factors than normally regulate embryo development. Nonetheless, this formula of factors may be quite distinct than the milieu found in the developing embryo.

Although Hh and Wnt signaling are abnormal in the *Ofd1*^{Gt} EBs, the effects of these two pathways on neurogenesis either alone or cooperatively is unclear. Wnt signaling is important for maintenance of neural precursors and specification of the dorsal spinal cord, whereas Shh signaling from the notochord induces neural proliferation and specifies subclasses of ventral interneurons [37,39,94,140,151,157-160]. There are, however, instances in which the Hh and Wnt pathways coordinately regulate neural development. For example, Gli3 repressor inhibits canonical Wnt-signaling by binding β-catenin during spinal cord patterning [140]. Additionally, GSK3, a component of both the Hh and Wnt pathways, is required for normal proliferation of neural progenitors [94]. Future studies are needed to dissect the molecular interactions of these two pathways during ES cell differentiation.

Another area of future study is to explore whether defects in primary cilia or defective centrioles may affect ES cell differentiation in distinct ways. *Ofd1* null mice show phenotypes resembling those of other mutant embryos lacking cilia, such as *Ift88* mutants, suggesting that Ofd1 functions primarily in cilium assembly. Indeed, *Ofd1* missense mutant cells lacking cilia, either partially or completely, all displayed increased neural induction. All *Ofd1* missense mutant lines also have abnormal centriolar structure, even in cases in which the mutant Ofd1 localizes properly to the distal centriole [54]. Thus, it is difficult to discern whether increased neural differentiation is a result of loss of cilia or defective centrioles. To distinguish these two possibilities, neural

differentiation will need to be assayed in ES cells that lack cilia but have normal centriolar structure.

The neural tube is patterned by signaling molecule gradients initiated in the notochord and roof plate. V3 interneurons are the most ventral subtype, requiring the highest levels of Hh signaling for specification. Indeed, mouse mutants without cilia lack V3 interneurons and form few motor neurons due to the absence of Hh signaling. Consistent with this finding and the loss of Hh responsiveness in *Ofd1*^{Gt} EBs, we observed few, if any, V3 interneurons in the *Ofd1*^{Gt} EBs. In contrast, *Ofd1*^{Gt} EBs do form motor neurons, which are induced more dorsally in the neural tube than are V3 interneurons. Given that inhibition of Gli3 repressor participates in motor neuron development, it is likely that formation of motor neurons in *Ofd1*^{Gt} EBs is caused by decreased Gli3 processing [39,52,161,162]. Hence, the defects in *Ofd1*^{Gt} EB neural specification recapitulate defects in neural tube patterning observed in mouse mutants with dysfunctional cilia.

This work demonstrates that ES cell differentiation is a useful way to study the developmental mechanisms underlying ciliopathies. Although we differentiated ES cells spontaneously in suspension culture, other protocols that direct differentiation to specific cell types such as pancreatic cells, cardiomyocytes, and motor neurons may be utilized to study the role of primary cilia in the development of specific cell types. For example, directed differentiation of *Ofd1*^{Gt} ES cells to cardiomyocytes or neural stem cells may allow one to study how primary cilia affect stem cell maintenance or the ability to respond to stress or damage [141]. This system also potentially alleviates the need to derive and immortalize cell lines, which may facilitate correlating cell biological and embryological findings. Taken together, these results examining the role of Ofd1 and primary cilia in neural differentiation demonstrate the utility of the ES cell system to study the intricate role of cilia in developmental signaling and patterning.

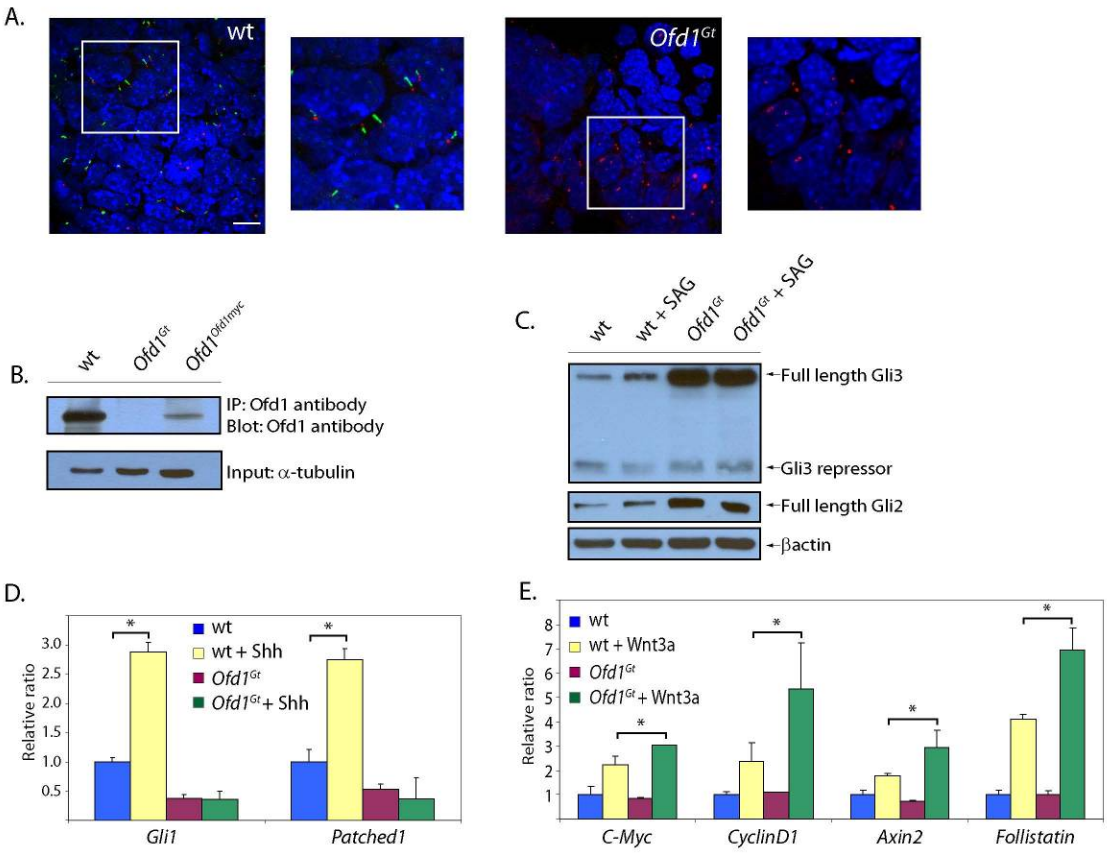


Figure 7. *Ofd1^{Gt} EBs lack cilia and display altered Hh and Wnt signaling.*

(A) Wild type and *Ofd1^{Gt}* EBs stained for primary cilia (Arl13b, green), basal bodies (γ-tubulin, red), and nuclei (DAPI, blue). Scale bar 10 μm. **(B)** Immunoblot of Ofd1 from wild type, *Ofd1^{Gt}*, *Ofd1^{Ofd1myc}* EBs after six days of differentiation. **(C)** Immunoblot of wild type and *Ofd1^{Gt}* EBs assayed for Gli3, Gli2, and β-actin following SAG or vehicle treatment for 8 hrs at seven days of differentiation. **(D)** Expression levels as determined by qRT-PCR of Hh target genes *Gli1* and *Patched1* in wild type and *Ofd1^{Gt}* EBs after addition of Shh for 18 hrs at eight days of differentiation. **(E)** Expression levels as determined by qRT-PCR of Wnt target genes *C-myc*, *CyclinD1*, *Axin2*, and *Follistatin* in wild type and *Ofd1^{Gt}* EBs grown for 2-4 days and treated with Wnt3a for 4 hrs. Asterisk indicates statistical significance of $p \leq 0.05$.

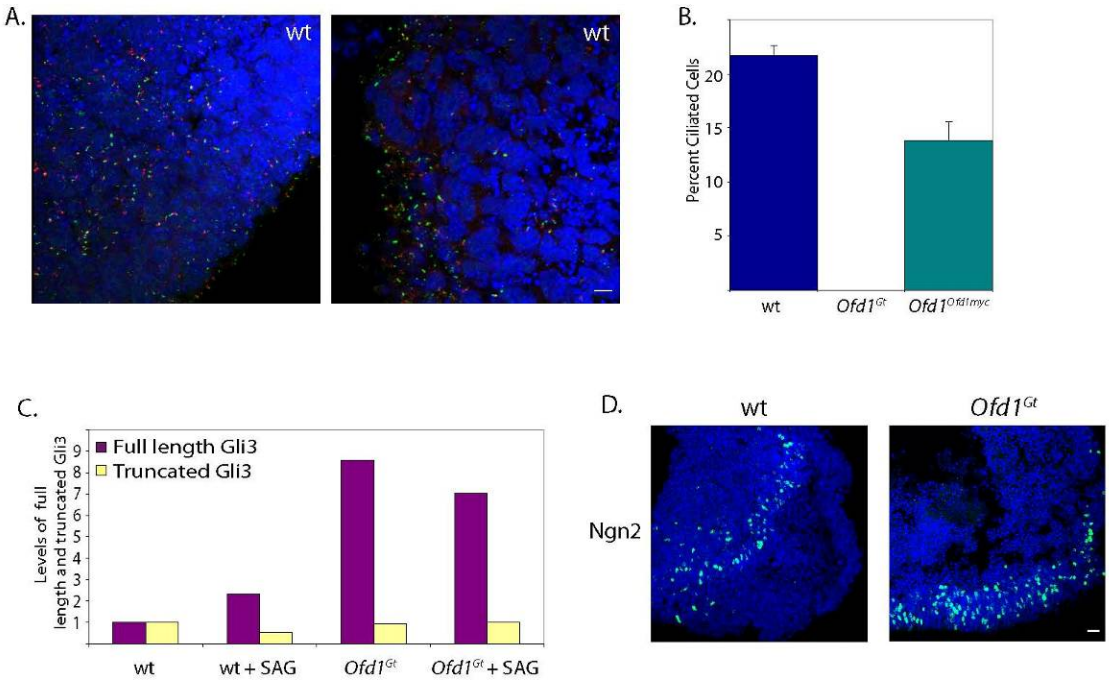


Figure 8. *Ciliation in wild type, Ofd1^{Gt}, Ofd1^{Ofd1myc} EBs.*

(A) Wild type EBs stained for Arl13b (green) marking cilia and γ-tubulin (red) marking basal bodies. Scale bar 10 μm. **(B)** Percent ciliated cells in synchronized wild type, *Ofd1^{Gt}*, *Ofd1^{Ofd1myc}* ES cells. **(C)** Quantification of full-length and truncated Gli3 by immunoblot in wild type and *Ofd1^{Gt}* control and SAG treated EBs. **(D)** Sections of wild type and *Ofd1^{Gt}* EBs stained with Ngn2 (green) after fifteen days of differentiation. Scale bar 20 μm.

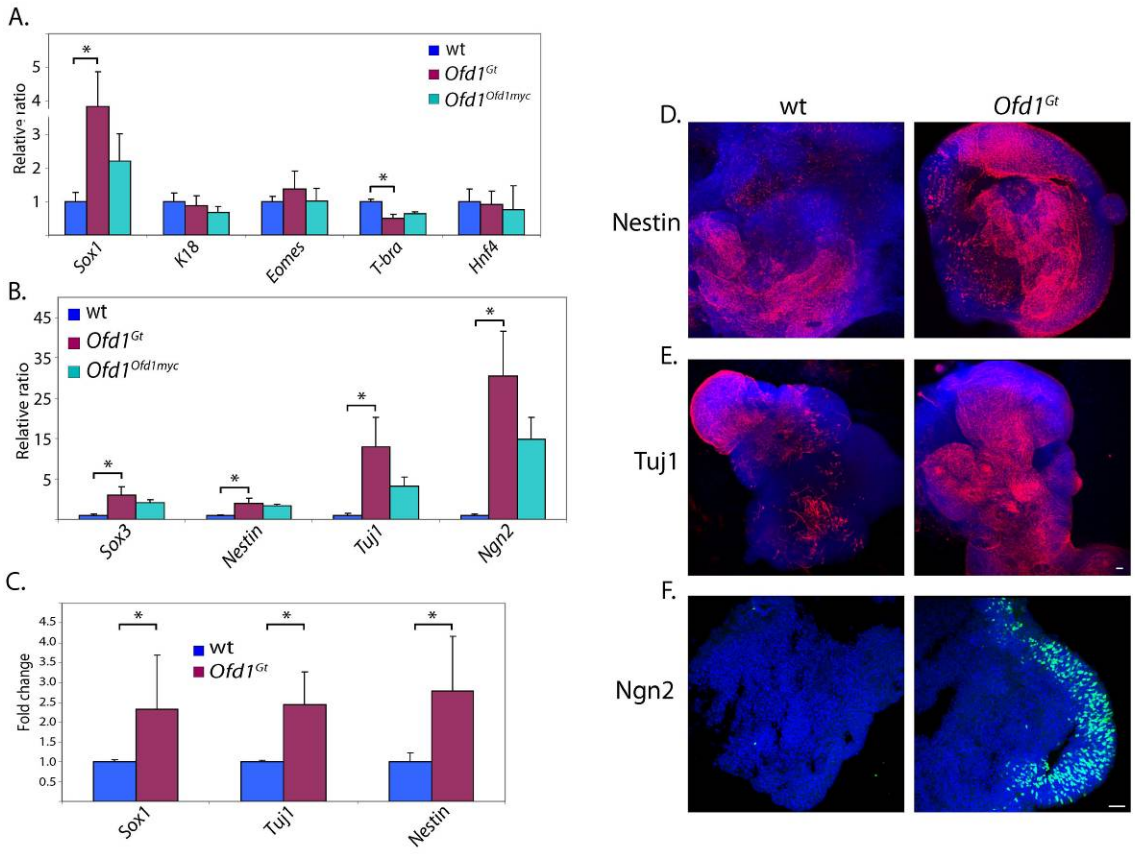


Figure 9. *Ofd1^{Gt} EBs exhibit increased neural differentiation.*

(A) Expression levels of tissue type markers *Sox1* (neuroectoderm), *Keratin 18* (epithelium), *Eomes* (trophectoderm), *T-brachyury* (mesoderm), and *Hnf4* (endoderm) in wild type, *Ofd1^{Gt}*, and *Ofd1^{Ofd1myc}* EBs. **(B)** Expression levels of neural markers *Sox3*, *Nestin*, *Tuj1*, and *Ngn2* in wild type, *Ofd1^{Gt}*, and *Ofd1^{Ofd1myc}* EBs. **(C)** Quantification of neural cells in wild type and *Ofd1^{Gt}* EBs by FACS analysis of Sox1, Tuj1, and Nestin stained cells. Asterisk indicates statistical significance of $p \leq 0.01$. **(D-E)** Wholemount wild type and *Ofd1^{Gt}* EBs stained for Nestin and Tuj1. Scale bar 20 μm. **(F)** Wild type and *Ofd1^{Gt}* EB sections stained for Ngn2 at day twelve of differentiation. Scale bar 20 μm.

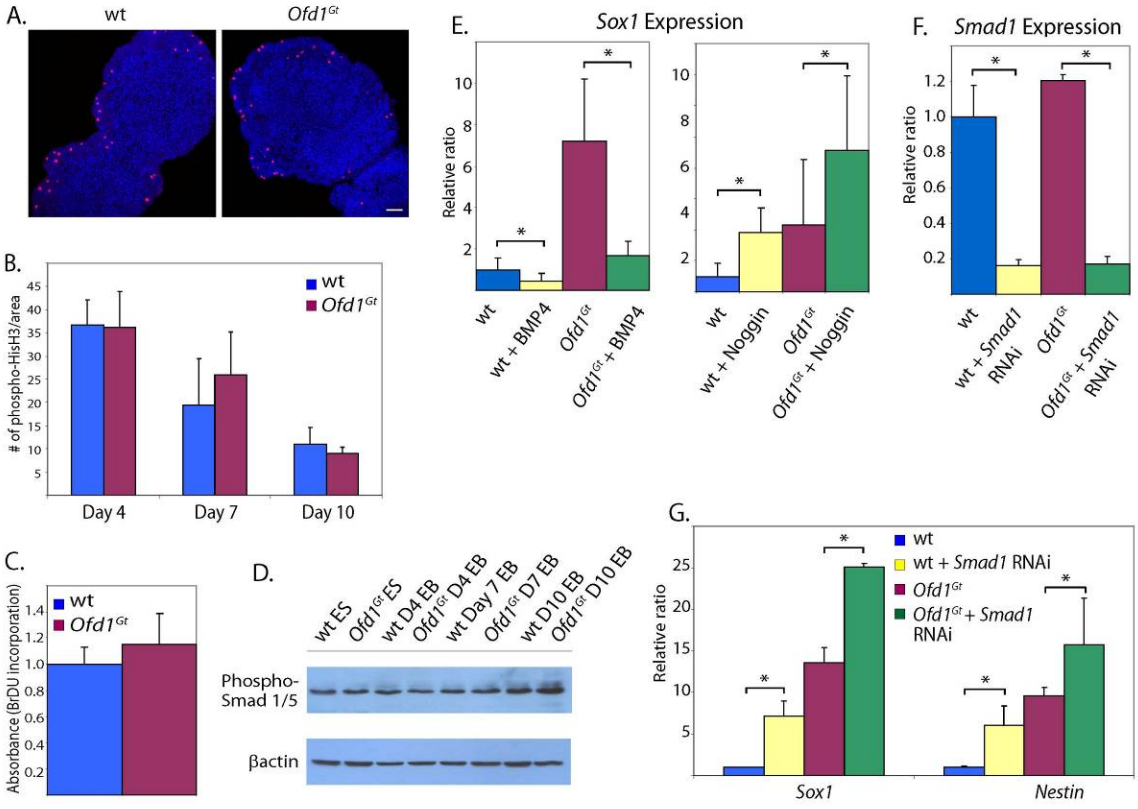


Figure 10. *Proliferation rates and BMP signaling are unaltered in Ofd1^{Gt} EBs.*

(A) Phospho-histone H3 staining in wild type and *Ofd1^{Gt}* EB sections differentiated for seven days. Scale bar 20 μm. **(B)** Quantification of phospho-histone H3 stained cells per given area of wild type and *Ofd1^{Gt}* EBs following 4, 7, and 10 days of differentiation. **(C)** BrdU incorporation as measured by colorimetric absorbance confirming similar proliferation in wild type and *Ofd1^{Gt}* EBs grown for 7 days. **(D)** Immunoblot of wild type and *Ofd1^{Gt}* cells during differentiation probed for phospho-Smad 1/5 and β-actin. **(E)** Neural specification as measured by *Sox1* expression levels in wild type and *Ofd1^{Gt}* EBs treated with BMP4 or Noggin, a BMP antagonist. **(F)** Assessment of *Smad1* knockdown in wild type and *Ofd1^{Gt}* EBs containing *Smad1* RNAi lentivirus. **(G)** Expression of *Sox1* and *Nestin* neural markers in wild type and *Ofd1^{Gt}* EBs treated with *Smad1* RNAi. Asterisk indicates statistical significance of $p \leq 0.05$.

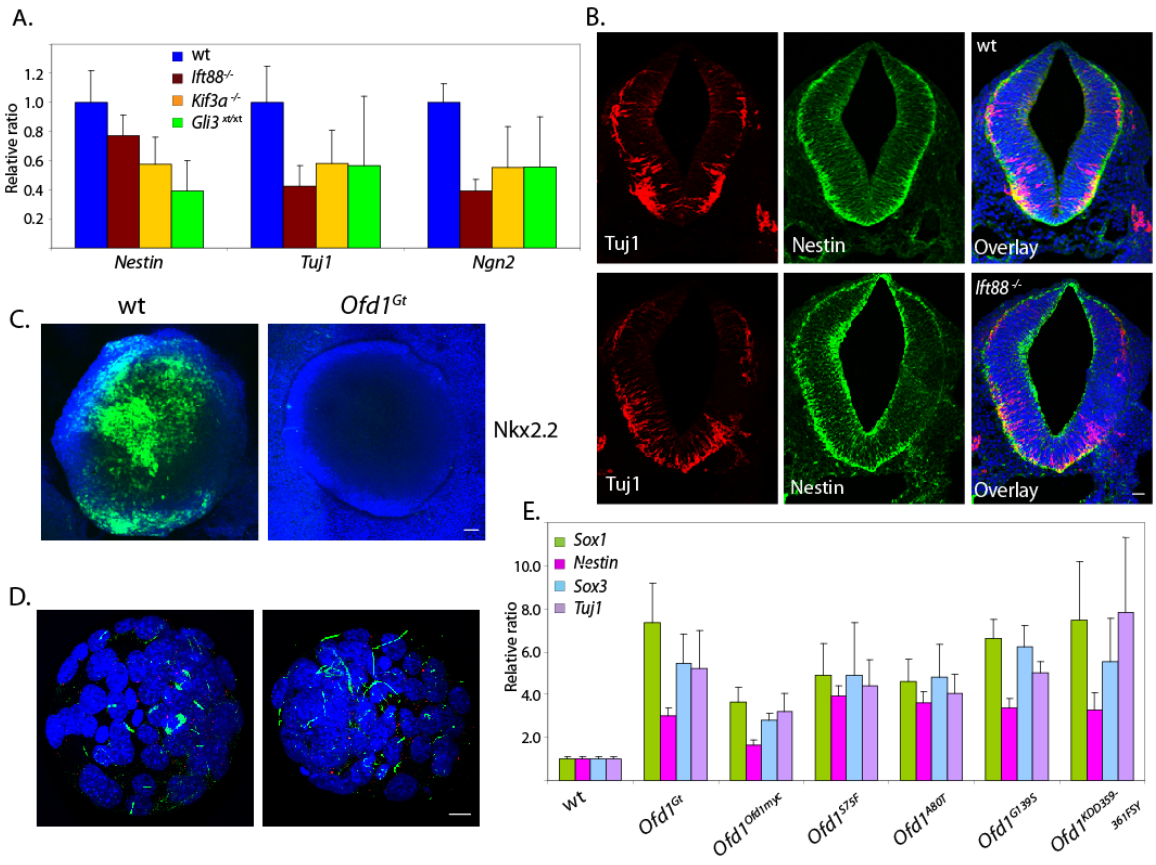


Figure 11. *Neural induction in mouse embryos lacking cilia and Ofd1 mutant missense EBs.*

(A) Expression of neural markers *Nestin*, *Tuj1*, and *Ngn2* in E9.5 wild type, *Ift88*, *Kif3a*, and *Gli3* mutant embryos. **(B)** Sections of E9.5 wild type and *Ift88* mutant embryo neural tubes stained for post-mitotic neurons (Tuj1, red), neural precursors (Nestin, green), and nuclei (DAPI, blue). Scale bar 20 μm. **(C)** Wholemount wild type and $Ofd1^{Gt}$ EBs stained for Nkx2.2. Scale bar 20 μm. **(D)** Wild type blastocysts stained for cilia (acetylated tubulin, green), basal bodies (Rootletin, red), and nuclei (DAPI, blue). Scale bar 10 μm. **(E)** qRT-PCR quantification of *Sox1*, *Nestin*, *Sox3*, and *Tuj1* expression in *Ofd1* missense mutant EBs. All expression changes between Ofd1 missense mutant and $Ofd1^{Ofd1myc}$ EBs are significant ($p \leq 0.05$).

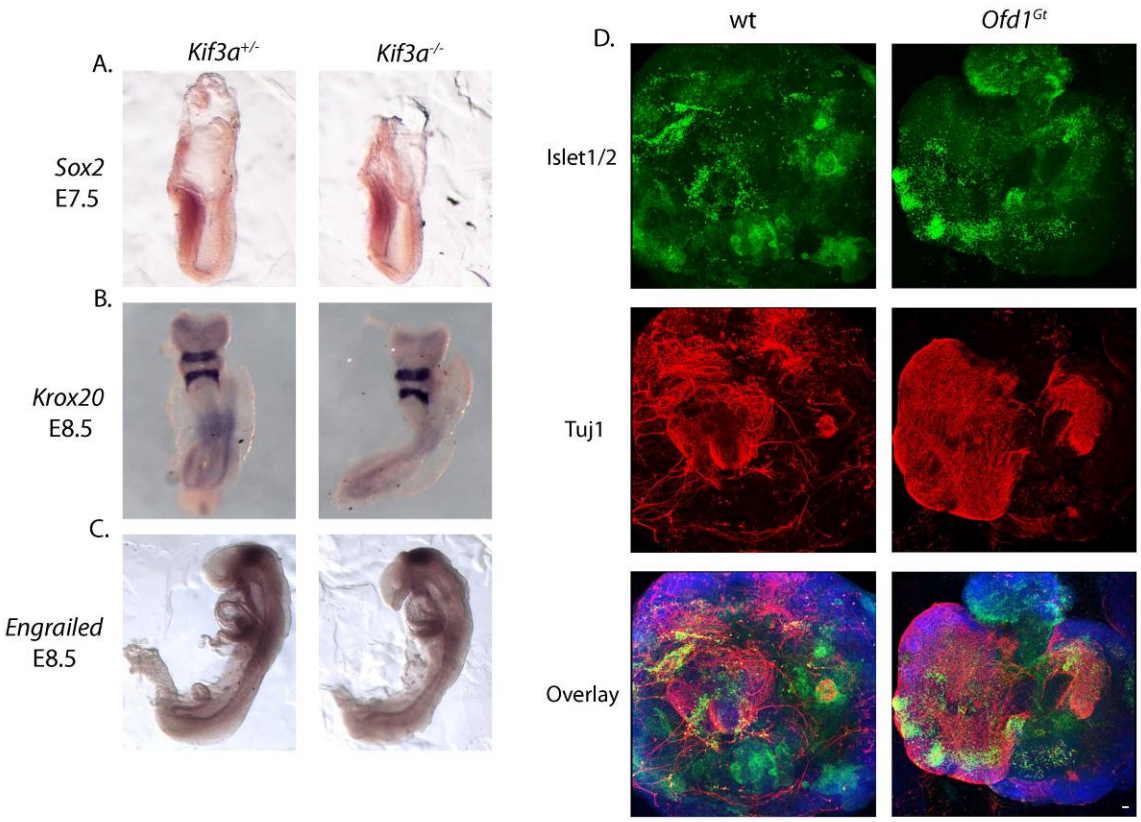


Figure 12. *Neural specification and patterning in wild type and Ofd1^{Gt} EBs and Kif3a mutant embryos.*

(A-C) Wholemount in situs of wild type and *Kif3a* mutant embryos at E7.5 or E8.5 of development probed for ectoderm (*Sox2*), hindbrain (*Krox20*), and midbrain (*Engrailed*) markers. **(D)** Wild type and *Ofd1^{Gt}* EBs stained for motor neurons (Islet1/2, green), post-mitotic neurons (Tuj1, red), and nuclei (DAPI, blue). Scale bar 20 μm.

METHODS

Tissue Culture

Embryonic stem cells and embryoid bodies were grown as described previously [20,54]. For Hh pathway activation, wild type and *Ofd1*^{Gt} (RRF427, Bay Genomics) EBs were cultured for seven days and incubated with recombinant Sonic Hh (1μM, R&D Systems) or SAG (0.1μM, Axxora) for 8, 18, or 24 hours. Similarly, Wnt pathway activation was performed by incubating embryoid bodies for 2-4 days and adding recombinant Wnt3a (0.1μg/ml, R&D Systems) for four hrs. BMP4 (0.1μM, R&D Systems) and Noggin (1μM, jCBS) were added to EBs for 48 hrs after being cultured for 3-6 days.

Lentiviral infection

ES cells were trypsinized and incubated for one hour with concentrated lentivirus for either pSico-Smad1-puro or pSico-puro empty vector (gifts of Drs. Rik Derynck and Michael McManus). The cells were plated overnight and placed under selection the following morning using 2uM puromycin. Resistant cells were selected for five days and knockdown was assayed. Two different Smad RNAi lentiviruses were used and gave similar results.

Immunofluorescence:

For wholemount embryoid bodies

EBs were grown for 10-12 days in low suspension culture and subsequently plated onto chamber slides coated with poly-lysine and Matrigel (BD) for 48 hrs. EBs were washed and fixed with 4% PFA for 20 minutes and washed three times with PBT. Subsequently, the EBs were exposed to primary block (2% BSA and 1% serum in PBT) for one hour. Primary antibody was added overnight in primary block at 4°. The next day EBs were

washed and secondary block (2% BSA and 10% serum in PBT) was added for 1 hr. The EBs were then incubated with secondary antibodies for 1-2 hrs followed by washes with PBT. Nuclei were stained with 4',6-Diamidino-2-phenylindole (DAPI) in primary block for 20 minutes and washed again with PBT three times. Slides were mounted and left overnight to dry. All Images were taken on a Nikon C1si Spectral confocal microscope.

For embryo and embryoid body sections

Embryos and EBs were fixed for 1 hr in 4% PFA, washed, and imbedded in OCT. The blocks were sectioned on a Microm HM 550 (Thermo-Fisher) at twelve µm thickness. The slides were washed three times with PBS and stained using the protocol above, except for the cilia staining protocol, which included a 2 min 100% methanol fixation.

For blastocysts

Blastocysts were flushed from the uterus at E3.5 and immunofluorescently analyzed similarly to the wholemount EBs.

Antibodies

The following antibodies were used for immunofluorescence and immunoblotting at the following dilutions : mouse anti-γ-tubulin (1:500, Abcam GTU488); mouse anti-Tuj1 (1:500, Covance MMS435); rabbit anti-Tuj1 (1:2000, Covance PRB435) ; mouse anti-Nestin (1:300; BD Pharmingen 556309); mouse anti-Islet1/2 (1:50, DSHB); mouse anti-Nkx2.2 (1:20, DSHB); rabbit anti-β-actin (1:5000, Abcam); rabbit anti-phospho-Smad 1/5 (1:1000, Cell Signaling); rabbit anti-phospho-Smad 1/5/8 (1:1000, Chemicon); rabbit anti-Rootletin (1:20,000, gift from Tiansen Li); mouse anti-acetylated tubulin (1:500, Sigma); rabbit anti-Ofd1 (1:5000); mouse anti-α-tubulin (1:5000). The mouse anti-Neurogenin2 antibody (1:100) was a gift of Dr. D. J. Anderson. Rabbit anti-Arl13b (1:500) was a gift of Dr. Tamara Caspary. Mouse anti-Gli3 and guinea pig anti-Gli2 were

gifts of Drs. Suzie Scales (Genentech) and Jonathan Eggenschwiler, respectively, and were used at (1:4000).

Quantitative RT-PCR

RNA was extracted from EBs at indicated time points using Qiagen RNeasy Plus Mini kit. RNA was then subjected to first-strand cDNA synthesis using iScript (Biorad or Fermentas). Expression levels were then analyzed in triplicate using a 7300 Real-time PCR machine (Applied Biosystems) and normalized to β-actin. The following primers were used:

β-actin F: CACAGCTTCTTTGCAGCTCCTT

β-actin R: CGTCATCCATGGCGAACTG

Nestin F: TTAAGGCCAGAACCCCCAC

Nestin R: CTCTGCATTTTTAGGATAGGGAGC

Hnf4 F: CAGACGTCCTCCTTTTCTTGTGATA

Hnf4 R: TGTTTGGTGTGAAGGTCATGATTA

T-brachyury F: CTGGGAGCTCAGTTCTTTCGA

T-brachyury R: GAGGACGTGGCAGCTGAGA

Keratin 18 F: CGCTTGCTGGAGGATGGA

Keratin 18 R: CTTCTGCACAGTTTGCATGGA

Sox1 F: TGAAGGAACACCCGGATTACA

Sox1 R: GCCAGCGAGTACTTGTCCTTCT

Gli1 F: CTTCACCCTGCCATGAAACT

Gli1 R: TCCAGCTGAGTGTTGTCCAG

Patched F: TGATTGTGGAAGCCACAGAAAA

Patched R: TGTCTGGAGTCCGGATGGA

CyclinD1 F: CCAAGTTCCCTAGCAAGCTG

CyclinD1 R: CTTTCATGTCACAGGGCAGA

C-myc F: CAACGTCTTGGAACGTCAGA

C-myc R: TCGTCTGCTTGAATGGACAG

Follistatin F: ACGTGTGAGAACGTGGACTG

Follistatin R: CATTCGTTGCGGTAGGTTTT

Axin2 F: CTCCCCACCTTGAATGAAGA

Axin2 R: ACTGGGTCGCTTCTCTTGAA

Ngn2 F: GCTGTGGGAATTTCACCTGT

Ngn2 R: AAATTTCCACGCTTGCATTC

Tuj1 F: TGAGGCCTCCTCTCACAAGT

Tuj1 R: CGCACGACATCTAGGACTGA

Sox3 F1: CGTAACTGTCGGGGTTTTGT

Sox3 F2: CACAACTCCGAGATCAGCAA

Sox3 R1: AACCTAGGAATCCGGGAAGA

Sox3 R2: GTCCTTCTTGAGCAGCGTCT

Western immunoblot

Cells were lysed in RIPA with protease and phosphatase inhibitor, and protein concentration was measured. For immunoprecipitate, 0.5μg of rabbit anti-Ofd1 antibody was added to 400μg protein lysate overnight and bound to rProtein G agarose (Invitrogen). Samples were run on a 9% polyacrylamide gel and transferred to a PVDF membrane. The membrane was blocked in 5% milk and incubated with primary antibody overnight at 4°. The next day the membrane was washed and probed with secondary for 1hr. Following three washes, the membrane was incubated with chemiluminescent substrate for 1min and exposed for 1-30 min.

FACS

Embryoid bodies were grown for 10-18 days and dissociated using collagenase/dispase (1mg/ml) for 10 min followed by 0.25% trypsin with Dnase I (1mg/ml) for 10 min. Cells were resuspended in EB media, washed with PBS, and fixed in 2% PFA for 15 min. permeabilized with 1% saponin for 15 min. Cells were resuspended in primary block (3% goat serum in PBS) for 15 min and incubated with primary for 2 hrs, followed by washing and incubation with secondary antibody for 1 hr. Samples were washed several more times and transferred to flow cytometry tubes. Analysis was performed using a FACSCalibur Flow Cytometer (BD) and FlowJo software (Treestar).

In situs

Protocol used for in situs has been described previously [163].

Cell proliferation by BrdU incorporation

ES cells were plated in 24-well low-suspension plates, grown for seven days, and analyzed using the Cell Proliferation ELISA BrdU kit (Roche). Absorbance measured in triplicate.

Chapter 3. Polycomb-like 3 promotes Polycomb repressive complex 2 binding to CpG islands and embryonic stem cell self-renewal

Julie Hunkapiller, Yin Shen, Aaron Diaz, Gerard Cagney, David McCleary, Miguel Ramalho-Santos, Nevan Krogan, Bing Ren, Jun S. Song, and Jeremy F. Reiter (2011). Polycomb-like 3 is a component of PRC2 that promotes embryonic stem cell self-renewal (Plos Genetics).

ABSTRACT

Polycomb repressive complex 2 (PRC2) trimethylates lysine 27 of histone H3 (H3K27me3) to regulate gene expression during diverse biological transitions in development, embryonic stem cell (ESC) differentiation, and cancer. Here, we show that Polycomb-like 3 (Pcl3) is a component of PRC2 that promotes ESC self-renewal. Using mass spectrometry, we identified Pcl3 as a Suz12 binding partner and confirmed Pcl3 interactions with core PRC2 components by co-immunoprecipitation. Knockdown of *Pcl3* in ESCs increases spontaneous differentiation, yet does not affect early differentiation decisions as assessed in teratomas and embryoid bodies, indicating that Pcl3 has a specific role in regulating ESC self-renewal. Consistent with Pcl3 promoting PRC2 function, decreasing Pcl3 levels reduces H3K27me3 levels while overexpressing Pcl3 increases H3K27me3 levels. Furthermore, chromatin immunoprecipitation and sequencing (ChIP-seq) reveals that Pcl3 co-localizes with PRC2 core component, Suz12, and depletion of Pcl3 decreases Suz12 binding at over 60% of PRC2 targets. Mutation of conserved residues within the Pcl3 Tudor domain, a domain implicated in recognizing methylated histones, compromises H3K27me3 formation, suggesting that the Tudor domain of Pcl3 is essential for function. We also show that Pcl3 and its

paralog, Pcl2, exist in different PRC2 complexes but bind many of the same PRC2 targets, particularly CpG islands regulated by Pcl3. Thus, Pcl3 is a component of PRC2 critical for ESC self-renewal, histone methylation, and recruitment of PRC2 to a subset of its genomic sites.

INTRODUCTION

The developmental plasticity of early embryos and embryonic stem cells (ESCs) requires the repression of cell-type specific genes. Two multiprotein complexes that participate in gene repression are Polycomb repressive complex 1 (PRC1) and Polycomb repressive complex 2 (PRC2) [75,164]. Core components of PRC2 include Suz12, Eed, and Ezh2, a methyltransferase that participates in di- and tri-methylation of lysine 27 on histone H3 (H3K27me2/3) [75,93,165-168]. Trimethylation of H3K27 can modulate the function of PRC1, which mono-ubiquitinates histone H2A on lysine 119 (H2AK119Ub) [93,165]. Both H3K27me3 and H2AK119Ub are early histone modifications involved in gene repression [168].

Whereas H3K27me3 is associated with repressed genes, H3K4me3 marks active genes. ESCs and a number of adult stem cells, however, contain a unique chromatin signature, termed bivalency, that is comprised of both H3K27me3 and H3K4me3 marks [70,169-175]. Many bivalent domains are at CpG islands, domains of DNA with elevated GC content that display low levels of DNA methylation. CpG islands are commonly found at vertebrate promoters and are associated with 70% of annotated genes including most housekeeping genes and many developmentally regulated genes [176-178]. CpG-rich domains commonly display H3K4me3, but GC-rich sequences also promote H3K27me3, creating opposing marks within the same domain [88,179]. By occupying CpG islands and marking them as bivalent domains in ESCs, PRC2 may

keep the associated genes repressed but poised for rapid activation upon differentiation [70]. How PRC2 is recruited to CpG islands is not known.

Disrupting core components of PRC2 causes a global reduction of H3K27me3 and misexpression of repressed genes, particularly bivalent genes [70,79-81,180]. This dysregulation of gene expression perturbs ESC maintenance and differentiation, and results in embryonic lethality in mice [3,77,82,180-183]. Furthermore, expression of PRC1 and PRC2 components is misregulated in diverse cancers, suggesting that PRC2-dependent gene regulation protects against neoplasia [112,184-186].

Beyond the core components of PRC2, accessory proteins such as Aebp2, Rbbp4/7, and Jarid2, influence PRC2 function [72,93-98]. Recently, Polycomb-like (Pcl) proteins, named for the similarity of the Drosophila *Pcl* mutant phenotype to that of the *Polycomb* mutant, have been found to modulate PRC2 activity [71,99-101,106,109,187,188]. Drosophila Pcl has three homologs in mammals: Pcl1 (also called PHD finger protein 1), Pcl2 (also called Metal response element binding transcription factor 2), and Pcl3 (also called PHD finger protein 19) [189]. *Pcl1* is expressed minimally in ESCs, but promotes PRC2 function in adult tissues and male germ cells [101,105,107,190]. Pcl2 regulates PRC2 differentially depending on cell context and target. In mouse embryonic fibroblasts (MEFS), Pcl2 inhibits PRC2 activity, whereas in ESCs, Pcl2 hinders H3K27me3 formation globally but promotes PRC2 activity at a subset of genes [71,101,109]. Human Pcl3 exists as two isoforms, which can bind Ezh2 and Eed [111]. Mammalian Pcl3 is expressed in ESCs, but it has been unclear how it contributes to PRC2 function and ESC biology.

Here, we show that mouse Pcl3 interacts with the core components of PRC2 and promotes complex function. By depleting Pcl3 in ESCs, we demonstrate that Pcl3 contributes to ESC self-renewal, but not differentiation, of the three germ layers. Using ChIP-seq of *Pcl3* shRNA-treated cells, we show that *Pcl3* knockdown causes decreased

H3K27me3 and Suz12 binding to the genome, indicating that Pcl3 regulates PRC2 binding at diverse target genes. Furthermore, Pcl3 localizes with Suz12 at a subset of PRC2 targets, including genes and microRNAs associated with differentiation and development. We also show that several Pcl3 Tudor domain residues are necessary for H3K27me3. Finally, we identify two GC-rich binding motifs that are enriched at Pcl3-dependent PRC2 targets, indicating that Pcl3 promotes PRC2 binding at CpG islands. Taken together, these results reveal that Pcl3 is an important regulator of PRC2 at a subset of target genes.

RESULTS

Pcl3 is a component of PRC2

To identify PRC2 binding partners that could contribute to its function, we used the recently developed Floxin system to create a tandem affinity purification (TAP) tagged allele of *Suz12* (Figure 14A) [20]. In brief, we reverted a *Suz12* gene trap ($Suz12^{Gt/+}$) allele generated in a mouse ESC line to produce an allele that re-expresses *Suz12* but that contains a loxP targeting site ($Suz12^{Rev/+}$). Via a modified Floxin shuttle vector, we inserted an exon encoding amino acids 277-741 of Suz12 fused to a carboxy-terminal 6xHis-3xFlag TAP tag. The resultant allele ($Suz12^{Suz12TAP/+}$) expressed the full-length TAP-tagged Suz12 from the endogenous locus (Figures 13A and 14B). We measured protein and mRNA levels of Suz12 in all ESC lines by immunoblot and quantitative reverse transcription PCR (qRT-PCR) (Figures 13A and 14C). As expected, $Suz12^{Gt/+}$ cells displayed reduced Suz12 expression, whereas $Suz12^{Rev/+}$ cells displayed levels restored to wild type amounts (Figures 13A and 14C). $Suz12^{Suz12TAP/+}$ cells displayed moderately increased mRNA and protein levels of Suz12 compared to wild type (Figures 13A and 14C).

To reveal novel binding partners of Suz12, we tandem affinity purified Suz12-TAP from *Suz12*$^{Suz12TAP/+}$ ESCs and identified co-purified proteins by mass spectrometry (Figure 13B) [101]. The PRC2 core components Eed, Ezh1, and Ezh2 were highly represented among Suz12 co-purified proteins, as were other known PRC2 interactors, including Rbbp4, Rbbp7, Aebp2, Jarid2, Pcl2, and esPRC2p48 [71,72,93-98,101,109]. In addition, we identified Pcl3 as co-purifying with Suz12. Human Pcl3 contains a long and short isoform, both of which are associated with PRC2 in HEK 293 cells by gel filtration chromatography [111]. To verify binding of mouse Pcl3 with PRC2, we confirmed that V5-tagged Pcl3 co-immunoprecipitated with Suz12-TAP in ESCs (Figures 13C and 14D). To determine whether Pcl3 interacts with all PRC2 core members, we immunoprecipitated Pcl3-V5 and probed for Suz12, Ezh2, and Eed. All core components of PRC2 were found to bind Pcl3 (Figure 13D). Thus, mass spectrometric and co-immunoprecipitation analyses indicated that Pcl3 interacts with PRC2.

Pcl3 promotes ESC self-renewal

Inhibiting PRC2 activity can affect both ESC self-renewal and differentiation by deregulating cell-type specific genes [3,79,80,82,181-183]. *Suz12*$^{-/-}$ ESCs cannot form neural lineages, whereas *Eed*$^{-/-}$ ESCs show an increased propensity to differentiate [80,82,180,191]. To determine if Pcl3 regulates ESC maintenance or differentiation, we tested whether *Pcl3* knockdown altered the ability of ESCs to self-renew or generate cell types derived from all three germ layers.

We depleted Pcl3 using multiple *Pcl3* shRNA targeting vectors (Figures 15A and 16A). Upon culturing multiple clones of *Pcl3* knockdown ESCs, we observed an increased percentage of cells that were larger, less dense and displayed morphologies consistent with differentiation (Figure 15B). These differentiated morphologies suggested that Pcl3 may play a role in ESC maintenance or self-renewal.

To assess whether Pcl3 contributes to ESC self-renewal, we examined ESC markers including Oct4, Nanog, and alkaline phosphatase. Consistent with the morphological changes, Oct4 and Nanog expression and protein levels were decreased in *Pcl3* knockdown cells compared to control ESCs (Figures 15C-D and 16B). We also assessed scramble and *Pcl3* shRNA ESCs for alkaline phosphatase activity. Alkaline phosphatase staining was slightly reduced in *Pcl3* shRNA-treated ESCs (Figure 16C). To quantitate this observation, we used a more sensitive colorimetric assay, which confirmed that *Pcl3* knockdown reduces ESC alkaline phosphatase activity (Figure 15E). To confirm that these phenotypes were specifically due to *Pcl3* knockdown, we overexpressed a TAP-tagged form of Pcl3 in wild type cells (Figure 16D-E). Overexpression of *Pcl3* in ESCs resulted in increased levels of *Oct4* and *Nanog*, further indicating that Pcl3 levels correlate with ESC gene expression (Figures 15C).

To determine whether decreased expression of *Oct4* and *Nanog* affects self-renewal, we assessed *Pcl3* knockdown ESCs for their ability to generate colonies and found that *Pcl3* shRNA ESCs formed significantly fewer colonies than control cells (Figures 15F and 16F). These data suggest that Pcl3 promotes ESC self-renewal. To substantiate this finding, we assayed the ability of wild type and *Pcl3*-overexpressing cells to self-renew and form colonies. We challenged wild type and *Pcl3*-overexpressing cells by growing them in media containing reduced LIF. *Pcl3*-overexpressing ESCs were able to form colonies modestly but significantly better than wild type cells, further indicating that Pcl3 enhances ESC self-renewal (Figure 15G).

In addition to its roles in self-renewal, PRC2 is critical for ESC differentiation and embryonic development [79,80,82,183]. Besides ESCs, *Pcl3* is expressed in a number of differentiated tissues from all three germ layers including bone, spleen, and prostate [101]. *Pcl3* is not well expressed in the adult nervous system, but it expressed in the head during development [101]. To ascertain whether Pcl3 is essential for ESC

differentiation, we performed teratoma and *in vitro* differentiation assays and assessed the presence of all three germ layers. Teratomas from scramble and *Pcl3* shRNA-expressing cells were weighed and histologically examined. While *Pcl3* knockdown was maintained, we did not observe differences in the size of the teratomas nor in the formation of endodermal, ectodermal, or mesodermal derivatives (data not shown and Figures 15H and 16G). Moreover, immunofluorescent staining revealed the presence of neural tissue (NeuN and Tuj1), muscle (Actin), and skin basal cells (K14) in both *Pcl3* expressing and *Pcl3* knockdown teratomas (Figure 16H).

In vitro differentiation of scramble and *Pcl3* knockdown ESCs to embryoid bodies resulted in a similar finding. *Pcl3* shRNA-treated EBs maintained *Pcl3* knockdown and expressed markers of neuroectoderm (*Nestin*), mesoderm (*T-brachyury*), and endoderm (*Hnf4*) at least as well as control EBs, indicating that depletion of *Pcl3* does not abrogate the ability of ESCs to differentiate to cell types of all three germ layers (Figure 16I). Thus, Pcl3 enhances ESC self-renewal, but is not critical for ESC differentiation to the three germ layers.

Pcl3 promotes trimethylation of H3K27

PRC2 accessory proteins, such as Aebp2, Pcl1, Pcl2 and Jarid2, can either promote or inhibit PRC2 function [71,72,94-98,101,105,109,190]. To elucidate whether Pcl3 regulates PRC2 function, we measured H3K27me3 levels in cells with diminished Pcl3. *Pcl3* shRNA-treated ESCs and EBs showed approximately 80% depletion of H3K27me3 levels compared to scramble shRNA controls, indicating that Pcl3 promotes PRC2 function (Figure 17A-B). To confirm this finding, we transfected ESCs with either *Suz12* or *Pcl3* siRNAs and assessed H3K27me3 levels [192,193]. qRT-PCR and immunoblot indicated that transfection of *Suz12* siRNAs diminished *Suz12* expression in a dose-dependent fashion (Figure 18A-B). Similarly, transfection of *Pcl3* siRNAs

resulted in decreased *Pcl3* expression and markedly reduced H3K27me3, further demonstrating that *Pcl3* promotes the formation of H3K27me3 (Figures 18C and 17C). As *Pcl3* knockdown decreased H3K27me3, we assessed whether increased Pcl3 resulted in a concomitant increase in H3K27me3. Indeed, overexpression of *Pcl3* in ESCs enhanced H3K27me3 levels (Figures 18D-E and 17D). Thus Pcl3 promotes trimethylation of H3K27.

Besides H3K27me3, histone modifications such as ubiquitination of lysine 119 of histone H2A (H2AK119Ub) and trimethylation of lysine 9 of histone H3 (H3K9me3) are involved in gene repression [194-196]. In addition, H3K27me3 is inversely correlated with acetylation on lysine 27 (H3K27ac) and marks bivalent domains with H3K4me3 [69,70]. To assess whether Pcl3 participates in histone modifications apart from H3K27me3, we analyzed histones from scramble and *Pcl3* shRNA-treated cells for H2AK119Ub, H3K9me3, H3K4me3, and H3K27ac. Levels of H2AK119Ub, H3K9me3, H3K4me3, and H3K27ac were all comparable between the control and *Pcl3* knockdown ESCs, indicating that Pcl3 is not involved in global histone modification, but has a specific role in generating H3K27me3 (Figure 17E).

To confirm that *Pcl3* shRNA-associated phenotypes were due to decreased *Pcl3* levels, we created a construct encoding a TAP-tagged Pcl3 not recognized by the *Pcl3* shRNAs. Pcl3-TAP was detected by immunoblot and immunofluorescence and was able to bind core PRC2 components as assessed by co-immunoprecipitation, indicating that Pcl3-TAP functions similarly to wild type Pcl3 (Figures 17F-G and 18D). *Pcl3* expression in *Pcl3-TAP*-expressing ESCs was nearly 60% of endogenous levels as measured by qRT-PCR (Figure 17H). Re-expression of *Pcl3-TAP* increased H3K27me3 levels in *Pcl3* knockdown ESCs (Figure 17I). These data indicate that Pcl3 promotes the formation of H3K27me3 by PRC2.

Pcl3 promotes PRC2 binding and function at a subset of PRC2 targets

To elucidate which genes depend upon Pcl3 for H3K27me3 formation and to investigate whether decreased H3K27me3 is a result of Pcl3-dependent PRC2 binding, we performed ChIP-seq for H3K27me3 and Suz12 in *Suz12*$^{Suz12TAP/+}$ cells expressing either *Pcl3* or control shRNA (GSE28325). Reads were aligned to the mouse reference genome (NCBI Build 37, MM9) using bowtie, and only uniquely aligned reads were considered for further analysis (Figure 20A-B) [197].

Suz12 ChIP-seq was performed using the Flag tag of Suz12-TAP. To validate this approach, we compared ChIP-seq of Suz12-TAP using anti-Flag antibodies with ChIP-seq using anti-Suz12 antibodies in *Suz12*$^{Suz12TAP/+}$ cells and found they were equivalent. We, thus, used FlagM2 to detect Suz12 binding in all the following experiments. ChIP-seq with FlagM2 antibodies detected 95% of previously reported Suz12 binding sites, indicating that this approach accurately measured Suz12-TAP localization in the genome (Figure 20C) [15,198]. In addition, we identified a number of novel Suz12 binding sites, typified by less significant Suz12 ChIP-seq peaks (Figure 20C). These additional sites may reflect weak Suz12 binding or may be an artifact of increased Suz12 levels.

Comparison of H3K27me3 ChIP-seq in scramble and *Pcl3*-shRNA treated ESCs revealed a widespread decrease in H3K27me3 levels in *Pcl3* shRNA ESCs (Figure 19A). The decrease in H3K27me3 as assessed by ChIP-seq was similar to, but not as profound as was detected biochemically. This may be attributable to the inability of ChIP-seq to detect H3K27me3 in regions of highly repetitive sequence, areas that are largely transcriptionally silent. At approximately 85% of sites showing decreased H3K27me3, *Pcl3* knockdown cells also displayed reduced Suz12 binding compared to scramble controls, thus correlating decreased Suz12 occupation with decreased PRC2 function (Figure 19A). Indeed, *Pcl3* knockdown significantly reduced Suz12 binding at

approximately 65% of PRC2 targets (Figure 19A, C). Sites with decreased Suz12 binding outnumbered those with reduced H3K27me3, indicating that *Pcl3* knockdown preferentially affected Suz12 (Figure 19A).

The regions with the most decreased Suz12 occupation following *Pcl3* knockdown were among the most significant Suz12 ChIP-seq peaks in control cells (Figure 20D). In contrast, less significant Suz12 peaks were more likely to be unaffected by *Pcl3* knockdown (Figure 20D). The finding that the role of Pcl3 is most pronounced on the most significant Suz12 binding sites is consistent with a role for Pcl3 in mediating the recruitment of additional Suz12 to regions of low level Suz12 binding. Alternatively, changes in Suz12 binding may be more difficult to detect at these less significant peaks.

Knockdown of *Pcl3* diminished Suz12 binding to chromatin particularly in areas of high gene density (Figure 19B). Accordingly, chromosome 11, the chromosome with the highest gene density, showed the most profound decrease in Suz12 binding and H3K27me3 upon *Pcl3* knockdown (Figures 19A and 20E). Consistent with PRC2 binding at repressed genes, sites with decreased Suz12 binding in *Pcl3* shRNA-treated ESCs were inversely correlated with sites bound by the activating transcription factors E2f1, c-Myc, Zfx, Klf4, and Ctcf, (Figure 20F and data not shown, Fisher test p-value = $9.5x10^{-53}$, $3.7x10^{-15}$, $1.2x10^{-13}$, $4.0x10^{-6}$, $3.6x10^{-6}$, respectively) [15]. This is in agreement with previous reports demonstrating nearly mutual exclusion of Suz12 binding with regions bound by this group of transcription factors [15].

As mentioned, Suz12 and H3K27me3 are present at bivalent genes, genes with both active (H3K4me3) and repressive (H3K27me3) histone marks in ESCs [70,199]. We found that more than 80% of bivalent genes showed decreased Suz12 binding upon *Pcl3* depletion (Figure 19C) [70,199]. Among bivalent genes are many developmental genes and microRNAs [198,200-204]. Using ChIP-qRT-PCR, we confirmed that Pcl3 is important for deployment of PRC2 at some of these developmental gene targets and

microRNAs-including *Hoxb7*, *Hoxb13*, *Hoxa3*, *Dusp4*, *Hand1*, *mir-196b*, *mir-196a1*, *mir-10a*, *mir-132/mir-212*, and *mir-34a* (Figure 19D-F) [200,201,203,205-208]. Suz12 ChIP-qRT-PCR also confirmed that some Suz12 targets (e.g., *Tle3*, *Tns1*, and *Tmem151a*) are unaffected by Pcl3 (Figure 20G). These data further demonstrate that Pcl3 promotes PRC2 binding and mediates PRC2 activity at a specific subset of PRC2 targets.

As H3K4me3 histone marks are also present at bivalent genes, we extended our ChIP-qRT-PCR analysis to H3K4me3 to assess whether *Pcl3* depletion affects H3K4me3 formation. At genes marked by H3K4me3 but not H3K27me3 (i.e., *Eaf1*, *Ash21*, *Dcaf8*, *Ift140*, *Glrx5*), H3K4me3 levels were unchanged by *Pcl3* depletion, consistent with our finding that global H3K4me3 levels are unchanged in *Pcl3* shRNA ESCs (Figures 20H and 17E). H3K4me3 levels at approximately half of the bivalent genes assayed showed modestly increased H3K4me3 levels (i.e., *Acta1*, *Hoxa3*, *Ebf2*, *mir196b*, *Hand1*, *Dusp4*, *Pmp22*), while other genes were unchanged (i.e., *mir196a1*, *Hoxb7*, *Matb2*, *Pcdh7*, *Otx2*, *Hoxb13*) (Figure 20H). Thus, *Pcl3* knockdown affects H3K4me3 levels specifically at a subset of bivalent genes.

Pcl3 co-localizes with Suz12 at many PRC2 targets to promote H3K27me3

As Pcl3 binds core components of PRC2 and promotes Suz12 localization to its target genes, we investigated whether Pcl3 co-localizes with Suz12 at target loci. To assess the distribution of Pcl3 binding to chromatin in ESCs, we performed ChIP-seq of Pcl3-TAP. Nearly all of the regions bound by Pcl3-TAP overlapped with Suz12 targets, suggesting that Pcl3 co-localizes at target genes as part of PRC2 (Figures 21A and 22A-B, Fisher test p-value = 1.5×10^{-306}). Quantification revealed that Pcl3-TAP bound nearly half of PRC2 targets, indicating that Pcl3 associates with a subset of PRC2 complexes (Figure 21A). To substantiate the co-localization analysis, we assessed which Suz12

sites overlapped with Pcl3. Pcl3 and Suz12 co-localized predominantly at the most significant Suz12 binding sites (Figure 21B). Suz12 sites at which Pcl3 did not co-localize were typified by less significant Suz12 binding (Figure 21B). In addition, 80% of Pcl3 and Suz12 sites overlapped with CpG islands compared to 65% of sites bound only by Suz12, indicating that Pcl3 and Suz12 preferentially bind at CpG islands (p-value = $2.2x10^{-16}$) [173].

To establish whether Pcl3 may directly contribute to PRC2 binding and function, we compared regions of Pcl3 binding to areas of Pcl3-dependent Suz12 chromatin occupation and H3K27me3 formation. Upon *Pcl3* knockdown, 84% of Pcl3 binding regions showed decreased Suz12 binding, whereas less than half of regions not bound by Pcl3 showed decreased Suz12 binding (Figure 21C-E). In addition, Pcl3 binding regions showed a near two-fold enrichment for decreased H3K27me3 upon *Pcl3* knockdown, as compared to regions not bound by Pcl3 (Figure 21E). ChIP-qRT-PCR for Pcl3 confirmed that Pcl3 bound to many of the same genes that exhibited decreased Suz12 binding and H3K27me3 upon *Pcl3* knockdown (Figures 19F and 21F). These data indicate that genes bound by Pcl3 are nearly twice as likely to have reduced Suz12 binding and H3K27me3 upon *Pcl3* knockdown. Thus, Pcl3 not only co-localizes with Suz12 at many PRC2 targets, but also promotes Suz12 binding and PRC2 function.

Pcl3 regulates gene expression at a subset of PRC2 targets

Inhibition of PRC2 core components and complete abrogation of H3K27me3 dramatically alter gene expression, whereas inhibition of accessory proteins such as Jarid2 result in more modest changes [72,79,82,183]. To test whether Pcl3 affects gene expression, multiple scramble and *Pcl3* shRNA clones were analyzed by microarray. Nearly 130 genes displayed altered expression upon *Pcl3* knockdown. Unexpectedly, given that PRC2 functions to repress gene expression, nearly three times as many

genes were down-regulated upon *Pcl3* knockdown as were up-regulated (Table 1). Notably, we identified more than half of the upregulated genes as PRC2 targets, which correlated well with depleted Suz12 binding and H3K27me3 (Figure 24B). We confirmed the microarray data by measuring gene expression by qRT-PCR for several genes (Figure 24A).

Given that *Pcl3* knockdown increases the proportion of differentiated cells, we tested whether the presence of differentiated cells masked changes in ESC gene expression by pre-plating twice to remove differentiated cells. Microarray analysis revealed that approximately 120 genes showed differential expression in the *Pcl3* shRNA cells compared to control after removal of differentiated cells (Table 2). qRT-PCR assessment of a subset of the affected genes confirmed the microarray results (Figure 24B). Of these Pcl3-regulated genes, almost two thirds were up-regulated upon Pcl3 depletion. While these results demonstrate that Pcl3 participates in the repression of genes in ESCs, only a few of these genes also displayed decreased Suz12 binding and H3K27me3 levels (i.e., *Hand1*, *Acta1*, and *Pmp22*) (Table 2, Figures 19E-F and 24B) which may be attributable to the necessity of culturing *Pcl3* shRNA ESCs for an extended time before subjecting them to microarray analysis.

Pcl3 regulates Suz12, but not complex stability

Loss of core PRC2 components can destabilize the complex [82,181]. To assess whether loss of Pcl3 affects complex stability, we tested whether the complex could form in the absence of Pcl3. All components, Suz12, Ezh2, and Eed, were found by co-immunoprecipitation to bind each other in both scramble and *Pcl3* shRNA-expressing cells, suggesting that Pcl3 is not required for assembly or stabilization of the core complex (Figure 23A).

To determine whether depleting Pcl3 influences levels of PRC2 components, we assessed the quantity of PRC2 core components in *Pcl3* shRNA-treated ESCs. Interestingly, suppression of *Pcl3* significantly lowered Suz12 mRNA and protein levels without affecting levels of Eed or Ezh2, suggesting that Pcl3 regulates *Suz12* (Figure 23B-C). Overexpressing Pcl3 caused a concomitant increase in Suz12 mRNA and protein levels (Figure 23D-E). As *Suz12* is not a known PRC2 target and Pcl3 is not enriched at the *Suz12* locus, effects on *Suz12* expression by Pcl3 are likely indirect (Figure 24D) [168,209,210].

The Tudor domain of Pcl3 is critical for its function

Polycomb-like proteins contain three domains, an amino-terminal Tudor domain and two carboxy-terminal PHD fingers (Figure 25A). Although the carboxy PHD finger of Pcl2 promotes PRC2 recruitment, the Tudor domain and amino PHD finger of Pcl2 are not required for its regulation of PRC2 [101,211]. In contrast, the Tudor domain and carboxy-terminal PHD finger of human Pcl3 are important for binding Ezh2 and self-association [111].

NMR studies indicate that the Drosophila Polycomb-like Tudor domain lacks a complete aromatic cage, needed to bind methylated lysines or arginines on histones [102]. Interestingly, a comparison of Drosophila Pcl Tudor domain sequence to those of mammalian Pcl Tudor domains suggests that mammalian Polycomb-like proteins may be able to form complete aromatic cages [102]. In particular, the conserved residues W48,Y54, F72, D74, and S76 are implicated in histone binding (Figure 25A-B) [102]. To ascertain if these Tudor domain residues are required for Pcl3 function, we mutated two sets (W48;Y54 and F72;D74), as well as a control set of residues (N75;Y78) within Pcl3-TAP (Figure 25A-B) [102]. These mutant forms of Pcl3 were expressed in *Pcl3* knockdown ESCs to see if they could rescue H3K27me3 levels as well as wild type Pcl3-

TAP (Figure 25C). Similar to wild type Pcl3-TAP, Pcl3-TAP N75S;Y78S was able to promote H3K27me3 formation (Figure 25D). In contrast, Pcl3-TAP W48A;Y54A and Pcl3-TAP F72S;D74S did not restore H3K27me3 formation (Figure 25D). Thus, W48;Y54 and F72;D74, two pairs of residues implicated in histone binding, are necessary for Pcl3 function, while the control N75;Y78 pair is dispensable.

To discern whether these Tudor domain residues are necessary for Pcl3 incorporation into PRC2, we assessed whether Suz12 co-immunoprecipitated with Pcl3-TAP W48A;Y54A, Pcl3-TAP F72S;D74S, and Pcl3-TAP W48A;Y54A. All mutant forms of Pcl3 associated with Suz12, indicating that these residues are not required for binding PRC2 (Figure 25C). Thus, W48;Y54 and F72;D74 are essential for Pcl3 promotion of H3K27me3 formation, but not for Pcl3 incorporation into PRC2.

Pcl3 and Pcl2 are part of distinct PRC2 complexes that have overlapping targets

While Pcl1 is minimally transcribed in ESCs, Pcl2 is highly expressed and binds a subset of PRC2 target genes [101]. To assess whether depletion of Pcl3 affects Pcl1 or Pcl2, we measured *Pcl1* and *Pcl2* expression in scramble and *Pcl3* knockdown ESCs. Pcl1 was expressed at extremely low levels but showed no difference upon *Pcl3* knockdown (Figure 26A). Likewise, we found that *Pcl2* levels were similar between scramble and *Pcl3* shRNA-treated ESCs (Figure 26A).

Although Pcl2 and Pcl3 are both expressed in ESCs, it is unclear if they participate in same complex or regulate the same genes. To assess whether Pcl3 and Pcl2 incorporate into the same complex or exist in distinct PRC2 complexes, we tested Pcl3 and Pcl2 association by co-immunoprecipitation. Flag immunoprecipitation confirmed that Pcl2 and Pcl3 both bind Suz12 (Figure 26B). However, Pcl3-TAP did not immunoprecipitate Pcl2, suggesting that Pcl2 and Pcl3 do not associate within the same PRC2 complex (Figure 26B). To substantiate these data, we performed the reciprocal

immunoprecipitation, and found that Pcl2 immunoprecipitated Suz12-TAP but not Pcl3-TAP. These findings indicate that Pcl2 and Pcl3 both interact with the core PRC2 complex, but not with each other, suggesting that Pcl2 and Pcl3 participate in distinct PRC2 complexes.

To elucidate whether Pcl3 and Pcl2 bind distinct or overlapping sets of targets, we compared our Pcl3 ChIP-seq data with previously generated Pcl2 ChIP-seq data and found that Pcl3 and Pcl2 binding overlaps at 60-70% of sites [101]. This extensive overlap raises the possibility that many genes may be regulated by both Pcl3 and Pcl2.

Like Pcl3, Pcl2 can promote PRC2 activity at some loci and in some cell types [101,109,110,211]. To test whether Pcl2 can compensate for Pcl3 function, we analyzed whether Suz12 binding at Pcl2 targets was affected by Pcl3 depletion. Of sites bound by both Suz12 and Pcl2, 86% showed decreased Suz12 binding upon Pcl3 depletion (Figure 26C-D), suggesting that Pcl2 cannot compensate for Pcl3 function at Pcl2 sites.

Unlike Pcl3, Pcl2 can either inhibit or promote PRC2 function [101,109,110,211]. To further investigate the functional relationship between Pcl2 and Pcl3, we assessed Suz12 binding and H3K27me3 at regions bound by both Pcl2 and Pcl3. If Pcl2 and Pcl3 act redundantly to promote Suz12 binding, sites bound by both would not show significant decreases in Suz12 binding upon *Pcl3* knockdown as Pcl2 would be functional at these sites. Conversely, if Pcl2 hinders PRC2 binding while Pcl3 promotes PRC2 binding, Suz12 binding at shared Pcl2 and Pcl3 sites would be more dramatically reduced upon *Pcl3* knockdown, as Pcl3 would no longer be able to counterbalance Pcl2 function. Among regions bound by Suz12, Pcl2, and Pcl3, 95% showed Suz12 depletion upon *Pcl3* knockdown (Figure 26C-D). Moreover, Pcl3-dependent Suz12 sites often co-localized with Pcl2 and Pcl3 binding, whereas Pcl3-independent Suz12 sites tended to be more distant from Pcl2 and Pcl3 binding sites (Figure 26E). In addition, 88% of Pcl2 and Pcl3 shared sites showed diminished H3K27me3. Thus, these data suggest that

not only does Pcl2 not compensate for Pcl3, but also that Pcl2 may hinder Suz12 binding and H3K27me3 at Pcl2 and Pcl3 shared binding sites.

Pcl2 and Pcl3 co-localize with Suz12 at CpG islands

Polycomb repressive elements (PREs), binding sites that recruit Polycomb proteins, have been well-described in Drosophila, but few have been found in mammals [83-87]. Nevertheless, mammalian PRC2 is known to preferentially bind CpG islands [88,173,212]. Similarly, we found that 86% of Pcl2 and Pcl3 co-localization sites were found in CpG islands [101,134,173]. To assess whether distinct DNA sequence motifs required for Pcl2 and Pcl3 recruitment exist, we searched for sequence features that can discriminate DNA sequences in genomic regions bound by both Pcl2 and Pcl3. We scanned 500 base pair regions surrounding Pcl3 ChIP-seq peaks for prevalent DNA sequences and identified two enriched 10- and 14-mer motifs (Figure 27A). Motif 1 consisted of three consecutive G-C dinucleotides with two nucleotides on either side containing modest GC enrichment. Motif 2 contained two G-C-rich regions separated by seven nucleotides with little to no GC enrichment. Thus, motifs commonly found in Pcl2 and Pcl3 binding regions contain short GC sequences surrounded by dissimilar sequence.

To assess whether these motifs were predictive of Pcl2 and Pcl3 binding sites, Pcl3-dependent Suz12 binding sites or CpG islands, we constructed a position specific scoring matrix (PSSM) for each motif. The occurrence of motifs 1 and 2 correlated more tightly with Pcl3-dependent Suz12 binding sites that co-localize with both Pcl2 and Pcl3 binding sites than with Suz12 sites that are not associated with Pcl2 and Pcl3 binding, nor depend on Pcl3 activity (Figure 27B-C). In addition to the maximum PSSM scores, we also computed the CpG density of each site and used these three sequence features to build a support vector machine classifier to help predict whether specific motifs would

co-localize with Pcl2 and Pcl3 binding and areas of Suz12 depletion following *Pcl3* knockdown (Figure 27D). The classifier had cross-validation accuracy of 75%, indicating that CpG density, Motif 1 and Motif 2 together account for the majority of Pcl3-dependent PRC2 binding. Even though the vast majority of Pcl2 and Pcl3 binding sites were found in CpG islands, they seemed to avoid regions with the highest CpG density, concordant with the moderate GC content of the motif sequences (Figure 27B-D). These data indicate that Pcl2 and Pcl3 may utilize specific binding motifs to recruit PRC2 to form H3K27me3 at CpG islands.

DISCUSSION

In an effort to understand the mechanisms that regulate PRC2, we have discovered that a PRC2 interacting protein, Polycomb-like 3, promotes ESC self-renewal and mediates PRC2 binding at a number of targets genes. In ESCs, Pcl3 promotes the expression of multiple markers of pluripotency. Consistent with a functional role for Pcl3 in PRC2, Pcl3 also promotes H3K27me3 formation. ChIP-seq analyses revealed that diminished Pcl3 decreased Suz12 binding, particularly in regions where Pcl3 co-localizes with Suz12, suggesting that one way that Pcl3 promotes PRC2 activity is by mediating PRC2 binding to chromatin. Pcl3 likely resides in a PRC2 complex distinct from that which incorporates its paralog, Pcl2, but both PRC2 complexes have overlapping targets predominantly at CpG-rich regions. Despite their overlap and homology, Pcl2 cannot compensate for Pcl3 and may inhibit PRC2 at shared genes, as regions of Pcl2 and Pcl3 co-binding show the most significant reduction in Suz12 binding upon *Pcl3* knockdown.

Pcl3 likely has a direct role in promoting PRC2 binding at select targets, as it 1) biochemically interacts with PRC2, 2) binds to genomic regions overlapping with Suz12,

and 3) is predominantly required for genomic Suz12 binding at Pcl3 binding sites. This co-occurrence and dependence of PRC2 on Pcl3 suggests that the biochemical interaction of Pcl3 with PRC2 underlies its role in PRC2 binding. In addition to this direct role for Pcl3 in PRC2 binding to chromatin, the reduction in Suz12 binding upon *Pcl3* knockdown may also be partially attributable to decreased Suz12 levels. Suz12 is not known to be regulated by PRC2 and Pcl3 does not bind the *Suz12* locus, indicating that this regulation of Suz12 by Pcl3 is likely to be indirect.

Many bivalent genes are involved in regulating ESC differentiation [70,169-175]. However, we observed that Pcl3 promotes ESC self-renewal, but is not required for differentiation into cell types of all germ layers. As we have found that Pcl3 promotes H3K27me3 at a subset of bivalent genes, it is possible that Pcl3 predominantly represses bivalent genes involved not in cell fate decisions in differentiating cells, but in genes that promote differentiation itself. Thus, Pcl3 depletion would derepress genes that promote differentiation, leading to specific defects in self-renewal. Conversely, Pcl3 overexpression may lead to increased self-renewal by repressing genes involved in differentiation that inhibit the expression of self-renewal genes such as *Nanog*.

PRC2 is recruited to chromatin by multiple mechanisms including interactions with CpG islands, non-coding RNAs, and histone binding proteins [95-97,213-215]. Our analysis of Suz12 and Pcl3 binding sites revealed that they associate with CpG-enriched regions, particularly two CpG-rich DNA sequence motifs. These motifs contained GC rich regions surrounded by areas with minimal GC content, indicating that Pcl proteins do not associate with the most GC-rich regions, but rather with regions of moderate GC content and containing motifs 1 or 2. Our support vector machine classifier supported these findings by showing that Pcl2 and Pcl3 binding sites that overlap with Pcl3-dependent Suz12 bound sites localize in areas with moderate CpG density. The finding that Pcl proteins associate with CpG islands is consistent with prior observations that

PRC2 binds to CpG islands and demonstrates that Pcl3 may participate in PRC2 recruitment to CpG islands [88,173,212].

Pcl proteins contain multiple domains by which they can associate with chromatin, including a Tudor domain and two PHD fingers implicated in recognizing methylated or unmethylated histone lysines or arginines [104,114,216]. Recent work has implicated several residues of the Pcl Tudor domain in histone binding [102]. Mutating two sets of these Pcl3 Tudor domain residues revealed that they are essential for Pcl3-dependent PRC2 activity. As these mutant forms were able to incorporate into PRC2, they may be acting as dominant negatives and poisoning the complex or as loss-of-function mutants by creating less-functional complexes. Thus, these mutants may perturb PRC2 function through inhibition of complex recruitment to histones or by decreasing PRC2 methyltransferase activity.

Two other paralogs of Pcl exist in mammals. Pcl1 is expressed minimally in ESCs, and Pcl2 has been implicated in either promoting or inhibiting PRC2 in a gene- and cell type-dependent manner [101]. In ESCs, Pcl2 restrains PRC2 activity globally, and in MEFs, Pcl2 inhibits PRC2 recruitment to certain loci [101,109,211]. In an effort to understand the relationship between these paralogs, we found that Pcl2 and Pcl3 do not associate and likely participate in separate PRC2 complexes, which both bind a subset of PRC2 targets. If Pcl2 and Pcl3 both promoted PRC2 binding and activity at these targets, we might expect Pcl2 to compensate in Suz12 binding and H3K27me3 upon loss of Pcl3. Instead, we observed the greatest dependence of Suz12 binding on Pcl3 at regions also bound by Pcl2, suggesting that Pcl2 may inhibit PRC2 binding at these genes (Figure 27E).

It is unlikely that Pcl2 requires Pcl3 for activity as knockdown of *Pcl2* and *Pcl3* in ESCs results in opposite phenotypes, further suggesting that these two paralogs have distinct and opposing functions. For example, Pcl2 destabilizes Ezh2, whereas Pcl3

enhances Suz12 protein levels without affecting Ezh2 [101]. Moreover, Pcl2 and Pcl3 have opposite effects on ESC self-renewal and differentiation. Pcl2 inhibits expression of *Oct4*, *Sox2*, and *Nanog* and promotes ESC differentiation [101]. In contrast, Pcl3 promotes *Oct4* and *Nanog* levels and ESC self-renewal. Finally, whereas the Tudor domain of Pcl2 is dispensable for its function, the Pcl3 Tudor domain is essential for mediating PRC2 activity [101].

Despite its global requirement to restrain H3K27me3 levels in ESCs, Pcl2 can promote H3K27me3 and PRC2 binding at certain sites [101,109]. Pcl2 and Pcl3 overlap at approximately 65% of sites, indicating that each binds separately at a third of their sites. While our data indicates that Pcl3 and Pcl2 function oppositely at Pcl2 and Pcl3 targets, one possibility is that Pcl2 promotes PRC2 activity at regions devoid of Pcl3 (Figure 27E) [101].

Based on these data, we propose that Pcl2 and Pcl3 have opposing functions on PRC2 binding and activity at sites they co-regulate (Figure 27E). At sites that Pcl2 and Pcl3 individually regulate, each may both promote PRC2 binding. Upon loss of Pcl3, sites dependent upon Pcl3 for promoting PRC2 function lose H3K27me3, and this loss may be exacerbated by negative regulation of PRC2 by Pcl2 at sites that Pcl2 and Pcl3 regulate together (Figure 27E). Thus, Pcl2 and Pcl3 may function antagonistically at shared sites, but promote H3K27me3 at sites regulated by a single Pcl (Figure 27E). This model predicts that the minority of sites at which Pcl2 promotes H3K27me3 levels will be regions not bound by Pcl3. The model does not resolve how Pcl2 could promote PRC2 activity at regions where Pcl3 does not bind, but could inhibit PRC2 activity at shared sites upon Pcl3 knockdown. One possibility is that regions of Pcl2 and Pcl3 co-occupancy are characterized by an inhibitory PRC2 activity independent of Pcl2.

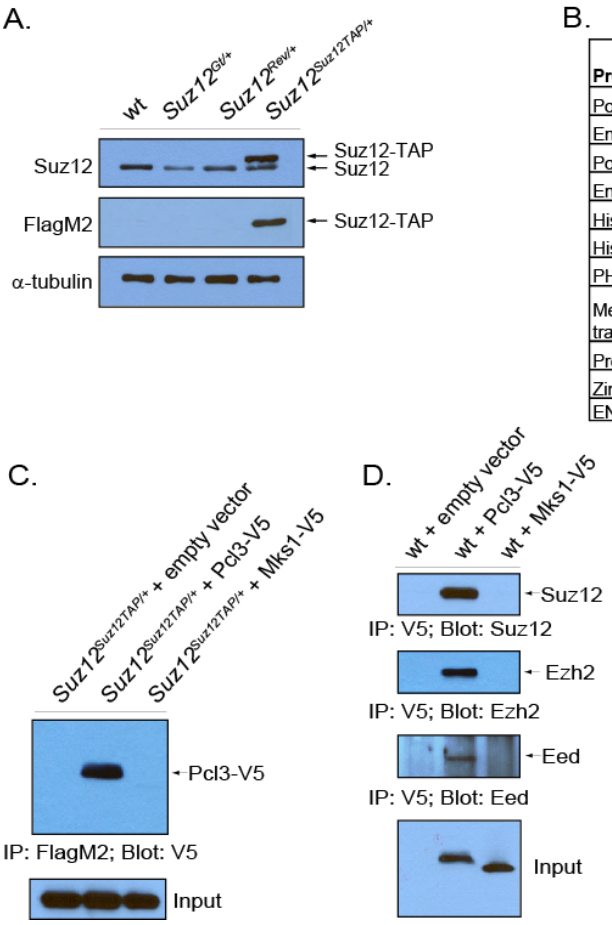
Both bivalency and gene repression at CpG promoters are dependent on PRC2 activity. In ESCs, CpG promoters may be poised for activation, reflected by their ability

to bind RNAPII even when inactive [217]. Pcl3-mediated deployment of PRC2 at CpG islands may be one mechanism by which these poised CpG promoters are kept in check in ESCs. When cell type-specific promoters are being repressed during development, Pcl3 may promote PRC2 binding at many of these same CpG islands in ways that recruit PRC1 to ensure more permanent transcriptional silencing.

Understanding how Pcl proteins regulate PRC2 at specific genes may also be critical to elucidating gene misregulation during cancer. Many cancers show a global silencing at gene promoters, particularly at CpG islands [218,219]. Ezh2 is one of the most commonly misregulated genes in cancer, and recent work has shown that Pcl3 is misregulated in diverse cancers as well [112,220]. Similar to our findings in ESCs, Pcl3 upregulation in cancer cells may promote PRC2 recruitment, inhibiting the expression of genes that inhibit self-renewal, and leading to inappropriate cell proliferation. Thus, studying the role of Pcl3 in PRC2 function in ESCs may provide insight into how Pcl3 up-regulation returns cancer cells to transcriptional states and self-renewal capacities similar to those of ESCs.

Accession Number

The ChIP-seq data are available from Gene Expression Omnibus under access number GSE28325.



B.

Protein Name	Symbol	Unique hits	% Coverage
Polycomb Protein Suz12	Suz12	26	47.1
Enhancer of zeste homolog 2	Ezh2	20	29.6
Polycomb Protein Eed	Eed	26	62.4
Enhancer of zeste homolog 1	Ezh1	11	19.1
Histone-binding protein RBBP4	Rbbp4	14	37.9
Histone-binding protein RBBP7	Rbbp7	8	15.5
PHD finger protein 19 (Phf19)	Pcl3	5	13.5
Metal-response element-binding transcription factor 2 (Mtf2)	Pcl2	18	34.4
Protein Jumonji	Jarid2	18	16.1
Zinc finger protein AEBP2	Aebp2	10	28.2
ENSP00000317905 homolog	esPRC2p48	11	49.9

Figure 13. *Pcl3 is a component of PRC2.*

(A) Protein levels of Suz12 and Suz12-TAP were measured in wild type, *Suz12*$^{Gt/+}$, *Suz12*$^{Rev/+}$, and *Suz12*$^{Suz12TAP/+}$ cell lines by immunoblot. **(B)** Proteins detected by mass spectrometry that specifically co-purified with Suz12-TAP, their symbol, unique hits, and percent coverage. **(C)** Pcl3-V5 binds to Suz12-TAP. *Suz12*$^{Suz12TAP/+}$ ESCs were transfected with empty vector, Pcl3-V5, or Mks1-V5 (control), and lysates were immunoprecipitated with FlagM2 and probed with anti-V5. **(D)** Pcl3-V5 binds Suz12, Ezh2, and Eed. Lysates from ESCs transfected with empty vector, Pcl3-V5, and Mks1-V5 (control) vectors were subjected to immunoprecipitation with anti-V5. Samples were then probed with anti-Suz12, anti-Ezh3, and anti-Eed. All westerns and co-immunoprecipitations were performed three times.

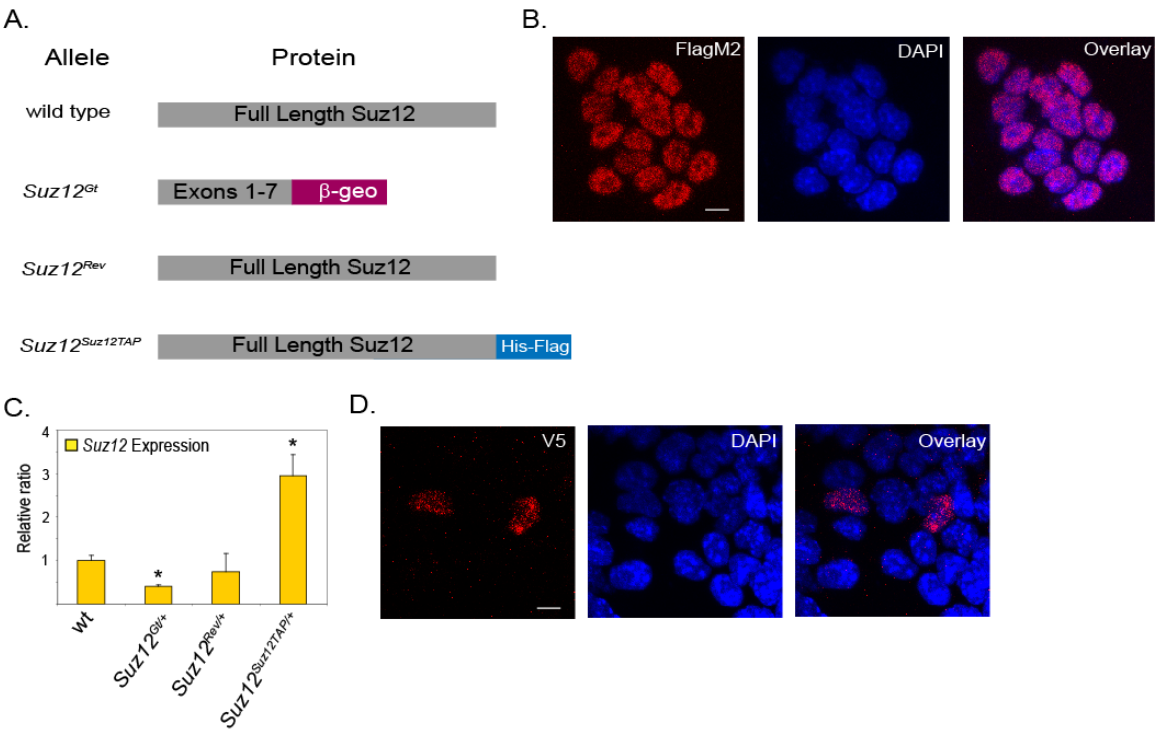


Figure 14. *Suz12-TAP and Pcl3-V5 localize to the nucleus.*

(A) Schematic of recombination events performed to create *Suz12*$^{Suz12TAP/+}$. The wild type endogenous allele is not pictured. Not to scale. **(B)** Staining of *Suz12*$^{Suz12TAP/+}$ with anti-FlagM2 (red), which overlays with nuclei stained with DAPI (blue). Scale bar 10μm. **(C)** Suz12 expression levels in wild type, *Suz12*$^{Gt/+}$, *Suz12*$^{Rev/+}$, and *Suz12*$^{Suz12TAP/+}$ cell lines measured by qRT-PCR. Error bars indicate standard deviation. Asterisk denotes statistical significance of $p < 0.0001$. Graph represents average expression of three different clones in three experiments assayed in quadruplet. **(D)** Wild type cells transfected with Pcl3-V5 and stained with anti-V5 (red). V5 staining overlays with nuclei stained with DAPI (blue). Scale bar 10μm. Stainings were performed two times.

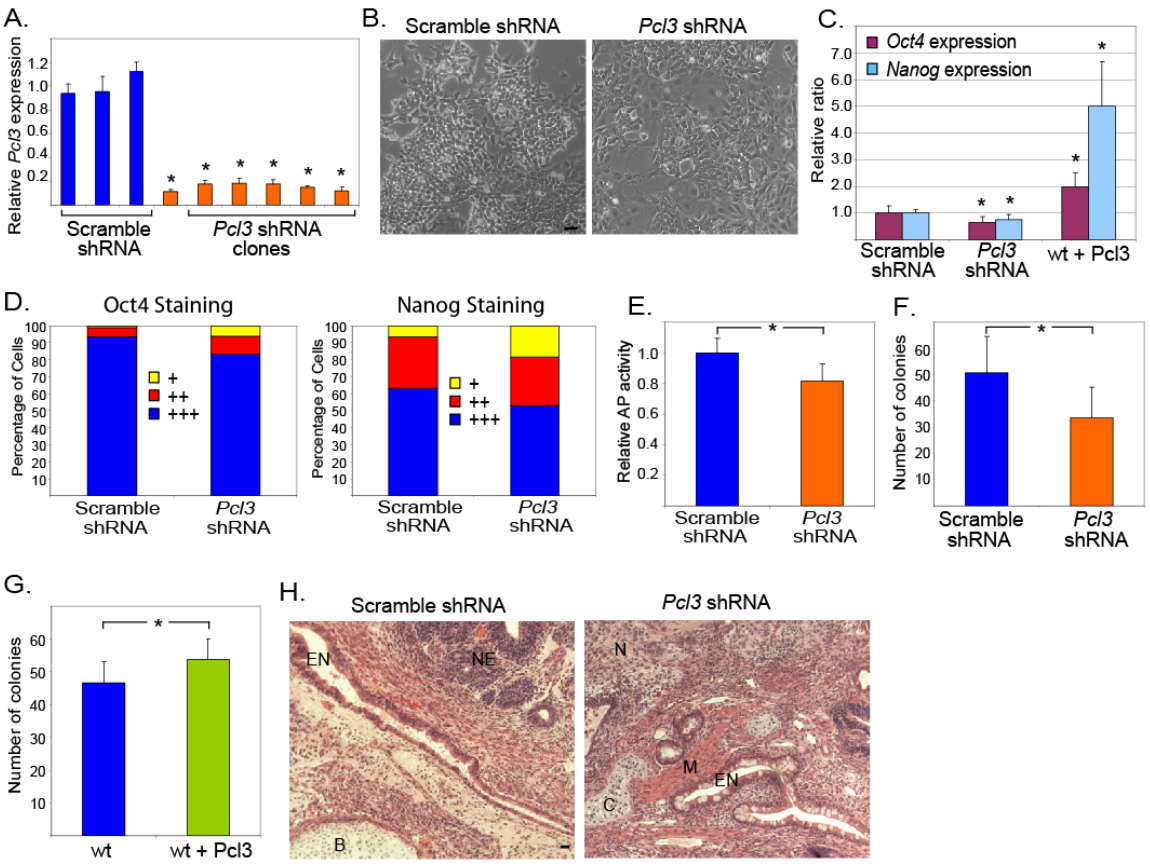


Figure 15. *Pcl3 promotes ESC self-renewal.*

(A) *Pcl3* expression levels measured by qRT-PCR in ESC clones transduced with scramble or multiple *Pcl3* shRNAs. Graph represents average expression from 3-6 different clones. **(B)** A portion of cells transduced with *Pcl3* shRNA, but not scramble shRNA, are larger, flatter, and less dense, signifying a decrease in ESC cell morphology. Scale bar 25μm. These pictures are representative of 2-3 different clones of scramble and *Pcl3* shRNA cells taken at three different time points. **(C)** Expression levels of *Oct4* and *Nanog* in scramble, *Pcl3* shRNA, and *Pcl3* overexpressing cells. **(D)** Quantification of Oct4 and Nanog staining in scramble and *Pcl3* shRNA treated ESCs. +++ indicates bright staining, ++ indicates less bright staining, and + indicates little or no staining as assessed by eye. Graphs are representative of two clones and between 5-10 fields of view at 10x magnification. **(E)** Alkaline phosphatase activity in scramble and *Pcl3* shRNA cells. Graph represents average activity from 3-6 different clones in three experiments assayed in duplicate. **(F)** Quantification of the number of colonies formed per well from scramble and *Pcl3* shRNA cells plated at 100 cells/well in a 6-well plate. Experiment was performed four times in duplicate with two clones each of scramble and *Pcl3* shRNA ESCs. **(G)** Quantification of colonies formed by plating 100 cells/well of wild type and Pcl3 overexpressing cells in a 6-well plate. LIF was reduced to 5% and was performed four times in duplicate. **(H)** Images of teratomas derived from scramble or *Pcl3* shRNA ESCs containing all three germ layers stained with hematoxylin and eosin. Abbreviations: EN-endoderm, NE-neuroectoderm, B-bone, C-cartilage, M-muscle, N-

neural tissue. Scale bar 25μm. Error bars indicate standard deviation. Expression analysis experiments represent 3-4 experiments assayed in quadruplet. For all experiments, asterisk denotes statistical significance of $p < 0.03$. Staining was performed 2-3 times in two or more clones.

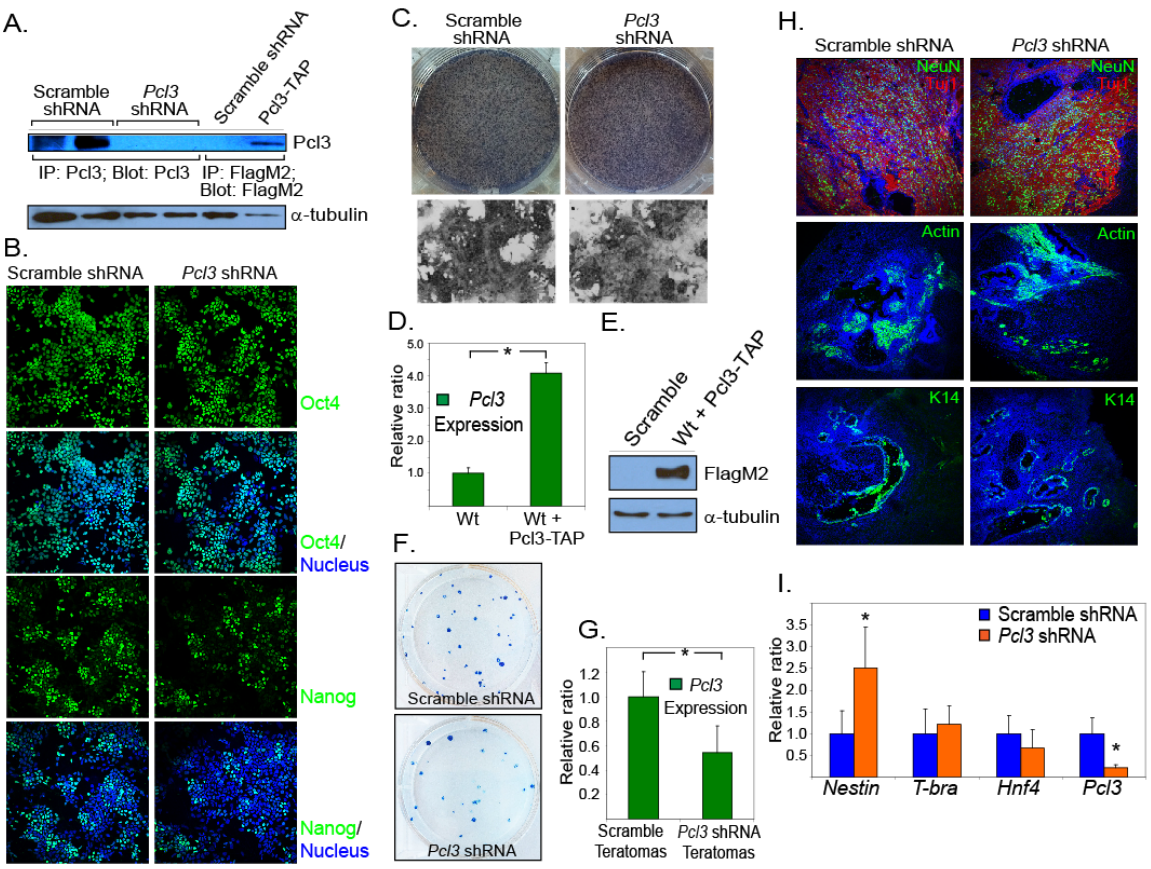


Figure 16. *Depletion of Pcl3 does not affect ESC differentiation to all three germ layers.*

(A) Protein levels of Pcl3 in scramble and *Pcl3* shRNA clones assessed by immunoblot of lysates immunoprecipitated and probed with anti-Pcl3. For a positive control, Pcl3-TAP was immunoprecipitated and probed for FlagM2. Each lane represents a different clone. **(B)** Immunofluorescent staining of scramble and Pcl3-shRNA treated ESCs with Oct4, and Nanog. DAPI marks nuclei. Taken at 20x. **(C)** Scramble and *Pcl3* shRNA ESCs stained for alkaline phosphatase activity. View of one well of a 6-well plate and magnified at 16x. **(D)** *Pcl3* mRNA levels as measured by qRT-PCR in wild type and *Pcl3* overexpressing cells. **(E)** Detection of Pcl3-TAP in wild type cells expressing *Pcl3-TAP* by immunoprecipitating and probing with anti-FlagM2. **(F)** Images of colonies stained with methylene blue formed from scramble and *Pcl3* shRNA cells. Experiment was performed five times with two clones each of scramble and *Pcl3* shRNA ESCs. **(G)** *Pcl3* expression in scramble and *Pcl3*-shRNA derived teratomas. Representative of eight teratomas. **(H)** Teratomas expressing scramble and *Pcl3* shRNA stained for the neural marker NeuN (green) and the neuron marker Tuj1 (red); muscle marker Actin (green); basal layer skin marker K14 (green). All images contain nuclear staining with DAPI (blue). Scale bar 20μm. **(I)** Expression levels of *Nestin*, *T-brachyury* (*T-bra*), *Hnf4*, and *Pcl3* which represent neuroectoderm, mesoderm, endoderm, and knockdown respectively, in scramble and *Pcl3* shRNA expressing EBs. Experiment was performed with 3-6 clones. All immunoblots were performed 2-4 times, and α-tubulin was used as a loading control. All stainings were performed 2-3 times in 2-4 clones or teratoma

samples. Expression analysis represents 3-4 experiments assayed in quadruplicate. Error bars indicate standard deviation and asterisks indicate statistical significance of $p < 0.005$.

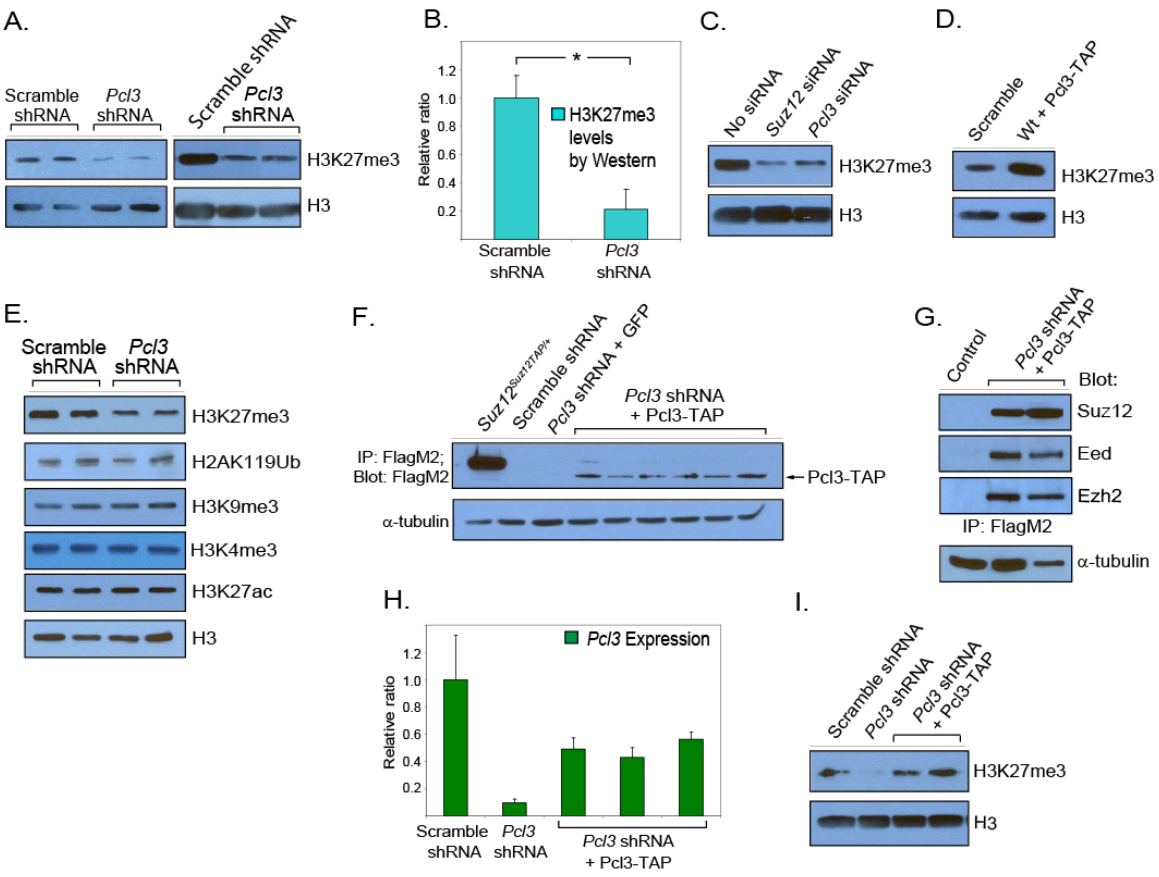


Figure 17. *Pcl3 promotes PRC2 function.*

(A) Immunoblot showing levels of H3K27me3 in multiple clones of scramble and *Pcl3* shRNA ESCs and EBs. **(B)** Quantification of H3K27me3 depletion in *Pcl3* shRNA treated cells. Graph shows an approximate 80% decrease in H3K27me3 and represents ten experiments of 2-6 clones each. **(C)** H3K27me3 levels in *Suz12* and *Pcl3* siRNA treated cells. **(D)** Increased levels of H3K27me3 as measured by immunoblot in cells overexpressing Pcl3. **(E)** Immunoblot of H3K27me3, H2AK119Ub, H3K9me3, H3K4me3, and H3K27ac levels in histones from scramble and *Pcl3* shRNA-expressing cells. **(F)** Pcl3-TAP resistant to *Pcl3* shRNA was reintroduced into *Pcl3* shRNA cells, immunoprecipitated, and detected with anti-FlagM2. *Suz12*$^{Suz12TAP/+}$ cells were used as a positive control. **(G)** Pcl3-TAP binds Suz12, Eed, and Ezh2. Lysates from scramble and *Pcl3* shRNA cells containing Pcl3-TAP were immunoprecipitated with FlagM2 and immunoblotted for Suz12, Eed, and Ezh2. **(H)** qRT-PCR shows partial rescue of *Pcl3* expression in *Pcl3* shRNA clones expressing *Pcl3-TAP*. Error bars indicate standard deviation. Graph represents average expression from 3-6 different clones in three experiments assayed in quadruplet. **(I)** Immunoblot showing restoration of H3K27me3 levels in *Pcl3* shRNA cells transduced with Pcl3-TAP. Histone H3 and α-tubulin were used as loading controls. All westerns and immunoprecipitations were performed three or more times with 2-6 clones.

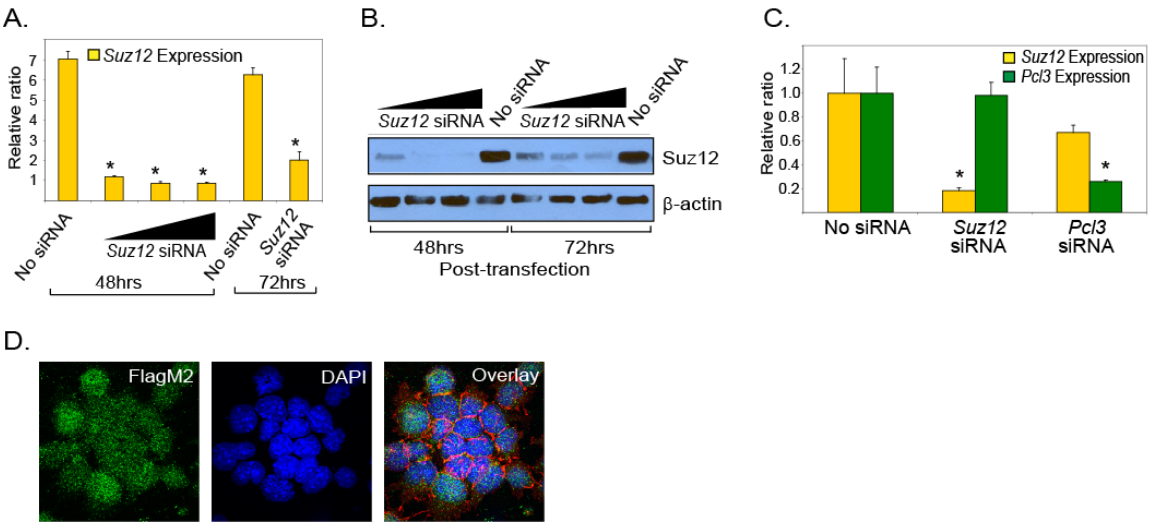


Figure 18. *Depletion of Suz12 and Pcl3.*

(A) qRT-PCR and **(B)** immunoblot indicating levels of Suz12 in cells treated with increasing amounts of *Suz12* siRNA 48hrs and 72hrs post-transfection. β-actin was used as a loading control. **(C)** Transfection with *Suz12* and *Pcl3* siRNAs causes decreased expression of *Suz12* and *Pcl3* respectively as measured by qRT-PCR. **(D)** Pcl3-TAP localizes to the nucleus as assessed by immunofluorescent overlay of FlagM2 (green) and DAPI (blue). E-cadherin (red) marks the cell membrane. Scale bar 10μm. Expression analysis was performed 2-4 times and assayed in quadruplet. Error bars indicate standard deviation, and asterisks indicate statistical significance of $p < 0.005$. Staining was performed twice with three different clones.

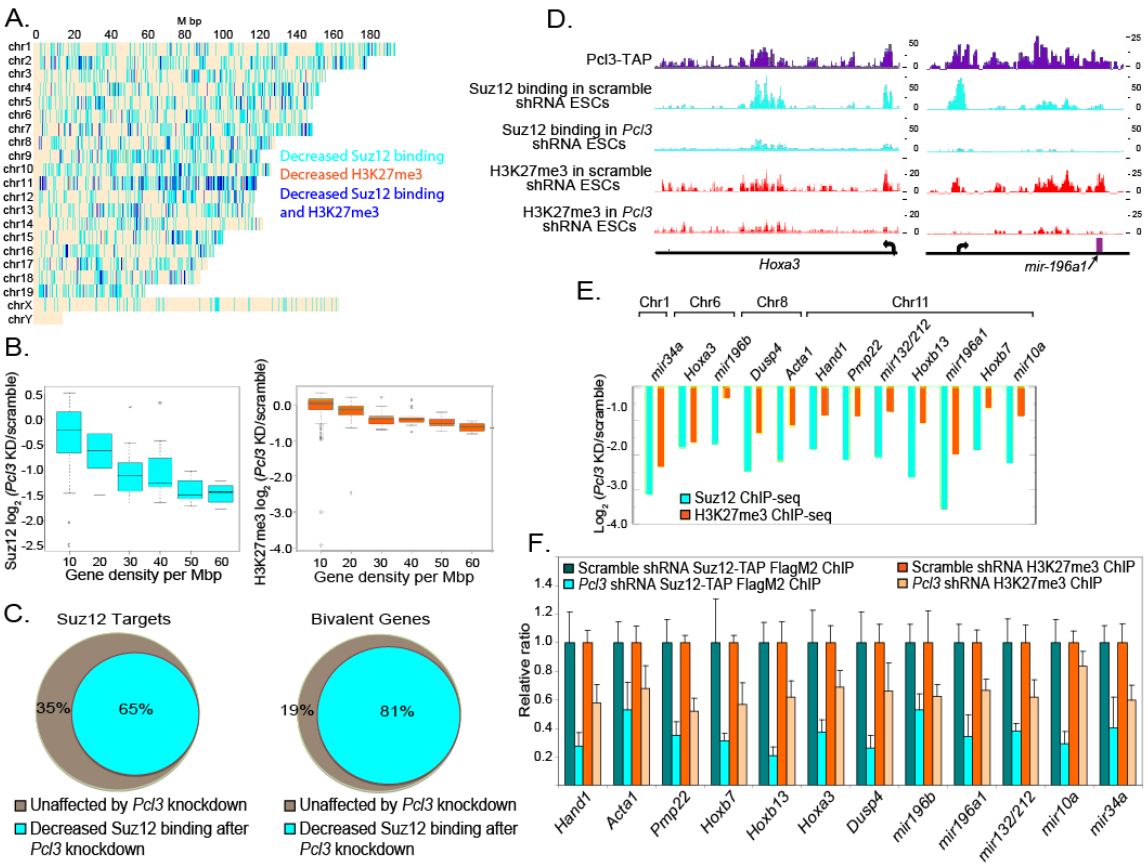


Figure 19. *Pcl3 affects Suz12 binding at a subset of PRC2 targets.*

(A) Graph depicts chromosome-wise distributions of decreased Suz12 and H3K27me3 ChIP-seq reads upon *Pcl3* knockdown (*Pcl3* KD) relative to scramble control. Many regions display both Suz12 and H3K27me3 depletion (blue). Independently decreased Suz12 occupancy and H3K27me3 are marked in cyan and orange, respectively. ChIP-seq peaks were called at a Skellam distribution p-value cutoff of 10^{-7}. Fisher test p-value was 3.2×10^{-41} for the overlap of H3K27me3 with Suz12 sites. **(B)** Reduced Suz12 and H3K27me3 ChIP-seq read density upon *Pcl3* knockdown is correlated with gene density. Pearson correlation is 0.40 for Suz12 and 0.26 for H3K27me3. **(C)** 65% of PRC2 targets show decreased Suz12 binding upon *Pcl3* knockdown, particularly bivalent genes (Fisher test $p=4.9 \times 10^{-250}$). **(D)** Binding of Pcl3 and Suz12, and histone mark profiles of H3K27me3 at the *Hoxa3* gene and in the region surrounding *mir-196a1*. Units are number of reads/250 bp. **(E)** Log fold-changes in Suz12 and H3K27me3 ChIP-seq read density at specific genes in *Pcl*3 knockdown cells compared to scramble control. **(F)** Levels of Suz12 binding and H3K27me3 in *Suz12*$^{Suz12TAP/+}$ cells expressing either scramble or *Pcl3* shRNA assessed by FlagM2 or H3K27me3 ChIP-qRT-PCR. ChIP-qRT-PCR experiments were performed at 3-4 times and assayed in quadruplicate. Error bars indicate standard deviation. Decreases are statistically significant, $p < 0.005$.

A.

bowtie parameters used: -n 2 -y --best -k 2 -m 1 --solexa1.3-quals		
	#Sequenced Reads	#Aligned Reads to MM9
H3K27me3 *Pcl3* shRNA	24433136	14089977
H3K27me3 scramble	21026083	8971449
Suz12 after *Pcl3* shRNA	20524163	12834457
Suz12 scramble	24965312	10403323
Input *Pcl3* shRNA	24611902	13154806
Input scramble	21899220	10441786
Input Pcl3-TAP clone 6	21026979	13706964
Input Pcl3-TAP clone 7	26905744	17961218
Pcl3 ChIP-seq clone 6	18847746	13148869
Pcl3 ChIP-seq clone 7	14145286	9566333

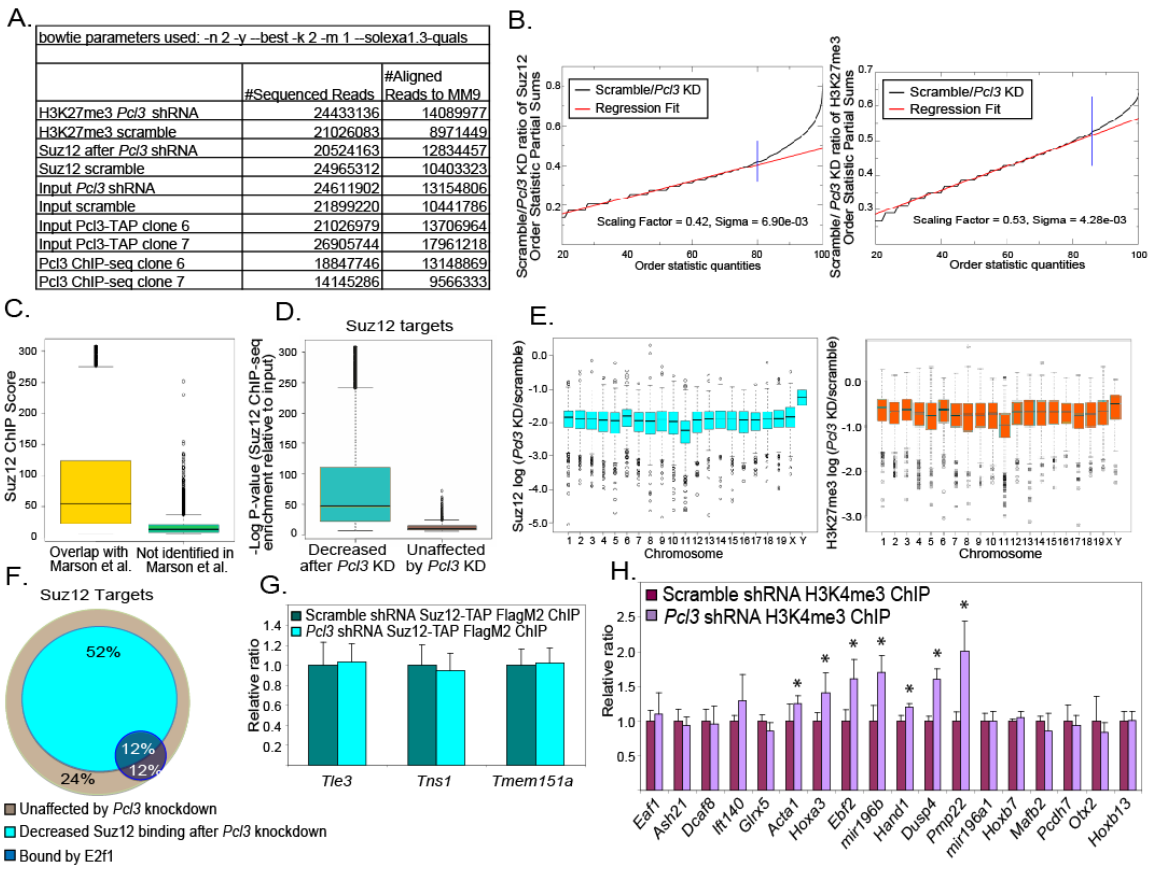


Figure 20. *Pcl3 knockdown causes the most significant depletion of Suz12 and H3K27me3 on chromosome 11.*

(A) Number of sequenced and aligned reads for ChIP-sequencing. **(B)** Partial sums of order statistics of binned read counts in the scramble and *Pcl3* shRNA cells were computed, and their ratios are plotted for the Suz12 FlagM2 ChIP and H3K27me3 ChIP. The theoretical ratio of partial sums is almost linear when two samples are identically distributed, as shown by the red line. The blue vertical bar marks the quantile at which the ratio begins to deviate from linearity, and it effectively separates the background bins from ChIP-enriched bins. The ratio at this quantile was used to scale the Pcl3 ChIP-seq counts. **(C)** Boxplot comparing Suz12 binding sites to Marson et al. [64]. The most significant Suz12 binding sites identified in our dataset overlap with Marson et al. while the less significant binding sites make up the majority of the remaining sites not identified in Marson et al. ChIP score = - log p-value for Suz12 ChIP-seq/Input. **(D)** Graph of –log p-values indicates that the most signficant Suz12 ChIP-seq binding sites are more likely to decrease following *Pcl3* knockdown, while sites unaffected by *Pcl3* knockdown are most often less signficant Suz12 ChIP-seq binding sites. **(E)** Boxplot of log fold-changes in ChIP-seq read density within Suz12 binding sites. Suz12 binding and H3K27me3 in *Pcl3* knockdown ESCs was decreased on all chromosomes. Chromosome 11 was the most significantly depleted (Pair-wise Wilcoxon rank sum test p-value < 1.3×10^{-15} for Suz12 binding sites; p-value < 5.2×10^{-14} for H3K27me3). **(F)** Sites with

decreased Suz12 binding upon *Pcl3* knockdown tend to be devoid of E2f1 [63] (Fisher test p-value = 9.5×10^{-53}). **(G)** Suz12 ChIP-qRT-PCR indicating that at some sites Pcl3 depletion does not affect Suz12 binding. **(H)** ChIP-qRT-PCR for H3K4me3 in scramble and *Pcl3* shRNA ESCs. Error bars indicate standard deviation, and asterisks indicate statistical significance of $p < 0.04$. ChIP-qRT-PCRs was performed 2-3 times and assayed in quadruplicate.

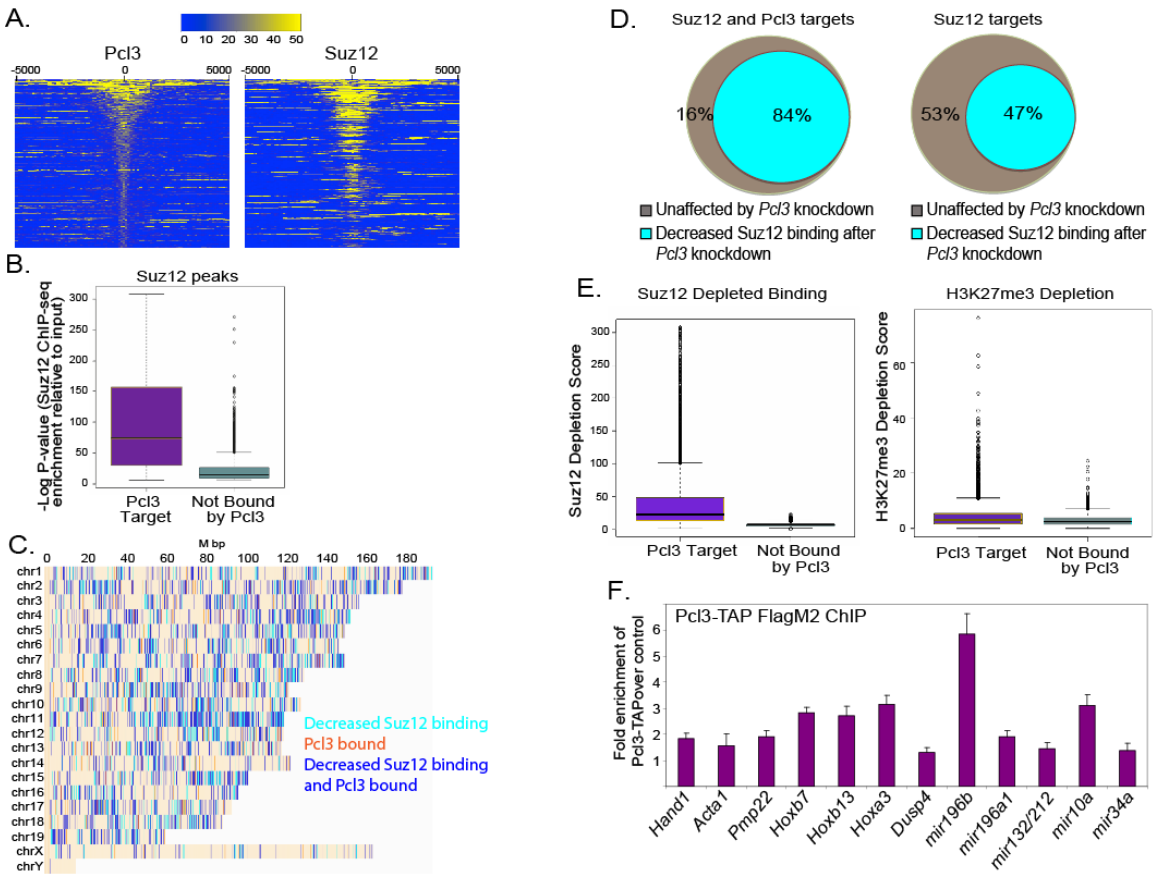


Figure 21. *Pcl3 localizes to PRC2 targets.*

(A) Heatmaps showing Pcl3 and Suz12 ChIP-seq read density in counts per 100bp around Pcl3 peak centers. Approximately 44% of Suz12 targets are bound by Pcl3-TAP. Each row corresponds to a Pcl3 ChIP-seq peak with rows ranked by Pcl3 peak significance (assessed by the Skellam distribution p-values). **(B)** Graph of –log p-values indicating that the most significant Suz12 ChIP-seq peaks overlap with Pcl3, while regions containing less significant Suz12 ChIP-seq peaks are not bound by Pcl3. Wilcoxon p-value < $1e^{-306}$ for the difference in the distribution of ChIP-seq –log p-values. **(C)** Pcl3 co-localizes with Suz12 depletion sites (blue). Regions containing only Suz12 depletion or only Pcl3 binding are indicated in cyan and orange respectively. **(D)** 84% of sites bound by Suz12 and Pcl3 show decreased Suz12 binding upon *Pcl3* (Fisher test p-value = $1.5x10^{-306}$). Two binding sites were considered to be overlapping if their peak centers were within 2kb from each other. **(E)** Suz12 (Wilcoxon p-value < $1e^{-306}$) and H3K27me3 (Wilcoxon p-value = $3.7e^{-36}$) co-localizing with Pcl3 show much more significant depletion compared to Suz12 and H3K27me3 targets not bound by Pcl3. Depletion score = -log p-value (read counts before/after Pcl3 KD). **(F)** Genes and microRNAs bound by Pcl3-TAP measured by FlagM2 ChIP-qRT-PCR. The graph depicts average fold enrichment over control levels for three different clones in three different experiments assayed in quadruplet. Error bars indicate standard deviation. All increases are statistically significant, $p < 0.04$.

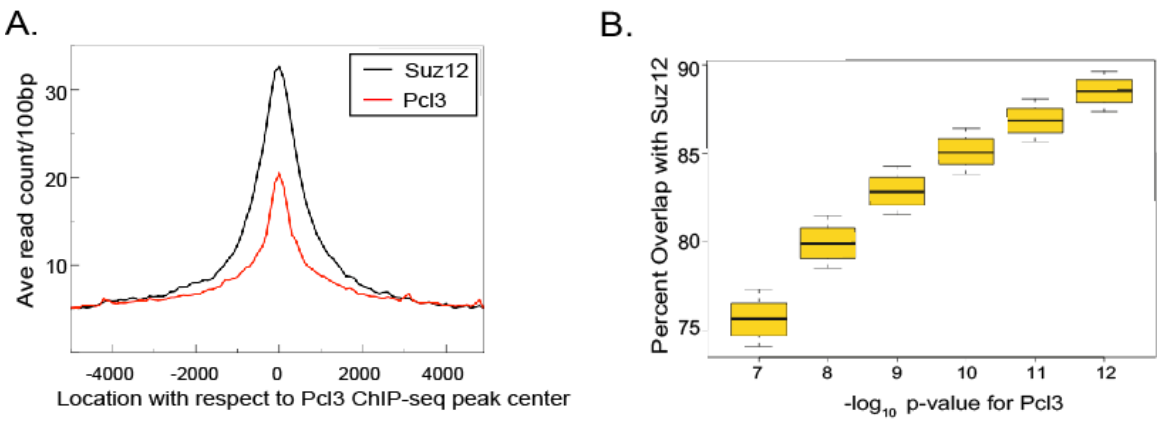


Figure 22. *Pcl3 co-localizes with Suz12.*

(A) Aligning Suz12 and Pcl3 ChIP-seq reads at Pcl3 peak centers shows genome-wide co-localization of Pcl3 with Suz12. **(B)** To estimate the percentage overlap between Suz12 and Pcl3 binding sites, a sensitivity analysis was performed by varying the Skellam distribution p-value cutoff for calling peaks, ranging between 10^{-7} to 10^{-12}. The boxplot shows the percentage of Pcl3 peaks at each p-value cutoff found to be overlapping with Suz12 binding sites that pass the p-value cutoffs 10^{-7}, 10^{-8}, 10^{-9}, 10^{-10}, 10^{-11}, and 10^{-12}.

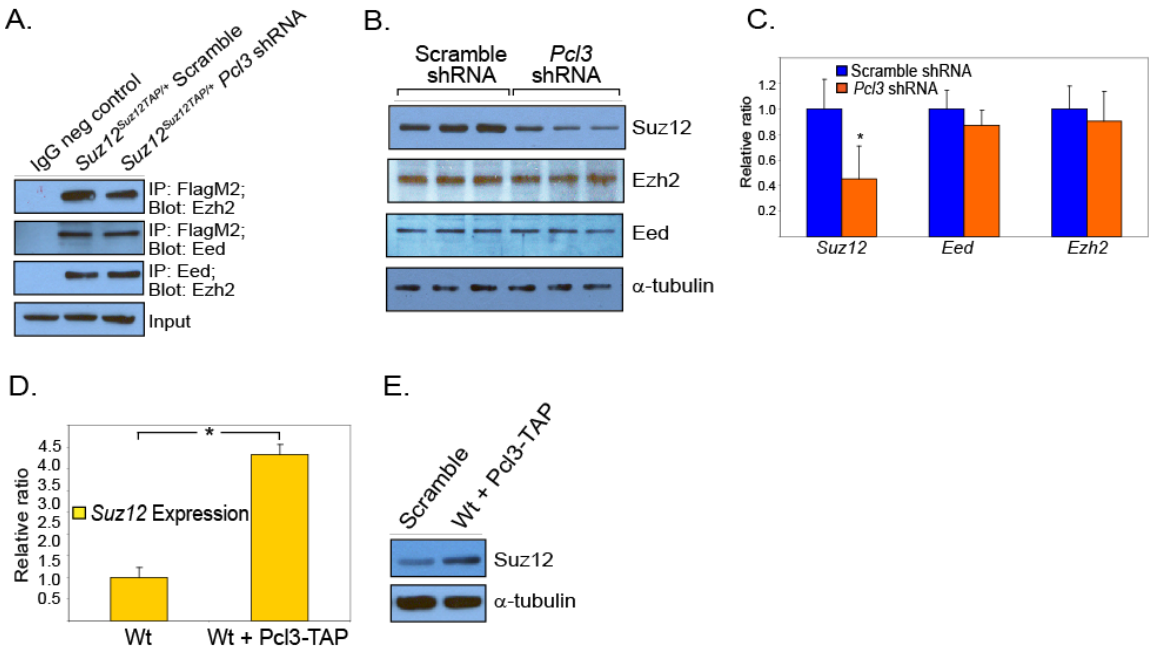
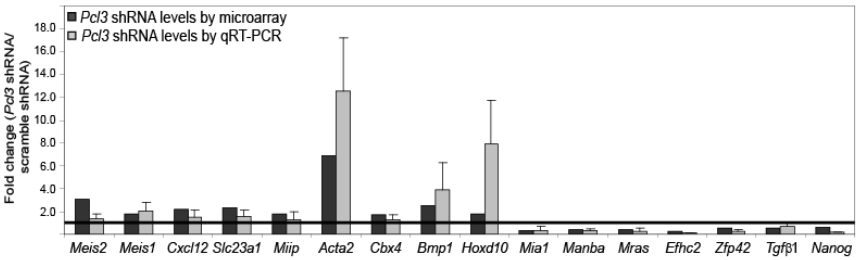


Figure 23. *Pcl3 regulates Suz12, but not complex stability*

(A) PRC2 components associate in the absence of Pcl3 based on co-immunoprecipitation. β-actin used as a loading control. **(B)** Protein levels of Suz12, Ezh2, and Eed in scramble and *Pcl3* shRNA clones as measured by immunoblot. Experiments were performed with 3-6 scramble and *Pcl3* shRNA clones each. **(C)** Expression levels of *Suz12*, *Ezh2*, and *Eed* in scramble and *Pcl3* shRNA measured by qRT-PCR. **(D)** *Suz12* mRNA levels measured by qRT-PCR in wild type and *Pcl3* overexpressing cells. **(E)** Suz12 protein levels in wild type cells and wild type cells overexpressing *Pcl3-TAP*. All immunoblot and co-immunoprecipitation experiments were performed 2-5 times. All expression analysis represents three experiments assayed in quadruplicate. Error bars represent standard deviation, and asterisk indicates statistical significance of $p < 0.005$.

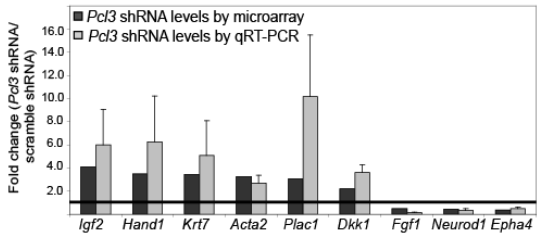
A.



B.

Depleted Suz12 binding and changed gene expression	Fold change Pcl3 KD /scramble
Sp8	5.608
Hand1	3.461
Mrg1 (Meis2)	3.106
Irx1	2.892
Itga11	2.7
Pmp22	2.528
Bmp1	2.525
Fgf5	2.445
Acta1	2.343
Cxcl12	2.201
Lrp2	1.98
Pdzrn3	1.928
E330016A19Rik	1.901
Meis1	1.821
Hoxd10	1.82
Cbx4	1.799
Usp25	1.621
Igf2r	1.577
Igf2bp2	1.497
Tgfb1	0.549
Mras	0.42

C.



D.

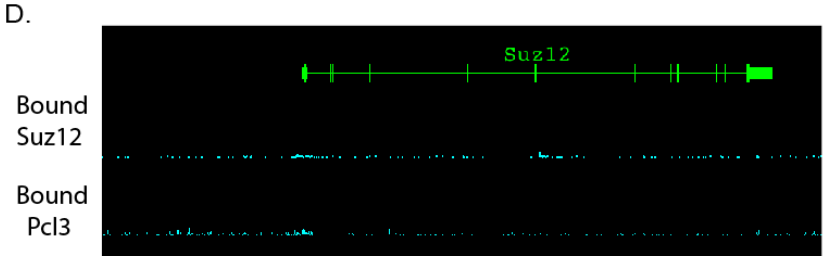


Figure 24. *Pcl3 misregulates a subset of genes.*

(A, C) Relative expression of genes that either increased or decreased upon *Pcl3* knockdown as measured by qRT-PCR and microarray. Cells were collected following selection for Pcl3 depletion, approximately 3-4 weeks. Microarray and qRT-PCR were performed with 2-6 clones each for control and *Pcl3* knockdown ESCs. qRT-PCR analysis represents 2-3 different experiments performed with multiple clones and assayed in quadruplicate. For both graphs, error bars indicate standard deviation and data represents statistical significance of $p < 0.05$. **(A)** Graph represents expression levels in ESCs that have not been pre-plated to remove differentiated cells. **(B)** Genes that show depleted Suz12 binding following *Pcl3* knockdown and that are misregulated by microarray analysis. Cells in **(C)** were pre-plated to remove differentiated cells and then expression was assessed. **(D)** Binding profile of Suz12 and Pcl3 at the *Suz12* locus. Turquoise puncta are background and not statistically significant.

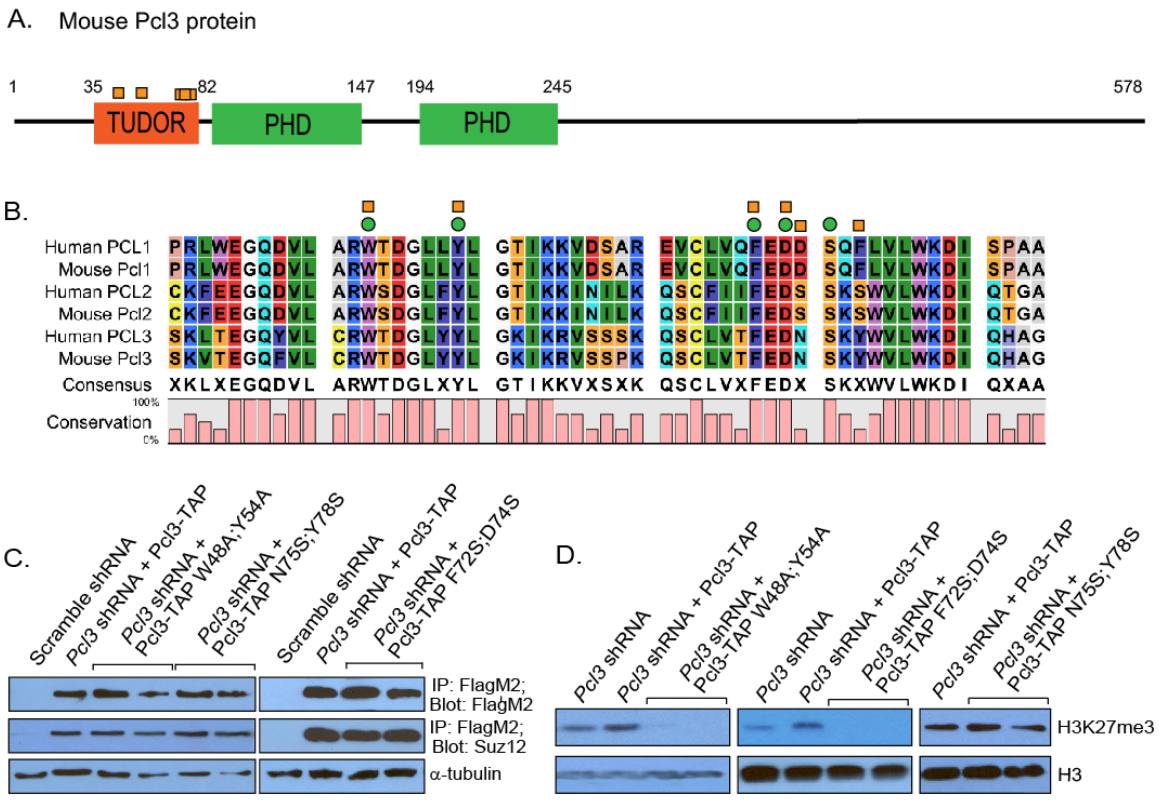


Figure 25. *Pcl3 requires Tudor domain residues for function.*

(A) Schematic of Pcl3 architecture including Tudor and PHD domains. Size of domains are indicated by amino acid number, but schematic is not to scale. Orange boxes indicate mutated residues. **(B)** Sequence alignment of the Tudor domain of human and mouse Polycomb-like homologues. Bottom graph indicates the degree of conservation at each residue. Green circles indicate putative histone binding sites. Orange squares indicate mutated residues. **(C)** Mutant Pcl3-TAP protein is detected at similar levels to wild type Pcl3-TAP by immunoprecipitating and probing with anti-FlagM2. Mutant Pcl3-TAP can bind Suz12 as well as wild type Pcl3-TAP by co-immunoprecipitation. **(D)** Pcl3-TAP W48A;Y48A and Pcl3-TAP F72S;D74S do not support H3K27me3 whereas Pcl3-TAP N75S;Y78S does. Immunoblot displaying H3K27me3 levels in histones purified from cells expressing *Pcl3* shRNA, *Pcl3* shRNA with Pcl3-TAP, and *Pcl3* shRNA with mutant Pcl3-TAP. Histone H3 and α-tubulin were used as loading controls. Each lane represents a different clone. All westerns and immunoprecipitations were performed 3-4 times with at least two clones from each mutant.

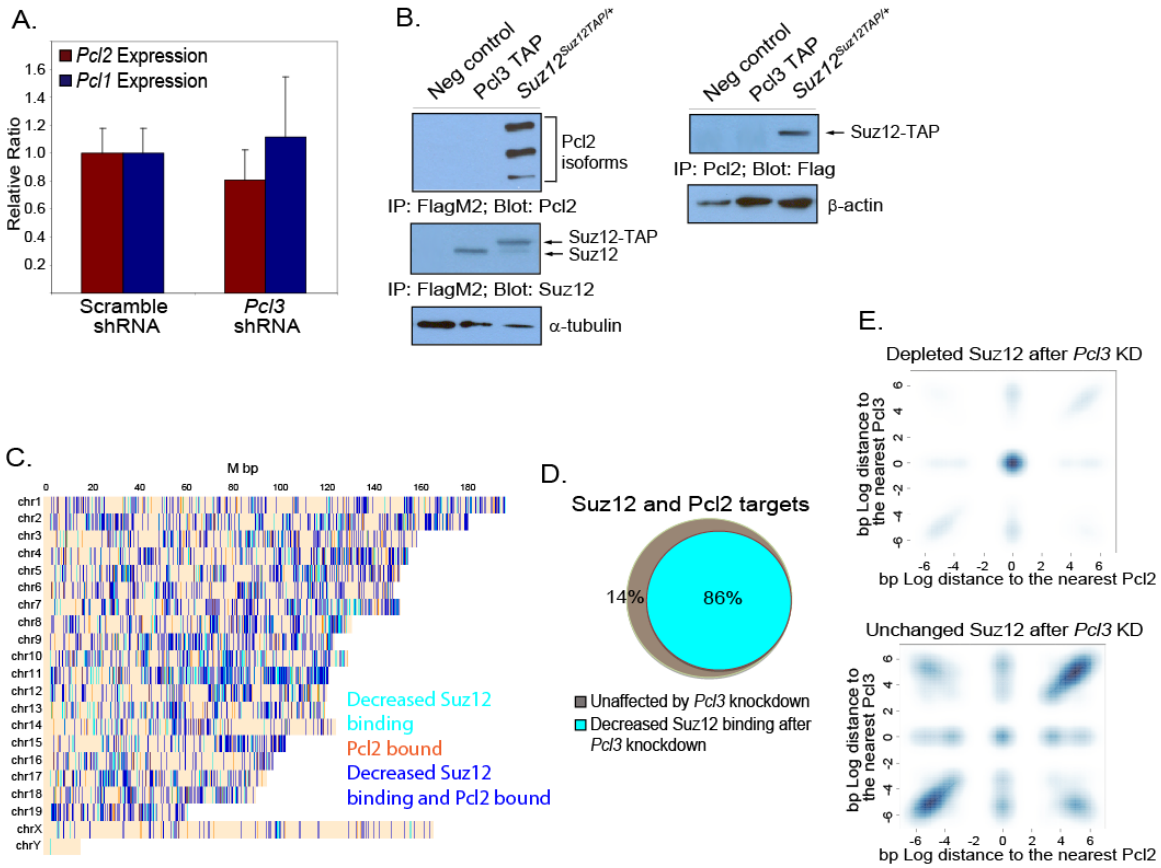


Figure 26. *Pcl3 and Pcl2 participate in separate PRC2 complexes but overlap at some PRC2 binding sites.*

(A) *Pcl2* and *Pcl1* mRNA levels are not significantly different between scramble and *Pcl3* shRNA cells. Graph represents average expression from three different experiments in 2-6 different clones and assayed in quadruplet. Error bars indicate standard deviation. **(B)** Co-immunoprecipitation showing that immunoprecipitated Suz12-TAP associates with Pcl2 whereas immunoprecipitated Pcl3-TAP does not. Pcl3-TAP does bind Suz12 as seen in the lower blot. α-tubulin was used as a loading control. Reciprocal co-immmunoprecipitations showing a Pcl2-specific association with Suz12-TAP but not an association between Pcl2 and Pcl3-TAP. β-actin was used as a loading control. Each blot is representative of three different experiments. **(C)** Graph depicts chromosome-wise distributions of Pcl2 co-localization with sites depleted of Suz12 following *Pcl3* knockdown (blue). Regions containing only Suz12 depletion and only Pcl2 binding are indicated in cyan and orange respectively. **(D)** Decreased Suz12 binding upon *Pcl3* knockdown occurs at 86% of targets bound by Suz12 and Pcl2 (Fisher test p-value < $p<10^{-300}$). Of sites bound by Pcl2 and Pcl3, 95% show decreased Suz12 binding upon *Pcl3* knockdown. **(E)** Pcl2 and Pcl3 binding sites co-localize with regions of Suz12 depletion upon *Pcl3* knockdown (top), whereas Pcl2 and Pcl3 binding sites reside far from unaffected Suz12 binding sites (bottom). The figure shows the joint probability density of the nearest Pcl2 and Pcl3 locations relative to Suz12. Darker color represents

higher probability density. Bivariate normal kernel density estimates were obtained by using the smoothScatter function in R. The distance coordinates were transformed to base pairs sign (Pcl2−Suz12) $\log_{10}$ |Suz12−Pcl2|, and similarly for Pcl3.

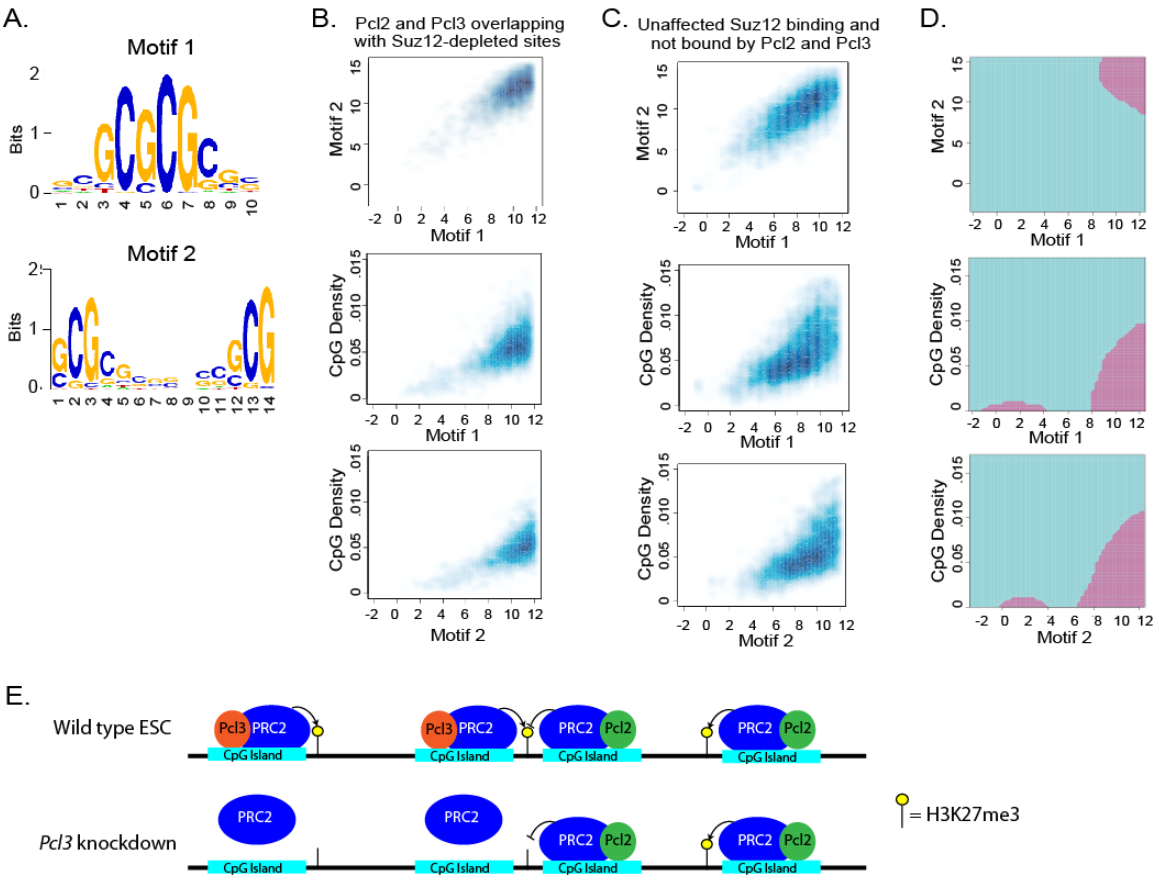


Figure 27. *Pcl2 and Pcl3 localize to CpG islands.*

(A) The 500bp central regions of Pcl3 ChIP-seq peaks were scanned for enriched motifs by using a 9th order Markov background dependence model [86]. Two examples of 10- and 14-mer enriched motifs are shown. **(B-C)** Smoothed scatter plots of maximum position specific-scoring matrix (PSSM) scores for the two motifs and CpG density are shown for **(B)** Suz12 binding sites depleted upon *Pcl3* knockdown overlapping with Pcl2 and Pcl3 and **(C)** Suz12 binding sites unaffected upon *Pcl3* knockdown and that do not overlap with Pcl2 and Pcl3. **(D)** Shown are the decision boundaries of a support vector machine classifier using these three features, where the purple regions correspond to Suz12 co-localizing with Pcl2 and Pcl3. The predictor had a cross validation accuracy of 75%. **(E)** A model of Pcl3 and Pcl2 regulation of PRC2 binding and activity. In wild type ESCs, Pcl3 promotes PRC2 binding and H3K27me3. Pcl2 antagonizes Pcl3-mediated Suz12 binding at sites bound by both but promotes PRC2 function at sites solely regulated by Pcl2. Knockdown of *Pcl3* causes decreased PRC2 binding and H3K27me3. Pcl2 does not compensate at Pcl2 and Pcl3 targets and continues to inhibit or promote PRC2 function depending on the gene.

Table 1. Gene expression changes in *Pcl3* shRNA ESCs by microarray

ProbeID	Short Name	Long Name	Fold change *Pcl3*/scramble
A_52_P282058	Col8a1	procollagen, type VIII, alpha 1	6.954
A_52_P420504	Acta2	actin, alpha 2, smooth muscle, aorta	6.894
A_51_P515605	Sp8	procollagen, type III, alpha 1	5.608
A_51_P180747	Ctla2a (Ctla2a| Ctla2b)	cytotoxic T lymphocyte-associated protein 2 alpha	3.249
A_52_P277641	Mrg1 (Meis2)	myeloid ecotropic viral integration site-related gene 1	3.106
A_52_P1092823	Irx1	Iroquois related homeobox 1 (Drosophila)	2.892
A_51_P417720	Itga11	integrin, alpha 11	2.7
A_52_P299921	Bmp1	bone morphogenetic protein 1	2.525
A_52_P569375	Fgf5	fibroblast growth factor 5	2.445
A_51_P140751	Slc23a1	solute carrier family 23 (nucleobase transporters), member 1	2.385
A_51_P342549	Hand1	heart and neural crest derivatives expressed transcript 1	2.244
A_51_P172502	Cxcl12	chemokine (C-X-C motif) ligand 12	2.201
A_52_P285470	Lrp2	low density lipoprotein receptor-related protein 2	1.98
A_52_P49321	Adamts9	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 9	1.978
A_52_P327176	Pdzrn3	PDZ domain containing RING finger 3	1.928
A_51_P289647	E330016A19Rik	C6orf150	1.901

A_52_P26195	Meis1	myeloid ecotropic viral integration site 1	1.821
A_52_P407049	Hoxd10	homeo box D10	1.82
A_51_P133582	Cbx4	chromobox homolog 4 (Drosophila Pc class)	1.799
A_51_P248609	D4Wsu114e	Miip; migration and invasion inhibitory protein	1.777
A_51_P384515	Indol1	indoleamine-pyrrole 2,3 dioxygenase-like 1	1.771
A_51_P130203	Pbx1	pre B-cell leukemia transcription factor 1	1.696
A_51_P351384	Upk2	uroplakin 2	1.681
A_51_P176377	Usp25	ubiquitin specific peptidase 25	1.621
A_51_P119429	Nckap1l	NCK associated protein 1 like	1.615
A_51_P125467	Igf2r	insulin-like growth factor 2 receptor	1.577
A_51_P510441	Suv39h1	suppressor of variegation 3-9 homolog 1 (Drosophila)	1.548
A_51_P315682	Igf2bp2	insulin-like growth factor 2 mRNA binding protein 2	1.497
A_51_P302336	Igf2bp3	insulin-like growth factor 2 mRNA binding protein 3	1.481
A_51_P340255	Nbea	neurobeachin	1.422
A_52_P292632	Wdr20b	WD repeat domain 20b	0.689
A_51_P511707	2310079N02Rik	Nucleoside-triphosphatase C1orf57 (NTPase)(EC 3.6.1.15)(Nucleoside triphosphate phosphohydrolase)	0.683
A_51_P421780	Myo18b*	myosin XVIIIb	0.682
A_51_P389017	Tulp2	tubby-like protein 2	0.654
A_52_P531651	Parvg	parvin, gamma	0.648
A_51_P150653	Ptprv*	OST-PTP; protein tyrosine phosphatase, receptor type, V	0.646
A_51_P294233	Nanog	Nanog homeobox	0.642

A_51_P189442	Adh4	alcohol dehydrogenase 4 (class II), pi polypeptide	0.631
A_51_P480311	F2	coagulation factor II	0.627
A_51_P431823	Hmha1	histocompatibility (minor) HA-1	0.621
A_52_P452689	Atf3	activating transcription factor 3	0.619
A_51_P254262	Nxph3	neurexophilin 3	0.617
A_51_P430620	Slc25a12	solute carrier family 25 (mitochondrial carrier, Aralar), member 12	0.616
A_51_P183894	Fbxo15	F-box protein 15	0.611
A_51_P151070	ENSMUSG00000055440	predicted gene, ENSMUSG00000055440	0.607
A_51_P517105	Aoah	acyloxyacyl hydrolase	0.604
A_51_P295034	Klk1b4 (Klk1b5)	kallikrein 1-related pepidase b4	0.6
A_51_P264495	Pgam2	phosphoglycerate mutase 2	0.596
A_51_P204831	Crip1	cysteine-rich protein 1 (intestinal)	0.593
A_52_P73475	A130092J06Rik	FAM78A	0.591
A_51_P139651	Nos3	nitric oxide synthase 3, endothelial cell	0.586
A_51_P454309	BC023179	zinc finger protein 772	0.583
A_52_P413394	Ppef2	protein phosphatase, EF hand calcium-binding domain 2	0.583
A_52_P629849	Zfp42	zinc finger protein 42; Rex1	0.579
A_52_P361323	Spic	Spi-C transcription factor (Spi-1/PU.1 related)	0.573
A_51_P413130	Myl4	myosin, light polypeptide 4	0.573
A_52_P108850	St8sia1	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 1	0.569
A_51_P519251	Nupr1	nuclear protein 1	0.569
A_52_P176659	V00829	Klk1b1; kallikrein 1-related peptidase b1	0.559

A_51_P338072	Myh4	myosin, heavy polypeptide 4, skeletal muscle	0.557
A_52_P115565	4930500J02Rik*	RIKEN cDNA 4930500J02 gene	0.555
A_51_P334072	Cyct	cytochrome c, testis	0.552
A_51_P201709	Ggtla1 (Ggt5)	gamma-glutamyltransferase-like activity 1	0.551
A_51_P390715	Tgfb1	transforming growth factor, beta 1	0.549
A_51_P386142	Cabp1	calcium binding protein 1	0.548
A_51_P512293	EG665798*	Aurkc; aurora kinase C	0.547
A_52_P519865	Centa1 (Adap1)	Acap1; centaurin, beta 1; ArfGAP with coiled-coil, ankyrin repeat and PH domains 1	0.539
A_51_P206445	2810409K11Rik	zinc finger protein 773	0.537
A_51_P157406	Ptgds	prostaglandin D2 synthase (brain)	0.536
A_51_P485731	Sytl1	synaptotagmin-like 1	0.536
A_51_P183025	Oas1d	2'-5' oligoadenylate synthetase 1D	0.534
A_51_P471458	Sult5a1	sulfotransferase family 5A, member 1	0.532
A_51_P328889	Tekt1	tektin 1	0.523
A_52_P313217	Gpr133	G protein-coupled receptor 133	0.522
A_51_P241577	Spats1	spermatogenesis associated, serine-rich 1	0.521
A_51_P352357	Ldhal6b	lactate dehydrogenase A-like 6B	0.519
A_52_P390685	OTTMUSG00000001246	Gm11487	0.518
A_51_P136441	Nudt12	nudix (nucleoside diphosphate linked moiety X)-type motif 12	0.512
A_51_P262515	Phf11	PHD finger protein 11	0.511
A_51_P302576	Spn	sialophorin	0.501

A_51_P423880	Smarcd3	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 3	0.5
A_51_P438278	C130038G02Rik	Mtus2; microtubule associated tumor suppressor candidate 2	0.487
A_51_P497875	Hormad1	HORMA domain containing 1	0.486
A_51_P408172	1110001D15Rik	dual von Willebrand factor A domains; collagen, type VI, alpha 6	0.484
A_51_P501858	Plekha4	pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 4	0.481
A_51_P384230	Sgk2	serum/glucocorticoid regulated kinase 2	0.479
A_52_P310548	Nlrp4a	NLR family, pyrin domain containing 4A	0.477
A_51_P121891	Rac2	RAS-related C3 botulinum substrate 2	0.473
A_52_P30803	TC1610785	Gm973	0.472
A_52_P260069	CN716893	melanoma antigen family B	0.47
A_52_P350344	Timm8a2	translocase of inner mitochondrial membrane 8 homolog a2 (yeast)	0.469
A_51_P245414	Klk1	kallikrein 1	0.468
A_52_P458319	Centb1 (Acap1)	Acap1; centaurin, beta 1; ArfGAP with coiled-coil, ankyrin repeat and PH domains 1	0.468
A_51_P451377	Kng1	kininogen 1	0.464
A_51_P316311	Ppfia4	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 4	0.459
A_51_P382970	Itga9	integrin alpha 9	0.459

A_51_P255875	Padi4	peptidyl arginine deiminase, type IV	0.459
A_51_P147766	Cpne9	copine family member IX	0.458
A_51_P403814	C330016K18Rik (Slc4a5)	Slc4a5; solute carrier family 4, sodium bicarbonate cotransporter, member 5	0.457
A_51_P141926	Fxyd4	FXYD domain-containing ion transport regulator 4	0.456
A_51_P130095	Fcgr2b	Fc receptor, IgG, low affinity IIb	0.446
A_51_P450469	Pnoc	prepronociceptin	0.441
A_51_P410576	Ccdc113	coiled-coil domain containing 113	0.44
A_52_P154741	Lgals6	Galectin-6; lectin, galactose binding, soluble 6	0.435
A_51_P114094	Clstn3	calsyntenin 3	0.432
A_51_P497741	4930434E21Rik	Wdr49; WD repeat domain 49	0.431
A_51_P140690	Stmn3	stathmin-like 3	0.431
A_51_P519696	Mras	muscle and microspikes RAS	0.42
A_52_P274960	Arhgap30	Rho GTPase activating protein 30	0.416
A_52_P602669	OTTMUSG00000000712	Serpinb6; serpin peptidase inhibitor, clade B (ovalbumin), member 6	0.412
A_51_P228604	Ceacam20	CEA-related cell adhesion molecule 20	0.41
A_51_P166434	Atp6v1e1	VATPase, H+ transporting, lysosomal V1 subunit E1	0.41
A_52_P584154	Manba	mannosidase, beta A, lysosomal	0.407
A_51_P186476	Slc11a1	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 1	0.394
A_51_P281778	2210010C17Rik	RIKEN cDNA 2210010C17 gene	0.394

A_51_P407494	Lefty2	Left-right determination factor 2	0.382
A_52_P508700	C230055K05Rik	RIKEN cDNA C230055K05 gene	0.382
A_51_P480861	Pga5	pepsinogen 5, group I	0.367
A_51_P282227	As3mt	arsenic (+3 oxidation state) methyltransferase	0.361
A_52_P305230	Igsf21	immunoglobin superfamily, member 21	0.352
A_52_P300533	2310043M15Rik	dopey family member 2	0.344
A_51_P443403	Smoc1	SPARC related modular calcium binding 1	0.343
A_51_P333253	Myo1g	myosin IG	0.334
A_51_P258359	Hus1b	Hus1 homolog b (S. pombe)	0.332
A_52_P174346	Mia1	melanoma inhibitory activity 1	0.328
A_51_P278464	3100002J23Rik	FAM183A	0.319
A_51_P182257	1700019N12Rik	C11orf20	0.311
A_52_P127184	Efhc2	EF-hand domain (C-terminal) containing 2	0.292
A_51_P204442	Phf19	PHD finger protein 19	0.279
A_52_P204331	D630039A03Rik	C9orf152	0.245
A_51_P377094	Col1a1	procollagen, type I, alpha 1	0.228

Table 1. *Gene expression changes upon Pcl3 knockdown by microarray.*

Microarray data indicating fold changes in gene expression in ESCs expressing *Pcl3* shRNA as compared to scramble shRNA ESCs. Comparison included six control samples and six *Pcl3* shRNA clones.

Table 2. Gene expression changes in pre-plated *Pcl*3 shRNA ESCs by microarray

ProbeID	Short Name	Long Name	Fold change-Pcl3/scramble
A_51_P507471	2510022D24Rik*	RIKEN cDNA 2510022D24 gene	23.67
A_51_P516826	Igf2	insulin-like growth factor 2	4.18
A_51_P516833	Igf2	insulin-like growth factor 2	3.99
A_52_P572808	Agpat9	RIKEN cDNA A230097K15 gene	3.86
A_51_P312348	Krt7	keratin 7	3.80
A_51_P306731	Agpat9	RIKEN cDNA A230097K15 gene	3.66
A_51_P243514	4732474O15Rik	RIKEN cDNA 4732474O15 gene	3.61
A_51_P342549	Hand1	heart and neural crest derivatives expressed transcript 1	3.46
A_52_P229943	Ostb	organic solute transporter beta	3.37
A_52_P420504	Acta2	actin, alpha 2, smooth muscle, aorta	3.29
A_52_P210078	Acta2*	actin, alpha 2, smooth muscle, aorta	3.21
A_52_P163750	3830417A13Rik	RIKEN cDNA 3830417A13 gene	3.16
A_52_P410685	Krt7	keratin 7	3.10
A_51_P255238	Plac1	placental specific protein 1	3.04
A_51_P371867	3830417A13Rik	RIKEN cDNA 3830417A13 gene	2.96
A_52_P83479	Ifi204	interferon activated gene 204	2.95
A_51_P251687	Serpinb9g	serine (or cysteine) peptidase inhibitor, clade B, member 9g	2.82
A_52_P648524	Dio3*	deiodinase, iodothyronine type III	2.79
A_51_P219424	Serpinb9d	serine (or cysteine) peptidase inhibitor, clade B, member 9d	2.75
A_52_P413940	Phc3*	polyhomeotic-like 3 (Drosophila)	2.73
A_51_P353934	Nkx2-4	NK2 transcription factor related, locus 4 (Drosophila)	2.69
A_52_P867738	AK042489		2.68

A_52_P255724	1810009N02Rik*	RIKEN cDNA 1810009N02 gene	2.64
A_51_P334104	Dcn	decorin	2.59
A_51_P258529	Pmp22	peripheral myelin protein	2.53
A_52_P502226	2810429I04Rik*	RIKEN cDNA 2810429I04 gene	2.47
A_52_P489295	Adamts1	a disintegrin-like and metallopeptidase with thrombospondin type 1 motif, 1	2.44
A_51_P140237	Fhl2	four and a half LIM domains 2	2.40
A_51_P309307	(Bcl2l15)	BCLl2-like 15 Gene [Source:MGI (curated);Acc:Bcl2l15-002]	2.39
A_51_P424532	Vnn1	vanin 1	2.36
A_51_P246854	Acta1	actin, alpha 1, skeletal muscle	2.34
A_51_P248609	D4Wsu114e	DNA segment, Chr 4, Wayne State University 114, expressed	2.34
A_51_P517904	Nxt2	nuclear transport factor 2-like export factor 2	2.26
A_52_P536869	Per2	period homolog 2 (Drosophila)	2.23
A_52_P648433	Tex13	testis expressed gene 13	2.22
A_51_P410703	Prss8	protease, serine, 8 (prostasin)	2.21
A_51_P379069	Dkk1	dickkopf homolog 1 (Xenopus laevis)	2.19
A_52_P224760	Ptgs2	prostaglandin-endoperoxide synthase 2	2.17
A_52_P165654	Star	steroidogenic acute regulatory protein	2.12
A_51_P290826	Calca	calcitonin/calcitonin-related polypeptide, alpha	2.12
A_52_P397909	2810429I04Rik*	RIKEN cDNA 2810429I04 gene	2.09
A_51_P254855	Ptgs2	prostaglandin-endoperoxide synthase 2	2.08
A_52_P778761	Vgll3*	vestigial like 3 (Drosophila)	2.08
A_51_P257938	Tacstd2	tumor-associated calcium signal transducer 2	2.03
A_52_P141628	Slc23a1	solute carrier family 23 (nucleobase transporters), member 1	2.00

A_52_P93467	Sdc4	syndecan 4	2.00
A_51_P426195	Nppb	natriuretic peptide precursor type B	1.99
A_51_P450248	Esx1	extraembryonic, spermatogenesis, homeobox 1	1.97
A_52_P46447	Crim1*	cysteine rich transmembrane BMP regulator 1 (chordin like)	1.95
A_51_P383991	Sept4	septin 4	1.95
A_51_P150710	Igj	immunoglobulin joining chain	1.93
A_51_P491329	Pdzk1ip1	PDZK1 interacting protein 1	1.93
A_52_P360259	Zfp451*	zinc finger protein 451	1.93
A_52_P333529	Slco2a1	solute carrier organic anion transporter family, member 2a1	1.92
A_52_P113830	1600025M17Rik*	RIKEN cDNA 1600025M17 gene	1.92
A_51_P478098	Epb4.1l5*	erythrocyte protein band 4.1-like 5	1.91
A_52_P382886	Gjb2	gap junction membrane channel protein beta 2	1.91
A_51_P289249	(Mpzl3)	myelin protein zero-like 3 Gene [Source:MGI (curated);Acc:Mpzl3-002]	1.90
A_52_P106482	Cryab	crystallin, alpha B	1.88
A_52_P593110	Ptbp2*	polypyrimidine tract binding protein 2	1.88
A_52_P195839	Ctsc	cathepsin C	1.87
A_51_P253019	BC010981*	cDNA sequence BC010981	1.87
A_51_P333518	Sh3rf2	SH3 domain containing ring finger 2	1.84
A_52_P348189	Krtcap3	keratinocyte associated protein 3	1.84
A_51_P122707	Ldoc1	leucine zipper, down-regulated in cancer 1	1.83
A_52_P322421	Eva1 (Mpzl2)	epithelial V-like antigen 1	1.82
A_52_P282838	Spin2	spindlin family, member 2	1.82
A_52_P867545	LOC674367*	region containing RIKEN cDNA C630010D07 gene; RIKEN cDNA 1110019K23 gene	1.81
A_51_P312268	1810062O18Rik*	RIKEN cDNA 1810062O18 gene	1.81
A_52_P432949	Erbb3*	v-erb-b2 erythroblastic leukemia viral oncogene	1.81

		homolog 3 (avian)	
A_51_P491350	Col4a2	procollagen, type IV, alpha 2	1.79
A_51_P506236	Clcn5	chloride channel 5	1.78
A_52_P378827	2810429I04Rik*	RIKEN cDNA 2810429I04 gene	1.78
A_51_P112405	Plaur	plasminogen activator, urokinase receptor	1.77
A_51_P428372	Ppbp	pro-platelet basic protein	1.76
A_51_P487298	Vasn*	vasorin	1.72
A_51_P150530	Tm2d1	TM2 domain containing 1	1.72
A_51_P476538	5430417L22Rik*	RIKEN cDNA 5430417L22 gene	1.72
A_52_P149577	Rab25	RAB25, member RAS oncogene family	1.72
A_52_P503927	Tmem181	transmembrane protein 181	1.71
A_51_P488340	Tpte2 (Crlf2)	transmembrane phosphoinositide 3-phosphatase and tensin homolog 2	1.71
A_51_P479818	Lonrf3	LON peptidase N-terminal domain and ring finger 3	1.71
A_52_P562661	Mal	myelin and lymphocyte protein, T-cell differentiation protein	1.70
A_52_P418489	Tinagl (Tinagl1)	tubulointerstitial nephritis antigen-like	1.68
A_51_P336391	Tmem18	transmembrane protein 18	1.67
A_51_P292073	Haghl (Haghl| Narfl)	hydroxyacylglutathione hydrolase-like	1.62
A_52_P196458	Dzip1	DAZ interacting protein 1	0.62
A_51_P193794	Lrp1	low density lipoprotein receptor-related protein 1	0.62
A_52_P472660	Smurf2*	SMAD specific E3 ubiquitin protein ligase 2	0.61
A_52_P678650	NAP123837-1		0.61
A_51_P216742	Arhgap23	RIKEN cDNA 4933428G20 gene	0.61
A_52_P400425	A530088H08Rik	RIKEN cDNA A530088H08 gene	0.60
A_51_P137913	Slc17a7	solute carrier family 17, member 7	0.59
A_51_P338713	Ass1	argininosuccinate synthetase 1	0.59

A_51_P373142	AI854703	expressed sequence AI854703	0.58
A_51_P366413	8030498J20Rik*	RIKEN cDNA 8030498J20 gene	0.58
A_51_P240384	Trps1	trichorhinophalangeal syndrome I (human)	0.58
A_51_P138876	Mapkap1	mitogen-activated protein kinase associated protein 1	0.58
A_52_P403443	Tns3*	tensin 3	0.57
A_52_P421220	Dzip1	DAZ interacting protein 1	0.57
A_52_P367520	Nexn	nexilin	0.56
A_51_P450955	1110036O03Rik	RIKEN cDNA 1110036O03 gene	0.55
A_52_P108845	Clip3	CAP-GLY domain containing linker protein 3	0.55
A_52_P644452	Dock9	dedicator of cytokinesis 9	0.54
A_51_P259876	Dock9	dedicator of cytokinesis 9	0.53
A_52_P373982	Grhl2	grainyhead-like 2 (Drosophila)	0.52
A_52_P538673	Fgf1	fibroblast growth factor 1	0.52
A_51_P373696	Maml2	mastermind like 2 (Drosophila)	0.52
A_51_P148828	Fgf1	fibroblast growth factor 1	0.51
A_51_P305230	Elavl3	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 3	0.50
A_51_P214796	Suz12	suppressor of zeste 12 homolog (Drosophila)	0.50
A_52_P401473	Ntn1	netrin 1	0.50
A_52_P432715	BB054706		0.50
A_51_P152550	Iqcg	IQ motif containing G	0.48
A_51_P280893	Pla2g1b	phospholipase A2, group IB, pancreas	0.46
A_51_P382970	Itga9	integrin alpha 9	0.46
A_52_P408391	Neurod1	neurogenic differentiation 1	0.45
A_51_P234692	2310043N10Rik*	RIKEN cDNA 2310043N10 gene	0.43
A_52_P42255	Crmp1	collapsin response mediator protein 1	0.43
A_52_P663296	A330008L17Rik	RIKEN cDNA A330008L17 gene	0.42
A_51_P166434	Atp6v1e1	VATPase, H+ transporting, lysosomal V1 subunit E1	0.42
A_51_P408172	1110001D15Rik	RIKEN cDNA	0.41

		1110001D15 gene	
A_52_P288177	Epha4	Eph receptor A4	0.40
A_51_P309095	AK050362		0.39
A_52_P1069478	OTTMUSG00000003947	predicted gene, OTTMUSG00000003947	0.38
A_52_P303388	Atp6v1e1	VATPase, H+ transporting, lysosomal V1 subunit E1	0.38
A_51_P204442	Phf19	PHD finger protein 19	0.35
A_51_P258359	Hus1b	Hus1 homolog b (S. pombe)	0.34
A_52_P302316	Kcnv1	potassium channel, subfamily V, member 1	0.32
A_51_P115953	Ctxn3	predicted gene, ENSMUSG00000069372	0.25
A_52_P240706	Ceacam2	CEA-related cell adhesion molecule 2	0.24
A_52_P300533	2310043M15Rik*	RIKEN cDNA 2310043M15 gene	0.21

Table 2. *Gene expression changes in pre-plated Pcl3 shRNA ESCs by microarray.*

Microarray data indicating fold changes in gene expression in pre-plated ESCs expressing *Pcl3* shRNA as compared to scramble shRNA ESCs. Comparison included two scramble clones and two *Pcl3* shRNA clones.

METHODS

Tissue culture

Mouse E14 embryonic stem cells were grown as described previously [20,54]. Bay Genomics gene trap line XG122 was used for the establishment of *Suz12*$^{Suz12TAP/+}$ ESCs [20].

Lentiviral infection

ESCs were trypsinized and 1ml of cells at 10^5/ml were plated into one well of a 6-well plate along with 15µl of concentrated lentivirus (Open Biosystems RMM4534 and Sigma SHC002). The plate was shaken every ten minutes for one hour and then incubated overnight at 37℃. Media was changed the next day and selection was started (2µg/ml puromyecin, 50µg/ml zeocin). Cells were selected for six days, clones were picked, and knockdown was assessed. Five different *Pcl3* shRNA lentiviral constructs were tested and all gave similar results.

Transfection

Cells were transfected using Lipofectamine 2000 (Invitrogen 11668-019). Briefly 0.5µg of DNA was combined with 1µl of Lipofectamine in 50µl of Optimem and incubated for 20min. ESCs were trypsinized and $2x10^5$ were plated into a well of a 6-well along with Lipofectamine mixture. Cells were incubated for 18hrs followed by media replacement. For siRNAs, 300ng siRNA and 3µl Lipofectamine in 300µl of Optimem was used for the Lipofectamine mixture.

In vitro prepared siRNA pools

Libraries of siRNAs were made as previously described [192,193].

Purification and mass spectrometry of Suz12

Suz12-TAP and bound proteins were purified and subjected to mass spectrometry as previously described [110].

Histone purification

Histones were purified from ESCs as previously described [221].

Immunoprecipitation and western immunoblot

Cells were treated as previously described [222]. For Flag immunoprecipitation, 40µl of FlagM2 resin (Sigma A2220) was added to 800µg protein lysate overnight. All experiments were performed three or more times. *Pcl3* knockdown experiments contained at least two scramble and *Pcl3* shRNA clones.

Antibodies

The following antibodies were used for immunoblotting and immunofluorescence: goat anti-Suz12 (1:500, Santa Cruz sc-46264), rabbit anti-FlagM2 (1:600 for IF and 1:1000 for Western blot, Cell Signaling 2368), mouse anti-α-tubulin (1:5000, Sigma T9026), goat anti-V5 (IF 1:400, Abcam ab9137), rabbit anti-V5 (Western blot 1:5000, Sigma V8137), mouse anti-Ezh2 (1:400, Hamer et al. 2002 Ezh2 M18), mouse anti-Eed (1:1000, gift from Barbara Panning), rabbit anti-β-actin (1:5000, Abcam ab8227), rabbit anti-H3K27me3 (1:3000, Millipore 07-449), mouse anti-H3 (1:5000, Abcam ab10799), mouse anti-H2A119Ub (1:1000, Upstate 05-678 clone E6C5), rabbit anti-H3K9me3 (1:5000, Abcam ab8898), rabbit anti-H3K27ac (1:5000, gift from Barbara Panning), anti-Pcl3 (1:1000, Proteintech), mouse anti-NeuN (1:800, Millipore MAB377), rabbit anti-Tuj1 (1:2000, Covance PRB-435P), mouse anti-Actin (1:100, Sigma A4700), rabbit anti-K14

(1:1000, Covance PRB-155P), mouse anti-E cadherin (1:50, BD Biosciences 610181), rabbit anti-Pcl2 (1:1000, Proteintech), mouse anti-Oct3/4 (C-10) (1:200, Santa Cruz sc-5279), rabbit anti-Nanog (1:200, abcam ab80892), rabbit anti-H3K4me3 (1:1000, Millipore clone MC315 04-745).

Immunofluorescence and staining

Cells were plated on coverslips coated with poly-lysine and Matrigel (BD Biosciences) and stained as previously described [222].

Expression analysis

Expression levels are representative of three or more experiments using 2-6 different clones and assayed in quadruplet. Statistical significance was determined using an unpaired student's *t*-test. Primers for qRT-PCR and ChIP are as following:

qRT-PCR primers

β-actin	CACAGCTTCTTTGCAGCTCCTT
	CGTCATCCATGGCGAACTG
Suz12	GTGCACTCTGAACTGCCGTA
	CCGGTCCATTTCGACTAAAA
Pcl3 set 1	TGTGTGTACAGCCACCACCT
	TGTCCCCAAAGTGATTGTCA
Pcl3 set 2	AGAGGAAGCTGACAGCCAAA
	CGTCTCCACTGATGCTTTCA
Pcl3-TAP	AGCACAGACCACCACCTAGC
	AATCACCGTCATGGTCTTTG
Eed	CTGGCAAAATGGAGGATGAT
	TGGGTCAGTGTTGTGCATTT
Ezh2	AGACGTCCAGCTCCTCTGAA
	ATCCTCAGTGGGAACAGGTG
Oct4	CCAATCAGCTTGGGCTAGAG
	CTGGGAAAGGTGTCCCTGTA
Nanog	GCTCAGCACCAGTGGAGTATCC
	TCCAGATGCGTTCACCAGATAG
Igf2	GTCCCCACATTTGCAGTTCT

	CTGGATGACATGGACAGTGG
Hand1	GTTCCCATTCGTTGCTGAAT
	CTGCGAGTGGTCACACTGAT
Krt7	ACGGCTGCTGAGAATGAGTT
	CGTGAAGGGTCTTGAGGAAG
Acta2	CTGACAGAGGCACCACTGAA
	CATCTCCAGAGTCCAGCACA
Plac1	CCTCCATCATGGGACAGAGT
	GCCCTTACATCTGGGCACTA
Dkk1	CAGCTCAATCCCAAGGATGT
	CAGGGGAGTTCCATCAAGAA
Fgf1	TAAACCCAGGGCTTTCAATG
	TTGCAGTCAGCTTGGAATTG
Neurod1	CAAAGCCACGGATCAATCTT
	CCCGGGAATAGTGAAACTGA
Epha4	AGGTGGGAGGAGGTCAACTT
	GTGAAACGGCTTGATTTGGT
Pcl2	CATTTGCGGAGAAGAAGAGG
	TGATTGAGCTGCAACTCCTG
Pcl1	GAGATCCTCCCCTTCACCTC
	AAATGAAACGGTCCTTGTGG
Meis2	CATGGTCGTTGCACCATAAG
	GTGAGGAACACGTGCTGAGA
Meis1	ACAGGAGACCCGACAATGAG
	TTGACTTCGGATGGTTCTCC
Cxcl12	TTTGGCCTCCTGTAGAATGG
	TCAGAGCCCATAGAGCCACT
Slc23a1	TGCTCTGCTACGTGTTGACC
	GGATCCAGGGAGAGATAGCC
Miip	GGTCACACCACCTCTCTGGT
	GGGACAATGGAGGAGAGACA
Cbx4	GCCCTTGAGTGAGTTCAAGC
	TCCCACTACACCGTCACGTA
Bmp1	CACTCCACAGCAGGAAGTGA
	CTCAGTGAAAGCTCCGGTTC
Hoxd10	GAAGAGGTGCCCTTACACCA
	TCGATTCTCTCGGCTCATCT
Mia1	GGGCCAAGTGGTGTATGTCT
	GGACAATGCTACTGGGGAAA
Manba	AGTGGACACTGGGGACAAAG

	GTCTTCTGGCCGTGCTTTAG
Mras	AGACAGGGCTACAGCTTCCA
	GTCTGGTGCGTTGTATGTGG
Efhc2	ACTTTACGCCCATTTCATGC
	GCTTCAGGCCTACACAGGAG
Zfp42	CTTCACGGAGAGCTCGAAAC
	CTTTGCGTGGGTTAGGATGT
Tgfβ1	TGCTTCAGCTCCACAGAGAA
	TGGTTGTAGAGGGCAAGGAC
ChIP primers	
mir196b	AAGGTGGATCCCAACAACAG
	CCTCCTAGGCGTACCTTTCC
mir34a	GGCAATGGCTCAGAGAAGAC
	AGCGAGAGCTCAGAGTTTGC
mir10a	CCCTTCCAACTCGTCCATAA
	GCCACAGGTTGCTAGTGTGA
mir132/212	GCAGACGCAGACACTATGGA
	CTAGGCTGAGTCACCGCTCT
mir196a	GCTGCAGAGGAGATCAAAGG
	TCTGCCAAGCAAGAGAAACA
Hoxb7	ACCGAGTTCCTTCAACATGC
	CAGGGGTAGATCCGGAAGTT
Hoxb13	CTCGTTAAACCTGCCAAAGC
	CATAAAATCCTCCCGCAGAA
Pmp22	CCTAGGACTGCCTCTGCATC
	CCAGACACGTCCATTTTCCT
Hoxa3	TCACAGGCTGGTGAGAACAC
	AGGGTCATTTGGCAAGACTG
Hand1	GGTGTGAGTGGTGATGATGG
	TTTGGACGTCTGAACCCTTC
Dusp4	GCCTCTACTCGGCTGTCATC
	CCGCCAAGTCCCTTACACTA
Acta1	TTGTCGATTGTCGTCCTGAG
	CAGGCACGTGACACTCTTGT
Tle3	TCCGCAAGGCAGACATCCGGT
	CACGCGCGCGCACCATAAAG
Tns1	AGCGGACCCACCACCCTCTG
	GGAGTCGGCTGACACGGTGC
Tmem151a	TCCGTGTCCTCCCCTCTGCG

	GCCAGCTCTGGGGCCTACCT
Eaf1	CCTCTGGAACTTCGATCTGC
	CAAGAGCCCTGATGGAAGAA
Ash21	CAGAGAGCCCGAGGTAGTTG
	GCACTCAGTCCAGGGGTTTA
Dcaf8	CCCGATCCCTTTCCTAAAAC
	CTAGGAGACGCCAACAGAGC
Ift140	TAGGAGTCCGTCACCGAAAC
	AAGCGTGGATGAAAAAGGTG
Glrx5	CGATTTGTCCAATCAACACG
	CACTCCAGGGCTTACAGAGC
Ebf2	CCACCCCTGAACTTGTCACT
	CTGACGCCTCTTTTCTCACC
Mafb2	AGAAGCGGTCCTCCACACTA
	AGCAGGTGTGACTCACGATG
Pcdh7	TTCTTTGGAAATCCGGACAG
	GGGAACTCGAGCTGAACTTG
Otx2	TTTCAAAGCGAGAGGGAAGA
	CCACGCACATACTGTGGTTC
Actin control	GGAGGTGCAATGGCTGTCTTGTCC
	CTGCCTTGGGTCACCTTACACCTCAC

DNA mutagenesis

Point mutations were created using the Agilent site-directed mutagenesis kit (Quikchange II XL Site-Directed Mutagenesis, Agilent Technologies, 200522) and Quikchange primer design application.

Alignment

Polycomb-like paralog alignment was performed using CLC sequence viewer (CLC Bio).

Alkaline phosphatase activity assay

Alkaline phosphatase activity was determined by staining and using StemTag Alkaline Phosphatase Activity Kit (Cell Biolabs CBA-301). ESCs were plated with ES media

containing LIF and allowed to grow to 80% confluency. For staining, ESCs were fixed in 0.5% glutaraldehyde and washed three times in HBS followed by three washes in AP buffer (100mM Tris pH 9.5, 100mM NaCl, 50mM $MgCl_2$). NBT/BCIP substrate was added and samples were incubated at 37° for 10 minutes. Reaction was stopped with 50mM EDTA pH5.1 in PBS, and pictures were taken. For the StemTag colorimetric assay, cells were lysed, incubated with assay substrate, and assayed for absorbance. Experiment was performed in duplicate three independent times.

Colony formation assay

ESCs were trypsinized, counted, and plated at 100 cells/well of a 6-well. After 8 days, colonies were stained with methylene blue and quantitated. For wild type and Pcl3-overexpressing cells, ESCs were grown in 5% of normal LIF concentration. Experiment was performed in duplicate four independent times.

Teratoma formation

ESCs were trypsinized, counted, and prepared to a concentration of $6.7x10^6$ cells/ml. 300µl ($2x10^6$) cells were injected subcutaneously into severely combined immunodeficient (SCID) mice and allowed to form teratomas for 2-3 weeks. Tumors were removed and divided for paraffin sections, frozen sections, and RNA. Teratomas were formed from two clones each of scramble and *Pcl3* shRNA ESCs.

Mircoarray

Sample preparation, labeling, and array hybridizations were performed according to standard protocols from the UCSF Shared Microarray Core Facilities and Agilent Technologies (http://www.arrays.ucsf.edu and http://www.agilent.com). For microarray #2, ESCs were pre-plated twice for one hour on untreated TC plates, counted, and

plated at equal densities on gelatinized plates. Microarray #1 was performed similarly but without pre-plated. The next day cells were collected and RNA extracted from two scramble and two *Pcl3* shRNA clones. Total RNA quality was assessed using a Pico Chip on an Agilent 2100 Bioanalyzer. RNA was amplified and labeled with Cy3-CTP using low RNA input fluorescent linear amplification kits following manufacturer's protocol (Agilent). Concentration of labeled cRNA was assessed using a Nanodrop ND-100, and equal amounts of Cy3 labeled target were hybridized to whole human genome 4x44K Ink-jet arrays (Agilent). Hybridizations were performed for 14 hrs, according to the manufacturer's protocol (Agilent). Arrays were scanned using an Agilent microarray scanner, and raw signal intensities were extracted with Feature Extraction v10.1 software (Agilent). This dataset was normalized using the *quantile* normalization method that is proposed by Bolstad et al. [223]. No background subtraction was performed, and the median feature pixel intensity was used as the raw signal before normalization. The microarray was performed in 2-6 clones from scramble and *Pcl3* shRNA ESCs.

ChIP

ChIP-seq experiments were done as previously described with 500µg of chromatin and 5µg of antibody [224]. ChIP-qRT-PCR experiments were performed at least two times with 10^7 ESCs and 10µg of antibody. Antibodies used for ChIP immunoprecipitation include anti-H3K27me3 (Upstate, 07-449), anti-Flag M2 (Sigma, 088k6018), anti-Flag M2 (Cell Signaling, 2368), anti-Suz12 (Upstate, 07-379), and anti-H3K4me3 (Millipore clone MC315 04-745) antibodies. Chromatin immunoprecipitation reactions were performed by washing 50µL of the appropriate Dynal magnetic beads (Dynabeads M-280 Sheep anti-Rabbit IgG, Invitrogen Cat # 112.04D, or Dynabeads M-280 Sheep anti-Mouse IgG, Invitrogen Cat # 112.02D) in 1mL PBS with 5mg/mL BSA three times.

Washes were aspirated while tubes were on Dynal MPC-S magnetic strips to retain magnetic beads. Beads were resuspended in 1mL BSA/PBS with 5µg of antibody. The antibodies used were anti-H3K27me3 (Upstate, 07-449), anti-Flag M2 (Sigma, 088k6018), and anti-Suz12 (Upstate, 07-379). After rotating overnight at 4℃, the beads were washed three more times with BSA/PBS and resuspended in 1mL chromatin, 100µL BSA/PBS, 1% Triton-X, 0.1% Sodium Deoxycholate (DOC), 0.1% TE, and 1X EDTA-free Complete (tablets). Samples rotated overnight at 4℃. Beads were then washed 8 times with RIPA buffer (50mM Hepes, pH 8.0, 1mM EDTA, 1% NP-40, 0.7% DOC, 0.5M LiCl, 1X EDTA-free Complete) and washed once with 1X TE. Beads resuspended in 170µL of Elution Buffer (10mM Tris, pH 8.0, 1mM EDTA, 1% SDS) and shaken for 20 minutes at 65℃. Samples were centrifug ed at high speed and the supernatant was removed to a new tube and reverse cross-linked overnight at 65℃. 50µL of untreated chromatin was also reverse cross-linked with 120µL Elution Buffer overnight at 65℃. 140 µL TE, 7µL Proteinase K, and 3µL Glycogen (20µg/µL) were added to each sample and incubated for 2 hours at 37℃. Samples were extracted twice with 300µL Phenol in a single Phase-Lock tube (5Prime, Cat # 2900308, Lot #42793024), and once with chloroform/isoamyl alcohol. After the aqueous layer was removed to a new tube, NaCl was added to a final concentration of 200mM, followed by 2 volumes of 100% ETOH. Samples were vortexed and incubated at -80℃ for 30 minutes, centrifuged for 15 minutes at 15,000 RPM at 4℃, and washed with 1mL 70% ETOH. After centrifuging for 5 minutes at 15,000 RPM at 4℃, pellets were airdried and resuspended in 30µL TE with 10µg RNase A. After incubating 1-2 hours at 37℃, samples were purified using Qiagen MinElute kit and eluted to 34µL with Qiagen EB buffer (50µL for inputs).

Library Preparation

34µL IP samples were treated with End Repair kit (Epicentre, Cat # ER81050, Lot # E85-91206), with 5µL of 10X End Repair buffer, 5µL of 2.5mM dNTPs, 5µL of 10mM ATP, and 1µL of END-IT enzyme mix added to each IP. Samples incubated at room temperature for 45 minutes. Samples were purified using Qiagen MinElute kit and eluted to 32µL. An additional 'A' base was added to the 3' end of DNA fragments by adding 5µL 10X NEBuffer 2 (New England Biolabs, Cat # B7002S, Lot # 0010901), 10µL of 1mM dATP, and 3µL of 5U/µL 3'-5' exo-Klenow Fragment (New England Biolabs, Cat #M0212L, Lot # 0180912). Samples were incubated at 37℃ for 30 minutes and purified with Qiagen MinElute kit, eluting to 20µL. Samples were then evaporated using a speedvac to 4µL, and sequencing adapters were ligated to the ends of the DNA fragments by adding 5µL of 2X Quick Ligase Reaction Buffer (New England Biolabs, Cat # B2200S, Lot # 0010910), 0.5µL of Adapter Oligo mix, 1:10 in H_2O, and 0.5µL of 1U/µL DNA Ligase (New England Biolabs, Cat # M2200L, Lot # 0980910). Samples were incubated at room temperature for 15 minutes and purified with Qiagen MinElute kit, eluting to 30µL. Samples were evaporated in a speedvac for 2 minutes and loaded with 6X Bromophenol blue/xylene cyanol loading dye on a 15 cm 8% polyacrylamide gel in TBE. 100bp DNA ladder loaded in far right and left lanes. Gel run at 200V for 2 hours and stained with 1X SYBR Gold (Invitrogen Cat # S11494) in 120µL TBE, shaking for 15 minutes. DNA between 200bps and 400bps were excised from gel with a clean scalpel, minced by centrifuging through a sieve, and shook vigorously overnight at 4℃ in 500 µL of Qiagen EB buffer. The next day, samples were mixed vigorously at 50℃ for 15 minutes and centrifuged at 15,000 RPM for 2 minutes. The supernatant was transferred to a 0.2 µm Nanoseq column (Pall, Cat # ODM02C35, Lot # 09551479) and centrifuged at 15,000 RPM for 2 minutes and the flow-through was collected. Repeat if necessary. 1/10 volume of 3M NaOAC (pH 5.2), 4µL of glycogen (20µg/µL), and 2 volumes of 100% ETOH were added to each sample, vortexed, and incubated at -80℃ for 30

minutes. Samples centrifuged at 15,000 RPM for 15 minutes at 4℃, washed with 1mL 70% ETOH, and centrifuged again for 5 minutes. Pellets were airdried and resuspended in 20μL of Qiagen EB buffer. 10μL 5X Phusion HF Buffer (Finnzymes, Cat # F-518, Lot # 82), 1μL 10mM dNTPs, 13.5μL PCR grade H_2O, 0.5μL HotStart Phusion Taq (Finnzymes, Cat # F540L, Lot #40), and 2.5μL each of forward and reverse 10mM Solexa PCR primers were added to each sample, and run with the following PCR program:

Step 1: 98℃ for 30 seconds

Step 2: 98℃ for 10 seconds

Step 3: 65℃ for 30 seconds

Step 4: 72℃ for 30 seconds

Step 5: Go to step 2, 17 cycles

Step 6: 72℃ for 5 minutes

Step 7: 4℃, forever

Samples purified with Qiagen MinElute kit, eluting to 40μL. Samples were evaporated in speedvac for 2 minutes and loaded on an 8% polyacrylamide gel exactly as before. DNA was excised from same region of gel as before and treated as before. The following day, the samples were treated exactly the same as after the first gel and overnight shaking, except the purified DNA was ultimately resuspended in 10μL of Qiagen EB buffer. Sample DNA concentration was determined using an Invitrogen Qubit Flourimeter and the samples were diluted to 10nM for sequencing.

ChIP-seq analysis

The data consisted of short reads from single-end 30-nucleotide sequencing. The short reads were aligned to the reference mouse genome (version MM9) by using bowtie with the parameters -n 2 -y --best -m 1[197]. The summary of alignment results is provided in

Figure. 20A. Paired samples were normalized by equalizing the number of reads in the background: partial sums of order statistics of binned read counts in paired samples were computed, and the ratios of the partial sums were plotted as a function of the quantile cutoff in the sum. The point where the curve deviates significantly from linearity was found by fitting a linear regression using half of the data between the 10th and 60th quantile and by finding the first point that deviates from the linear regression line by more than 3 standard deviations of the estimated error. The ratio of partial sums at this critical point was used to scale the read density in one of the samples. The significance of the difference in read counts was assessed by using the Skellam distribution as a null model. We used 400 bp running windows to scan the genome. A *p*-value cutoff of 10^{-7} was used to call significant windows, because this cutoff roughly corresponds to an adjusted *p*-value of 0.05 based on the method of Poisson clumping heuristic for approximating the probability of locally correlated rare events. The relation between the Skellam and adjusted *p*-values is given by $p_{Skellam} = -\log(1-p_{adjusted})\, E[C]/L$, where L is the genome size, and E[C] the expected correlation length of Skellam *p*-values across adjacent windows. Overlapping significant windows were concatenated, and we report the new *p*-values recomputed in the joined windows.

Normalization of Paired ChIP-seq Data

NGS experiments may produce vastly different numbers of total reads for paired data, e.g. Suz12 ChIP after Pcl3 shRNA vs. scramble, or Suz12 ChIP vs. Input. Suppose there are *M* reads for IP and *N* reads for Input. Then, many researchers currently scale the Input counts per genomic window by *M*/*N*, in order to equalize the total number of reads between the two samples. This method, however, does not take into account that the IP data should accumulate a significantly greater number of reads in regions targeted by the antibody relative to the background; as a result, this scaling approach

will artificially inflate the background noise captured by Input. A better alternative approach is to find a scaling factor that equalizes just the background between IP and Input channels without considering the IP peak regions.

In order to separate the antibody-targeted loci from background noise, we will apply the theory of order statistics. Our motivation comes from the following observation for normally distributed random variables: let Y_1, Y_2, ... , Y_N be independent identically distributed normal random variables with mean μ and variance σ^2, and let $Y_{(1)}$, $Y_{(2)}$, ... , $Y_{(N)}$ be their order statistics, i.e. the rearrangements of Y_i such that $Y_{(1)} \le Y_{(2)} \le ... \le Y_{(N)}$. If we partition the reference genome into N equal-sized bins, then for sufficiently large μ, we can think of Y_i as counting reads in the i-th bin. Define the partial mean $S_{n/N}$ as

$$S_{n/N} = \sum_{k=1}^{n} Y_{(k)} / n .$$

Because the order statistics are ranked in an increasing order, it can be seen that the partial mean is an increasing function of n. In fact, in the limit of large sample size N, the partial mean is almost a linear function of n with a positive slope. More precisely, for large N,

$$S_\alpha \to \mu - \frac{\sigma}{\alpha} f(Y_\alpha),$$

where f is the probability density of Y_i and $\alpha = n/N$ *[225]*. Expanding this asymptotic form around $\alpha = 1/2$, one can show that the partial sum satisfies

$$S_\alpha \overset{N\to\infty}{\sim} \mu - \frac{\sigma}{\alpha} \frac{(1 - \pi(\alpha - 1/2)^2)}{\sqrt{2\pi}},$$

which is **almost linear in α** and can be fitted with linear regression with $R^2 > 0.99$.

Similarly, consider a set of bivariate normal random variables $Z_i = (Y_i, X_i,)$, where Y_i are defined as above and $X_1, X_2, ..., X_N$ are independent identically distributed normal random variables that are uncorrelated with Y_i. We can think of Y_i as binned

immunoprecipitated DNA counts and X_i as binned Input DNA counts. Then, we define the *i*-th order statistic $Z_{(i)}$ to be the pair (Y_k,X_k), such that Y_k corresponds to the *i*-th order statistic $Y_{(i)}$; i.e., the order statistics are obtained by sorting Z_i with respect to the first entry. Because we have assumed that *X* and *Y* are uncorrelated, $\sum_{k=1}^{n} X_{(k)} / n$ is an unbiased estimate of the expectation $E[X]$, and for sufficiently large *n*, the ratio $R_{n/N} = \sum_{k=1}^{n} Y_{(k)} \Big/ \sum_{k=1}^{n} X_{(k)}$ of partial sums thus approaches $\sum_{k=1}^{n} Y_{(k)} \Big/ nE[X]$, which is proportional to the partial mean $S_{n/N} = \sum_{k=1}^{n} Y_{(k)} / n$. Consequently, for large *N*, the above analysis shows that $R_{n/N}$ can be **approximated by a linear function** of α=*n*/*N*.

In a more realistic situation, the distribution of the IP channel data Y_i can be modeled as a mixture of two Poisson distributions, e.g. one component representing the basal level of background noise and the second component representing the enrichment of actual immunoprecipitated DNA. For sufficiently large mean, Poisson distributions approach normal distributions, and the above analysis still holds; but, the ratio R_α in this case begins to diverge from linearity at α roughly equal to the mixing probability. By computing this critical value of α, we can thus approximate the proportion of background noise in the IP channel data Y_1, Y_2, ... , Y_N. The corresponding ratio R_α also provides the optimal scaling factor for normalizing the Input channel *X* in order to equalize the read counts in background regions that are not directly targeted by the antibody. This method can be applied to general paired ChIP-seq data, e.g. Suz12 ChIP-seq with Pcl3 shRNA vs. scramble. Figure 19B provides an example of this method, where the mouse genome was partitioned into 2kb bins.

Detection of Peaks and Differentially Enriched ChIP-seq Regions

We have assessed the statistical significance of the differences between paired normalized ChIP-seq data as follows: let $n_1(x)$ and $n_2(x)$ be the normalized tag counts in a window centered at genomic location x in Sample 1 and Sample 2, respectively. Then, assuming the null hypothesis that $n_1(x)$ and $n_2(x)$ are independent and locally Poisson with common mean λ, the difference $Y=n_1(x)-n_2(x)$ in tag counts follows the Skellam distribution

$$f(Y = n; \lambda) = e^{-2\lambda} I_{|n|}(2\lambda)$$

where the maximum likelihood estimate (MLE) of the mean is $\lambda=(n_1(x)+n_2(x))/2$. Because we have only two samples, the MLE may underestimate the true mean and increase the false positive rate. In our analysis, we have estimated the 95% confidence interval and used the more conservative value $\lambda=0.5$ qpois(0.975, $n_1(x)+n_2(x)$), where qpois is the quantile function in R for the Poisson random variable $n_1(x)+n_2(x)$. For $n > 0$, the right-tail p-value is

$$P(Y \geq n) = e^{-2\lambda} \sum_{k=n}^{\infty} I_k(2\lambda)$$

$$= e^{-2\lambda} \sum_{k=n}^{\infty} \sum_{j=0}^{\infty} \frac{\lambda^{2j+k}}{j!(j+k)!}$$

$$= \sum_{k=n}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!(j+k)!}$$

$$= \sum_{k=n}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!\Gamma(j+k+1)}$$

$$= \sum_{k=n}^{\infty} \frac{\lambda^k}{k!\Gamma(k+1)} e^{-\lambda} {}_0F_1(k+1, \lambda^2/4)$$

$$= \sum_{k=n}^{\infty} \frac{\lambda^k}{k!\Gamma(k+1)} e^{-\lambda} \frac{\gamma(k+1, \lambda^2/4)}{(\lambda^2/4)^{k+1}}$$

where the incomplete γ function is defined as

$$\gamma(n+k,\lambda) = \int_0^{\lambda} e^{-t} t^{n+k-1}\, dt.$$

The final expression of the p-value is equal to the cumulative distribution of χ^2 random variable with 2n degrees of freedom and non-centrality parameter 2λ evaluated at 2λ.

This novel method of detecting differential enrichment is more rubust than other approaches using Poisson *p*-values, which do not take into account the over-dispersion of count data and are thus more prone to false positives.

Gene density analysis

The MM9 mouse genome was partitioned into disjoint 1 Mbp bins. The log fold-change of normalized counts in paired ChIP-seq data with *Pcl3* knockdown or scramble was computed in each bin. Gene density was computed by counting the number of RefSeq genes in each 1 Mbp bin. Pearson correlation was computed between gene density and log fold-change across bins. In Figure 19B, we grouped the bins into units of 10; e.g., 10 denotes bins with gene density greater than or equal 0 and less than 10.

Suz12 binding sites in miRNA promoters

To find miRNAs targeted by Suz12, we used a prior annotation of miRNA promoters [198]. The differential Suz12 binding levels shown in Figure 19D-E were computed within Suz12 ChIP-seq peaks overlapping with the annotated promoters.

Motif analysis

We used a greedy algorithm called LeitMotif which models the background nucleotide distribution with the 9th order Markov dependence [134]. Motif lengths ranging between 6 and 15 were used for scanning. Top scoring motifs with length 10 and 14 were chosen for further analysis, because other motifs were either contained in or contained those

two motifs. Sequences obtained from LeitMotif were aligned to produce the logos in Figure 27A and to generate the PSSM matrices used in Fig. 27B-C. A custom C code was used to scan the Suz12 sites with those PSSM matrices, assuming independent bases in the background. Support vector machine implemented in the e1071 R-package was used to classify the Suz12 binding sites with 10-fold cross validation.

Chapter 4: Pluripotency Factor Sall4 Associates with the Mi-2/Nurd Complex in Embryonic Stem Cells

INTRODUCTION

Sall4 is a member of the Spalt family transcription factors conserved throughout metazoa. Until recently, Sall4 was principally studied as the cause of many overlapping congenital disorders such as Okihiro syndrome and Holt-Oram syndrome [121]. However, current studies have shown that Sall4 plays a critical role in ESC pluripotency and the establishment of the inner cell mass [121,226]. Indeed, embryos lacking *Sall4* die at peri-implantation stages [126]. While *Sall4* mutant ES cells can be derived, it is achieved at a very low efficiency [129].

Three important transcription factors involved in maintaining ESC pluripotency are Oct4, Sox2, and Nanog [4]. *Sall4* is expressed in ESCs in a manner identical to *Oct4* and binds both Oct4 and Nanog [132]. The Sall4-Nanog complex auto-regulates itself as well as promotes *Oct4* and *Sox2* expression [132,134]. Loss of *Sall4* in ES cells leads to upregulation of trophectoderm genes and decreased proliferation [129]. Sall4 also acts as an enhancer of cell reprogramming during ESC-somatic cell fusion [119].

In contrast to its role in gene activation, Sall4 can localize with heterochromatin and repress genes [135]. Indeed, Sall4 functions as a transcriptional repressor when bound to CyclinD1, a cell cycle regulator [135]. In addition, Sall4 can inhibit *Sall1* and *Sall3* expression by competing with Oct4, which activates both genes in ESCs [131]. Thus, Sall4 is unique in that it can act as both a transcriptional activator and repressor. However, the molecular mechanisms that determine its mode of regulation are unknown.

Given that Sall4 interacts with a wide array of proteins, we hypothesized that protein partners may confer distinct functions and influence transcriptional regulation of

target genes. By purifying and analyzing Sall4-bound proteins by mass spectrometry, we found that Sall4 associates with all components of Nucleosome Remodeling and Deacetylation (NuRD) complex in ESCs and differentiated cells. The NuRD complex is primarily involved in gene repression and has the dual capacity to remodel and deactylate nucleosomes [227,228]. The binding of Sall4 to the NuRD complex suggests that they cooperate in gene inhibition in ESCs and during embryonic development.

RESULTS

To identify novel binding partners, we tagged Sall4 within the endogenous locus and affinity purified associated proteins. To efficiently modify Sall4 in ESCs, we utilized the recently developed Floxin system [20]. In a two-step process, we used a Sall4 gene trap ESC line and DNA recombination to knockin a C-terminal tandem affinity purification (TAP) tagged form of Sall4 into the endogenous locus (Figure 28) [20].

Briefly, addition of Cre-Recombinase to a Sall4 gene trap (*Sall4*$^{Gt/+}$) cell line induced a recombination event, which removed the exogenous splice acceptor and allowed expression of the endogenous *Sall4* allele. Of the clones analyzed, 39% reverted (*Sall4*$^{Rev/+}$) and were neomycin sensitive and stained negative for X-gal (Figure 29A and data not shown). The *Sall4*$^{Rev/+}$ ESCs were exposed once more to Cre-Recombinase along with a specifically designed Floxin shuttle vector. The Sall4 shuttle vector contains a LoxP site, the remaining two Sall4 exons linked to a TAP tag (6xHis-3xFlag), and a β-actin promoter. Incorporation of the shuttle vector into the endogenous locus occurred with 89% efficiency, and the formation of a dead Lox site prevented reversal of the reaction (Figure 29A). *Neomycin* and *β-galactosidase* driven by the β-actin promoter were used for selection and detection of positive clones (Figure 28).

The *Sall4* gene can be alternatively spliced to create three isoforms (a-c) with isoforms a and b predominantly found in ESCs. Isoform a is full-length Sall4, while Sall4b is produced by alternative splicing to exon three after transcribing the first half of exon two. The *Sall4*$^{Gt/+}$ ESC line used for this study (XE027) has the gene trap in the second half of exon two. Thus, the Floxin allele specifically expresses *Sall4b-TAP*, and *Sall4a* can only be transcribed from the wild type allele.

To determine whether the Sall4-TAP shuttle vector successfully recombined and produced stable protein, we immunoprecipitated and probed wild type, *Sall4*$^{Gt/+}$, *Sall4*$^{Rev/+}$, and *Sall4*$^{Sall4TAP/+}$ ESC lysates for FlagM2 and TAP respectively. We found that Sall4-TAP was stably and specifically expressed in the *Sall4*$^{Sall4TAP/+}$ ESCs (Figure 29B). Thus, Sall4-TAP is expressed from the endogenous Sall4 promoter using the inserted slice acceptor in *Sall4*$^{Sall4TAP/+}$ ESCs.

Since overexpression can sometimes cause spurious interactions, we wanted to assess whether *Sall4-TAP* is expressed at levels comparable to wild type. Therefore, we measured transcript and protein levels by quantitative reverse transcription PCR (qRT-PCR) and by immunoblot. *Sall4* expression was similar between wild type, *Sall4*$^{Gt/+}$, *Sall4*$^{Rev/+}$, and *Sall4*$^{Sall4TAP/+}$ ESCs, which supports data indicating that Sall4 autoregulates itself (Figure 29B-C). Notably, several Sall4 bands were seen in the *Sall4*$^{Gt/+}$, *Sall4*$^{Rev/+}$, and *Sall4*$^{Sall4TAP/+}$ ESCs, suggesting that spurious forms of Sall4 may be expressed from the gene trap allele (Figure 29B). These spurious forms may be attributable to combined alternative splicing within the endogenous Sall4 locus and the gene trap.

Sall4 is found in many different complexes in ESCs and during development [121,123,132,136]. To identify Sall4-interacting proteins in ESCs and differentiated cells, we affinity purified Sall4 with its binding partners from *Sall4*$^{Sall4TAP/+}$ ESCs and *Sall4*$^{Sall4TAP/+}$ cells grown without LIF for seven days (Figure 29D). The purified proteins

were subjected to MALDI TOF and MuDPIT MS/MS, a form of mass spectrometry which couples 2D-LC to MS/MS. As expected the protein with the most hits by mass spectrometry was Sall4 with just over 150 hits in both samples (Figure 29E). Sall4 was found to associate with the other three SAL family homologues (Sall1-3) (Figure 29E). Sall4 is known to heterodimerize with Sall1, but dimerization with Sall2 and Sall3 had not previously been reported [129,229]. In addition, we identified all components of the Mi-2/Nucleosome Remodeling and Deacetylase (NuRD) complex (Figure 29E).

While several components of the NuRD complex have multiple paralogs, the NuRD complex composition consists of six proteins with a 1:1 stoichiometry [227,228]. For example, while metastasis-associated (MTA) protein has three homologues (MTA1-3), only one MTA protein is found in a single NuRD complex. A similar situation exists for histone deacetylases (Hdac) 1 and 2, transcriptional repressor p66 alpha and beta, and retinoblastoma binding proteins (Rbbp) 4 and 7. Interestingly homologues exist for Mi-2 (Chd4) and methyl-CpG-binding domain protein 3 (Mbp3), Chd3 and Mbp2; however, these homologues were not associated with Sall4 in either ESCs or differentiated cells (Figure 29E). This suggests that Sall4 may bind a subset of NuRD complexes that specifically contain Chd4 and Mbp3.

DISCUSSION

Sall4 cooperates in a regulatory network in ESCs with Oct4 and Nanog, which promotes expression of pluripotency and self-renewal genes. Sall4 also has inhibitory properties in which it can compete with Oct4 or bind to CyclinD1 at heterochromatin to repress genes. We sought to identify novel Sall4 binding partners in an effort to elucidate how protein interactions can affect function. Using the Floxin system, we TAP-tagged Sall4 within the endogenous locus and purified Sall4 complexes from ESCs and

differentiated cells. The complexes were subjected to mass spectrometry and associated proteins were identified.

We found that Sall4 bound all SAL family paralogs, Sall1-3. Although, Sall4 has been shown to homodimerize and heterodimerize with Sall1, it was unknown that Sall4 could heterodimerize with Sall2 and Sall3 [129,229]. Sall4 has several isoforms, which can also dimerize. For example, Sall4a and Sall4b heterodimerize and bind many of the same chromatin sites in ESCs [230]. However, Sall4a and Sall4b homodimers also have isoform independent roles in ESCs and in the embryo [124,230].

The mass spectrometry data also revealed that Sall4 binds all components of the NuRD complex including Mi-2 (Chd4), MTA1-3, Hdac1/2, Mbd3, p66α/β, and Rbbp4/7. The NuRD complex is a unique chromatin remodeling complex that has both Mi-2 ATPase activity as well as Hdac1/2 histone deacetylase activity. Both of these enzymatic activities are associated with gene repression. Chromatin compaction requires ATPase activity, and histones lacking acetylation are typically found in condensed regions with repressed transcription. Thus, the NuRD complex can independently repress genes in two ways, deacetylating and compacting histones. The association of Sall4 with the NuRD complex suggests Sall4 cooperates with the NuRD complex to inhibit genes in ESCs and differentiated cells.

Despite all other NuRD components showed indiscriminate incorporation of their homologues, Chd4 and Mbd3 homologues, Chd3 and Mbd2, were absent from the mass spectrometry data in ESCs and differentiated cells. Mbd2 and Mbd3 have overlapping as well as independent roles in ESCs and during development [108,231,232]. The different roles of Chd3 and Chd4 have not been explored. It is possible that Sall4 only interacts with Chd4 and Mbd3 and that this specificity provides unique function to the complex.

Sall1 is known to recruit the NuRD complex through a twelve amino acid motif [233]. Since our finding, another group discovered that Sall4 interacts with the NuRD complex to inhibit genes such as *PTEN* and *Sall1 in vitro* and *in vivo* [125]. They show that both isoforms Sall4a and Sall4b associate with NuRD complex components in ESCs and in 293 cells [125]. Furthermore, they found that Sall4 transgenic mice have decreased expression of *PTEN* and *Sall1* in kidneys and bone marrow, which is correlated with cystic kidneys and leukemia [125]. Thus, gene repression by Sall4 and NuRD may protect certain tissues from carcinogenesis.

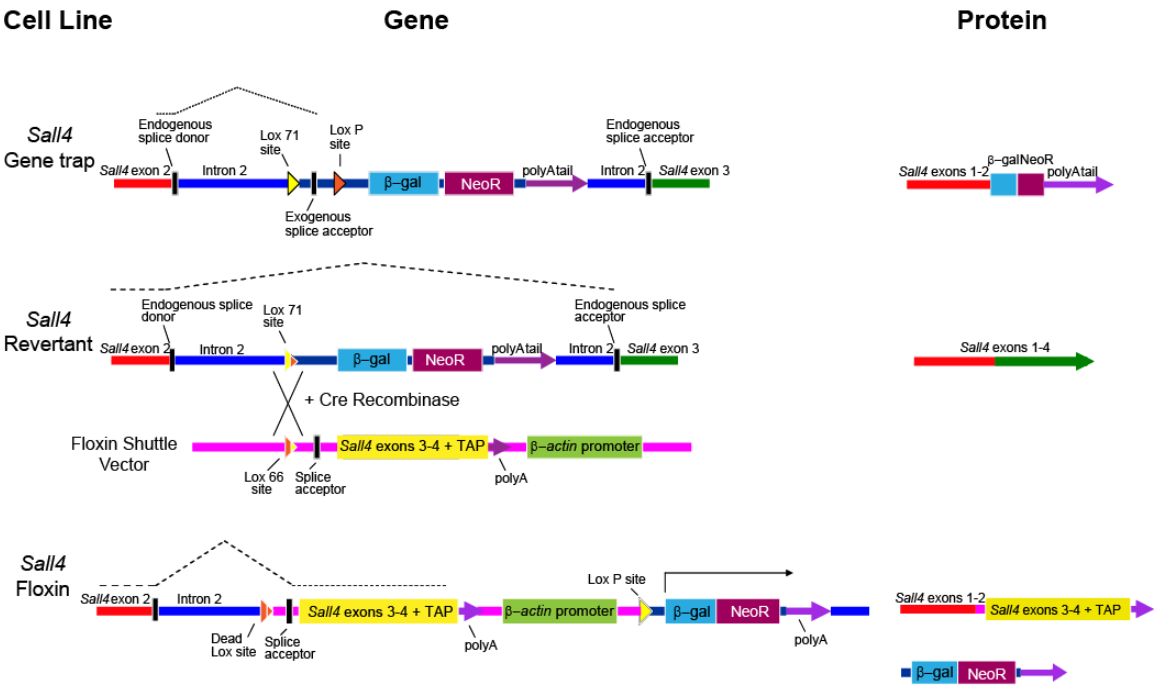
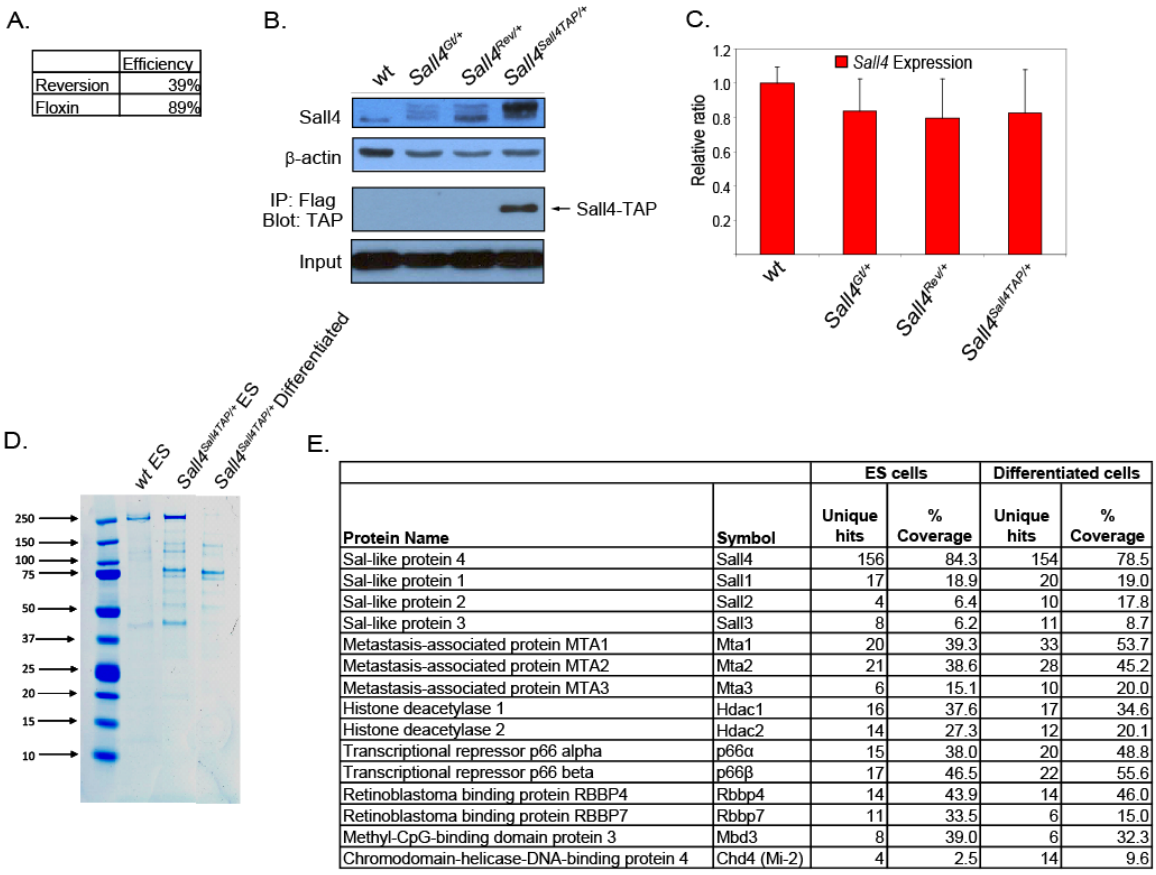


Figure 28. (Adapted from Singla et al 2010) *Floxin mediated manipulation of the Sall4 locus.*

The Floxin system uses Bay Genomics gene trap ESC lines, which contain an exogenous slice acceptor surrounded by two Lox sites. The resulting protein is a fusion of the upstream exons to neomycin and β-galactosidase. Addition of Cre-Recombinase recombines the Lox sites to remove the splice acceptor and allows normal splicing to the following exon. Using a specifically designed shuttle vector and Cre-Recombinase, an exogenous slice acceptor, the remaining exons, and a C-terminal TAP (6xHis-TEV-3xFlag) tag are incorporated into the endogenous locus, creating a tagged version of Sall4 expressed from the endogenous locus. The amino Lox site is defective upon the initial recombination so the reaction is irreversible. Moreover, incorporation of the *β-actin* promoter upstream of *Neomycin* and *β-galactosidase* allows selection and simple screening.



	Efficiency
Reversion	39%
Floxin	89%

		ES cells		Differentiated cells	
Protein Name	**Symbol**	**Unique hits**	**% Coverage**	**Unique hits**	**% Coverage**
Sal-like protein 4	Sall4	156	84.3	154	78.5
Sal-like protein 1	Sall1	17	18.9	20	19.0
Sal-like protein 2	Sall2	4	6.4	10	17.8
Sal-like protein 3	Sall3	8	6.2	11	8.7
Metastasis-associated protein MTA1	Mta1	20	39.3	33	53.7
Metastasis-associated protein MTA2	Mta2	21	38.6	28	45.2
Metastasis-associated protein MTA3	Mta3	6	15.1	10	20.0
Histone deacetylase 1	Hdac1	16	37.6	17	34.6
Histone deacetylase 2	Hdac2	14	27.3	12	20.1
Transcriptional repressor p66 alpha	p66α	15	38.0	20	48.8
Transcriptional repressor p66 beta	p66β	17	46.5	22	55.6
Retinoblastoma binding protein RBBP4	Rbbp4	14	43.9	14	46.0
Retinoblastoma binding protein RBBP7	Rbbp7	11	33.5	6	15.0
Methyl-CpG-binding domain protein 3	Mbd3	8	39.0	6	32.3
Chromodomain-helicase-DNA-binding protein 4	Chd4 (Mi-2)	4	2.5	14	9.6

Figure 29. *Sall4-TAP associates with all members of the NuRD complex in ESC and differentiated cells.*

(A) Reversion and floxin efficiency were measured as the number of clones that recombined/ the total number of clones analyzed. 39% and 89% of clones reverted from *Sall4*$^{Gt/+}$ to *Sall4*$^{Rev/+}$ and floxed from *Sall4*$^{Rev/+}$ to *Sall4*$^{Sall4TAP/+}$ respectively. **(B)** Sall4 protein levels were measured in wild type, *Sall4*$^{Gt/+}$, *Sall4*$^{Rev/+}$ and *Sall4*$^{Sall4TAP/+}$ ESC lysates. To detect Sall4-TAP, wild type, *Sall4*$^{Gt/+}$, *Sall4*$^{Rev/+}$ and *Sall4*$^{Sall4TAP/+}$ ESC lysates were probed with Flag and then blotted for TAP. As expected, Sall4-TAP was only present in *Sall4*$^{Sall4TAP/+}$ ESCs. β-actin was used as a loading control. **(C)** *Sall4* transcript levels were comparable between wild type, *Sall4*$^{Gt/+}$, *Sall4*$^{Rev/+}$ and *Sall4*$^{Sall4TAP/+}$ ESC lines as measured by qRT-PCR. **(D)** Coomassie stained gel of Flag-tagged purified proteins in wild type, *Sall4*$^{Sall4TAP/+}$ ESCs, and *Sall4*$^{Sall4TAP/+}$ differentiated cells. **(E)** Mass spectrometry results of Sall4 associated proteins in ESCs and differentiated cells. All SAL homologues are present as well as NuRD complex components.

METHODS

Tissue Culture

ESCs were grown as previously described [20,54]. The Sall4 gene trap line (XE027) was purchased from Bay Genomics and further manipulation was performed as previously described [20].

Immunoprecipitation and Western Blot

Cells were lysed in RIPA with protease and phosphatase inhibitor, and protein concentration was measured. For Western blot, 30µg of lysate was run on an 8% acrylamide gel and probed for Sall4 (Abcam, ab29112) and β-actin (Abcam, ab8227). For immunoprecipitation, 40µl of mouse anti-FlagM2 resin (Sigma A2220) was added to 500µg protein lysate overnight and subsequently washed three times. Samples were run on an 8% polyacrylamide gel and transferred to a PVDF membrane. The membrane was blocked in 5% milk and incubated with a TAP antibody (GenScript, A00683) overnight at 4°. The next day the membrane was washed and probed with secondary for 1hr. Following three washes, the membrane was incubated with chemiluminescent substrate for 1min and exposed for 1-30 min.

Quantitative RT-PCR

RNA was extracted from ESCs using the Qiagen RNeasy Mini kit. RNA was then subjected to first-strand cDNA synthesis using iScript (Biorad or Fermentas). Expression levels were then analyzed in triplicate using a 7300 Real-time PCR machine (Applied Biosystems) and normalized to β-actin. The following primers were used:

Sall4 F: CTCATGGGGCCAACAATAAC

Sall4 R: CGGAGATCTCGTTGGTCTTC

Purification of Sall4-TAP

Sall4-TAP was purified according to Krutchinsky's protocol (http://cms.biomslab.org/category/protocols/IP%20protocol/) with the following modifications. Wild type, Sall4$^{Sall4TAP/+}$ ESCs, and Sall4$^{Sall4TAP/+}$ differentiated cells were treated with RIPA buffer and protein concentration measured. Four milligrams of total cell lysate from each sample was homogenized. Following solubilization and centrifugation, lysates were incubated with FlagM2 Dynal beads (Invitrogen) for 1hr. Protein was eluted from the beads using 3xFLAG peptide, and the purified protein was run on an SDS-PAGE gel.

Identification of Sall4-interacting Proteins by Mass Spectrometry

After being separated by SDS-PAGE and Coomassie blue stained (GelCode Blue, Pierce), protein bands were excised and processed for digestion with trypsin (Promega) as described [234]. Peptides were fractionated using an Eksigent 1D NanoLC system and a 100 um ID x 150 mm self-packed column with Ultro120 (Peeke Scientific) C-18 resin. The flow rate was 350 nl/min, solvents were water/0.1% formic acid (Buffer A) and acetonitrile/0.1% formic acid (Buffer B), and the gradient was 5-30% B over 40 minutes. The LC system was coupled to an LTQ-Orbitrap (Thermo Fisher Scientific) operating in positive ion mode. The peaklist-generating software was Mascot Distiller v2.1 and the search engine used was Protein Prospector v4.25.4 (http://prospector.ucsf.edu/). The database searched was SwissProt.2007.04.19 (264492 entries). Database search parameters were as follows: trypsin enzyme specificity; three missed cleavages; 25 ppm mass tolerance for precursor ions; 0.8 Da mass tolerance for fragment ions; variable modifications considered were oxidation of methionine, N-terminal protein acetylation,

pyroglutamate formation from N-terminal glutamine residues, modification of cysteine by carbamidomethylation.

Chapter 5: Conclusions and Future Directions

The work described in this dissertation provides insights into multiple regulatory mechanisms of ESC self-renewal and differentiation. Because ESCs possess so many layers of regulation, it can be difficult to assess the role of different components in a clear manner. Using the Floxin system, we were able to investigate cell signaling, epigenetics, and transcription in ESCs and answer a number of biological questions. In Chapter 2, we showed that a hemizygous *Ofd1* gene trap cell line could be utilized to better understand human disease. By easily modifying the *Ofd1* gene, we were able to represent and study specific disease alleles. Using the Floxin system in Chapters 3 and 4, we identified novel binding partners of proteins known to be involved in ESC pluripotency. This was performed by tagging Suz12 or Sall4 within the endogenous allele and subsequently purifying and analyzing associated proteins. Notably, a number of interesting binding partners have yet to be confirmed and analyzed. Analysis of Sall4 in Chapter 4 illustrates how different protein isoforms can be easily isolated and studied, a technique that is often difficult without the use of specific peptides antibodies or overexpression. In short, the Floxin system provides a simple and efficient way of manipulating genes in ESCs that can be used in broad contexts.

The ciliogenic protein Ofd1 regulates the neuronal differentiation of embryonic stem cells

In this work, we revealed that differentiation of ESCs lacking primary cilia can serve as a model to study the role of cilia during early embryonic development. We demonstrated that ESCs lacking Ofd1 and cilia have defective Hh and Wnt signaling, phenotypes observed *in vivo*. We also observed a loss of ventral neural subtypes, which are characteristic of mispatterned neural tubes in embryos lacking cilia. In addition to

these defects, there are numerous other phenotypes associated with ciliopathies including polycystic kidneys, extra or loss of digits, mental retardation, as well as face and mouth anomalies. Indeed, *Ofd1*^{Gt} ESCs may function as a useful model for better understanding the molecular basis behind these malformations as well.

The mechanisms that drive kidney cyst formation in patients with ciliopathies are unknown, and currently, the prevailing treatment is kidney transplantation. Recently, protocols have been developed for the differentiation of ESCs to renal precursors [235]. There is still much progress to be made to create renal tubule cells *in vitro*, however, this future model may shed light on potential cystic kidney treatments in patients with defective cilia [224].

Several ciliopathies including OFD1 are characterized by facial malformations. Indeed, loss of cilia specifically in the neural crest causes severe craniofacial defects [154]. Studies have found that craniofacial malformations are due at least part to defective bone and skeletal development; however, there are likely other neural crest cell types affected by cilia loss. Protocols for ESC differentiation to neural crest cells have been developed [236-238]. Thus, investigating the potential of *Ofd1*^{Gt} ESCs to differentiate to neural crest cells and subsequent progenitors may help us better understand how cilia contribute to specifying cells of the face and mouth.

In addition to defective signaling, *Ofd1*^{Gt} ESCs had an increased capacity to differentiate into neurons, indicating that either Ofd1 or cilia restrains neural differentiation. We did not observe increased neural tissue in *Kif3a* or *Ift88* null embryos, indicating that our findings are either an Ofd1 or ESC specific phenotype. Investigating neural differentiation in ESCs lacking cilia by deletion of some other gene like *Ift88* would clarify this. Alternatively, neural development in *Ofd1*$^{-/-}$ mice has not been thoroughly characterized and may provide evidence that Ofd1 has a specific role in specifying neurons.

We investigated multiple signaling pathways in an effort to dissect how Ofd1 or cilia could be restraining neural development. Addition of BMP, Wnt, Hh, or Notch pathway activators or inhibitors never resulted in specific normalization of neural expression in *Ofd1*^{Gt} EBs. Nevertheless, it would be interesting to elucidate the molecular underpinnings behind this phenotype. Spontaneous EB differentiation produces extensive variability between duplicates and between experiments, which at times can make it difficult to assess significant rescue by signaling factors. Direct differentiation to neurospheres may provide a more controlled system to study signaling defects and discrepancies in gene expression. If neural differentiation was increased in *Ofd1*^{Gt} neurospheres, microarray analysis of a differentiation time course would likely provide clues as to which genes may be causing this phenotype. If neural differentiation was similar between wild type and *Ofd1*^{Gt} neurospheres, it would suggest that non-neural cells are promoting neural differentiation. In either case, this system could potentially elucidate a previously undescribed role of Ofd1 or cilia in development.

Polycomb-like 3 is a component of PRC2 that promotes embryonic stem cell self-renewal

A significant amount of research has been dedicated to understanding the function and role of PRC2 during development and in cancer, yet there are many aspects of PRC2 that are not well-understood. In an effort to identify novel proteins that bind and regulate PRC2, we identified Polycomb-like 3 (Pcl3). This work characterizes Pcl3 and demonstrates that Pcl3 promotes PRC2 function and ESC self-renewal.

Although we confirmed that Pcl3 associates with the core components of PRC2, it is unknown whether Pcl3 has PRC2-independent functions. Mass spectrometry on Pcl3 and fractionation experiments would reveal whether Pcl3 binds other complexes or

whether it can function alone. Studies on the two isoforms of human Pcl3 demonstrate that both isoforms can bind PRC2 components and that the full-length isoform can homodimerize [111]. Although the murine Pcl3 is not known to have additional isoforms, it is possible that it also homodimerizes.

Pcl3 contains a Tudor domain and two PHD zinc fingers [75,103]. Several residues within the Tudor domain have been implicated in histone binding. Indeed, we found mutation of these residues caused disruption of PRC2 function [102]. Whether these residues are necessary for histone binding or whether they required for some other function is unclear. We found that the mutant forms of Pcl3 still bound Suz12, suggesting these residues are not essential for incorporation into PRC2. The regions of Pcl3 required for PRC2 binding and the residue interactions necessary for integration have not been characterized. Additional Pcl3 mutation and binding studies will be needed to address these questions.

We found that knockdown of *Pcl3* caused a significant decrease in Suz12 binding, suggesting that Pcl3 promotes recruitment of PRC2 at some target genes. Even so, we cannot discern whether Pcl3 association with PRC2 facilitates binding to target genes or whether Pcl3 binds chromatin and subsequently recruits PRC2. ChIP-sequencing of Pcl3 in *Suz12* or *Ezh2* null ESCs would distinguish between these two possibilities. In addition, PRC2 can be recruited through association with non-coding RNAs [75,213]. Thus, Pcl3 may also use RNAs for binding or recruitment, which could be assessed by sequencing Pcl3-bound RNAs.

Upon depletion of *Pcl3*, we observed increased differentiation of ESCs, suggesting that Pcl3 promotes ESC self-renewal. *Oct4* and *Nanog* expression were decreased in *Pcl3* knockdown ESCs, yet it is unclear whether this is a direct effect of *Pcl3* knockdown. In-depth analysis of global expression patterns upon *Pcl3* knockdown as well as extensive cell cycle analysis may shed light on this question. Understanding

Pcl3-enhanced self-renewal may be useful in studying tumorigenesis as *Pcl3* expression is misregulated in many types of cancer [112]. In addition to promoting ESC self-renewal, we found that Pcl3 supports Suz12 expression and stability. Although we do not know the manner by which this occurs, Suz12 is not known to be autoregulated by PRC2, suggesting that it is an indirect mechanism.

Pcl3 is important for ESC maintenance, but its role in development has not been investigated. Knockout of core PRC2 components causes peri-implantation lethality, however, *Pcl2* knockout mice are born and exhibit only minor skeletal defects [3,118,180,182]. The mild phenotypes seen in $Pcl2^{-/-}$ mice may be attributable to redundancy by Pcl1 or Pcl3. On the other hand, ESC studies on mammalian Pcl proteins indicate that they may have independent roles as well. Mouse mutants of Pcl1 and Pcl3 will need to be made to answer these questions. These mice would also provide useful tools to elucidate how Pcl proteins cooperate or compete in different cell types and tissues.

Although we found Pcl2 and Pcl3 to bind the same regions of chromatin, we observed many contrasting results in the *Pcl3* knockdown ESCs compared to *Pcl2* knockdown ESCs. Knockdown of Pcl2 caused a global increase in H3K27me3, increased Ezh2 levels, and enhanced ESC self-renewal [101]. Nevertheless, several Pcl2 target genes exhibited decreased PRC2 binding and H3K27me3 and upregulated expression [101,109]. Indeed, Pcl2 appears to play both active and repressive roles on PRC2 regulation in ESCs and differentiated cells [101,109]. We demonstrate that Pcl2 and Pcl3 do not interact, suggesting that they each bind a distinct subset of PRC2. Performing dual knockdown or knockout experiments on Pcl2 and Pcl3 would help clarify whether these two paralogs cooperate or antagonize each other in their regulation of PRC2.

PRC2 is an essential regulatory complex during development and is often hijacked by cancer for plasticity and metastasis. As a result, extensive studies have explored PRC2 as a potential cancer target [239-241]. Further investigation into Pcl3 will not only help us better understand the molecular function of PRC2, but may also reveal novel mechanisms to treat cancer.

Pluripotency Factor Sall4 Associates with the Mi-2/Nurd Complex in Embryonic Stem Cells

Sall4 associates with many different protein partners to regulate pluripotency and developmental genes. To identify novel binding partners of Sall4, we tagged the endogenous allele and purified and identified bound proteins. In ESCs and differentiated cells, we found Sall4 to interact with Sall1-3 and all components of the NuRD complex. Although Sall4 was known to heterodimerize with Sall1, heterodimerization with Sall2 and Sall3 had not previously been reported [129,229]. Even so, it is unclear whether different dimers have specific targets or binding partners. Dissecting whether homodimers or heterodimers are interchangeable and whether they confer distinct function will provide interesting insight into the role of SAL family members in ESCs and during development.

The NuRD complex is involved in chromatin remodeling and histone deacetylation, typically resulting in gene repression [227,228]. Numerous NuRD target genes exist, yet it is unclear whether Sall4 associates at all target genes or only a subset. ChIP-sequencing of Sall4 and NuRD would provide a map of Sall4 and NuRD binding sites within the genomic landscape. Perhaps there are certain types of genes which are regulated by the Sall4-NuRD complex or perhaps it is a small subset of NuRD targets. Alternatively, Sall4 may associate with NuRD to regulate an extensive gene

network. As Sall4 promotes ESC pluripotency and self-renewal, it is possible that the Sall4-NuRD complex represses differentiation genes in ESCs and pluripotency genes upon differentiation. Comparing the ChIP-sequencing data with expression analysis in *Sall4* or *Chd4* mutant cells would help answer these questions. The NuRD complex is also known to compact chromatin, so it would be interesting to assess whether knockout of *Sall4* affected chromatin structure [227,228].

SAL family proteins have multiple zinc fingers, which can bind DNA as well as other proteins [121,123]. It is unclear whether Sall4 binds specific DNA targets and subsequently recruits the NuRD complex or whether they form a complex and bind together. Assessing whether the NuRD complex has decreased binding in the absence of Sall4 or vice versa would address this. Furthermore, *in vitro* binding studies on mutated Sall4 would clarify what domains may be responsible for binding NuRD.

While we specifically investigated Sall4 isoform b, another group found the NuRD complex to bind both Sall4a and Sall4b. Knockout of either Sall4 isoforms in embryos has revealed distinct phenotypes, indicating that there is some redundancy as well as isoform-independent functions. Whether both isoforms function similarly within the NuRD complex or target the same genes is still ambiguous. Moreover, the NuRD complex has many possible combinations due to all components having multiple homologues, most of which were apparent in the mass spectrometry data. Chd3 and Mbd2 were the only two homologues not identified as associating with Sall4. It is possible that Sall4 specifically binds Chd4 and Mbd3, creating a specialized complex that can target specific genes. In any case, investigation into the Sall4-NuRD complex will provide data on ESC and developmental gene regulation.

BIBLIOGRAPHY

1. Martin GR. (1981). 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 78:7634-8.
2. Evans MJ and MH Kaufman. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature 292:154-6.
3. O'Carroll D, S Erhardt, M Pagani, SC Barton, MA Surani and T Jenuwein. (2001). The polycomb-group gene Ezh2 is required for early mouse development. Mol Cell Biol 21:4330-6.
4. Chambers I. (2004). The molecular basis of pluripotency in mouse embryonic stem cells. Cloning Stem Cells 6:386-91.
5. Liu N, M Lu, X Tian and Z Han. (2007). Molecular mechanisms involved in self-renewal and pluripotency of embryonic stem cells. J Cell Physiol 211:279-86.
6. Keller G. (2005). Embryonic stem cell differentiation: emergence of a new era in biology and medicine. Genes Dev 19:1129-55.
7. Keller GM. (1995). In vitro differentiation of embryonic stem cells. Curr Opin Cell Biol 7:862-9.
8. Doetschman TC, H Eistetter, M Katz, W Schmidt and R Kemler. (1985). The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. J Embryol Exp Morphol 87:27-45.
9. Desbaillets I, U Ziegler, P Groscurth and M Gassmann. (2000). Embryoid bodies: an in vitro model of mouse embryogenesis. Exp Physiol 85:645-51.
10. Sachlos E and DT Auguste. (2008). Embryoid body morphology influences diffusive transport of inductive biochemicals: a strategy for stem cell differentiation. Biomaterials 29:4471-80.
11. Gaulden J and JF Reiter. (2008). Neur-ons and neur-offs: regulators of neural induction in vertebrate embryos and embryonic stem cells. Hum Mol Genet 17:R60-6.
12. Zhou JM, JX Chu and XJ Chen. (2008). An improved protocol that induces human embryonic stem cells to differentiate into neural cells in vitro. Cell Biol Int 32:80-5.
13. Vidarsson H, J Hyllner and P Sartipy. (2010). Differentiation of human embryonic stem cells to cardiomyocytes for in vitro and in vivo applications. Stem Cell Rev 6:108-20.
14. Dambrot C, R Passier, D Atsma and CL Mummery. (2011). Cardiomyocyte differentiation of pluripotent stem cells and their use as cardiac disease models. Biochem J 434:25-35.
15. Chen X, H Xu, P Yuan, F Fang, M Huss, VB Vega, E Wong, YL Orlov, W Zhang, J Jiang, YH Loh, HC Yeo, ZX Yeo, V Narang, KR Govindarajan, B Leong, A Shahab, Y Ruan, G Bourque, WK Sung, ND Clarke, CL Wei and HH Ng. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133:1106-17.
16. Van Hoof D, AD Mendelsohn, R Seerke, TA Desai and MS German. (2011). Differentiation of human embryonic stem cells into pancreatic endoderm in patterned size-controlled clusters. Stem Cell Res 6:276-85.
17. Trounson A. (2006). The production and directed differentiation of human embryonic stem cells. Endocr Rev 27:208-19.
18. Chimge NO and D Bayarsaihan. (2009). Generation of neural crest progenitors from human embryonic stem cells. J Exp Zool B Mol Dev Evol 314:95-103.

19. Luo Z, G Wang, W Wang, Q Xiao and Q Xu. (2011). Signalling pathways that regulate endothelial differentiation from stem cells. Front Biosci 16:472-85.
20. Singla V, J Hunkapiller, N Santos, AD Seol, AR Norman, P Wakenight, WC Skarnes and JF Reiter. (2010). Floxin, a resource for genetically engineering mouse ESCs. Nat Methods 7:50-2.
21. Christophersen NS and K Helin. (2010). Epigenetic control of embryonic stem cell fate. J Exp Med 207:2287-95.
22. Corbit KC, P Aanstad, V Singla, AR Norman, DY Stainier and JF Reiter. (2005). Vertebrate Smoothened functions at the primary cilium. Nature 437:1018-21.
23. Corbit KC, AE Shyer, WE Dowdle, J Gaulden, V Singla, MH Chen, PT Chuang and JF Reiter. (2008). Kif3a constrains beta-catenin-dependent Wnt signalling through dual ciliary and non-ciliary mechanisms. Nat Cell Biol 10:70-6.
24. Singla V and JF Reiter. (2006). The primary cilium as the cell's antenna: signaling at a sensory organelle. Science 313:629-33.
25. Gerdes JM, EE Davis and N Katsanis. (2009). The vertebrate primary cilium in development, homeostasis, and disease. Cell 137:32-45.
26. Fliegauf M, T Benzing and H Omran. (2007). When cilia go bad: cilia defects and ciliopathies. Nat Rev Mol Cell Biol 8:880-93.
27. Bisgrove BW and HJ Yost. (2006). The roles of cilia in developmental disorders and disease. Development 133:4131-43.
28. Ishikawa H and WF Marshall. (2011). Ciliogenesis: building the cell's antenna. Nat Rev Mol Cell Biol 12:222-34.
29. Nonaka S, Y Tanaka, Y Okada, S Takeda, A Harada, Y Kanai, M Kido and N Hirokawa. (1998). Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. Cell 95:829-37.
30. Schneider L, CA Clement, SC Teilmann, GJ Pazour, EK Hoffmann, P Satir and ST Christensen. (2005). PDGFRalphaalpha signaling is regulated through the primary cilium in fibroblasts. Curr Biol 15:1861-6.
31. Haycraft CJ, B Banizs, Y Aydin-Son, Q Zhang, EJ Michaud and BK Yoder. (2005). Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function. PLoS Genet 1:e53.
32. Huangfu D and KV Anderson. (2005). Cilia and Hedgehog responsiveness in the mouse. Proc Natl Acad Sci U S A 102:11325-30.
33. Liu A, B Wang and LA Niswander. (2005). Mouse intraflagellar transport proteins regulate both the activator and repressor functions of Gli transcription factors. Development 132:3103-11.
34. May SR, AM Ashique, M Karlen, B Wang, Y Shen, K Zarbalis, J Reiter, J Ericson and AS Peterson. (2005). Loss of the retrograde motor for IFT disrupts localization of Smo to cilia and prevents the expression of both activator and repressor functions of Gli. Dev Biol 287:378-89.
35. Ruiz i Altaba A. (1999). Gli proteins encode context-dependent positive and negative functions: implications for development and disease. Development 126:3205-16.
36. Sasaki H, Y Nishizaki, C Hui, M Nakafuku and H Kondoh. (1999). Regulation of Gli2 and Gli3 activities by an amino-terminal repression domain: implication of Gli2 and Gli3 as primary mediators of Shh signaling. Development 126:3915-24.
37. Wong SY and JF Reiter. (2008). The primary cilium at the crossroads of mammalian hedgehog signaling. Curr Top Dev Biol 85:225-60.

38. Wang B, JF Fallon and PA Beachy. (2000). Hedgehog-regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb. Cell 100:423-34.
39. Litingtung Y and C Chiang. (2000). Specification of ventral neuron types is mediated by an antagonistic interaction between Shh and Gli3. Nat Neurosci 3:979-85.
40. Logan CY and R Nusse. (2004). The Wnt signaling pathway in development and disease. Annu Rev Cell Dev Biol 20:781-810.
41. Fuerer C, R Nusse and D Ten Berge. (2008). Wnt signalling in development and disease. Max Delbruck Center for Molecular Medicine meeting on Wnt signaling in Development and Disease. EMBO Rep 9:134-8.
42. Wallingford JB and B Mitchell. (2011). Strange as it may seem: the many links between Wnt signaling, planar cell polarity, and cilia. Genes Dev 25:201-13.
43. Zaghloul NA and N Katsanis. (2009). Mechanistic insights into Bardet-Biedl syndrome, a model ciliopathy. J Clin Invest 119:428-37.
44. Badano JL, N Mitsuma, PL Beales and N Katsanis. (2006). The ciliopathies: an emerging class of human genetic disorders. Annu Rev Genomics Hum Genet 7:125-48.
45. Parisi MA. (2009). Clinical and molecular features of Joubert syndrome and related disorders. Am J Med Genet C Semin Med Genet 151C:326-40.
46. Pazour GJ, BL Dickert, Y Vucica, ES Seeley, JL Rosenbaum, GB Witman and DG Cole. (2000). Chlamydomonas IFT88 and its mouse homologue, polycystic kidney disease gene tg737, are required for assembly of cilia and flagella. J Cell Biol 151:709-18.
47. Ferrante MI, A Zullo, A Barra, S Bimonte, N Messaddeq, M Studer, P Dolle and B Franco. (2006). Oral-facial-digital type I protein is required for primary cilia formation and left-right axis specification. Nat Genet 38:112-7.
48. Marszalek JR, P Ruiz-Lozano, E Roberts, KR Chien and LS Goldstein. (1999). Situs inversus and embryonic ciliary morphogenesis defects in mouse mutants lacking the KIF3A subunit of kinesin-II. Proc Natl Acad Sci U S A 96:5043-8.
49. Huangfu D, A Liu, AS Rakeman, NS Murcia, L Niswander and KV Anderson. (2003). Hedgehog signalling in the mouse requires intraflagellar transport proteins. Nature 426:83-7.
50. Lupo G, WA Harris and KE Lewis. (2006). Mechanisms of ventral patterning in the vertebrate nervous system. Nat Rev Neurosci 7:103-14.
51. Wilson L and M Maden. (2005). The mechanisms of dorsoventral patterning in the vertebrate neural tube. Dev Biol 282:1-13.
52. Persson M, D Stamataki, P te Welscher, E Andersson, J Bose, U Ruther, J Ericson and J Briscoe. (2002). Dorsal-ventral patterning of the spinal cord requires Gli3 transcriptional repressor activity. Genes Dev 16:2865-78.
53. Macca M and B Franco. (2009). The molecular basis of oral-facial-digital syndrome, type 1. Am J Med Genet C Semin Med Genet 151C:318-25.
54. Singla V, M Romaguera-Ros, JM Garcia-Verdugo and JF Reiter. (2010). Ofd1, a human disease gene, regulates the length and distal structure of centrioles. Dev Cell 18:410-24.
55. Ferrante MI, A Barra, JP Truong, S Banfi, CM Disteche and B Franco. (2003). Characterization of the OFD1/Ofd1 genes on the human and mouse sex chromosomes and exclusion of Ofd1 for the Xpl mouse mutant. Genomics 81:560-9.
56. Bibikova M, LC Laurent, B Ren, JF Loring and JB Fan. (2008). Unraveling epigenetic regulation in embryonic stem cells. Cell Stem Cell 2:123-34.

57. Sawarkar R and R Paro. (2010). Interpretation of developmental signaling at chromatin: the Polycomb perspective. Dev Cell 19:651-61.
58. Zhou VW, A Goren and BE Bernstein. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nat Rev Genet 12:7-18.
59. Beisel C and R Paro. (2011). Silencing chromatin: comparing modes and mechanisms. Nat Rev Genet 12:123-35.
60. Cedar H and Y Bergman. (2009). Linking DNA methylation and histone modification: patterns and paradigms. Nat Rev Genet 10:295-304.
61. Jaenisch R and A Bird. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 33 Suppl:245-54.
62. Feng S, SE Jacobsen and W Reik. (2010). Epigenetic reprogramming in plant and animal development. Science 330:622-7.
63. Li E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet 3:662-73.
64. Campos EI and D Reinberg. (2009). Histones: annotating chromatin. Annu Rev Genet 43:559-99.
65. Razin SV, OV Iarovaia, N Sjakste, T Sjakste, L Bagdoniene, AV Rynditch, ER Eivazova, M Lipinski and YS Vassetzky. (2007). Chromatin domains and regulation of transcription. J Mol Biol 369:597-607.
66. Marmorstein R. (2001). Protein modules that manipulate histone tails for chromatin regulation. Nat Rev Mol Cell Biol 2:422-32.
67. Spivakov M and AG Fisher. (2007). Epigenetic signatures of stem-cell identity. Nat Rev Genet 8:263-71.
68. Ruthenburg AJ, H Li, DJ Patel and CD Allis. (2007). Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol 8:983-94.
69. Pasini D, M Malatesta, HR Jung, J Walfridsson, A Willer, L Olsson, J Skotte, A Wutz, B Porse, ON Jensen and K Helin. (2010). Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. Nucleic Acids Res 38:4958-69.
70. Bernstein BE, TS Mikkelsen, X Xie, M Kamal, DJ Huebert, J Cuff, B Fry, A Meissner, M Wernig, K Plath, R Jaenisch, A Wagschal, R Feil, SL Schreiber and ES Lander. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125:315-26.
71. Zhang Z, A Jones, CW Sun, C Li, CW Chang, HY Joo, Q Dai, MR Mysliwiec, LC Wu, Y Guo, W Yang, K Liu, KM Pawlik, H Erdjument-Bromage, P Tempst, Y Lee, J Min, TM Townes and H Wang. (2011). PRC2 Complexes with JARID2, MTF2, and esPRC2p48 in ES Cells to Modulate ES Cell Pluripotency and Somatic Cell Reprograming. Stem Cells 29:229-40.
72. Landeira D, S Sauer, R Poot, M Dvorkina, L Mazzarella, HF Jorgensen, CF Pereira, M Leleu, FM Piccolo, M Spivakov, E Brookes, A Pombo, C Fisher, WC Skarnes, T Snoek, K Bezstarosti, J Demmers, RJ Klose, M Casanova, L Tavares, N Brockdorff, M Merkenschlager and AG Fisher. (2010). Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators. Nat Cell Biol 12:618-24.
73. Vastenhouw NL, Y Zhang, IG Woods, F Imam, A Regev, XS Liu, J Rinn and AF Schier. (2010). Chromatin signature of embryonic pluripotency is established during genome activation. Nature 464:922-6.

74. Stock JK, S Giadrossi, M Casanova, E Brookes, M Vidal, H Koseki, N Brockdorff, AG Fisher and A Pombo. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat Cell Biol 9:1428-35.
75. Margueron R and D Reinberg. (2011). The Polycomb complex PRC2 and its mark in life. Nature 469:343-9.
76. Schwartz YB and V Pirrotta. (2008). Polycomb complexes and epigenetic states. Curr Opin Cell Biol 20:266-73.
77. Cao R and Y Zhang. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. Mol Cell 15:57-67.
78. Gould A. (1997). Functions of mammalian Polycomb group and trithorax group related genes. Curr Opin Genet Dev 7:488-94.
79. Chamberlain SJ, D Yee and T Magnuson. (2008). Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. Stem Cells 26:1496-505.
80. Lee TI, RG Jenner, LA Boyer, MG Guenther, SS Levine, RM Kumar, B Chevalier, SE Johnstone, MF Cole, K Isono, H Koseki, T Fuchikami, K Abe, HL Murray, JP Zucker, B Yuan, GW Bell, E Herbolsheimer, NM Hannett, K Sun, DT Odom, AP Otte, TL Volkert, DP Bartel, DA Melton, DK Gifford, R Jaenisch and RA Young. (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell 125:301-13.
81. Leeb M, D Pasini, M Novatchkova, M Jaritz, K Helin and A Wutz. (2010). Polycomb complexes act redundantly to repress genomic repeats and genes. Genes Dev 24:265-76.
82. Pasini D, AP Bracken, JB Hansen, M Capillo and K Helin. (2007). The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol Cell Biol 27:3769-79.
83. Kerppola TK. (2009). Polycomb group complexes--many combinations, many functions. Trends Cell Biol 19:692-704.
84. Chan CS, L Rastelli and V Pirrotta. (1994). A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. Embo J 13:2553-64.
85. Muller J and JA Kassis. (2006). Polycomb response elements and targeting of Polycomb group proteins in Drosophila. Curr Opin Genet Dev 16:476-84.
86. Sing A, D Pannell, A Karaiskakis, K Sturgeon, M Djabali, J Ellis, HD Lipshitz and SP Cordes. (2009). A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. Cell 138:885-97.
87. Woo CJ, PV Kharchenko, L Daheron, PJ Park and RE Kingston. (2010). A region of the human HOXD cluster that confers polycomb-group responsiveness. Cell 140:99-110.
88. Mendenhall EM, RP Koche, T Truong, VW Zhou, B Issac, AS Chi, M Ku and BE Bernstein. (2010). GC-rich sequence elements recruit PRC2 in mammalian ES cells. PLoS Genet 6:e1001244.
89. Schuettengruber B and G Cavalli. (2009). Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice. Development 136:3531-42.
90. Margueron R, N Justin, K Ohno, ML Sharpe, J Son, WJ Drury, 3rd, P Voigt, SR Martin, WR Taylor, V De Marco, V Pirrotta, D Reinberg and SJ Gamblin. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. Nature 461:762-7.

91. Simon JA and RE Kingston. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol 10:697-708.
92. Bender LB, R Cao, Y Zhang and S Strome. (2004). The MES-2/MES-3/MES-6 complex and regulation of histone H3 methylation in C. elegans. Curr Biol 14:1639-43.
93. Kuzmichev A, K Nishioka, H Erdjument-Bromage, P Tempst and D Reinberg. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev 16:2893-905.
94. Kim H, K Kang and J Kim. (2009). AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2. Nucleic Acids Res 37:2940-50.
95. Li G, R Margueron, M Ku, P Chambon, BE Bernstein and D Reinberg. (2010). Jarid2 and PRC2, partners in regulating gene expression. Genes Dev 24:368-80.
96. Pasini D, PA Cloos, J Walfridsson, L Olsson, JP Bukowski, JV Johansen, M Bak, N Tommerup, J Rappsilber and K Helin. (2010). JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. Nature 464:306-10.
97. Peng JC, A Valouev, T Swigut, J Zhang, Y Zhao, A Sidow and J Wysocka. (2009). Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. Cell 139:1290-302.
98. Shen X, W Kim, Y Fujiwara, MD Simon, Y Liu, MR Mysliwiec, GC Yuan, Y Lee and SH Orkin. (2009). Jumonji modulates polycomb activity and self-renewal versus differentiation of stem cells. Cell 139:1303-14.
99. Lonie A, R D'Andrea, R Paro and R Saint. (1994). Molecular characterisation of the Polycomblike gene of Drosophila melanogaster, a trans-acting negative regulator of homeotic gene expression. Development 120:2629-36.
100. Duncan IM. (1982). Polycomblike: a gene that appears to be required for the normal expression of the bithorax and antennapedia gene complexes of Drosophila melanogaster. Genetics 102:49-70.
101. Walker E, WY Chang, J Hunkapiller, G Cagney, K Garcha, J Torchia, NJ Krogan, JF Reiter and WL Stanford. (2010). Polycomb-like 2 associates with PRC2 and regulates transcriptional networks during mouse embryonic stem cell self-renewal and differentiation. Cell Stem Cell 6:153-66.
102. Friberg A, A Oddone, T Klymenko, J Muller and M Sattler. (2010). Structure of an atypical Tudor domain in the Drosophila Polycomblike protein. Protein Sci 19:1906-16.
103. Campagna-Slater V, AG Arrowsmith, Y Zhao and M Schapira. (2010). Pharmacophore screening of the protein data bank for specific binding site chemistry. J Chem Inf Model 50:358-67.
104. Kim J, J Daniel, A Espejo, A Lake, M Krishna, L Xia, Y Zhang and MT Bedford. (2006). Tudor, MBT and chromo domains gauge the degree of lysine methylation. EMBO Rep 7:397-403.
105. Nekrasov M, T Klymenko, S Fraterman, B Papp, K Oktaba, T Kocher, A Cohen, HG Stunnenberg, M Wilm and J Muller. (2007). Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. Embo J 26:4078-88.
106. O'Connell S, L Wang, S Robert, CA Jones, R Saint and RS Jones. (2001). Polycomblike PHD fingers mediate conserved interaction with enhancer of zeste protein. J Biol Chem 276:43065-73.
107. Sarma K, R Margueron, A Ivanov, V Pirrotta and D Reinberg. (2008). Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. Mol Cell Biol 28:2718-31.

108. Hong Z, J Jiang, L Lan, S Nakajima, S Kanno, H Koseki and A Yasui. (2008). A polycomb group protein, PHF1, is involved in the response to DNA double-strand breaks in human cell. Nucleic Acids Res 36:2939-47.
109. Li X, K Isono, D Yamada, TA Endo, M Endoh, J Shinga, Y Mizutani-Koseki, AP Otte, M Casanova, H Kitamura, T Kamijo, J Sharif, O Ohara, T Toyada, BE Bernstein, N Brockdorff and H Koseki. (2011). Mammalian Polycomb-Like Pcl2/Mtf2 Is a Novel Regulatory Component of PRC2 That Can Differentially Modulate Polycomb Activity both at the Hox Gene Cluster and at Cdkn2a Genes. Mol Cell Biol 31:351-64.
110. Walker E, JL Manias, WY Chang and WL Stanford. (2011). PCL2 modulates gene regulatory networks controlling self-renewal and commitment in embryonic stem cells. Cell Cycle 10.
111. Boulay G, C Rosnoblet, C Guerardel, PO Angrand and D Leprince. (2011). Functional characterization of hPCL3 (human Polycomb-like 3) isoforms identifies them as components of distinct EZH2 protein complexes. Biochem J.
112. Wang S, GP Robertson and J Zhu. (2004). A novel human homologue of Drosophila polycomblike gene is up-regulated in multiple cancers. Gene 343:69-78.
113. Li B, M Gogol, M Carey, D Lee, C Seidel and JL Workman. (2007). Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. Science 316:1050-4.
114. Lan F, RE Collins, R De Cegli, R Alpatov, JR Horton, X Shi, O Gozani, X Cheng and Y Shi. (2007). Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. Nature 448:718-22.
115. Yoshitake Y, TL Howard, JL Christian and SM Hollenberg. (1999). Misexpression of Polycomb-group proteins in Xenopus alters anterior neural development and represses neural target genes. Dev Biol 215:375-87.
116. Kitaguchi T, K Nakata, T Nagai, J Aruga and K Mikoshiba. (2001). Xenopus Polycomblike 2 (XPcl2) controls anterior to posterior patterning of the neural tissue. Dev Genes Evol 211:309-14.
117. Reijnen MJ, KM Hamer, JL den Blaauwen, C Lambrechts, I Schoneveld, R van Driel and AP Otte. (1995). Polycomb and bmi-1 homologs are expressed in overlapping patterns in Xenopus embryos and are able to interact with each other. Mech Dev 53:35-46.
118. Wang S, F He, W Xiong, S Gu, H Liu, T Zhang, X Yu and Y Chen. (2007). Polycomblike-2-deficient mice exhibit normal left-right asymmetry. Dev Dyn 236:853-61.
119. Wong CC, A Gaspar-Maia, M Ramalho-Santos and RA Reijo Pera. (2008). High-efficiency stem cell fusion-mediated assay reveals Sall4 as an enhancer of reprogramming. PLoS One 3:e1955.
120. Takahashi K and S Yamanaka. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126:663-76.
121. Sweetman D and A Munsterberg. (2006). The vertebrate spalt genes in development and disease. Dev Biol 293:285-93.
122. Kuhnlein RP, G Frommer, M Friedrich, M Gonzalez-Gaitan, A Weber, JF Wagner-Bernholz, WJ Gehring, H Jackle and R Schuh. (1994). spalt encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the Drosophila embryo. Embo J 13:168-79.

123. de Celis JF and R Barrio. (2009). Regulation and function of Spalt proteins during animal development. Int J Dev Biol 53:1385-98.
124. Uez N, H Lickert, J Kohlhase, MH de Angelis, R Kuhn, W Wurst and T Floss. (2008). Sall4 isoforms act during proximal-distal and anterior-posterior axis formation in the mouse embryo. Genesis 46:463-77.
125. Velu CS, AM Baktula and HL Grimes. (2009). Gfi1 regulates miR-21 and miR-196b to control myelopoiesis. Blood 113:4720-8.
126. Elling U, C Klasen, T Eisenberger, K Anlag and M Treier. (2006). Murine inner cell mass-derived lineages depend on Sall4 function. Proc Natl Acad Sci U S A 103:16319-24.
127. Nishinakamura R, Y Matsumoto, K Nakao, K Nakamura, A Sato, NG Copeland, DJ Gilbert, NA Jenkins, S Scully, DL Lacey, M Katsuki, M Asashima and T Yokota. (2001). Murine homolog of SALL1 is essential for ureteric bud invasion in kidney development. Development 128:3105-15.
128. Parrish M, T Ott, C Lance-Jones, G Schuetz, A Schwaeger-Nickolenko and AP Monaghan. (2004). Loss of the Sall3 gene leads to palate deficiency, abnormalities in cranial nerves, and perinatal lethality. Mol Cell Biol 24:7102-12.
129. Yuri S, S Fujimura, K Nimura, N Takeda, Y Toyooka, Y Fujimura, H Aburatani, K Ura, H Koseki, H Niwa and R Nishinakamura. (2009). Sall4 is essential for stabilization, but not for pluripotency, of embryonic stem cells by repressing aberrant trophectoderm gene expression. Stem Cells 27:796-805.
130. Karantzali E, V Lekakis, M Ioannou, C Hadjimichael, J Papamatheakis and A Kretsovali. (2011). Sall1 regulates embryonic stem cell differentiation in association with nanog. J Biol Chem 286:1037-45.
131. Xu C, C Bian, W Yang, M Galka, H Ouyang, C Chen, W Qiu, H Liu, AE Jones, F MacKenzie, P Pan, SS Li, H Wang and J Min. (2010). Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). Proc Natl Acad Sci U S A 107:19266-71.
132. Zhang J, WL Tam, GQ Tong, Q Wu, HY Chan, BS Soh, Y Lou, J Yang, Y Ma, L Chai, HH Ng, T Lufkin, P Robson and B Lim. (2006). Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nat Cell Biol 8:1114-23.
133. Wu Q, X Chen, J Zhang, YH Loh, TY Low, W Zhang, W Zhang, SK Sze, B Lim and HH Ng. (2006). Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. J Biol Chem 281:24090-4.
134. Carroll JS, CA Meyer, J Song, W Li, TR Geistlinger, J Eeckhoute, AS Brodsky, EK Keeton, KC Fertuck, GF Hall, Q Wang, S Bekiranov, V Sementchenko, EA Fox, PA Silver, TR Gingeras, XS Liu and M Brown. (2006). Genome-wide analysis of estrogen receptor binding sites. Nat Genet 38:1289-97.
135. Bohm J, FJ Kaiser, W Borozdin, R Depping and J Kohlhase. (2007). Synergistic cooperation of Sall4 and Cyclin D1 in transcriptional repression. Biochem Biophys Res Commun 356:773-9.
136. Koshiba-Takeuchi K, JK Takeuchi, EP Arruda, IS Kathiriya, R Mo, CC Hui, D Srivastava and BG Bruneau. (2006). Cooperative and antagonistic interactions between Sall4 and Tbx5 pattern the mouse limb and heart. Nat Genet 38:175-83.
137. Netzer C, SK Bohlander, L Rieger, S Muller and J Kohlhase. (2002). Interaction of the developmental regulator SALL1 with UBE2I and SUMO-1. Biochem Biophys Res Commun 296:870-6.
138. Ramirez J and J Hagman. (2009). The Mi-2/NuRD complex: a critical epigenetic regulator of hematopoietic development, differentiation and cancer. Epigenetics 4:532-6.

139. Lauberth SM, AC Bilyeu, BA Firulli, KL Kroll and M Rauchman. (2007). A phosphomimetic mutation in the Sall1 repression motif disrupts recruitment of the nucleosome remodeling and deacetylase complex and repression of Gbx2. J Biol Chem 282:34858-68.
140. Ulloa F, N Itasaki and J Briscoe. (2007). Inhibitory Gli3 activity negatively regulates Wnt/beta-catenin signaling. Curr Biol 17:545-50.
141. Clement CA, SG Kristensen, K Mollgard, GJ Pazour, BK Yoder, LA Larsen and ST Christensen. (2009). The primary cilium coordinates early cardiogenesis and hedgehog signaling in cardiomyocyte differentiation. J Cell Sci 122:3070-82.
142. Jaenisch R and R Young. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell 132:567-82.
143. Chen JK, J Taipale, KE Young, T Maiti and PA Beachy. (2002). Small molecule modulation of Smoothened activity. Proc Natl Acad Sci U S A 99:14071-6.
144. Ocbina PJ, M Tuson and KV Anderson. (2009). Primary cilia are not required for normal canonical Wnt signaling in the mouse embryo. PLoS One 4:e6839.
145. Veland IR, A Awan, LB Pedersen, BK Yoder and ST Christensen. (2009). Primary cilia and signaling pathways in mammalian development, health and disease. Nephron Physiol 111:p39-53.
146. Jiang J and CC Hui. (2008). Hedgehog signaling in development and cancer. Dev Cell 15:801-12.
147. Munoz-Sanjuan I and AH Brivanlou. (2002). Neural induction, the default model and embryonic stem cells. Nat Rev Neurosci 3:271-80.
148. Beddington RS. (1994). Induction of a second neural axis by the mouse node. Development 120:613-20.
149. Hemmati-Brivanlou A and DA Melton. (1992). A truncated activin receptor inhibits mesoderm induction and formation of axial structures in Xenopus embryos. Nature 359:609-14.
150. Smith WC and RM Harland. (1992). Expression cloning of noggin, a new dorsalizing factor localized to the Spemann organizer in Xenopus embryos. Cell 70:829-40.
151. Bachiller D, J Klingensmith, C Kemp, JA Belo, RM Anderson, SR May, JA McMahon, AP McMahon, RM Harland, J Rossant and EM De Robertis. (2000). The organizer factors Chordin and Noggin are required for mouse forebrain development. Nature 403:658-61.
152. Di-Gregorio A, M Sancho, DW Stuckey, LA Crompton, J Godwin, Y Mishina and TA Rodriguez. (2007). BMP signalling inhibits premature neural differentiation in the mouse embryo. Development 134:3359-69.
153. Liu A and LA Niswander. (2005). Bone morphogenetic protein signalling and vertebrate nervous system development. Nat Rev Neurosci 6:945-54.
154. Kolpakova-Hart E, M Jinnin, B Hou, N Fukai and BR Olsen. (2007). Kinesin-2 controls development and patterning of the vertebrate skeleton by Hedgehog- and Gli3-dependent mechanisms. Dev Biol 309:273-84.
155. Dessaud E, AP McMahon and J Briscoe. (2008). Pattern formation in the vertebrate neural tube: a sonic hedgehog morphogen-regulated transcriptional network. Development 135:2489-503.
156. Sulik K, DB Dehart, T Iangaki, JL Carson, T Vrablic, K Gesteland and GC Schoenwolf. (1994). Morphogenesis of the murine node and notochordal plate. Dev Dyn 201:260-78.
157. Mukhopadhyay M, S Shtrom, C Rodriguez-Esteban, L Chen, T Tsukui, L Gomer, DW Dorward, A Glinka, A Grinberg, SP Huang, C Niehrs, JC Izpisua Belmonte

and H Westphal. (2001). Dickkopf1 is required for embryonic head induction and limb morphogenesis in the mouse. Dev Cell 1:423-34.
158. Kelly OG, KI Pinson and WC Skarnes. (2004). The Wnt co-receptors Lrp5 and Lrp6 are essential for gastrulation in mice. Development 131:2803-15.
159. Muroyama Y, M Fujihara, M Ikeya, H Kondoh and S Takada. (2002). Wnt signaling plays an essential role in neuronal specification of the dorsal spinal cord. Genes Dev 16:548-53.
160. Shimizu T, T Kagawa, T Inoue, A Nonaka, S Takada, H Aburatani and T Taga. (2008). Stabilized beta-catenin functions through TCF/LEF proteins and the Notch/RBP-Jkappa complex to promote proliferation and suppress differentiation of neural precursor cells. Mol Cell Biol 28:7427-41.
161. Oh S, X Huang, J Liu, Y Litingtung and C Chiang. (2009). Shh and Gli3 activities are required for timely generation of motor neuron progenitors. Dev Biol 331:261-9.
162. Wijgerde M, JA McMahon, M Rule and AP McMahon. (2002). A direct requirement for Hedgehog signaling for normal specification of all ventral progenitor domains in the presumptive mammalian spinal cord. Genes Dev 16:2849-64.
163. Henrique D, J Adam, A Myat, A Chitnis, J Lewis and D Ish-Horowicz. (1995). Expression of a Delta homologue in prospective neurons in the chick. Nature 375:787-90.
164. Surface LE, SR Thornton and LA Boyer. (2010). Polycomb group proteins set the stage for early lineage commitment. Cell Stem Cell 7:288-98.
165. Cao R, L Wang, H Wang, L Xia, H Erdjument-Bromage, P Tempst, RS Jones and Y Zhang. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298:1039-43.
166. Czermin B, R Melfi, D McCabe, V Seitz, A Imhof and V Pirrotta. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell 111:185-96.
167. Muller J, CM Hart, NJ Francis, ML Vargas, A Sengupta, B Wild, EL Miller, MB O'Connor, RE Kingston and JA Simon. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell 111:197-208.
168. Kirmizis A, SM Bartley, A Kuzmichev, R Margueron, D Reinberg, R Green and PJ Farnham. (2004). Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. Genes Dev 18:1592-605.
169. Mozzetta C, S Consalvi, V Saccone, SV Forcales, PL Puri and D Palacios. (2011). Selective control of Pax7 expression by TNF-activated p38alpha/polycomb repressive complex 2 (PRC2) signaling during muscle satellite cell differentiation. Cell Cycle 10:191-8.
170. Khromov T, DV Pantakani, J Nolte, M Wolf, R Dressel, W Engel and U Zechner. (2011). Global and gene-specific histone modification profiles of mouse multipotent adult germline stem cells. Mol Hum Reprod 17:166-74.
171. Oguro H, J Yuan, H Ichikawa, T Ikawa, S Yamazaki, H Kawamoto, H Nakauchi and A Iwama. (2010). Poised lineage specification in multipotential hematopoietic stem and progenitor cells by the polycomb protein Bmi1. Cell Stem Cell 6:279-86.
172. Jorgensen HF, S Giadrossi, M Casanova, M Endoh, H Koseki, N Brockdorff and AG Fisher. (2006). Stem cells primed for action: polycomb repressive complexes restrain the expression of lineage-specific regulators in embryonic stem cells. Cell Cycle 5:1411-4.

173. Ku M, RP Koche, E Rheinbay, EM Mendenhall, M Endoh, TS Mikkelsen, A Presser, C Nusbaum, X Xie, AS Chi, M Adli, S Kasif, LM Ptaszek, CA Cowan, ES Lander, H Koseki and BE Bernstein. (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet 4:e1000242.
174. Azuara V, P Perry, S Sauer, M Spivakov, HF Jorgensen, RM John, M Gouti, M Casanova, G Warnes, M Merkenschlager and AG Fisher. (2006). Chromatin signatures of pluripotent cell lines. Nat Cell Biol 8:532-8.
175. Lim DA, YC Huang, T Swigut, AL Mirick, JM Garcia-Verdugo, J Wysocka, P Ernst and A Alvarez-Buylla. (2009). Chromatin remodelling factor Mll1 is essential for neurogenesis from postnatal neural stem cells. Nature 458:529-33.
176. Zhu J, F He, S Hu and J Yu. (2008). On the nature of human housekeeping genes. Trends Genet 24:481-4.
177. Saxonov S, P Berg and DL Brutlag. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci U S A 103:1412-7.
178. Larsen F, G Gundersen, R Lopez and H Prydz. (1992). CpG islands as gene markers in the human genome. Genomics 13:1095-107.
179. Thomson JP, PJ Skene, J Selfridge, T Clouaire, J Guy, S Webb, AR Kerr, A Deaton, R Andrews, KD James, DJ Turner, R Illingworth and A Bird. (2010). CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature 464:1082-6.
180. Pasini D, AP Bracken, MR Jensen, E Lazzerini Denchi and K Helin. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. Embo J 23:4061-71.
181. Montgomery ND, D Yee, A Chen, S Kalantry, SJ Chamberlain, AP Otte and T Magnuson. (2005). The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. Curr Biol 15:942-7.
182. Faust C, A Schumacher, B Holdener and T Magnuson. (1995). The eed mutation disrupts anterior mesoderm production in mice. Development 121:273-85.
183. Shen X, Y Liu, YJ Hsu, Y Fujiwara, J Kim, X Mao, GC Yuan and SH Orkin. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol Cell 32:491-502.
184. Sauvageau M and G Sauvageau. (2010). Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. Cell Stem Cell 7:299-313.
185. Valk-Lingbeek ME, SW Bruggeman and M van Lohuizen. (2004). Stem cells and cancer; the polycomb connection. Cell 118:409-18.
186. Sparmann A and M van Lohuizen. (2006). Polycomb silencers control cell fate, development and cancer. Nat Rev Cancer 6:846-56.
187. Tie F, J Prasad-Sinha, A Birve, A Rasmuson-Lestander and PJ Harte. (2003). A 1-megadalton ESC/E(Z) complex from Drosophila that contains polycomblike and RPD3. Mol Cell Biol 23:3352-62.
188. Savla U, J Benes, J Zhang and RS Jones. (2008). Recruitment of Drosophila Polycomb-group proteins by Polycomblike, a component of a novel protein complex in larvae. Development 135:813-7.
189. Coulson M, S Robert, HJ Eyre and R Saint. (1998). The identification and localization of a human gene with sequence similarity to Polycomblike of Drosophila melanogaster. Genomics 48:381-3.

190. Cao R, H Wang, J He, H Erdjument-Bromage, P Tempst and Y Zhang. (2008). Role of hPHF1 in H3K27 methylation and Hox gene silencing. Mol Cell Biol 28:1862-72.
191. Boyer LA, K Plath, J Zeitlinger, T Brambrink, LA Medeiros, TI Lee, SS Levine, M Wernig, A Tajonar, MK Ray, GW Bell, AP Otte, M Vidal, DK Gifford, RA Young and R Jaenisch. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 441:349-53.
192. Fazzio TG, JT Huff and B Panning. (2008). An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. Cell 134:162-74.
193. Foley E and PH O'Farrell. (2004). Functional dissection of an innate immune response by a genome-wide RNAi screen. PLoS Biol 2:E203.
194. de Napoles M, JE Mermoud, R Wakao, YA Tang, M Endoh, R Appanah, TB Nesterova, J Silva, AP Otte, M Vidal, H Koseki and N Brockdorff. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev Cell 7:663-76.
195. de la Cruz CC, A Kirmizis, MD Simon, K Isono, H Koseki and B Panning. (2007). The polycomb group protein SUZ12 regulates histone H3 lysine 9 methylation and HP1 alpha distribution. Chromosome Res 15:299-314.
196. Rea S, F Eisenhaber, D O'Carroll, BD Strahl, ZW Sun, M Schmid, S Opravil, K Mechtler, CP Ponting, CD Allis and T Jenuwein. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406:593-9.
197. Langmead B, C Trapnell, M Pop and SL Salzberg. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25.
198. Marson A, SS Levine, MF Cole, GM Frampton, T Brambrink, S Johnstone, MG Guenther, WK Johnston, M Wernig, J Newman, JM Calabrese, LM Dennis, TL Volkert, S Gupta, J Love, N Hannett, PA Sharp, DP Bartel, R Jaenisch and RA Young. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell 134:521-33.
199. Mikkelsen TS, M Ku, DB Jaffe, B Issac, E Lieberman, G Giannoukos, P Alvarez, W Brockman, TK Kim, RP Koche, W Lee, E Mendenhall, A O'Donovan, A Presser, C Russ, X Xie, A Meissner, M Wernig, R Jaenisch, C Nusbaum, ES Lander and BE Bernstein. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448:553-60.
200. Tzur G, A Levy, E Meiri, O Barad, Y Spector, Z Bentwich, L Mizrahi, M Katzenellenbogen, E Ben-Shushan, BE Reubinoff and E Galun. (2008). MicroRNA expression patterns and function in endodermal differentiation of human embryonic stem cells. PLoS One 3:e3726.
201. Barber BA and M Rastegar. (2010). Epigenetic control of Hox genes during neurogenesis, development, and disease. Ann Anat 192:261-74.
202. Huang H, C Xie, X Sun, RP Ritchie, J Zhang and YE Chen. (2010). miR-10a contributes to retinoid acid-induced smooth muscle cell differentiation. J Biol Chem 285:9383-9.
203. Tarantino C, G Paolella, L Cozzuto, G Minopoli, L Pastore, S Parisi and T Russo. (2010). miRNA 34a, 100, and 137 modulate differentiation of mouse embryonic stem cells. Faseb J 24:3255-63.
204. Kim YJ, SW Bae, SS Yu, YC Bae and JS Jung. (2009). miR-196a regulates proliferation and osteogenic differentiation in mesenchymal stem cells derived from human adipose tissue. J Bone Miner Res 24:816-25.

205. Brown JL, M Snir, H Noushmehr, M Kirby, SK Hong, AG Elkahloun and B Feldman. (2008). Transcriptional profiling of endogenous germ layer precursor cells identifies dusp4 as an essential gene in zebrafish endoderm specification. Proc Natl Acad Sci U S A 105:12337-42.
206. Mallo M, DM Wellik and J Deschamps. (2010). Hox genes and regional patterning of the vertebrate body plan. Dev Biol 344:7-15.
207. Ucar A, V Vafaizadeh, H Jarry, J Fiedler, PA Klemmt, T Thum, B Groner and K Chowdhury. (2010). miR-212 and miR-132 are required for epithelial stromal interactions necessary for mouse mammary gland development. Nat Genet 42:1101-8.
208. Sato F, S Tsuchiya, SJ Meltzer and K Shimizu. (2011). MicroRNAs and epigenetics. Febs J.
209. Pizzatti L, R Binato, J Cofre, BE Gomes, J Dobbin, ME Haussmann, D D'Azambuja, LF Bouzas and E Abdelhay. (2010). SUZ12 is a candidate target of the non-canonical WNT pathway in the progression of chronic myeloid leukemia. Genes Chromosomes Cancer 49:107-18.
210. Kirmizis A, SM Bartley and PJ Farnham. (2003). Identification of the polycomb group protein SU(Z)12 as a potential molecular target for human cancer therapy. Mol Cancer Ther 2:113-21.
211. Casanova M, T Preissner, A Cerase, R Poot, D Yamada, X Li, R Appanah, K Bezstarosti, J Demmers, H Koseki and N Brockdorff. (2011). Polycomblike 2 facilitates the recruitment of PRC2 Polycomb group complexes to the inactive X chromosome and to target loci in embryonic stem cells. Development.
212. Tanay A, AH O'Donnell, M Damelin and TH Bestor. (2007). Hyperconserved CpG domains underlie Polycomb-binding sites. Proc Natl Acad Sci U S A 104:5521-6.
213. Zhao J, BK Sun, JA Erwin, JJ Song and JT Lee. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 322:750-6.
214. Gupta RA, N Shah, KC Wang, J Kim, HM Horlings, DJ Wong, MC Tsai, T Hung, P Argani, JL Rinn, Y Wang, P Brzoska, B Kong, R Li, RB West, MJ van de Vijver, S Sukumar and HY Chang. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 464:1071-6.
215. Tsai MC, O Manor, Y Wan, N Mosammaparast, JK Wang, F Lan, Y Shi, E Segal and HY Chang. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science 329:689-93.
216. Sprangers R, MR Groves, I Sinning and M Sattler. (2003). High-resolution X-ray and NMR structures of the SMN Tudor domain: conformational variation in the binding site for symmetrically dimethylated arginine residues. J Mol Biol 327:507-20.
217. Guenther MG, SS Levine, LA Boyer, R Jaenisch and RA Young. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell 130:77-88.
218. Ohm JE, KM McGarvey, X Yu, L Cheng, KE Schuebel, L Cope, HP Mohammad, W Chen, VC Daniel, W Yu, DM Berman, T Jenuwein, K Pruitt, SJ Sharkis, DN Watkins, JG Herman and SB Baylin. (2007). A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 39:237-42.
219. Widschwendter M, H Fiegl, D Egle, E Mueller-Holzner, G Spizzo, C Marth, DJ Weisenberger, M Campan, J Young, I Jacobs and PW Laird. (2007). Epigenetic stem cell signature in cancer. Nat Genet 39:157-8.

220. Chase A and NC Cross. (2011). Aberrations of EZH2 in cancer. Clin Cancer Res 17:2613-8.
221. Shechter D, HL Dormann, CD Allis and SB Hake. (2007). Extraction, purification and analysis of histones. Nat Protoc 2:1445-57.
222. Hunkapiller J, V Singla, A Seol and JF Reiter. (2010). The Ciliogenic Protein Ofd1 Regulates the Neuronal Differentiation of Embryonic Stem Cells. Stem Cells Dev.
223. Bolstad BM, RA Irizarry, M Astrand and TP Speed. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19:185-93.
224. Hawkins RD, GC Hon, LK Lee, Q Ngo, R Lister, M Pelizzola, LE Edsall, S Kuan, Y Luu, S Klugman, J Antosiewicz-Bourget, Z Ye, C Espinoza, S Agarwahl, L Shen, V Ruotti, W Wang, R Stewart, JA Thomson, JR Ecker and B Ren. (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. Cell Stem Cell 6:479-91.
225. Burrows PM. (1972). Expected selection differentials for directional selection. Biometrics 28:1091-100.
226. Yang J, L Chai, TC Fowles, Z Alipio, D Xu, LM Fink, DC Ward and Y Ma. (2008). Genome-wide analysis reveals Sall4 to be a major regulator of pluripotency in murine-embryonic stem cells. Proc Natl Acad Sci U S A 105:19756-61.
227. Bowen NJ, N Fujita, M Kajita and PA Wade. (2004). Mi-2/NuRD: multiple complexes for many purposes. Biochim Biophys Acta 1677:52-7.
228. Denslow SA and PA Wade. (2007). The human Mi-2/NuRD complex and gene regulation. Oncogene 26:5433-8.
229. Sakaki-Yumoto M, C Kobayashi, A Sato, S Fujimura, Y Matsumoto, M Takasato, T Kodama, H Aburatani, M Asashima, N Yoshida and R Nishinakamura. (2006). The murine homolog of SALL4, a causative gene in Okihiro syndrome, is essential for embryonic stem cell proliferation, and cooperates with Sall1 in anorectal, heart, brain and kidney development. Development 133:3005-13.
230. Rao S, S Zhen, S Roumiantsev, LT McDonald, GC Yuan and SH Orkin. (2010). Differential roles of Sall4 isoforms in embryonic stem cell pluripotency. Mol Cell Biol 30:5364-80.
231. Hendrich B, J Guy, B Ramsahoye, VA Wilson and A Bird. (2001). Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes Dev 15:710-23.
232. Martin Caballero I, J Hansen, D Leaford, S Pollard and BD Hendrich. (2009). The methyl-CpG binding proteins Mecp2, Mbd2 and Kaiso are dispensable for mouse embryogenesis, but play a redundant function in neural differentiation. PLoS One 4:e4315.
233. Lauberth SM and M Rauchman. (2006). A conserved 12-amino acid motif in Sall1 recruits the nucleosome remodeling and deacetylase corepressor complex. J Biol Chem 281:23922-31.
234. Shevchenko A, M Wilm, O Vorm and M Mann. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem 68:850-8.
235. Ren X, J Zhang, X Gong, X Niu, X Zhang, P Chen and X Zhang. (2010). Differentiation of murine embryonic stem cells toward renal lineages by conditioned medium from ureteric bud cells in vitro. Acta Biochim Biophys Sin (Shanghai) 42:464-71.
236. Jiang X, Y Gwye, SJ McKeown, M Bronner-Fraser, C Lutzko and ER Lawlor. (2009). Isolation and characterization of neural crest stem cells derived from in vitro-differentiated human embryonic stem cells. Stem Cells Dev 18:1059-70.

237. Lee G, SM Chambers, MJ Tomishima and L Studer. (2010). Derivation of neural crest cells from human pluripotent stem cells. Nat Protoc 5:688-701.
238. Aihara Y, Y Hayashi, M Hirata, N Ariki, S Shibata, N Nagoshi, M Nakanishi, K Ohnuma, M Warashina, T Michiue, H Uchiyama, H Okano, M Asashima and MK Furue. (2010). Induction of neural crest cells from mouse embryonic stem cells in a serum-free monolayer culture. Int J Dev Biol 54:1287-94.
239. Iliopoulos D, M Lindahl-Allen, C Polytarchou, HA Hirsch, PN Tsichlis and K Struhl. (2010). Loss of miR-200 inhibition of Suz12 leads to polycomb-mediated repression required for the formation and maintenance of cancer stem cells. Mol Cell 39:761-72.
240. Crea F, EM Hurt, LA Mathews, SM Cabarcas, L Sun, VE Marquez, R Danesi and WL Farrar. (2011). Pharmacologic disruption of Polycomb Repressive Complex 2 inhibits tumorigenicity and tumor progression in prostate cancer. Mol Cancer 10:40.
241. Tsang DP and AS Cheng. (2011). Epigenetic regulation of signaling pathways in cancer: role of the histone methyltransferase EZH2. J Gastroenterol Hepatol 26:19-27.
242. Thomson JA, J Itskovitz-Eldor, SS Shapiro, MA Waknitz, JJ Swiergiel, VS Marshall and JM Jones. (1998). Embryonic stem cell lines derived from human blastocysts. Science 282:1145-7.
243. Brustle O, AC Spiro, K Karram, K Choudhary, S Okabe and RD McKay. (1997). In vitro-generated neural precursors participate in mammalian brain development. Proc Natl Acad Sci U S A 94:14809-14.
244. Wichterle H, JM Garcia-Verdugo, DG Herrera and A Alvarez-Buylla. (1999). Young neurons from medial ganglionic eminence disperse in adult and embryonic brain. Nat Neurosci 2:461-6.
245. Temple S. (2001). The development of neural stem cells. Nature 414:112-7.
246. Zhang SC. (2006). Neural subtype specification from embryonic stem cells. Brain Pathol 16:132-42.
247. Merkle FT and A Alvarez-Buylla. (2006). Neural stem cells in mammalian development. Curr Opin Cell Biol 18:704-9.
248. Spemann HMH. (1924). Über Induktion von Embryonanlagen durch Implantation artfremder Organisatoren. Wilhelm Roux Arch. Entw. Mech. Organ. 100:599-638.
249. Waddington CH. (1932). Experiments on the Development of Chick and Duck Embryos, Cultivated in vitro. Phil. Trans. Roy. Soc. Lond. B.:179-230.
250. Oppenheimer J. (1936). Structures developed in amphibians by implantation of living fish organizer. Proc. Soc. Exp. Biol. Med. 34:461-463.
251. Lamb TM, AK Knecht, WC Smith, SE Stachel, AN Economides, N Stahl, GD Yancopolous and RM Harland. (1993). Neural induction by the secreted polypeptide noggin. Science 262:713-8.
252. Zimmerman LB, JM De Jesus-Escobar and RM Harland. (1996). The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4. Cell 86:599-606.
253. Sasai Y, B Lu, H Steinbeisser, D Geissert, LK Gont and EM De Robertis. (1994). Xenopus chordin: a novel dorsalizing factor activated by organizer-specific homeobox genes. Cell 79:779-90.
254. Born J, J Janeczek, W Schwarz, H Tiedemann and H Tiedemann. (1989). Activation of masked neural determinants in amphibian eggs and embryos and their release from the inducing tissue. Cell Differ Dev 27:1-7.
255. Godsave SF and JM Slack. (1989). Clonal analysis of mesoderm induction in Xenopus laevis. Dev Biol 134:486-90.

256. Grunz H and L Tacke. (1989). Neural differentiation of Xenopus laevis ectoderm takes place after disaggregation and delayed reaggregation without inducer. Cell Differ Dev 28:211-7.
257. Barth LG. (1941). Neural differentiation without organizer. J. Exp. Zool.:371-384.
258. Brunet LJ, JA McMahon, AP McMahon and RM Harland. (1998). Noggin, cartilage morphogenesis, and joint formation in the mammalian skeleton. Science 280:1455-7.
259. Matzuk MM, N Lu, H Vogel, K Sellheyer, DR Roop and A Bradley. (1995). Multiple defects and perinatal death in mice deficient in follistatin. Nature 374:360-3.
260. Gazzerro E and E Canalis. (2006). Bone morphogenetic proteins and their antagonists. Rev Endocr Metab Disord 7:51-65.
261. Khokha MK, D Hsu, LJ Brunet, MS Dionne and RM Harland. (2003). Gremlin is the BMP antagonist required for maintenance of Shh and Fgf signals during limb patterning. Nat Genet 34:303-7.
262. Reversade B, H Kuroda, H Lee, A Mays and EM De Robertis. (2005). Depletion of Bmp2, Bmp4, Bmp7 and Spemann organizer signals induces massive brain formation in Xenopus embryos. Development 132:3381-92.
263. Kuroda H, L Fuentealba, A Ikeda, B Reversade and EM De Robertis. (2005). Default neural induction: neuralization of dissociated Xenopus cells is mediated by Ras/MAPK activation. Genes Dev 19:1022-7.
264. Stern CD. (2006). Neural induction: 10 years on since the 'default model'. Curr Opin Cell Biol 18:692-7.
265. Pelegri F and HM Maischein. (1998). Function of zebrafish beta-catenin and TCF-3 in dorsoventral patterning. Mech Dev 77:63-74.
266. Nojima H, T Shimizu, CH Kim, T Yabe, YK Bae, O Muraoka, T Hirata, A Chitnis, T Hirano and M Hibi. (2004). Genetic evidence for involvement of maternally derived Wnt canonical signaling in dorsal determination in zebrafish. Mech Dev 121:371-86.
267. Wessely O, E Agius, M Oelgeschlager, EM Pera and EM De Robertis. (2001). Neural induction in the absence of mesoderm: beta-catenin-dependent expression of secreted BMP antagonists at the blastula stage in Xenopus. Dev Biol 234:161-73.
268. Tao Q, C Yokota, H Puck, M Kofron, B Birsoy, D Yan, M Asashima, CC Wylie, X Lin and J Heasman. (2005). Maternal wnt11 activates the canonical wnt signaling pathway required for axis formation in Xenopus embryos. Cell 120:857-71.
269. Erter CE, TP Wilm, N Basler, CV Wright and L Solnica-Krezel. (2001). Wnt8 is required in lateral mesendodermal precursors for neural posteriorization in vivo. Development 128:3571-83.
270. Lekven AC, CJ Thorpe, JS Waxman and RT Moon. (2001). Zebrafish wnt8 encodes two wnt8 proteins on a bicistronic transcript and is required for mesoderm and neurectoderm patterning. Dev Cell 1:103-14.
271. Baker JC, RS Beddington and RM Harland. (1999). Wnt signaling in Xenopus embryos inhibits bmp4 expression and activates neural development. Genes Dev 13:3149-59.
272. Heeg-Truesdell E and C LaBonne. (2006). Neural induction in Xenopus requires inhibition of Wnt-beta-catenin signaling. Dev Biol 298:71-86.
273. Wilson SI, A Rydstrom, T Trimborn, K Willert, R Nusse, TM Jessell and T Edlund. (2001). The status of Wnt signalling regulates neural and epidermal fates in the chick embryo. Nature 411:325-30.

274. Yoshikawa Y, T Fujimori, AP McMahon and S Takada. (1997). Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm development in the mouse. Dev Biol 183:234-42.
275. Ciruna BG, L Schwartz, K Harpal, TP Yamaguchi and J Rossant. (1997). Chimeric analysis of fibroblast growth factor receptor-1 (Fgfr1) function: a role for FGFR1 in morphogenetic movement through the primitive streak. Development 124:2829-41.
276. Hurtado C and EM De Robertis. (2007). Neural induction in the absence of organizer in salamanders is mediated by MAPK. Dev Biol 307:282-9.
277. Lamb TM and RM Harland. (1995). Fibroblast growth factor is a direct neural inducer, which combined with noggin generates anterior-posterior neural pattern. Development 121:3627-36.
278. Storey KG, A Goriely, CM Sargent, JM Brown, HD Burns, HM Abud and JK Heath. (1998). Early posterior neural tissue is induced by FGF in the chick embryo. Development 125:473-84.
279. Alvarez IS, M Araujo and MA Nieto. (1998). Neural induction in whole chick embryo cultures by FGF. Dev Biol 199:42-54.
280. Kengaku M and H Okamoto. (1993). Basic fibroblast growth factor induces differentiation of neural tube and neural crest lineages of cultured ectoderm cells from Xenopus gastrula. Development 119:1067-78.
281. Pera EM, A Ikeda, E Eivers and EM De Robertis. (2003). Integration of IGF, FGF, and anti-BMP signals via Smad1 phosphorylation in neural induction. Genes Dev 17:3023-8.
282. Fuentealba LC, E Eivers, A Ikeda, C Hurtado, H Kuroda, EM Pera and EM De Robertis. (2007). Integrating Patterning Signals: Wnt/GSK3 Regulates the Duration of the BMP/Smad1 Signal. Cell 131:980-93.
283. Aubin J, A Davy and P Soriano. (2004). In vivo convergence of BMP and MAPK signaling pathways: impact of differential Smad1 phosphorylation on development and homeostasis. Genes Dev 18:1482-94.
284. Wilson SI, E Graziano, R Harland, TM Jessell and T Edlund. (2000). An early requirement for FGF signalling in the acquisition of neural cell fate in the chick embryo. Curr Biol 10:421-9.
285. Furthauer M, J Van Celst, C Thisse and B Thisse. (2004). Fgf signalling controls the dorsoventral patterning of the zebrafish embryo. Development 131:2853-64.
286. Delaune E, P Lemaire and L Kodjabachian. (2005). Neural induction in Xenopus requires early FGF signalling in addition to BMP inhibition. Development 132:299-310.
287. Linker C and CD Stern. (2004). Neural induction requires BMP inhibition only as a late step, and involves signals other than FGF and Wnt antagonists. Development 131:5671-81.
288. Bain G, D Kitchens, M Yao, JE Huettner and DI Gottlieb. (1995). Embryonic stem cells express neuronal properties in vitro. Dev Biol 168:342-57.
289. Schuldiner M, O Yanuka, J Itskovitz-Eldor, DA Melton and N Benvenisty. (2000). Effects of eight growth factors on the differentiation of cells derived from human embryonic stem cells. Proc Natl Acad Sci U S A 97:11307-12.
290. Rohwedel J, K Guan and AM Wobus. (1999). Induction of cellular differentiation by retinoic acid in vitro. Cells Tissues Organs 165:190-202.
291. Finley MF, S Devata and JE Huettner. (1999). BMP-4 inhibits neural differentiation of murine embryonic stem cells. J Neurobiol 40:271-87.

292. Kawasaki H, K Mizuseki, S Nishikawa, S Kaneko, Y Kuwana, S Nakanishi, SI Nishikawa and Y Sasai. (2000). Induction of midbrain dopaminergic neurons from ES cells by stromal cell-derived inducing activity. Neuron 28:31-40.
293. Ying QL, M Stavridis, D Griffiths, M Li and A Smith. (2003). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. Nat Biotechnol 21:183-6.
294. Tropepe V, S Hitoshi, C Sirard, TW Mak, J Rossant and D van der Kooy. (2001). Direct neural fate specification from embryonic stem cells: a primitive mammalian neural stem cell stage acquired through a default mechanism. Neuron 30:65-78.
295. Xu RH, X Chen, DS Li, R Li, GC Addicks, C Glennon, TP Zwaka and JA Thomson. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat Biotechnol 20:1261-4.
296. Sonntag KC, R Simantov, L Bjorklund, O Cooper, J Pruszak, F Kowalke, J Gilmartin, J Ding, YP Hu, MM Shen and O Isacson. (2005). Context-dependent neuronal differentiation and germ layer induction of Smad4-/- and Cripto-/- embryonic stem cells. Mol Cell Neurosci 28:417-29.
297. Smith JR, L Vallier, G Lupo, M Alexander, WA Harris and RA Pedersen. (2008). Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm. Dev Biol 313:107-17.
298. Vallier L, D Reynolds and RA Pedersen. (2004). Nodal inhibits differentiation of human embryonic stem cells along the neuroectodermal default pathway. Dev Biol 275:403-21.
299. Camus A, A Perea-Gomez, A Moreau and J Collignon. (2006). Absence of Nodal signaling promotes precocious neural differentiation in the mouse embryo. Dev Biol 295:743-55.
300. Stavridis MP, JS Lunn, BJ Collins and KG Storey. (2007). A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. Development 134:2889-94.
301. Conley BJ, S Ellis, L Gulluyan and R Mollard. (2007). BMPs regulate differentiation of a putative visceral endoderm layer within human embryonic stem-cell-derived embryoid bodies. Biochem Cell Biol 85:121-32.
302. Pera MF, J Andrade, S Houssami, B Reubinoff, A Trounson, EG Stanley, D Ward-van Oostwaard and C Mummery. (2004). Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin. J Cell Sci 117:1269-80.
303. Smukler SR, SB Runciman, S Xu and D van der Kooy. (2006). Embryonic stem cells assume a primitive neural stem cell fate in the absence of extrinsic influences. J Cell Biol 172:79-90.
304. Sonntag KC, J Pruszak, T Yoshizaki, J van Arensbergen, R Sanchez-Pernaute and O Isacson. (2007). Enhanced yield of neuroepithelial precursors and midbrain-like dopaminergic neurons from human embryonic stem cells using the bone morphogenic protein antagonist noggin. Stem Cells 25:411-8.
305. Brons IG, LE Smithers, MW Trotter, P Rugg-Gunn, B Sun, SM Chuva de Sousa Lopes, SK Howlett, A Clarkson, L Ahrlund-Richter, RA Pedersen and L Vallier. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature 448:191-5.
306. Tesar PJ, JG Chenoweth, FA Brook, TJ Davies, EP Evans, DL Mack, RL Gardner and RD McKay. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature 448:196-9.

307. Aubert J, H Dunstan, I Chambers and A Smith. (2002). Functional gene screening in embryonic stem cells implicates Wnt antagonism in neural differentiation. Nat Biotechnol 20:1240-5.
308. Verani R, I Cappuccio, P Spinsanti, R Gradini, A Caruso, MC Magnotti, M Motolese, F Nicoletti and D Melchiorri. (2007). Expression of the Wnt inhibitor Dickkopf-1 is required for the induction of neural markers in mouse embryonic stem cells differentiating in response to retinoic acid. J Neurochem 100:242-50.
309. Kim WY, X Wang, Y Wu, BW Doble, S Patel, JR Woodgett and WD Snider. (2009). GSK-3 is a master regulator of neural progenitor homeostasis. Nat Neurosci.
310. Lindsley RC, JG Gill, M Kyba, TL Murphy and KM Murphy. (2006). Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. Development 133:3787-96.
311. Otero JJ, W Fu, L Kan, AE Cuadra and JA Kessler. (2004). Beta-catenin signaling is required for neural differentiation of embryonic stem cells. Development 131:3545-57.
312. Ying QL, J Nichols, I Chambers and A Smith. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell 115:281-92.
313. Zhang SC, M Wernig, ID Duncan, O Brustle and JA Thomson. (2001). In vitro differentiation of transplantable neural precursors from human embryonic stem cells. Nat Biotechnol 19:1129-33.
314. Pera EM, O Wessely, SY Li and EM De Robertis. (2001). Neural and head induction by insulin-like growth factor signals. Dev Cell 1:655-65.
315. Di Giorgio FP, MA Carrasco, MC Siao, T Maniatis and K Eggan. (2007). Non-cell autonomous effect of glia on motor neurons in an embryonic stem cell-based ALS model. Nat Neurosci 10:608-14.
316. Martinat C, S Shendelman, A Jonason, T Leete, MF Beal, L Yang, T Floss and A Abeliovich. (2004). Sensitivity to oxidative stress in DJ-1-deficient dopamine neurons: an ES- derived cell model of primary Parkinsonism. PLoS Biol 2:e327.
317. Lorincz MT, PJ Detloff, RL Albin and KS O'Shea. (2004). Embryonic stem cells expressing expanded CAG repeats undergo aberrant neuronal differentiation and have persistent Oct-4 and REST/NRSF expression. Mol Cell Neurosci 26:135-43.

Appendix: Neural induction and differentiation during embryonic development and in embryonic stem cells

Julie Hunkapiller (nee Gaulden) and Jeremy F. Reiter. (2008). Neur-ons and Neur-offs: Regulators of neural induction in vertebrate embryos and embryonic stem cells. *Human Molecular Genetics*, 17 (R1): R60-66.

ABSTRACT

Although the spatial and temporal orchestration of early vertebrate embryogenesis is missing from cell culture systems, recent work suggests that many of the same signals affecting neural induction in vertebrate embryos also regulate embryonic stem (ES) cell neurogenesis. One key regulatory mechanism involved in both *in vivo* and *in vitro* neural induction is the inhibition of BMP signals. Wnts and FGFs represent additional regulatory influences which may affect the adoption of neural fates through both BMP-dependent and -independent mechanisms. Insights into neural induction *in vivo* help to guide paradigms for promoting neural differentiation by ES cells. Conversely, insights into the mechanisms by which ES cells adopt neural fates may provide an improved understanding of neural induction in the early embryo.

INTRODUCTION

It is a truism that ES cells hold tremendous promise as an inexhaustible source for novel cell-based therapies for diseases for which treatments are imperfect or nonexistent. Work over the past twenty-five years has aimed at understanding the properties of stem cells, especially the means to derive specific differentiated cell types

from stem cells. One of the first lineages to be generated from ES cells was neural [1,2,242], and many subsequent studies have succeeded in promoting the differentiation of ES cells towards the neural lineage to create various kinds of neuronal and glial subtypes. This work reached an important milestone with the demonstration that neural cells derived from ES cells can be transplanted back into vertebrate embryos and integrate into many areas of the central nervous system [243,244].

Beyond the therapeutic implications, neural differentiation by ES cells can also provide insights into the mechanisms of neural induction, the initial steps by which uncommitted cells begin differentiating along the neural lineage. Several recent reviews have focused on the properties of neural stem cells and neural differentiation along defined pathways [245-247]. Here, we contrast neural induction in vertebrate model organisms with neural induction in ES cells. Similarities and differences between *in vivo* and *in vitro* neural induction may lend insights into the molecular signals that direct neural induction in the embryo. Additionally, a better understanding of the signaling and tissue dynamics that underlie neural induction *in vivo* may lead to more efficient ways of deriving neural subtypes *in vitro.*

Neural induction through BMP antagonism

Neural induction has fascinated scientists since Spemann and Mangold established the presence of an organizer in amphibians capable of conferring neural identity to ectoderm [248]. The presence of similar organizers in other species, such as the embryonic shield in teleosts, Hensen's node in birds, and the node in mice, suggests that the mechanisms by which the organizer imparts neural identity are at least partly conserved in vertebrates [148,249,250]. The first clues to the molecular nature of the organizer's inductive influence came from studies of a receptor for transforming growth factor-beta (TGFβ) superfamily members, and Noggin, a secreted factor expressed by

the organizer [149,150,251]. The ectodermal region of the early Xenopus embryo animal pole normally gives rise to both epidermis and neurectoerm (Figure 30A). However, if this region is physically isolated as an animal cap, it does not form neural tissue, though expression of either a dominant negative form of a TGFβ receptor, or Noggin is sufficient to generate neural cells in Xenopus animal caps [149,150,251]. Noggin, as well as other factors that can induce neural fates in animal caps such as Chordin and Follistatin, share the ability to bind to and inhibit the activity of Bone Morphogenetic Proteins (BMPs), TGFβ superfamily members [149,252,253]. The finding that these factors specify neural identity by inhibiting the repressive influence of BMPs suggests that it is the absence of repressive influences, and not an active promoting influence, that may be essential for neural induction [149]. This so-called default model of neural induction is further supported by the finding that dissociated Xenopus animal caps, deprived of extensive intercellular contacts and communication, adopt neural fates [208,254-256] and that isolated animal caps of some amphibians differentiate into neural tissue [257].

Given the central role that inhibition of BMP plays in neural induction, it is somewhat surprising that mouse embryos mutant for *Chordin*, *Noggin*, or *Follistatin* do not display gross defects in neural plate formation [151,258,259]. Indeed, individual loss of these BMP inhibitors, as well as others including *Twisted gastrulation*, *Gremlin* and *Cerberus-like*, does not dramatically affect neural allocation in mice [151,260,261]. However, loss of multiple BMP antagonists, such as both Chordin and Noggin, results in neural deficits, indicating functional overlap among BMP inhibitors [151]. Consistent with this interpretation, Xenopus or zebrafish embryos depleted of multiple BMPs, or mouse embryos lacking *BMPR1a*, a component of the principle BMP receptor during early embryonic development, display increased and early neural differentiation [152,262]. The similarity of these phenotypes suggests that inhibition of BMP activity through the

overlapping functions of several BMP antagonists and the resulting BMP activity gradient play a critical role in vertebrate neural induction (Figure 30).

Beyond the Default Model

The default model may be too simple to completely describe neural induction [263,264]. For example, additional influences, such as Wnt signals, have complex roles in promoting, inhibiting, and patterning neural tissue at discrete times during embryogenesis. Xenopus and zebrafish both require an initial activation of the β-catenin-dependent Wnt signaling pathway during the blastula stage for specification of the organizer, and therefore for subsequent BMP antagonism [158,265-267]. In Xenopus, this early Wnt signaling depends on Wnt11 [268].

While this initial Wnt signaling is required early in embryo development for subsequent neural induction, both the Wnt and BMP pathways must be inhibited later in development for neural induction to occur, consistent with the central tenant of the default model. At gastrulation stages, Wnt signals collaborate with BMPs to ventralize frog and fish embryos and inhibit neural fates (Figure 30) [269-272]. Similarly, Wnt signaling inhibits neural fates in the chick epiblast, and inhibition of Wnt signaling is sufficient to induce neural differentiation in regions normally fated to become epidermis [273].

This inhibitory role for Wnts may be conserved in mammals, as genetic studies support an inhibitory function for Wnt signaling in mouse neural induction. Mouse mutants lacking effectors of β-catenin-dependent Wnt signaling, such as *Wnt3a,* display increased neural tissue and even ectopic neural tubes [274]. Similarly, mutation of the Wnt co-receptors, *Lrp5* and *Lrp6,* results in an expansion of anterior neuroectoderm [158]. Consistent with this negative regulatory role, loss of *Dickkopf (Dkk)*, a Wnt inhibitor, prevents forebrain development [157].

However, mammalian Wnt signaling is involved in both anteroposterior patterning and cell movement in the primitive streak, raising the possibility that the neural phenotypes of Wnt pathway mutants may reflect indirect consequences of disregulated Wnt signaling. Indeed, ectopic neural tubes similar to those caused by loss of Wnt3a can result from defective cell movement in the primitive streak [275].

In addition to BMP inhibitors such as Chordin, Noggin, and Follistatin, the Xenopus organizer expresses Frzb-1, Crescent, SFRP2, Dkk1, Cerberus, and Lefty (Figure 30), inhibitors of the Wnt and Nodal pathways [151]. The production of so many signaling pathway inhibitors by the organizer substantiates the underlying supposition of the default model. So, is there no role for positive signaling in neural induction?

At least in non-mammalian vertebrates, there is substantial evidence that activation of the mitogen-activated protein kinase (MAPK) pathway promotes neural induction. For example, neural induction in isolated salamander animal caps depends on the MAPK pathway [276], and the members of one class of MAPK activators, Fibroblast Growth Factors (FGFs), induce early neural plate markers, such as Sox3 and ERNI [277-280]. In support of the involvement of FGFs in neural induction, FGFs can act cooperatively with BMP inhibition to promote Xenopus neural induction [262]. One possible molecular mechanism for this functional cooperation is through MAPK pathway convergence on BMP signaling through the differential phosphorylation of Smad1, an important BMP effector [281]. GSK3, a component and inhibitor of the Wnt pathway, promotes additional phosphorylation and degradation of Smad1 after a priming phosphorylation by MAPK, which may similarly explain the anti-neuralizing properties of Wnts [282]. However, mutation of the MAPK phosphorylation site of Smad1 does not overtly abrogate neural induction in mice, suggesting that this interaction is not essential for the effects of FGFs on neural induction [283]. Another possibility is that early FGF signals downregulate expression of BMPs in the prospective neural domain, allowing

neural differentiation to proceed [284,285]. A third possibility is that early FGF signaling promotes neural fate through a parallel, BMP-independent mechanism [286,287]. One BMP-independent mechanism may involve Wnts which, as described above, can inhibit neural induction during gastrulation stages [273]. However, addition of multiple FGFs, BMP antagonists, and Wnt antagonists to the chick embryonic epiblast is not sufficient to induce the expression of the neural marker Sox2, suggesting that still other pathways regulate neural induction [287].

Although these data from amphibians and birds suggest that FGFs are positive effectors of vertebrate neural induction, to date there has not been genetic confirmation in fish or mice. For example, mouse epiblast cells lacking the FGF receptor, *Fgfr1,* can adopt neural fates [275]. Also, inhibition of FGF signaling in the pre-gastrula mouse embryo does not block neural induction [152]. It is possible that the involvement of an early FGF signal in neural induction does not pertain to mammals, or may be redundant with other inputs into the MAPK pathway.

ES cells: "All cases are unique, and very similar to others"

Partially due to ease of manipulation, ES cells represent an attractive model for studying early embryonic differentiation in general, and neural induction in particular. ES cells can be differentiated spontaneously to embryoid bodies or along specific lineages by addition of extrinsic factors or by genetic manipulation.

The relevance of ES cell differentiation to *in vivo* events is often unclear. For example, retinoic acid is one of the most potent inducers of neural differentiation in both mouse and human embryonic stem cells [288,289], but there is little evidence for a role for retinoic acid in embryonic neural induction. Nevertheless, in ES cells, retinoic acid has dosage dependent effects whereby high concentrations drive ES cells towards a neural lineage and low retinoic acid concentrations cause cardiac differentiation [290].

In contrast, other neural influences are conserved between ES cells and *in vivo* systems (Figure 31). Given that mouse and human ES cells have markedly different growth factor requirements for maintaining pluripotency, it is interesting that they both respond relatively similarly to BMP and BMP inhibitors with regards to neural induction. In both mouse and human ES cells, high level BMP signaling blocks neural differentiation [291-295]. For these reasons, most neural differentiation protocols require that ES cells be grown in low or no serum conditions to mitigate the effect of BMPs found in serum. Further evidence that BMPs inhibit ES cell neural induction comes from a recent study showing that mouse ES cells lacking Smad4 preferentially differentiate down the neural lineage [296]. As Smad4 is required for all signaling by TGFβ superfamily members, and not just by BMPs, it is possible that this finding reflects inhibitory influences by Nodal or other TGFβ superfamily members in addition to BMPs. Indeed, Nodal signaling may inhibit neural differentiation as mouse ES cells mutant for *Cripto*, a Nodal co-receptor, exhibit enhanced neural differentiation, and human ES cells overexpressing *Lefty* or a truncated version of *Cerberus*, Nodal pathway inhibitors, display increased neural induction [296-298]. Also, mouse *Nodal* mutants show early and ectopic adoption of neural fates [299].

While it is clear that BMPs inhibit neural differentiation *in vitro*, it is ambiguous whether adding BMP inhibitors increases neural induction beyond that caused by serum starvation. For example, treatment of mouse ES cells with Noggin does not affect neural induction in ES cells differentiated as embryoid bodies or as an adherent monolayer [291,293,300,301]. Moreover, neither addition of Chordin nor Follistatin significantly affects ES cell neural differentiation [292,302]. On the other hand, Noggin increases the differentiation of mouse ES cells to neurospheres [294] and can enhance neural differentiation by mouse and human ES cells [12,302-304].

The requirement for BMPs in maintaining mouse ES cell pluripotency is not shared by human ES cells, for which even low level BMP signaling promotes trophoblast differentiation [295]. These different requirements may reflect different cells of origin for mouse and human ES cells; most mouse ES cell lines are derived at the late inner cell mass stage, whereas human ES cell lines are derived at the blastocyst stage. Indeed, mouse ES cells that, like human ES cells, are derived from blastocysts require growth factors and display gene expression profiles similar to human ES cells [305,306].

Like the BMPs, Wnts appear to affect neural induction by ES cells in ways similar to those observed *in vivo*. For example, activating the Wnt pathway in differentiating mouse ES cells, either with Wnt1 or with an inhibitor of GSK3, inhibits neural differentiation [307]. Consistent with a negative role for Wnts in neural induction, genetic manipulations that increase Wnt signaling, such as loss of the extracellular Wnt inhibitor Dkk1, or both GSK3α and GSK3β, abrogate neural differentiation [308,309]. Conversely, treatment with Dkk1, or overexpression of another Wnt inhibitor, SFRP2, enhances mouse ES cell neural differentiation [307,310]. These data suggest that high level Wnt signaling prevents neural differentiation whereas the inhibition of Wnt signaling enhances neural differentiation, closely paralleling *in vivo* mouse models.

However, others have found that inducing Wnt signaling can increase neural differentiation in mouse ES cells [311]. The discrepant effects of both Wnts and BMP inhibitors among various protocols may be attributable to differences in culture media, which can contain BMPs, Wnts and/or additional signaling inhibitors. Another possible explanation is that BMP inhibitors or Wnts may influence neural differentiation only under specific conditions, such as the presence of FGF signaling or the absence of Nodal signaling. A third possibility may reflect the complex involvement of these factors in both ES maintenance and differentiation. For example, in addition to inhibiting neural induction, intermediate levels of BMP signaling maintain mouse ES cell pluripotency

[312]. Similarly, Wnt signaling may promote ES cell pluripotency [208]. Because some of the factors required for ES cell pluripotency and survival are the same factors that promote neural fates, it is difficult to discern if these factors act primarily to induce neural differentiation or to simply enhance progenitor survival or proliferation. This uncertainty reflects the artificiality inherent in ES cell culture systems, as maintenance of pluripotency does not occur outside of the germline *in vivo.*

Given the diverse roles that FGFs play in regulating proliferation, self-renewal, survival, and differentiation, it is perhaps unsurprising that FGFs similarly have complex effects on ES cell neural induction. FGF2 and FGF4 are used in several neural differentiation protocols [293,294], [313], but it is unclear whether they promote differentiation, proliferation and/or survival of neural precursor cells. This ambiguity is especially relevant to human ES cells, as FGF2 is used to maintain human ES cell pluripotency.

Consistent with a role for FGF signaling in promoting neural induction, treatment of differentiating mouse ES cells with a pharmacological inhibitor of FGF signaling, PD184352, arrests neural differentiation at an early stage [300]. In contrast, treatment with a different pharmacological inhibitor of FGF signaling, SU5402, does not affect early neural differentiation (but may affect survival of neural precursors), and mouse ES cells lacking Fgfr1 are able to differentiate down the neural lineage similarly to wild-type cells [303]. PD184352 inhibits the activity of MEK, a component of the MAPK cascade, whereas SU5402 inhibits the tyrosine kinase activity of Fgfr1, raising the possibility that MAPK pathway activation through an Fgfr1-independent mechanism may be important for ES cell neural induction. Insulin-like growth factors (IGFs) can also activate the MAPK pathway and promote neural fates in Xenopus [314], raising the possibility that IGFs are the activators of the MAPK pathway relevant to neurogenesis in mammals.

CONCLUSIONS AND PERSPECTIVES

Two of the remarkable properties of ES cells are pluripotency and a capacity for genetic modification. The potential to use ES cells to create cell-based therapies has received much attention, and neural and glial derivatives of ES cells are likely to be important for treating traumatic injuries and degenerative diseases affecting the nervous system. The same remarkable properties are also valuable for creating cellular models of human diseases critical both for elucidating disease pathogenesis and for testing potential therapies. Neural and glial derivatives of ES cells are already providing insights into the molecular and cellular bases of neurodegenerative disorders such as amyotrophic lateral sclerosis [315]. ES cells can also be used to create genetic models of human disease and to test the linkage between genes and associated disorders. For example, expression of alleles associated with Huntington's and Parkinson's diseases in ES cells has provided insights into underlying molecular pathogenesis [316,317].

One important step toward these potential biomedical advances is the development of a better understanding of neural induction by ES cells. The study of neural induction has a long history that provides extensive knowledge of early neural differentiation in several different vertebrate embryos. Common themes have emerged from these developmental model organisms, suggesting that early low-level FGF signals, combined with protective influences from the dorsal organizer shielding the ectoderm from Wnt and BMP signals, are critical for neural induction. Several of these influences are reflected in the requirements of ES cells for neural differentiation, although important differences between neural induction *in vivo* and *in vitro* are demonstrable. Better understanding of the regulatory influences that control neural induction, especially in mammals, may provide additional means of promoting neural induction and differentiation in ES cells. Conversely, better understanding of the

molecular mechanisms by which ES cells adopt neural fates may provide new insights into the regulation of neural induction in the early vertebrate embryo.

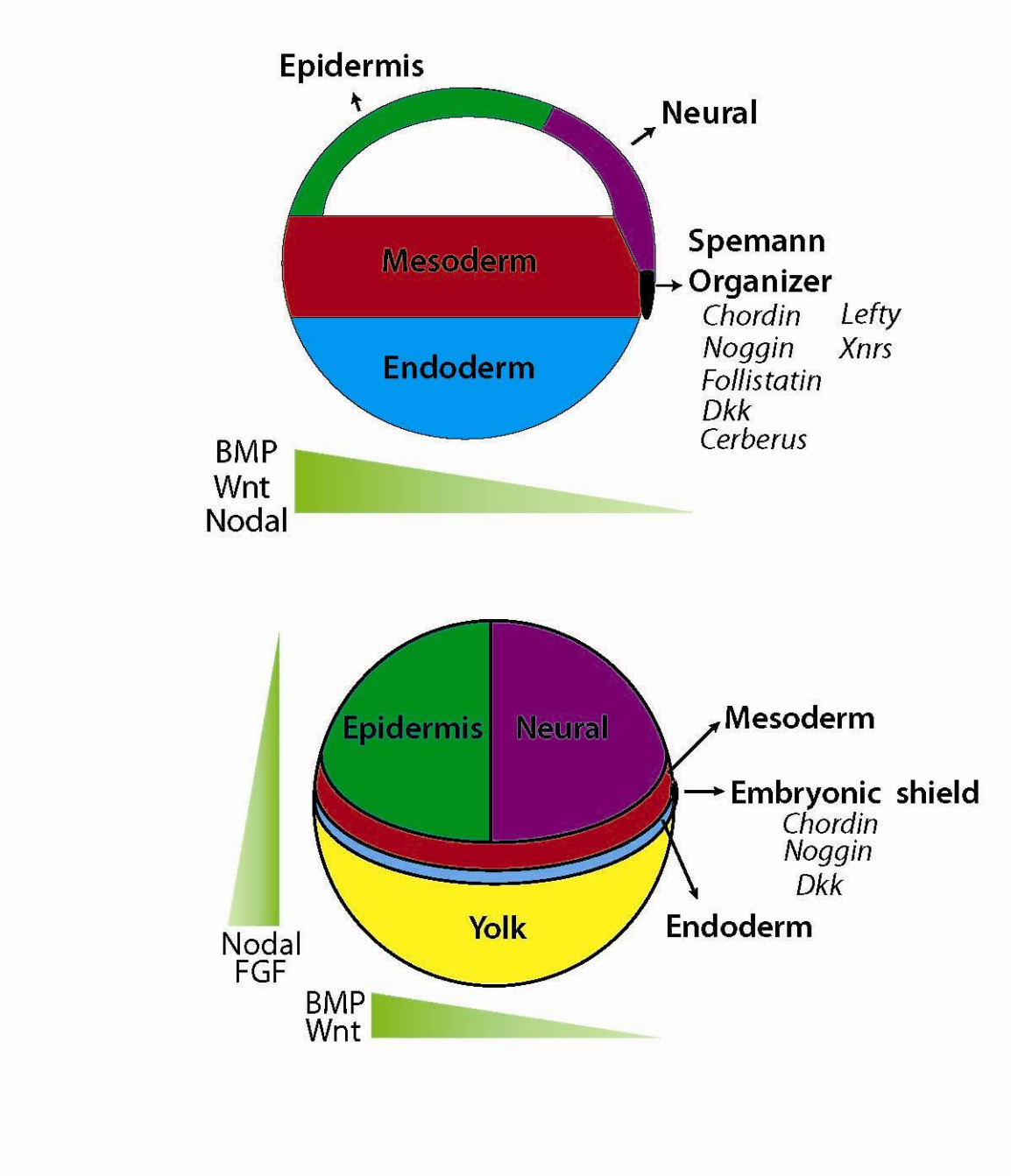


Figure 30. *In vivo neural induction.*

The dorsal organizer in Xenopus (A.) and similarly the shield in zebrafish (B.) express many factors that protect the ectodermal domain that will give rise to neural fates from activators of the BMP, Nodal, and Wnt pathways. The combined functions of secreted agonists of these pathways and the inhibitors produced by the dorsal organizer give rise to gradients that both pattern the embryo and allow for neural induction.

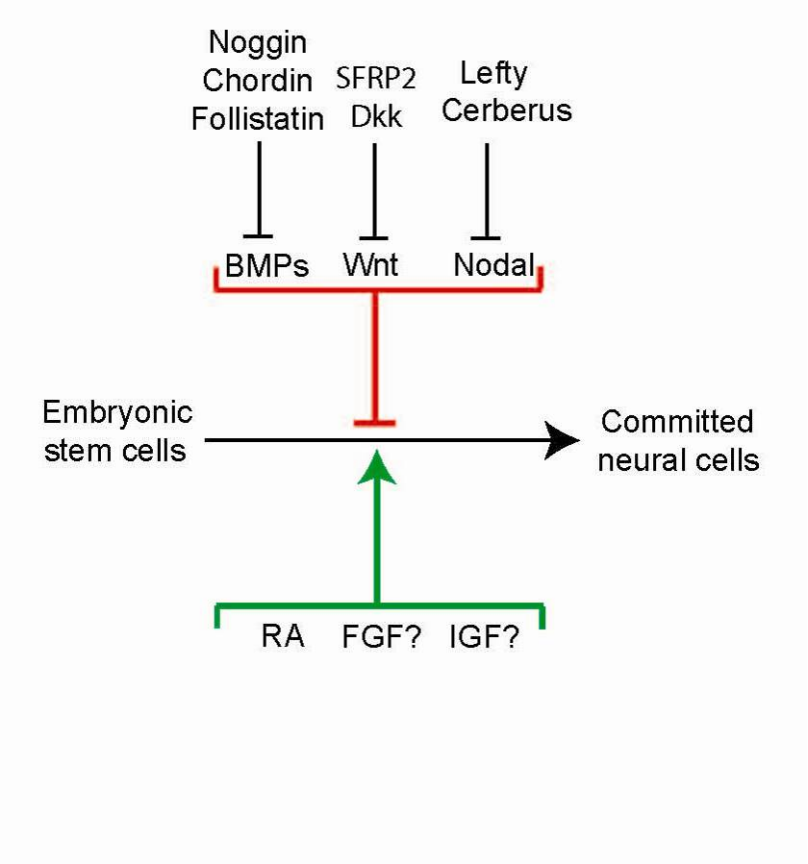


Figure 31. *In vitro neural induction.*

Many in vivo neural inducers that act as inhibitors of BMPs, Nodal, and Wnt signaling also promote ES cell differentiation to committed neural cells. In contrast, RA, which promotes neural induction in ES cells, is not known to be important for neural induction in vivo.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Julie Hamaguchi
Author Signature

12/13/11
Date