UNIVERSITY OF CALIFORNIA,
IRVINE

Dynamic interplay between transcription factors and epigenome during early *Xenopus* embryogenesis

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Jin Sun Cho

Dissertation Committee:
Professor Ken W.Y. Cho, Chair
Associate Professor Rahul Warrior
Professor Kyoko Yokomori

2019

DEDICATION

To

my family and my parents

for their love and support

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my mentor, Dr. Ken Cho, for his patience, support, and guidance throughout my Ph.D. He has always taught me how to be an active person, critical thinker, and productive scientist. Through his great insight about the future direction of developmental biology, he has pushed me to try new approaches and collaborate with experts in other fields.

I am very thankful to my committee members Drs. Kyoko Yokomori and Rahul Warrior for their helpful guidance and patience. And I also thank my advancement committee member Dr.Zeba Wunderlich. Their expertise and critical questions and suggestions have always helped me to improve the way to do my research.

I would also like to thank Dr. Gert Jan C. Veenstra and Dr. Ila van Kruijsbergen for their feedback and suggestions during our collaboration.

I am grateful to all the Cho lab members for their tremendous support and passion. They made the lab a great place to be in every day. I thank Dr. Ira Blitz for his guidance and expertise in everything about *Xenopus*. I thank all the former graduate students Dr. William Chiu, Dr. Anna L. Javier, Dr. Mui Luong, Dr. Soledad Reyes de Mochel, Dr. Rebekah Charney and Dr. Kitt Paraiso for their support and encouragement. I also thank Dr. Margaret Fish for her support and expertise. I also thank Jeff Zhou, Paula Pham and Jessica Cheung for their support. I learned a lot from all the Cho lab members.

Finally, I extremely thank my family for all their love and support in my pursuit of Ph.D. My husband, Rock Hur, my son, Teddy Hur, my mother in law, Minja Hur and my parents, June-young Cho and Chune-oak Kim, have been a source of strength to get through my life journey. I couldn't have done anything without them.

Some material of this dissertation is a reprint of the material as it appears in Cold Spring Harbor Laboratory Press. The co-author, Ken W.Y. Cho, listed in this publication directed and supervised research which forms the basis for the dissertation. I would also like to thank Ira L. Blitz for his contributions to this publication.

CURRICULUM VITAE

Jin Sun Cho

EDUCATION

2012-2019 Ph.D. in Biological Sciences, University of California, Irvine
Dissertation: Dynamic Interplay between maternal transcription factors and epigenome during early *Xenopus* embryogenesis

1999-2001 M.S. in Molecular Life Science, Ewha Woman's University, South Korea

1995-1999 B.S. in Chemistry, Ewha Woman's University, South Korea

PUBLICATIONS

Cho, J.S., Blitz, I.L., and Cho, K.W.Y. (2019). DNase-seq to Study Chromatin Accessibility in Early *Xenopus* tropicalisEmbryos. Cold Spring Harbor Protocols 2019, pdb.prot098335.

Charney, R.M., Forouzmand, E., **Cho, J.S.**, Cheung, J., Paraiso, K.D., Yasuoka, Y., Takahashi, S., Taira, M., Blitz, I.L., Xie, X., Cho, K.W.Y. (2017). Foxh1 Occupies *cis*-Regulatory Modules Prior to Dynamic Transcription Factor Interactions Controlling the Mesendoderm Gene Program. Developmental Cell 40, 595–607.e4.

Seok, H., **Cho, J.S.**, Cheon, M. and Park, I. (2002). Biochemical Characterization of Apoptotic Cleavage of KH-Type splicing Regulatory Protein (KSRP)/Far Upstream Element-Binging Protein 2(FBP2). Protein and Peptide Letters, Vol.9, No. 6, 511-519.

POSTER PRESENTATIONS AND COURSES ATTENDED

Cho, J.S., Paraiso, K., Cheung, J., Fish M.B., Blitz, I.L., Zorn, A., Cho, K.W.Y., "The dynamic interplay of Sox transcription factors regulating the activity of endoderm specific *cis*-regulatory modules," 17th International *Xenopus* Conference, University of Washington, Seattle, WA, August 12-16, 2018. Poster.

Cho, J.S., Charney, R.M., Cheung, J., Forouzmand, E., Fish M.B., Blitz, I.L., Cho, K.W.Y., "Foxa2 is recruited to Foxh1-primed *cis*-regulatory sequences to refine endoderm specification during *Xenopus* gastrulation," Society for Developmental Biology 75th Annual Meeting, Boston, MA, August 4-8, 2016. Poster.

Cho, J.S., Blitz, I. L., Cho, K. W. Y., "The transcriptional regulation of the BMP signaling pathway in the early *Xenopus* tropicalis embryos using genome-wide approaches," 15th International *Xenopus* Conference, Pacific Grove, CA, August 24th-28th, 2014. Poster.

Xenopus Bioinformatics Course, *Xenopus* National Resource Center, Marine Biological Laboratory, Woods Hole, MA, May 2012

PROFFESSIONAL AND TEACHING EXPERIENCE

Teaching assistant for *Biological Sciences D104*: Developmental Biology, University of California, Irvine, Summer 2014, Winter 2018

Teaching assistant for *Biological Sciences 75*: Human Development, Conception to Birth, University of California, Irvine, Spring 2016 and Spring 2017

Teaching assistant for *Biological Sciences D137*: Human and Eukaryotic Genetics, University of California, Irvine, Fall 2014, Fall 2015 and Fall 2016

Teaching assistant for *Biological Sciences D170*: Applied Human Anatomy, University of California, Irvine, Spring 2016

Teaching assistant for *Biological Sciences D103*: Cell Biology, University of California, Irvine, Fall 2013

Assistant Specialist for Department of Developmental and Cell Biology, University of California, Irvine, Jan.2011 - Aug.2012

Junior Specialist for Department of Developmental and Cell Biology, University of California, Irvine, Oct.2009 - Dec.2011

ABSTRACT OF THE DISSERTATION

Dynamic interplay between transcription factors and epigenome during early *Xenopus* embryogenesis

By

Jin Sun Cho

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2019

Professor Ken W. Y. Cho, Chair

After fertilization, the unified genome from the egg and sperm must go through reprogramming to reset the newly formed zygotic genome for the onset of the embryonic development. Maternally-deposited transcriptions factors (TFs) initiate this process that includes modification of the inherited epigenetic landscape. The subsequent differential expression of zygotic genes drives the formation of distinctive cell types - ectoderm, mesoderm, and endoderm - known as germ layer specification. However, the regulation of maternal TFs with epigenetic remodeling during germ layer specification is not well known.

To investigate the role of TFs and its relationship with chromatin during early embryogenesis, I have optimized Deoxyribonuclease I (DNase I) hypersensitive sites sequencing (DNase-seq) in early *Xenopus* embryos. DNase-seq identifies the genome-wide open chromatin regions that are accessible to regulatory factors.

I also used Foxh1-deficient *Xenopus* embryos to study its function in chromatin remodeling. I found that Foxh1 is required for the recruitment of a core subunit of polycomb repressive complex 2 (PRC2), Ezh2, and the regional activity of H3K27me3. Differential enrichment of H3K27me3 between ectodermal and endodermal germ layers of the *Xenopus* embryos suggests that Foxh1 directs Ezh2 recruitment to regulate histone modification both temporally and spatially. Ezh2 binding is also associated with another maternal TF, Sox3, which suggests that the combinatorial bindings of multiple maternal TFs fine-tune the temporal and spatial gene expression to induce the correct cell types in an embryo.

My findings provide a more comprehensive understanding of how maternal TFs regulate the epigenetic landscape during early vertebrate embryogenesis. I propose a model where Foxh1 recruits Ezh2 to mark epigenetically the regulatory regions of germ layer-specific genes to promote and maintain embryonic germ layer specification.

CHAPTER 1

INTRODUCTION

Early embryonic development is coordinated by maternal transcription factors and epigenetic regulation

Embryonic development starts from a fertilized egg - one single cell fused of maternal and paternal gametes - to make a multicellular organism. The first distinct cell types formed during embryogenesis are the three germ layers - ectoderm, mesoderm, and endoderm - which generate the tissues and organs of an adult body. Highly conserved gene regulatory networks (GRNs), comprised of interactive groups of molecular regulators include transcription factors, signaling proteins, and co-regulators which govern the formation of these primary germ layers in most metazoans (Davidson and Levine, 2005; reviewed in Davidson and Erwin, 2006).

Transcription factors (TFs) play central roles in GRNs by binding on the *cis*-regulatory elements (CREs) of the genome to regulate activation and repression of target gene expression. In addition, the presence of other transcriptional regulators (co-regulators) and the chromatin state surrounding these CREs influences their gene expression (Levine, 2010; Bannister and Kouzarides, 2011). However, the functional relationship among TFs, co-regulators and the epigenetic landscape is generally not well understood.

The clawed frog, *Xenopus tropicalis* is a model organism well-suited to investigate this critical question. First, the high tolerance of *Xenopus* embryos toward experimental manipulations allows the injection of macromolecules such as RNAs and DNAs to examine the knockdown or overexpression phenotypes of a given gene. Many of the core regulators such as Vegt and Ctnnb1/β-catenin are involved in germ layer specification was identified using such approaches (White and Heasman, 2008; Kiecker et al., 2016). Second, its phylogenetic distance between human and *Xenopus* is approximately 360 million years ago (MYA), providing an opportunity for comparative genomic analysis to uncover the

conserved regulatory mechanism regulating human development and diseases (Wheeler & Brändli, 2009). The availability of the complete genome sequence of true diploid *Xenopus tropicalis*, coupled together with the ability to obtain synchronously developing *Xenopus* embryos from a single fertilization allows genome-wide high-throughput sequencing (HTS) studies (Hellsten et al., 2010).

In this chapter, I will provide the current status of the role of maternal TFs and the epigenetic regulation during germ layer specification. The research described in this thesis will provide a deeper understanding of the relationship between maternal TFs and epigenetic regulation during *Xenopus* germ layer specification, which can aid research into stem cell therapy and organoid generation.

Role of maternal transcription factors during the germ layer specification

After fertilization, most vertebrate embryos undergo similar developmental processes: cleavage, gastrulation, and organogenesis. During cleavage stages, the fertilized egg is divided into smaller cells (called blastomeres) without increasing the size of the embryo. At this stage, the cells of the embryo are still pluripotent but some maternally deposited mRNAs are asymmetrically distributed which determine the animal-vegetal axis of the embryo (Heasman, 2006) (Figure 1.1). These asymmetrically localized maternal gene products (also known as determinants) regulate the formation of the three germ layers during gastrulation that follows the pluripotent blastula stages (Borchers and Pieler 2010; Paranjpe and Veenstra 2015). In *Xenopus*, the time period between blastula and early gastrula embryos (Nieuwkoop-Faber stage 9-10.5, Nieuwkoop and Faber, 1994) is when ectoderm is formed in the animal cap (top side of the embryo), while endoderm is located

in the vegetal mass (Figure 1.1A-B). Mesoderm is induced at the equatorial region between the animal and the vegetal poles of the *Xenopus* embryo. After gastrulation endodermal cells give rise to the gastrointestinal and respiratory tracts, the lungs, the liver, and the pancreas. Next, mesodermal cells give rise to the heart, the muscle system, the bones, and the bone marrow (the blood). Lastly, ectodermal cells give rise to the epidermis (skin) and nervous system.

One of the most well-studied maternal determinants of germ layer specification in the *Xenopus* embryo is T-box TF, Vegt (previously known as Xombi, Antipodean, or Brat, Lustig et al., 1996; Zhang and King, 1996; Stennard et al., 1996; Horb and Thomsen, 1997). *Vegt* mRNA is localized in the vegetal mass of mature eggs and early embryos. Zygotic transcription of *vegt* begins in dorsal mesoderm and then extends to lateral and ventral mesoderm. Vegt-depleted embryos do not form endoderm and express a reduced amount of mesoderm-inducing signals (Zhang et al., 1998; Kofron et al., 1999). Vegt regulates the expression of Nodal ligands, *nodal1, 2, 4, 5*, and *6*, which are one of the earliest transcribed zygotic genes of Nodal signaling pathway – part of the TGF-β superfamily – which is necessary for the initiation of both the endoderm and mesoderm formation in vertebrates (reviewed in Schier, 2003). Even though Vegt is a core TF for the initiation of the Nodal signaling pathway in *Xenopus*, its mammalian orthologs have not been identified (White and Heasman, 2008) and other vertebrate vegt orthologs (zebrafish *spadetail/tbx16*, and chick *tbx6L/tbx6*) are not expressed maternally (Fukuda et al., 2010).

F-type Sox (SRY-related high motility group (HMG)-box), Sox7 is another vegetally localized maternal TF in *Xenopus* (Zhang et al., 2005). Sox7 binds to the nodal5 promoter, and induces activation of Nodal ligands - *nodal1, 2, 4, 5*, and *6*, and other endoderm markers

(zygotic endodermal TFs) such as mixer and sox17b. Thus, vegetal-specific maternal TFs, Vegt and Sox7, regulate Nodal ligands expression and endoderm formation (Zhang and Klymkowsky, 2007). Vegt depends upon Sox7 activity for activation of *mixer* and *endodermin* in animal caps. However, the depletion of Sox7 does not cause any phenotypic effect on endoderm formation. This suggests the combinatorial function of these localized maternal TFs may refine the endodermal gene expression in the *Xenopus* embryo. Since Vegt also regulates zygotic *sox7* expression, the functional study of maternal Sox7 in endoderm formation is required to comprehend the combinatorial role of maternal TFs.

Forkhead protein Foxh1 is the master TF of Nodal signaling. Foxh1 activates Nodal target genes by recruiting phosphorylated Smad2/3 (pSmad2/3), co-effectors of Nodal signaling (Shen, 2007). Foxh1 is maternally supplied and expressed ubiquitously in the *Xenopus* blastula embryo (Chiu et al., 2014; Charney et al., 2017; Paraiso et al., 2019), whereas Vegt and Sox7 are localized in the vegetal mass. Morpholino oligonucleotide (MO) knockdown of Foxh1 in *Xenopus* embryos results in defects in endoderm and mesoderm formation, delayed gastrulation, reduced head structures, and a shortened anteroposterior (A-P) body axis, similar to inhibition of Nodal signaling (Chiu et al., 2014).

Foxh1 also functions independently of Nodal signaling (Chiu et al., 2014; Charney et al., 2017). First, Chiu et al. have identified dozens of genes, whose expression is positively and negatively regulated by Foxh1 (Chiu et al., 2014). For example, zygotic TF, *hand2* is negatively regulated by Foxh1 in a Nodal-independent manner. Second, inhibition of Nodal signaling using a chemical inhibitor, SB-431542 failed to abolish the binding of Foxh1 to DNA, suggesting that Foxh1 binding occurs in a Nodal-independent manner (Chiu et al., 2014; Charney et al., 2017). Lastly, in early blastula embryos, Foxh1 binding regions

overlap with those of Tle/Groucho (co-repressor), which are attenuated on certain endodermal *cis*-regulatory modules (CRMs) upon Foxh1-MO knockdown (Charney et al., 2017). Taken together, these results suggest that maternal TF can have a role not only as an activator but also as a repressor before zygotic gene activation (ZGA).

Another maternally expressed forkhead domain TF, Foxi2, is highly enriched in the animal region of the *Xenopus* blastula embryo (Cha et al., 2012). Foxi2 activates a zygotic forkhead TF, *foxi1e*, and regulates the ectoderm formation. Since its expression is restricted to the animal cap (ectoderm), Foxi2-depleted embryos are still able to respond normally to mesoderm-inducing signals from vegetal cells. The genome-wide approach is required to understand the regulatory modules of the ectodermal maternal TF.

Mesoderm is induced by the production of Nodal ligands, which are regulated by the activity of both Nodal and Wnt signaling. Wnt signaling regulates the activity of the lymphoid enhancer factor/T-cell factor (Lef/Tcf) TFs (Novaka and Dedhar, 1999). Tcf1, 3, and 4 are maternally deposited and ubiquitously expressed in *Xenopus* blastula embryos (Roel et al., 2002). They repress target gene transcription when co-bound with co-repressor Tle/Groucho. Upon Wnt signaling, nuclear β-catenin, a maternally loaded Wnt signaling co-activator localized in the dorsal side of the *Xenopus* embryo (Heasman et al., 1994), replaces Tle so that the β-catenin-Tcf complex can activate target genes (Hurlstone and Clevers, 2002; Daniels and Weis, 2005). Dorsal activity of Wnt/β-catenin signaling results in high concentrations of Nodal ligands in the dorsal marginal zone and contributes to the formation of the dorsoventral (D-V) body axis of embryos. β-catenin together with Tcf induces the expression of dorsal target genes *sia1* and *sia2*, and the expression of a pan-mesodermal marker, *t* (*Xenopus brachyury*, *Xbra*) where T is a core regulator for FGF

signaling (Schohl and Fagotto, 2003). Taken together, multiple signaling pathways are involved in regulating the mesoderm formation and D-V patterning by affecting downstream key regulators.

To distinguish between ectoderm and mesoderm layers, zygotic mesodermal genes are not expressed, and mesoderm-inducing signals are absent in the animal cap. This hypothesis is supported by the finding that no pSmad2/3 and no nodal transcripts found in the ectoderm (Osada and Wright, 1999; Chen, 2007). Maternal B1-type SOX TF, Sox3, has been proposed as a mesodermal suppressor in the animal cap since antibody inhibition experiments yielded results of a negative regulator of *nodal5* (Zhang et al., 2003; Zhang et al., 2004; Zhang and Klymkowsky, 2007). However, *sox3* mRNA expression is not restricted to the animal cap and is also expressed in the vegetal mass of early blastula embryos. Therefore its early role in gene regulation needs to be carefully investigated.

The interactions between the maternal TFs, TF-regulating signaling factors, and the zygotic targets genes generate a GRN, which underlies the formation of the three germ layers. To investigate these TFs within the GRN, TFs are bound to active regulatory regions, which can be detected through chromatin accessibility assays. In the next section, I will review the structure and function of chromatin within the context of early embryogenesis.

Epigenetic regulation during early embryogenesis

Chromatin is a complex of DNA and histones present in all eukaryotic cells. In the eukaryotic nucleus, 147bp of DNA is wrapped around an octamer of two copies of each core histone protein (H2A, H2B, H3, and H4) to form a nucleosome (Luger et al. 1997; Li, 2002). Nucleosomes are coiled into fibers and looped into higher-order chromatin

structures to fit into a nucleus. This higher-order chromatin structure and histone modification regulates gene expression and developmental programs during embryogenesis by modulating the accessibility of the transcriptional machinery to DNA.

Initially, the embryonic genome is transcriptionally silent and is reprogrammed into the pluripotent state from the union of two fully differentiated germ cells, the egg and the sperm. To create a new genome, chromatin from two distinct germ cells has to be remodeled before zygotic gene activation (ZGA). The onset of ZGA varies significantly among different animals (Jukam et al., 2017), and ZGA is not a single temporal event but occurs broadly during a time window where new transcription gradually begins. In mice, this process begins right after the first cleavage cycle (2-cell stage embryo; 24 hours post-fertilization). In *Drosophila melanogaster*, ZGA occurs at the 14th nuclear cycle (2.5 hours post-fertilization), when the division cycle slows dramatically and when ~6,000 nuclei become cellularized (Hamm and Harrison, 2018). In *Xenopus*, it occurs during the first 12 cleavage divisions, which is also known as the mid-blastula transition (MBT) (Nieuwkoop-Faber stage 8.5; 5 hours post-fertilization). However, recent high-resolution transcriptome profiling of *Xenopus* has revealed that zygotic transcripts of *pri-mir427* are detected as early as the 32-cell stage (at the fifth cleavage), which is significantly earlier than the classically defined MBT (Collart et al., 2014; Owens et al., 2016). Additionally, dozens of zygotic transcripts are first detected at the 128- and 256-cell stages included the known early expressed genes like *nodal5* and *nodal6*. The dynamics of the epigenetic landscape and chromatin remodeling will provide a better understanding of the mechanism of ZGA.

Transcriptional activation is tightly linked with relatively nucleosome-free regions (i.e. promoters, enhancers, silencers, and insulators) due to the binding of transcription

factors or co-regulators. This regulatory DNA coincides with open or accessible chromatin. This accessible genome can be measured by quantifying the enzymatically or chemically isolated accessible DNA using next-generation sequencing (NGS) platforms - MNase-seq, FAIRE-seq, DNase-seq, and ATAC-seq (reviewed by Tsompana and Buck, 2014). Chromatin accessibility during early embryogenesis has been measured in flies (Blythe and Wieschaus, 2016), zebrafish (Liu et al., 2018), mice (Lu et al., 2016; Wu et al., 2016) and humans (Guo et al., 2018; Wu et al., 2018), which revealed that accessible regulatory regions are established concomitantly with ZGA. Profiling in early mouse embryos revealed broad regions of open chromatin in late 1-cell and early 2-cell stages, followed by more narrow peaks on promoters at 8-cell stage when major waves of ZGA occurs (Wu et al., 2016). These assays have not been performed in early-stage *Xenopus* embryos. In this thesis, I will describe how I adapted the DNase I hypersensitive site sequencing (DNase-seq; Neph et al., 2012) method for studying chromatin accessibility in early *Xenopus* embryos.

The epigenetic landscape is also affected by post-translational modifications on histone tails, which impact nucleosome stability and the recruitment of transcriptional regulators. Lysine residues on histone tails can be acetylated, methylated, sumoylated or ubiquitylated (Kouzarides, 2007). Chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq) is a method to detect histone modifications. ChIP-seq peaks of histone H3 lysine 4 trimethylation (H3K4me3) on the transcription start sites (TSSs) are associated with permissive gene expression. The emergence of H3K4me3 on the promoters of permissive genes has been monitored before ZGA in flies (Li et al., 2014), zebrafish (Vastenhouw et al., 2010; Lindeman et al., 2011), and frogs (Akkers et al., 2009;

Hontelez et al., 2015). In the case of mice, unusually broad, non-canonical H3K4me3 domains (wider than 5kb) were observed in matured oocytes, after which these domains are mostly restricted to the conventional TSSs of transcriptionally active genes during ZGA (Dahl et al., 2016; Liu et al., 2016; Zhang et al., 2016). Histone H3 lysine 27 trimethylation (H3K27me3), on the other hand, is associated with repression of target genes (Bannister and Kouzarides, 2011). In most species, the increase of H3K27me3 occurs during or after ZGA, which is later than the appearance of H3K4me3. This implies that transcriptional quiescence before ZGA is not imposed by H3K27me3-marked repression. In embryonic stem cells (ESCs), the co-occurrence of active (H3K4me3) and repressive (H3K27me3) chromatin modifications has been described as a bivalent mode on promoters of poised developmental genes (Bernstein et al., 2006; Mikkelsen et al., 2007). However, this bivalent mode has not been detected in mice (Liu et al., 2016; Zhang et al., 2016), flies (Takayama et al., 2014) or frogs (Akkers et al., 2009) during ZGA. In frogs, Akkers et al. performed sequential ChIP experiments for H3K4me3 and H3K27me3 on both marked genes in the whole embryos to test whether this bivalency occurs in the same cells or not. They revealed that bivalent marking of genes is not a prevalent configuration in *Xenopus* embryos and further showed that H3K27me3 marks are on some endodermal genes on the animal cap side of *Xenopus* embryos. This suggests that H3K27me3 may function in a spatial manner for epigenetic repression.

The maternal TFs involvement in epigenetic regulation during ZGA

Transcription factors regulate gene expression by recruiting the transcriptional machinery to particular genes by binding to a specific DNA sequence that each TF detects. Therefore

the chromatin state of those TF-bound specific DNA regions during ZGA is critical to comprehend the relationship between epigenetic regulation and gene expression in early embryonic developmental processes.

Pioneer factors are a special class of transcription factors that can access their DNA target sites in closed chromatin and presumably bind to the genome before the binding of other factors (Iwafuchi-Doi and Zaret, 2014). Two general features of pioneer TFs have been described - “passive” and “active” roles to endow transcriptional competence (Zaret and Carroll, 2011). In the passive role, a pioneer factor can bind alone, and then recruit other TFs and additional factors to create an active enhancer. This priming can increase the rapidity of the transcriptional response. In the active role, pioneer factors can directly facilitate other factors to bind to nucleosomal DNA or open the local chromatin.

One of the key regulators of ZGA in flies, maternal TF, Zelda (Zld) has been studied for its epigenetic role. Maternal mutants of Zld fail to complete cellularization and die before the end of the maternal-to-zygotic transition (MZT) (Liang et al., 2008). The maternal Zld is required for the expression of hundreds of genes during ZGA. This maternal TF primes enhancers by lowering the high nucleosome barrier to assist another TF, Dorsal (Dl) to the enhancer elements during dorsoventral specification (Sun et al., 2015). Chromatin accessibility in the maternal Zld-depleted embryos is reduced at a subset of Zld-binding sites, even though most of the Zld-bound regions retained chromatin accessibility (Schulz et al., 2015). This suggests that Zld functions as a pioneer factor. However, Zld orthologues are limited to the insect clade (Ribeiro et al., 2017). It is still unclear whether there are other pioneer factors in vertebrate maternal TFs.

Nanog, SoxB1 (Sox19b), and Pou5f3 are core maternal TFs and are required to initiate the zygotic developmental program in zebrafish (Lee et al., 2013). These TFs are homologues of the mammalian pluripotency factors NANOG, SOX2, and OCT4, respectively, which are known for their ability to reprogram differentiated cells to a pluripotent state similar to embryonic stem cells (Takahashi and Yamanaka, 2016). Pou5f3 knockdown by MO injection in zebrafish embryos results in decreased chromatin accessibility at their binding sites (Liu et al., 2018). In frogs, there are no known maternal homologs of Nanog, while Sox3 is the homolog of SoxB1 TF. There are three Oct-related genes - Oct 60, 25, and 91 - in frogs. Only Oct 60 and 25 are maternally expressed. Thus, it would be interesting to study their role in early *Xenopus* embryonic development.

Foxh1 has been recognized as a pioneer factor since it bookmarks mesendoderm enhancers before ZGA and chromatin modifications (Charney et al., 2017). Foxh1 occupies enhancers as early as the 32-cell cleavage stage, in the presence of nucleosome-dense chromatin (Bogdanovic et al., 2011). Foxh1 bindings in *Xenopus* occur before zygotic forkhead TFs, Foxas bindings, which are the well-known hepatic pioneer factors. Foxh1 can also be associated with repressive chromatin states or epigenetic regulation since Foxh1-dependent Tle/Groucho binding on endodermal CREs was found in early blastula embryo. However, the previous studies focused primarily on Foxh1 pre-binding before ZGA and did not address howFoxh1 is functionally regulating chromatin dynamics.

Overview of this thesis

Maternal TFs are required to initiate gene regulatory programs but their roles in epigenetic regulation during early embryogenesis are not yet fully revealed. The research described in

this thesis is focused on the epigenetic regulation mediated by maternal TFs during early *Xenopus* embryonic development and germ layer specification.

Chapter 2 is focused on a method to measure global chromatin accessibility in early *Xenopus* embryos. This method can help to understand the properties of the epigenome of early embryos along with the genome-wide binding locations of TFs, RNAPII, and other co-regulators from ChIP-seq experiments.

In chapter 3, I will describe the essential function of Foxh1 in early *Xenopus* development using Foxh1 mutants generated by CRISPR-Cas9 system in *Xenopus tropicalis* (Blitz et al., 2013; Nakayama et al., 2013). Furthermore, I have generated F0 females, which have frameshift mutations in the Foxh1 transcription start site. Eggs derived from these Foxh1 F0 mutant mothers completely lacked *foxh1* transcripts. These Foxh1 null F1 embryos were subsequently used for genome-wide analysis to study the epigenetic roles of Foxh1. Additionally, mass spectrometry analysis uncovered the interaction of Foxh1 together with various epigenetic regulators in mouse embryonic stem cells.

With the combined work from chapter 2 and chapter 3, I have focused on the relationship of Foxh1 and Ezh2 (a histone modifier that deposits H3K27me3, Cao and Zhang 2004; Schuettengruber et al. 2009; Margueron and Reinberg, 2011) to affect the chromatin landscape. This highlights the epigenetic roles of maternal Foxh1 by recruiting histone modifiers to regulate spatiotemporal gene expression during the germ layer specification. In Chapter 4, I will discuss the work presented in this thesis and the possible future direction of the research about the epigenetic roles of maternal TFs.

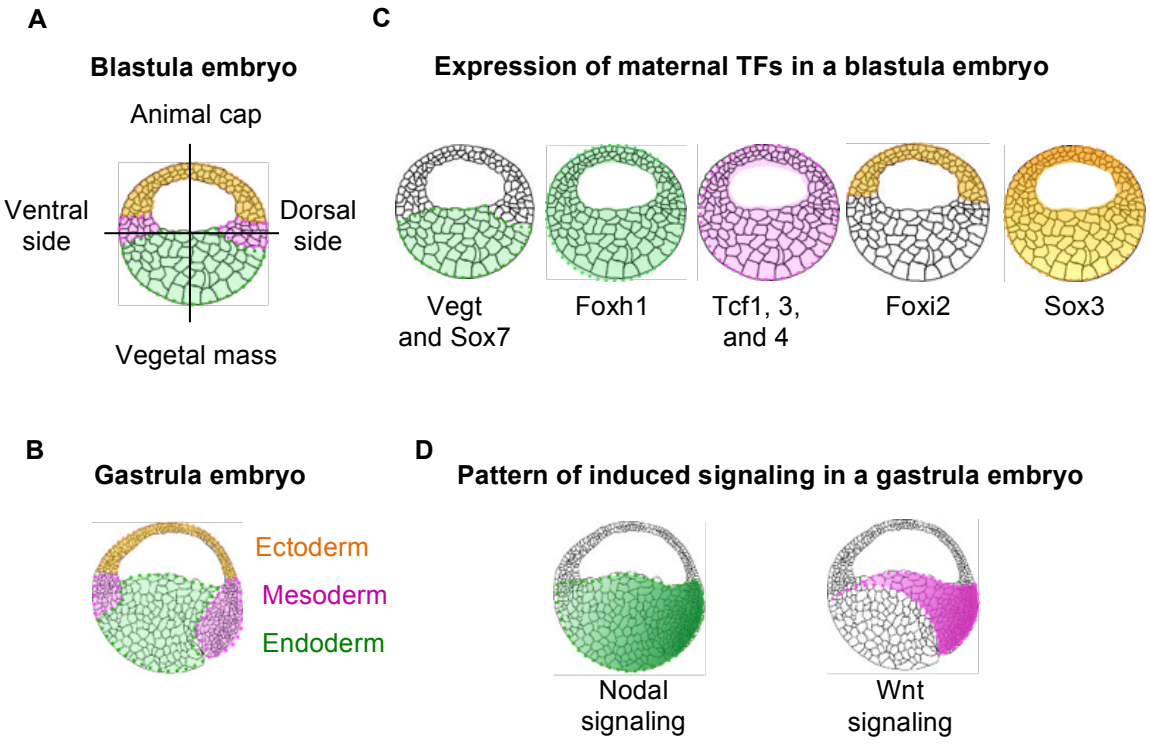

Figure 1.1 Expression patterns of maternal transcription factors (TFs) and spatial activity of the associated signaling pathway in the *Xenopus* embryo Drawings of cross-sections of *Xenopus* embryos at A) blastula stage and B) gastrula stage. A) The animal-vegetal and dorsal-ventral axes are determined in the fertilized embryo. B) The three germ layers are formed at the gastrula stage. C) The spatial distribution of maternal TFs effects on the three germ layer specification. D) Maternal TFs regulate the activity of induced signaling for the germ layer-specific gene expression

CHAPTER 2

DNase-seq to Study Chromatin Accessibility in Early *Xenopus tropicalis* Embryos

(This Chapter has been published in Cold Spring Harbor Protocols)

Abstract

Transcriptional regulatory elements are typically found in relatively nucleosome-free genomic regions, often referred to as “open chromatin.” Deoxyribonuclease I (DNase I) can digest nucleosome-depleted DNA (presumably bound by transcription factors), but DNA in nucleosomes or higher-order chromatin fibers is less accessible to the nuclease. The DNase-seq method uses high-throughput sequencing to permit the interrogation of DNase hypersensitive sites (DHSs) across the entire genome and does not require prior knowledge of histone modifications, transcription factor binding sites, or high quality antibodies to identify potentially active regions of chromatin. Here, discontinuous iodixanol gradients are used as a gentle preparation of the nuclei from *Xenopus* embryos. Short DNase I digestion times are followed by size selection of digested genomic DNA, yielding DHS fragments. These DNA fragments are subjected to real-time quantitative polymerase chain reaction (qPCR) and sequencing library construction. A library generation method and pipeline for analyzing DNase-seq data are also described.

MATERIALS

It is essential that you consult the appropriate Material Safety Data Sheets and your institution's Environmental Health and Safety Office for proper handling of equipment and hazardous materials used in this protocol.

RECIPES: Please see the end of this protocol for recipes indicated by <R>.

Reagents

Agarose

Buffer A for DNase-seq <R> (4˚C)

Chloroform

DNase I digestion buffer (1×) <R> (freshly prepared, equilibrated to 37˚C)

DNase I stock solution (10 U/µL) <R>

Ethanol

Ethidium bromide

Gel extraction kit (e.g., NucleoSpin Gel and PCR Clean-up kit [Macherey-Nagel 740609.250])

High sensitivity DNA kit (Agilent Technologies 5067-4626)

Iodixanol solutions (20%, 25%, 30%) <R> (freshly prepared, at 4˚C)

Library quantification kit (e.g., KAPA KK4824 [Roche 07960140001])

LightCycler 480 SYBR Green I Master Mix (Roche 04707516001)

NaCl (5 M)

NEXTflex ChIP-Seq Barcodes (Illumina-compatible barcode adaptors) (Perkin Elmer NOVA-514121)

NEXTflex ChIP-seq Library Prep kit (Perkin Elmer NOVA-5143-01)

Phenol:chloroform:isoamyl alcohol (25:24:1 [v/v])

Primers for qPCR validation (Figure 2.1B)

eef1a1o forward: 5'-GCTGGAATTTAAAGGGATGGA-3'

eef1a1o reverse: 5'-CCGGCGTTTTATTGGAACT-3'

hbe1 forward: 5'-TTGCATTTGGTTCAGTGCTC-3'

hbe1 reverse: 5'-TGTCAGATGCTGGTTCTCCA-3'

otx2 forward: 5'-CAGAAAGGGCTTTGTTTTCG-3'

otx2 reverse: 5'-AAACTTGATTGGGGCCATTT-3

Proteinase K (20 mg/mL)

Qubit dsDNA HS Assay kit (Invitrogen Q32851)

RNase A (10 mg/mL)

Stop buffer for DNase I <R> (freshly prepared, equilibrated to 37°C)

Equipments

Bioanalyzer (Agilent 2100)

Centrifuge (low-speed, refrigerated), with a swinging-bucket rotor

Centrifuge tubes (50-mL)

Centrifuge tubes, polyallomer (38.5-mL)

Dounce homogenizer (15-mL) with pestle B (0.025- to 0.076-mm clearance) (prechilled to 4°C) Gel electrophoresis equipment

Microcentrifuge tubes (2-mL)

Nutator at 4°C

Nylon mesh (100-μm pore size) (Ted Pella 41-12115)

Pipette tips (P1000), made with a wide bore by clipping with scissors

Polypropylene tubes (15-mL, conical)

Qubit fluorometer and assay tubes

Real-time quantitative PCR (qPCR) system (Lightcycler 480 II [Roche])

Sequencer (Illumina)

Syringe (15-mL)

Tygon tubing

Ultracentrifuge, with swinging-bucket rotor (e.g., Beckman SW32Ti)

Water baths at 37˚C and 55˚C

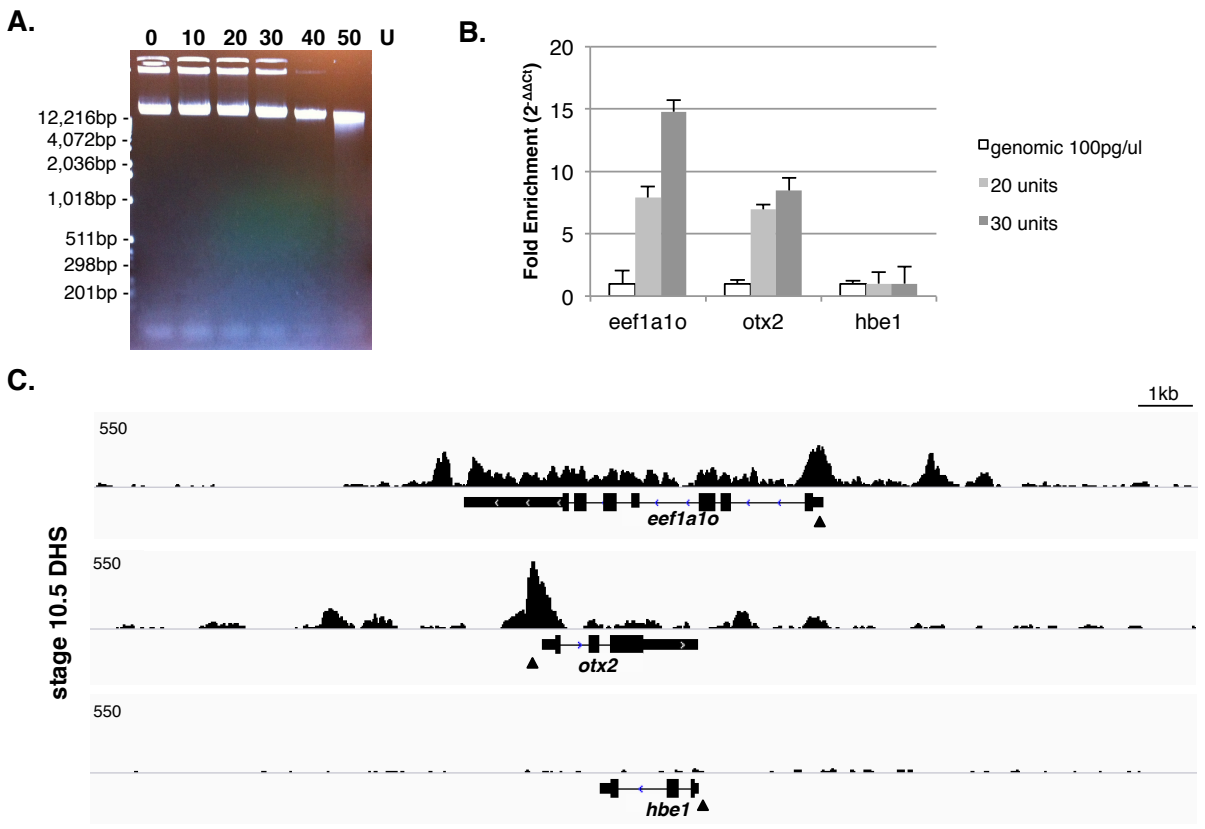

FIGURE 2.1. Methods to detect DHSs using *Xenopus* early gastrulae. A) Gel picture of DNase-digested genomic DNA extracted from nuclei isolated from stage 10.5 *Xenopus tropicalis* embryos. The amount of high-molecular-weight DNA fragments are gradually decreased as DNase concentrations increase. B) DNase hypersensitivity monitored by qPCR. DHSs are normalized by Cp value from 100 pg of genomic DNA from *Xenopus* liver using $2^{-\Delta\Delta Ct}$ method. Reference regions are selected from nonexpressing genes' promoters (e.g., hbe1, encoding hemoglobin subunit epsilon 1, which is not expressed until tailbud stages). C) A genome-wide profile of DHSs at gastrula stage by DNase-seq on *Xenopus* tropicalis genome version 9.0. The number of sequence reads from a stage 10.5 DNase- seq library was 45.5 million. Bowtie aligned 43.7% of these reads after discarding multiply aligned reads. Samtools (Step 32) removed duplicates, and peaks were called using Homer findPeaks or by MACS2 (Zhang et al. 2008). DHSs were detected on the promoter and enhancer regions of eefla1o and otx2 but no DHSs were shown around hbe1 (also used as a reference gene for qPCR in Figure 2.1B). Regions of qPCR are marked by arrowheads.

METHOD

Isolating Nuclei from *Xenopus* Embryos

This nuclear isolation protocol was modified from Farzaneh and Pearson (1978) and uses iso-osmotic iodixanol gradients instead of hyperosmotic sucrose. This reduces centrifugation time and avoids the extreme depletion of water from the nucleoplasm (Graham 2002).

During Steps 1-11, keep all materials and solutions on ice.

1. Collect dejellied embryos at the desired stage and transfer to a pre-chilled 15-mL Dounce homogenizer.
2. Gently wash the embryos twice in 10 mL of ice-cold 0.3 M SS per wash. Remove as much solution as possible after the second wash.
3. Add 4 mL of SS containing 0.4% (v/v) Triton N-101.
4. Homogenize the embryos using 5-7 strokes of the pestle to release intact nuclei.
5. Filter the homogenate through nylon mesh into a 50-mL tube. Rinse the mesh with 1 mL of SS containing 0.4% (v/v) Triton N-101, to collect the nuclei in mesh in to the tube.
6. Mix the filtered homogenate with one volume of 30% iodixanol to make 15% iodixanol.
7. In a 38.5-mL polyallomer centrifuge tube, layer 12 mL each of 25% and 20% iodixanol solutions followed by 14 mL of the 15% iodixanol with homogenate.

    To make sharply separated discontinuous gradients (25%; 20%; 15% iodixanol), add 20% iodixanol solution to the centrifuge tube. Then, gently add the 25% iodixanol solution underneath the 20% iodixanol using a syringe attached to Tygon tubing. The

final layer of 15% iodixanol solution containing the homogenized embryos is gently pipetted onto the top of the 20% gradient.

8. Centrifuge the sample at 20,000g for 20 min at 4oC using a SW32Ti rotor with maximum braking.

9. Using a pipette, harvest ~5 mL of the solution above the visible interface between the 25% iodixanol cushion and the 20% iodixanol layer; this solution contains the nuclei. Transfer this solution to a 15-mL polypropylene tube.

10. Dilute the nuclei with two volumes of SS and mix by gently inverting several times. Centrifuge at 1,000g for 5 min at 4°C using a swinging-bucket rotor. Remove the solution by pipetting.

11. Gently resuspend the pellet in 5 mL of buffer A. Centrifuge at 1,000g for 5 min at 4°C using a swinging-bucket rotor. Remove as much of the solution as possible by gentle pipetting.

An aliquot of the nuclei can be stained with DAPI and quantitated using a hemocytometer.

Digesting Isolated Nuclei with DNase I

This DNase I digestion procedure was modified from Neph et al. (2012).

For early gastrulae, nuclei from ~500 embryos are used in each digestion reaction. Thus, if starting from 2000 embryos, four reaction tubes are prepared in Step 12.

12. During one of the centrifugations above (Step 10 or 11), prepare tubes for DNase digestion as follows.

i. Add 200 µL of 1X digestion buffer to each of four 2-mL microcentrifuge tubes.

The number of reaction tubes can be adjusted according to the starting number of embryos.

ii. Add the required volumes of DNase stock solution (10 U/μL) (e.g., 1, 2 or 4 μL for 10, 20 or 40 U/reaction, respectively).

The DNase digestion conditions that permit the maximal release of DHS regions should be determined empirically (see Steps 24–25).

iii. Gently flick to mix.

13. Resuspend the nuclear pellet from Step 11 in 1200 μL of 1X digestion buffer (N x 300 μL, where N = number of reaction tubes) by gentle pipetting with a wide-bore pipette tip.

14. Transfer 300-μL volumes of nuclear suspension to each reaction tube prepared in Step 12 with a wide-bore pipette tip. Mix the samples by gentle pipetting.

15. Incubate the DNase digestion reactions in a water bath for exactly 3 minutes at 37°C.

16. Add 500 μL of stop buffer to each reaction. Mix by gently inverting the tubes and then incubate for 15 minutes at 37°C.

Recovering the DNase I Hypersensitive Fragments

17. Transfer the samples to 15-mL conical tubes and add 2 mL of TE to each sample.

18. Add 60 μL of RNase A (10 mg/mL) to each tube. Incubate for 1 hour at 37°C.

19. Add 40 μL of proteinase K (20 mg/mL) to each tube. Incubate for 2 hours to overnight at 55°C.

20. Extract the DNA using phenol:chloroform:isoamyl alcohol (25:24:1 [v/v]). Remove the organic layer, perform one chloroform extraction, and then recover the DNA by adding 1/10 volume of 5 M NaCl and two volumes of ethanol.

To prevent shearing of genomic DNA during the organic extraction, rock on a nutator for 30 min at 4˚C.

21. Resuspend the DNA pellet in 30 µL of TE.

22. Electrophoretically fractionate the DNA through a 1% agarose gel and visualize using ethidium bromide.

Run the gel at ~10 Volts/cm to permit better size resolution. The vast majority of the DNA should be more than 10 kb. DNA liberated in the 50- to 500-bp size range will not be visible to the eye (Figure 2.1A).

See Troubleshooting.

23. Isolate the gel region corresponding to 50- to 500-bp DHS fragments. Use a gel extraction kit (e.g., NucleoSpin) to recover the DNA. Elute in 20-30 µL of 10 mM Tris-HCl (pH 8.0).

[Optional] Validating DHS Fragments by qPCR

The success of DNase digestion conditions can be monitored by qPCR (Figure 2.1B and C). The promoter regions of highly expressed genes serve as positive controls. Negative controls include genes not expressed at the desired stage. *Xenopus* liver DNA can be used as an external standard.

24. Dilute a small aliquot of the DNA recovered from Step 23 10-fold in 10 mM Tris-HCl (pH 8.0) for use in qPCR.

25. Perform qPCR using standard protocols and primers for your known target regions or the recommended reference genes (e.g., *eef1a1o, otx2, hbe1*).

Constructing the DNase-seq Library

26. Quantitate the total amount of DNase-digested DNA from Step 23 using a Qubit fluorometer. Use ~10 ng of DNA as input for library construction.

27. Build the library using a Nextflex ChIP-seq kit together with Illumina-compatible barcode adaptors under the direction of the kit manual.

We use 11 cycles of PCR amplification.

28. Determine the size distribution of the library using a Bioanalyzer 2100.

29. Quantitate the library concentration by qPCR (e.g., using a KAPA Library Quantification kit).

30. Sequence the library on an Illumina platform.

Analyzing Data

31. Align the sequencing reads to the *Xenopus* tropicalis v 9.0 genome assembly (ftp://ftp.xenbase.org/pub/Genomics/JGI/Xentr9.0/) using Bowtie v1.0.0 (Langmead et al. 2009) with the following command: "bowtie -S -v 2 -k 1 -m 1 --best –strata."

32. Remove the duplicate reads from a sorted BAM file using the 'rmdup' command in Samtools (Li et al. 2009).

33. Create a BigWig file using deepTool2 bamCoverage (Ramírez et al. 2016) and then visualize it using the Broad Institute's Integrative Genomics Viewer genome browser (Robinson et al. 2011).

34. Call DHS peaks using Homer (Heinz et al. 2010) with the following command: "findPeaks -style dnase -gsize 1.42e09."

TROUBLESHOOTING

Problem (Step 22): DNA laddering with multiple bands in ~150-bp increments is apparent after gel electrophoresis.

Solution: The DNA is over-digested. Reduce the amount of DNase I or increase the number of nuclei.

Recipes

Buffer A for DNase-seq

Reagent	Final concentration
Tris-HCl (1 M [pH 8.0])	15 mM
NaCl (5 M)	15 mM
KCl (1 M)	60 mM
EDTA (0.5 M [pH 8.0])	1 mM
EGTA (50 mM [pH 8.0])	0.5 mM
Spermine (500 mM)	0.5 mM
Pefabloc SC PLUS (Roche 11873601001) (20 mg/mL)	0.1 mg/mL
Dithiothreitol (DTT) (1 M)	2 mM
Protease inhibitor tablet (cOmplete, Mini, EDTA-free; Roche 04693159001)	1 tablet for 10 mL

Combine Tris-HCl, NaCl, KCl, EDTA, EGTA, and spermine. Store the buffer at 4˚C.

Immediately before use, add Pefabloc SC, DTT, and protease inhibitor.

DNase I Digestion Buffer (1X)

To make 5 mL of 1× DNase I digestion buffer, add 500 µL of DNase I Digestion Buffer (10×) <R> to 4.5 mL of Buffer A for DNase-seq <R>. Prepare fresh and equilibrate to 37˚C prior to use.

DNase I Digestion Buffer (10X)

Reagent	Final concentration
NaCl (5 M)	750 mM
$CaCl_2$ (1 M)	60 mM

Store for up to 1 year at room temperature.

DNase I stock solution (10 U/µL)

Reagent	Final concentration
Tris-HCl (1 M [pH 7.6])	20 mM
NaCl (5 M)	50 mM
$MgCl_2$ (1 M)	2 mM
$CaCl_2$ (1 M)	2 mM
Pefabloc SC PLUS (Roche 11873601001) (20 mg/mL)	0.1 mg/mL
Dithiothreitol (1 M)	1 mM
Glycerol (100%)	50%

Combine the reagents listed above to prepare storage buffer. On ice, solubilize an entire bottle of deoxyribonuclease I (DNase I [Sigma-Aldrich D4527; 10,000 U]) with 1 mL of ice-cold storage buffer. Prepare 50- to 100-µL aliquots and store at −20˚C.

Iodixanol solutions (20%, 25%, 30%)

Reagent	Final concentration
Tris-HCl (1 M [pH 8.0])	10 mM
Sucrose (1 M)	0.3 M
MgCl2 (1 M)	5 mM
KCl (1 M)	25 mM
NaF (500 mM)	10 mM
β-glycerophosphate (1 M)	5 mM
Sodium pyrophosphate (100 mM)	5 mM
Spermine (500 mM)	0.5 mM
Spermidine (500 mM)	0.5 mM
Pefabloc SC PLUS (Roche 11873601001) (20 mg/mL)	0.1 mg/mL
Dithiothreitol (1 M)	2 mM
Protease inhibitor tablet (cOmplete, Mini, EDTA-free)(Roche 04693159001)	1 tablet for 10 mL
Iodixanol (60%) (OptiPrep Density Gradient Medium;	20%; 25%; 30%

Sigma-Aldrich D1556)

Prepare fresh and keep at 4˚C.

Stop Buffer for DNase I

Reagent	Final concentration
Tris-HCl (1 M [pH 8.0])	50 mM
NaCl (5 M)	100 mM
SDS (20%)	0.10%
EDTA (0.5 M [pH 8.0])	100 mM
Spermine (500 mM)	1 mM
Spermidine (500 mM)	0.3 mM

Prepare fresh and equilibrate to 37oC prior to use.

Sucrose Solution (SS) (0.3 M)

Reagent	Final concentration
Tris-HCl (1 M [pH 8.0])	10 mM
Sucrose (1 M)	0.3 M
$MgCl_2$ (1 M)	5 mM
KCl (1 M)	25 mM
NaF (500 mM)	10 mM
β-Glycerophosphate (1 M)	5 mM
Sodium pyrophosphate (100 mM)	5 mM
Spermine (500 mM)	0.5 mM
Spermidine (500 mM)	0.5 mM
Pefabloc SC PLUS (Roche 11873601001) (20 mg/mL)	0.1 mg/mL
Dithiothreitol (DTT) (1 M)	2 mM
Protease inhibitor tablet (cOmplete, EDTA-free) (Roche 4693132001)	1 tablet for 50 mL

Combine the first nine ingredients and store the buffer at 4oC. Immediately before use, add Pefabloc SC, DTT and protease inhibitor.

REFERENCES

Farzaneh F, Pearson CK. 1978. A method for isolating uncontaminated nuclei from all stages of developing *Xenopus laevis* embryos. *J Embryol Exp Morphol* **48:** 101-108.

Graham J. 2002. Rapid purification of nuclei from animal and plant tissues and cultured cells. *Scientific World J* **2:** 1551–1554.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38:** 576–589.

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

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25:** 2078–2079.

Nieuwkoop PD, Faber J. 1967. *Normal Table of* Xenopus laevis *(Daudin): A Systematical & Chronological Survey of the Development from the Fertilized Egg till the End of Metamorphosis.* North-Holland Pub. Co., Amsterdam.

Neph S, Vierstra J, Stergachis AB, Reynolds AP, Haugen E, Vernot B, Thurman RE, John S, Sandstrom R, Johnson AK, et al. 2012. An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* **489:** 83–90.

Ogino H, McConnell WB, Grainger RM. 2006. Highly efficient transgenesis in *Xenopus tropicalis* using I-SceI meganuclease. *Mech Dev* **123:** 103–113.

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44:** W160–W165.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29:** 24–26.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9:** R137.

CHAPTER 3

The epigenetic regulation of the chromatin modifier Ezh2 requires the maternal transcription factor Foxh1 during *Xenopus* germ layer specification

Abstract

The maternal transcription factor, Foxh1, is known for its bookmarking role on selected mesendodermal CRMs before zygotic genome activation. However, it is unknown whether Foxh1 is recruited based on the chromatin state, or upon Foxh1 recruitment, it functions to affect the epigenetic remodeling. To address these questions, we initially identified Foxh1-associated proteins in mouse embryonic stem cells. Subsequently, I determined that critical components of Polycomb Repressive Complex 2 (PRC2) are all present in *Xenopus tropicalis* embryos and they interact physically with Foxh1 *in vivo*. The interaction is Foxh1-dependent as revealed by the loss of the PRC2 complex formation on mesendodermal CRMs using Foxh1-deficient *Xenopus* embryos. The physical interaction between Ezh2 and Foxh1 occurs well before the onset of Ep300 recruitment and the accumulation of the repressive histone mark, H3K27me3. These data suggest that Foxh1-like maternal TFs orchestrate the epigenetic states of target genes to ensure that specific gene regulatory programs are turned on in specific lineages.

Introduction

The three primary germ layers - ectoderm, mesoderm, and endoderm - are the first distinct cell types formed during vertebrate embryogenesis. To specify the three germ layers, gene expression of each cell in the embryo is tightly regulated by highly conserved gene regulatory programs (Loose and Patient, 2004; Koide et al., 2005). Molecular regulators, such as transcription factors (TFs), co-factors in signaling pathways, and co-regulators of the gene expression, bind the *cis*-regulatory modules (CRMs) of target genes and initiate germ layer-specific gene regulatory programs (Kiecker et al., 2016; Charney et al. 2017). Maternally deposited TFs are particularly critical regulators of germ layer specification. One such factor is T-box protein, Vegt (previously known as Xombi, Antipodean, or Brat, Lustig et al., 1996; Zhang and King, 1996; Stennard et al., 1996; Horb and Thomsen, 1997), which shows localized expression in the vegetal region and plays a pivotal role in the mesendodermal gene activation (Zhang et al., 1998; Kofron et al., 1999).

The forkhead protein, Foxh1 is another core maternal TF, which recruits phosphorylated Smad2/3, co-effectors of Nodal/TGF-b signaling to activate mesendodermal genes (Shen, 2007; Chiu et al., 2014). Unlike Vegt, Foxh1 is expressed ubiquitously in the *Xenopus* blastula embryo (Chiu et al., 2014; Charney et al., 2017; Paraiso et al., 2019) and binds mesendodermal CRMs (e.g., *gsc, cer1, nodal1* and *nodal2* enhancers). The binding occurs as early as at the 32-cell stage, which is considerably earlier than the onset of zygotic genome activation (Owen et al, 2016). Importantly, inhibition of Nodal signaling using a chemical inhibitor, SB-431542 does not block the binding of Foxh1 on CRMs, which indicates that Foxh1 binding to DNA is independent of Nodal signaling (Chiu et al., 2014; Charney et al., 2017). Foxh1 also mediates transcriptional regulation without

Smad2/3 and serves as both an activator and a repressor of target genes, which include *hand2* and *ssh1* (Charney et al., 2017; Chiu et al., 2014). Charney et al. showed the correlation of Foxh1 and Tle/Groucho (co-repressor) binding genome-wide in early blastula embryos (stage 8), which is the stage before the phosphorylation of Smad2/3 by Nodal signaling, and that Tle binding is diminished on certain mesendodermal CRMs in Foxh1-Morpholino oligonucleotide (MO) knockdown early blastula embryos. While maternal TFs are critically important during germ layer specification and confer their DNA binding activity before the onset of zygotic gene activation (ZGA) (Charney et al., 2017; Paraiso et al., 2019), the spatial-temporal dynamics among maternal TFs and other co-regulators and the interplay between maternal TF binding to DNA and the epigenetic landscape of the genome during this period are still not well understood.

Several different approaches were used to examine the epigenetic state of chromatin during ZGA. ATAC-seq and histone ChIP-seq analyses revealed that the overall chromatin state of the early mouse and human embryos is relatively open and unmodified (Wu et al., 2016; Wu et al., 2018; Liu et al., 2016; Zhang et al., 2016). Similarly in *Xenopus*, the chromatin of embryos around the time of ZGA is relatively unmodified (Akkers et al., 2009; Gupta et al., 2014; van Heeringen et al., 2014; Hontelez et al., 2015). The H3K27me3 mark, which detects heterochromatin, and H3K4me3 and H3K4me1 marks, which detect active promoters and enhancers, respectively, are not found in embryos at the early blastula stage (Akkers et al., 2009; Gupta et al., 2014; van Heeringen et al., 2014; Hontelez et al., 2015). Since many maternal TFs are already bound to DNAs before significant epigenetic modifications occur, this raises an important question. How do maternal TFs bind to the genome on pre-selected sites and confer subsequent epigenetic modifications?

Can maternal TFs like Foxh1 recruit chromatin modifier, thus directly linking the role of Foxh1 in epigenetic regulation? I will address some of these questions by focusing on the role of Foxh1 in recruiting chromatin modifier complexes.

Several pieces of evidence suggest that maternal TF, Foxh1 is a critical molecule in coordinating the gene regulatory program controlling germ layer specification. In addition to its ability to bind DNA at 32-cell stage, the Foxh1 bound sites are subsequently recognized by RNA polymerase II (RNAPII), subjected to histone H3K4me1 modification, and recognized by Ep300 (E1A Binding Protein P300) (Charney et al., 2017; Paraiso et al., 2019). To uncover the critical link between Foxh1 binding to DNA and epigenetic modifications of the surrounding Foxh1 bound regions, we set out to identify the proteins that interact with Foxh1. We reasoned that in addition to mediating Nodal signaling via Smad2/3, Foxh1 interacts with chromatin modifierand these interactors can be identified using mass spectrometry. We show that Foxh1 pulls down the components of Polycomb Repressive Complex 2 (PRC2), which includes SUZ12, JARID2, EZH2, and RBBP7. Using CRISPR/Cas9 technology, I obtained Foxh1 null mutant lines and show that Foxh1 is required for the recruitment of the PRC2 complex and regulates the chromatin state of early embryos.

Results

Polycomb-associated proteins interact with Foxh1

To identify co-factors that interact with Foxh1, we first identified Foxh1-associated proteins by tandem affinity purification-mass spectrometry (TAP-MS) analysis (Wang et al., 2014; Li et al., 2016) in mouse embryonic stem cells (mESCs). Foxh1 cDNA was cloned downstream of a C-terminal SFB-tag (S-protein tag, 2X Flag tag, and streptavidin Binding protein) and placed into a lentiviral vector (pLV-EF1a-IRES-Puro; Hayer et al., 2016). Three independent populations of mESCs were harvested and subjected to immunoprecipitation analysis. Two immunoprecipitations used two independent single clonally expanded mESC lines harboring stably expressing mFoxh1-SFB cDNA and one immunoprecipitation was obtained from mESCs that were transiently transfected with mFoxh1-SFB cDNA. We performed the standard tandem affinity purification (TAP) steps; first affinity purification using streptavidin beads, followed by a second affinity purification step using S protein beads (Li et al., 2016). The affinity-purified proteins were subjected to SDS-PAGE and excised for mass spectrometry analysis. We focused on 371 proteins that are present among all three replicates (Figure 3.1A). We performed Panther GO analysis (Thomas et al., 2003; Mi et al., 2019) to narrow down the search to focus on proteins present in the nucleus, by excluding extranuclear, ribosomal related proteins and proteins that are involved in non-transcriptional biological processes (Figure 3.1B).

Following GO analysis-based filtering, eighty-four genes (Figure 3.1B, Table 3.2) were subjected to gene ontology analysis using Metascape (http://metascape.org; Zhou et al., 2019). Those proteins are mostly categorized under transcriptional regulation (Figure 3.1C). One category, “Nodal signaling pathway” in the list validates that our TAP-MS

analysis of mFoxh1-SFB is specific enough to isolate Foxh1 interactors. Foxh1 itself is detected by a high number of peptides and Smad2 and Smad3 are also detected from all three replicates (Table 3.2). Positive and negative regulation of gene expression and chromatin organization are top-ranked functions (Figure 3.1C). SMARCA4 and 5, and HELLS on the list are the subunits of a nucleosome remodeling complex, SWI/SNF (SWItch/Sucrose Non-Fermentable; also called BAF (Brg/Brahma-associated factors)) which is essential for the embryonic stem cell (ESC) pluripotency (Peterson and Workman, 2000; Ho et al., 2009). TET1, OGT, and HDAC1 are involved in regulating CpG island methylation in ESCs (Vella et al., 2013). This suggests Foxh1 associated proteins are involved in the epigenetic regulation of gene expression.

Among all the Foxh1 interactors, we noted that critical components of PRC2 are all consistently detected. These include SUZ12, JARID2, enhancer of zeste homolog 2 (EZH2) and embryonic ectoderm development (EED) (Table 3.2). Since the PCR2 complex contributes to chromatin compaction and is responsible for the deposition of H3K27me3 to repress target genes (Margueron and Reinberg, 2011), they are excellent candidates to investigate their interactions and involvement in mediating Foxh1-dependent chromatin modifications. Transcripts of *suz12, jarid2, ezh2,* and *eed* were all present in eggs and are also maintained during early *Xenopus* embryonic development (Figure 3.2C). To confirm the direct physical interactions between PRC2 subunits and Foxh1, we performed immuno-western blot analyses using mESCs overexpressing HA-tagged SUZ12 and EZH2 and 3xFlag-tagged FOXH1 proteins. As shown in figure 3.2A and B, we detected specific interactions of FOXH1- EZH2 and FOXH1- SUZ12, indicating that PCR2 subunits physically interact with Foxh1. Since the interactions of FOXH1-EZH2 and FOXH1-SUZ12 proteins in

transfected cells could be indirectly mediated via tethering to DNA, we performed western blot in the presence of ethidium bromide and benzonase. Ethidium bromide intercalates DNA and significantly disrupts the conformation of DNA. Thus, if protein-protein interaction is mediated via DNA, this will interfere with the complex formation. Benzonase is a nuclease and will digest DNA that is present in the immunopreciptiation reaction. Figure 3.2A and B show that the interaction between PRC2 subunits and Foxh1 is not dependent on the presence of DNA.

Ezh2 binding is dynamic during early *Xenopus* embryogenesis

Ezh2 is a core subunit of PRC2 and a SET-domain-containing histone methyltransferase (Cao and Zhang 2004). Its catalytic enzyme activity is responsible for depositing methylation on histone 3 lysine 27 (Cao and Zhang 2004; Nekrasov et al., 2005; Tie et al., 2007). In *Xenopus*, Ezh2 genome-wide binding was measured at the late blastula stage (stage 9) and only a small subset of its binding sites gained H3K27me3 during subsequent developmental stages (van Heeringen et al., 2014). To better understand the dynamic activity of Ezh2, I performed Ezh2 ChIP-seq analysis using embryos from early blastula (stage 7, ~256 cells in an embryo) to early gastrula (stage 10.5) stages (Figure 3.3). This developmental time window corresponds to the beginning stage of zygotic genome activation (stage 7), the stage forming three germ layers (stage 8-9), and the stage of gastrulation movements (stage 10.5). Ezh2 binding was detected through those stages.

To learn more about Ezh2 binding dynamics, we identified high confidence Ezh2 peaks from stage 8 and stage 10.5 samples using 'irreproducibility discovery rate' (IDR) analysis (Li et al., 2011). After implementing IDR analysis between two biological replicates

at each stage, 2,215 and 11,973 Ezh2 peaks survived the IDR analysis at stage 8 and 10.5, respectively. Forty-nine percent of stage 8 Ezh2 peaks (1,086 peaks) overlapped with stage 10.5 peaks (Figure 3.3B). Interestingly, the overlapping peaks (cluster II, Figure 3.3B-C) persist through early *Xenopus* developmental stages between stage 7-10.5 (Figure 3.3C). On the other hand, Ezh2 peaks in cluster I represent transient binding to DNA in the early blastula stage 7 and 8. Peaks in cluster III represent Ezh2 binding occurring after the late blastula stage into the gastrula stage, with increased binding at the early gastrula stage compared to the blastula stage. I examined the overlap between Ezh2 and Jarid2 peaks, and note that clusters II and I show extensive overlap with Jarid2 peaks, suggesting that the Ezh2-Jarid2 PRC2 complex is formed around these peaks at early stages. However, the weak peaks of Jarid2 in cluster III suggest that Ezh2 at later stages is in a PRC2-complex-independent of Jarid2. It has been demonstrated previously that Ezh2 can bind to DNA and activate the transcriptional response of some target genes (Kim et al., 2018). Interestingly, the peaks in cluster III show a strong correlation with the binding of Ep300 and DNaseI sensitive regions. This suggests that genes associated with cluster III peaks may be activated at later developmental stages. I also examined the presence of H3K27ac peaks around Ezh2, but could only identify weak signals (Gupta et al., 2014) (Figure 3.3C).

Ezh2 forms a complex with Foxh1

It has been shown that Ezh2 does not bind to DNA directly (Lynch et al., 2012; Schuettengruber et al., 2017), and requires the presence of other DNA-binding proteins. To identify a partner TF that mediates Ezh2 binding during the early gem layer specification and ZGA, I performed a motif enrichment search (MEME-ChIP) within a window size of

500bp surrounding Ezh2 peak summits (Figure 3.4A). Cluster I peaks that show co-binding of Ezh2 and Jarid2 are strongly enriched with Foxh1 motifs. Cluster 2 Ezh2 peaks also show enrichment of Foxh1 motifs, followed by Pou and Sox motifs. Interestingly, cluster 3 shows the strongest preference for the Sox motif. These data indicate that Ezh2 recruitment to DNA is dynamic and regulated by different transcription factors at different developmental stages.

Since Fox and Sox motifs are highly enriched surrounding Ezh2 peaks, and Foxh1 and Sox3 are major maternal TFs expressed at the stage, I examined the relationship between Ezh2-Foxh1 and Ezh2-Sox3 using ChIP-seq at stage 8 and stage 10.5 (Figure 3.4B, C). At stage 8, 66% of Ezh2 peak locations coincide with both Foxh1 and Sox3 peaks. At stage 10.5 there is a significant increase in the appearance of Ezh2 peaks, and 59% of Ezh2 peaks coincide with the location of Sox3 peaks, whereas only 6% of Ezh2 peaks overlap with Foxh1 peaks. This finding confirms the observation that Ezh2 peaks in gastrula stage are highly enriched with Sox motifs, and that the likely transcription factor partner is Sox3.

To demonstrate the direct physical interaction between Ezh2 and Foxh1, I performed a sequential ChIP-qPCR analysis. First-round ChIP was performed using the Ezh2 antibody, followed by a second ChIP using anti-Foxh1 or HA (as a mock control) antibodies (Figure 3.4E-F). I chose three positive CRMs (pitx2, gata2, and zic3 enhancers) that as active enhancer regions – pitx2 intron, gata2 0.7kb upstream, and zic3 5kb upstream- are enriched over the control. This data demonstrate that Ezh2 and Foxh1 form a complex in early *Xenopus* embryos.

Foxh1-deficient mutant embryos uncover the role of Ezh2 in H3K27me3 activity

Since Ezh2 interacts with Foxh1 in blastula embryos, we investigated whether maternal Foxh1 binding is required for the recruitment of Ezh2 during embryonic development. Previously we have successfully knocked down Foxh1 protein synthesis by microinjecting an antisense morpholino oligonucleotide (MO) targeting foxh1 mRNA in fertilized embryos (Chiu et al., 2014). While the approach is efficient, the morphant phenotypes are not full null phenotypes. In addition, the availability of thousands of Foxh1-deficient embryos is needed to perform ChIP-seq experiments. We, therefore, wished to generate a *Xenopus tropicalis* null mutant line deficient in foxh1 using a CRISPR-Cas9 genome editing approach (Blitz et al., 2013; Nakayama et al., 2013). We targeted the beginning of Foxh1 translation regions by designing a specific gRNA (Foxh1-G58 used for PAM site, Table 3.2). Together with hCas9 mRNA, gRNA was injected into fertilized embryos, and the resulting embryos were raised to mature female frogs. Eggs from Foxh1 F0 CRISPR females were *in vitro* fertilized using wild-type (WT) male sperm. F1 embryos resulting from two F0 female lines were 100% embryonic lethal displaying severe axial defects (Figure 3.5A), suggesting that all eggs resulting from these females have out-of-frame indels. All control sibling embryos (wild type) grew normally. We genotyped F1 embryos resulting from the CRISPR-targeted F0 females by sequencing the *foxh1* locus. All embryos resulting from the two independent F0 females had frame-shift mutations (Table 3.2). Microinjection of wild type *foxh1* mRNA into the Foxh1-deficient F1 embryos full rescued the embryonic lethal phenotype and these embryos grew normally (Figure 3.5A). This indicates that the lethality of Foxh1 CRISPR mutants is solely caused by the loss of Foxh1 transcripts in eggs.

I also performed immunoprecipitation and western blot (IP-western) analysis of Foxh1-deficient F1 embryos to ensure that the Foxh1 mutants are true null. I isolated

protein extracts from both Foxh1 F1 mutant embryos and control sibling WT control embryos. IP-western blot analysis of the extracts using anti-Foxh1 antibody showed that Foxh1 protein is absent in F0 heterozygous mutants, demonstrating that eggs derived from Foxh1-deficient F0 females are deficient in Foxh1 protein expression.

Next, I performed Ezh2 ChIP-seq analysis on Foxh1-deficient embryos. Ezh2 binding surrounding the enhancer regions of *gata2* and *pitx2* was undetectable (Figure 3.6A), and Ezh2 binding was significantly reduced in Foxh1 mutants at the whole genome level (Figure 3.6B). This data shows that Foxh1 is required for Ezh2 recruitment during early *Xenopus* embryonic development. Interestingly, Ep300 binding was not affected by Foxh1 depletion. The data suggest that Ep300 recruitment and Ezh2 recruitment via Foxh1 may be regulated by two independent events.

Ezh2 is the catalytic subunit of PRC2, which has histone methyltransferase activity to deposit trimethyl groups to lysine 27 of histone 3 (H3K27me3)(Cao and Zhang 2004; Schuettengruber et al. 2009). Since Ezh2 recruitment is dependent on Foxh1, I tested the possibility that H3K27me3 is mediated by Foxh1, and that in the absence of Foxh1, H3K27me3 modification is significantly compromised. Reported ChIP-seq analysis showed that the H3K27me3 mark is not prevalent in early developmental stages (van Heeringen et al., 2014; Hontelez et al., 2015). While I detected a reduction of H3K27me3 peaks in some regions (Figure 3.6C), overall the change was not statistically significant at the whole genome level (Figure 3.6 D).

H3K27me3 activity is regionally regulated

I examined the distribution of the H3K27me3 mark in different germ layers, as the PRC2 activity may be different among three germ layers. Wild type embryos were manually dissected into ectoderm and endoderm explants and H3K27me3 levels were examined in each specific germ layer by performing ChIP-seq analysis. H3K27me3 marks are detected in both ectoderm and endoderm layers, but clearly, the degrees of H3K27me3 were different. For example, H3K27me3 marks are more enriched in ectodermal cells in *Xenopus* embryos around the endodermal genes such as *gata4* and *sox17*s (*sox17a, sox17b1,* and *sox17b2*) (Figure 3.6C-D). Whereas the same endodermal genes are devoid of H3K27me3 modification in endodermal cells. The mesoderm specific gene, *eomes* was similarly strongly marked with H3K27me3 and its enrichment was also higher in the ectodermal layer than the endodermal layer (Figure 3.6B).

I also note that there were only a few ectodermal genes that were marked with H3K27me3. In the case of ectodermal genes, *dlx6* and *dlx5* (Figure 3.6A), H3K27me3 enrichment was seen in both germ layers. However, this analysis is somewhat limited since few ectodermal specific genes were subjected to H3K27me3 modification. Additional study is needed using Foxh1 mutant embryos to determine the link between Foxh1 and Ezh2, and the PCR2 complex activity. Nonetheless, from the limited data, I conclude that in general germ layer-specific H3K27me3 enrichment occurs on the endodermal genes, and this activity seems to be specifically restricted to the ectodermal territory. This suggests that chromatin remodeling is spatially regulated during early embryogenesis. I propose that this localized PRC2 activity on endodermal genes provides an attractive model to explain how endoderm and mesoderm germ layer-specific genes remain active in mesoderm and

endoderm, but at the same time, the same genes are silenced in another tissue or cell types (such as ectoderm) where these genes are not needed.

Discussion

How chromatin remodelers are recruited to their genomic target regions is one of the major questions about the epigenetic regulation of gene expression. Here we showed the maternal transcription factor, Foxh1 is required to recruit the histone modification enzyme, Ezh2 - a core subunit of the PRC2 complex which deposits the repressive histone mark, H3K27me3 - during early *Xenopus* embryogenesis. We showed the reduction of Ezh2 binding at early blastula embryos is Foxh1-dependent. First, using both mESCs and *Xenopus* embryos, Foxh1 was shown to physically interact with Ezh2. Second, the recruitment of Ezh2 to DNA requires the presence of Foxh1 as Ezh2 was unable to bind to DNA in Foxh1-deficient embryos. However, in the absence of Foxh1 pre-marking, and thus in the absence of Ezh2 recruitment, I was able to detect only a small fraction of genes showing differential H3K27me3 marks. I also observed the differential modification of endodermal genes in different germ layers and propose that germ layer-specific H3K27me3 marking is an important regulatory process conferring germ layer-specific gene expression.

Foxh1 recruits histone modifier Ezh2

Based on this data, I propose the following mechanism (Figure 3.8). Maternal TF, Foxh1 binds to CRMs at least as early as 32-cell stage and Ezh2 is recruited on Foxh1 binding sites at the early blastula stage (stage 7, about 250 cells per embryo). Before ZGA, Foxh1 is prebound to DNA together with Ezh2 throughout the whole embryo. Since Foxh1 is ubiquitously expressed in the blastula embryos, and Ezh2 is detected in both the animal and the vegetal side of 8-cell embryos, then its target genes are transcriptionally silent. After ZGA, in the endodermal cells of the embryo, Nodal signaling is activated via the

transcriptional activation of nodal genes by maternally-expressed Vegt (Zhang et al., 1998; Kofron et al., 1999). Since Foxh1 is a core component of Nodal signaling, Smad2/3 will be recruited to the site and Ezh2 will be evicted from the CRMs. This results in activation of endodermal genes when the Nodal signaling pathway is activated (Chen et al., 1996; Chen et al., 1997; Yoon et al., 2011; Chiu et al., 2014). This eviction mechanism by Smad2/3 predicts that the Ezh2-Foxh1 complex formed around endodermal genes will not be evicted in the ectoderm, as Nodal signaling is absent in ectoderm. Thus in the ectodermal cells, endoderm genes are actively silenced by the recruitment of PRC2 complex by Foxh1-Ezh2 complex. This model is attractive as it explains how the cell ensures activation of genes for appropriate lineage development while keeping inappropriate genes repressed to prevent alternate lineage commitment.

In the case of ectodermal genes, H3K27me3 marking is less prevalent than in endodermal genes. One possible explanation of this phenomenon is that PRC2 activity would be different in each germ layer of the gastrula embryos. We see stage 10.5 Ezh2 peaks are more overlapping with Sox3 than Foxh1 and Sox3 peaks also overlap with Ezh2 at stage 8 (Figure 3.4B-D). Since Sox2 and Ezh2 both are known for the involvement of ectoderm specification in embryonic stem cell differentiation (Amador-Arjona et al., 2015; Juan_2016; Shan et al., 2017) and both Sox2 and Sox3 are SOXB1 transcription factors, then maternally expressed Sox3 in *Xenopus* could play similar roles as mouse SOX2 for ectoderm specification. The conserved expression pattern of maternal SoxB1 in the animal pole (future ectodermal region) and zygotic SoxB1 in the nascent ectoderm in chordate embryos has been reported (Cattell et al., 2012). Therefore, at stage10.5 Ezh2 might have a stronger

interaction with Sox3 than Foxh1 in ectodermal cells. This Sox3 interaction with Ezh2 in ectodermal cells needs to be more carefully tested with zygotic Sox2 binding in the future.

While I prefer the above model described in Figure 3.8, an alternative model is that Ezh2/PRC2 activity is more stable or abundant in the ectodermal germ layer after blastula stages. This is supported by the view that Ezh2/PRC2 is more active in neural and neural crest cells during *Xenopus* neurulation (Reijnen et al., 1995; Aldiri and Vetter, 2009; Tien et al., 2015). However, events in neurulation at later, post-gastrulation developmental stages, would be less relevant to Foxh1-mediated regulatory programs and more related to zygotic ectodermal TFs like Sox2.

Foxh1 mediated PRC2.2 complex recruitment

Not only the recruitment of PRC2 but also the maintenance of PRC2 at its target site is required for PRC2 activity - deposition and spreading of H3K27me3 mark. In *Drosophila*, it was shown that the presence of H3K27me3 was not sufficient to maintain PRC2 at target sites and propagate repressive chromatin mark (Laprell et al., 2017; Coleman and Struhl, 2017). Continuous PRC2 recruitment would be required to maintain and propagate PRC2 and H3K27me3 to regulate gene expression. It has been suggested that PRC2 facultative subunits could have this function because the core subunits of PRC2 are considered not to possess this ability (Laugesen et al., 2016). Based on the components of the facultative subunits, PRC2 have at least two distinct multimeric protein complexes - PRC2.1 and PRC2.2 - and those facultative subunits are mutually exclusive to either complex (Hauri et al., 2016). PRC2.1 includes one of the PCL proteins (PHF1, MTF2, or PHF19) and EPOP or PALI1, and PRC2.2, by contrast, includes AEBP2 and JARID2 and these non-core PRC2

subunits distinctively regulate PRC2 recruitment and activity (Alekseyenko et al., 2014; Hauri et al., 2016; Grijzenhout et al., 2016). Our Foxh1 mESCs mass-spec data does not list any PRC2.1 facultative subunits but does show one of the PRC2.2 facultative subunits, Jarid2. From the Jarid2 ChiP-seq, we see Jarid2 binding overlapping with Foxh1 and Ezh2 binding in stage 9 *Xenopus* embryos (Figure3.3D). Therefore the PRC2.2 complex including Ezh2 and Jarid2 would be recruited on Foxh1-dependent target sites in early *Xenopus* embryos. It would be interesting to test PRC2.1 involvement after ZGA and later developmental stages when maternal Foxh1 is removed or to Foxh1 independent PRC2 recruitment.

Multiple TFs are involved for PRC2 recruitment on DNA

Motif enrichment analysis of Ezh2 binding regions (Figure 3.4A) indicates Ezh2-bound region recruits Sox and Pou-like TFs in addition to Foxh1. Jarid2 peaks are also enriched with Fox, Sox, and Pou motifs (van Heeringen et al., 2014). This suggests PRC2 is recruited to sites where Fox, Sox or Pou TF binds. Furthermore, ChIP-seq analysis of Sox3 shows a strong correlation between Sox3 and Ezh2-Foxh1 binding regions (Figure 3.4B). Together, these pieces of evidence suggest the co-occupancy of multiple TFs, forming enhanceosomes. Co-localization of PcG components with pluripotency factors Oct4, Sox2, and Nanog in ESCs was previously reported (Bernstein et al., 2006; Boyer et al., 2006; Lee et al., 2006), and those Oct4, Sox2, and Nanog co-bound regions in ESCs form super-enhancers (SE) to control the pluripotent state. The SE is distinguished from typical enhancers by size, transcription factor density and content, and the ability to activate transcription of the target genes (Whyte et al., 2013). Recently Paraiso et al. (2019)

identified *Xenopus* endodermal SEs, which are the large enhancer clusters with the co-occupancy of three maternal TFs - Otx1, Vegt, and Foxh1 (Paraiso et al., 2019). These endodermal SEs contain not only stronger signals of enhancer marks - H3K4me1, Ep300, and H3K27ac - but also have stronger signals of polycomb markers - H3K27me3, Jarid2, and Ezh2 - than regular enhancers (Paraiso et al., 2019). While highly speculative, it is tempting to propose that, Ezh2/PCR2 complex recruitment is mediated by multiple Foxh1/Sox/Pou TFs forming enhanceosomes to promote the PRC2 complex formation, and subsequent epigenetic modifications such as H3K27me3.

The interplay between maternal TFs and PRC2 for spatial regulation of gene expression

Here I showed the regionally differential H3K27me3 activity in early *Xenopus* gastrula embryos. Germ layer-specific H3K27me3 enrichment was especially monitored on the endodermal genes in the ectodermal tissue of the embryo. This localized PRC2 activity on endodermal genes suggests that chromatin remodeling is spatially regulated during early embryogenesis. Ezh2, the core catalytic subunit of PRC2, depends on Foxh1 for recruitment. Foxh1 binding sites are enriched with Sox and Pou motifs (Chiu et al., 2014; Charney et al., 2017), which suggests the interaction of multiple TFs on their *cis*-regulatory regions. I also detected the presence of Sox and Pou motifs besides the Foxh1 motif underneath Ezh2 peaks and the preference of Sox motif changed in the temporal binding dynamics of Ezh2 (Figure 3.4A and B). Moreover Sox3 bindings overlap with Ezh2-Foxh1 binding sites and *de novo* Ezh2 binding at stage 10.5 (cluster 3 in Figure 3.4B) is more strongly correlated with Sox3 binding than Foxh1. These results suggest the involvement of

multiple TFs on chromatin modifier recruitment. Although, Sox3 is a known ectodermal TF and its expression level has been reported in the animal cap side of the embryo (Cattell et al., 2012), RNA-seq data from dissected 8-cell embryos (De Dominico et al., 2015; Paraiso et al., 2019) showed the high transcription level of *sox3* in both animal cap and vegetal mass cells and therefore maternal Sox3 may have a distinctive role in early embryogenesis other than the known roles in ectodermal lineage specification and neurulation. Localized maternal TFs can give spatially different combinatorial TF bindings on their target sites and result in spatially differential regulation of gene expression and chromatin states. In-depth measurement of the spatial activity of these TFs as well as biochemical studies will help to understand the mechanism mediated by Ezh2/PRC2.

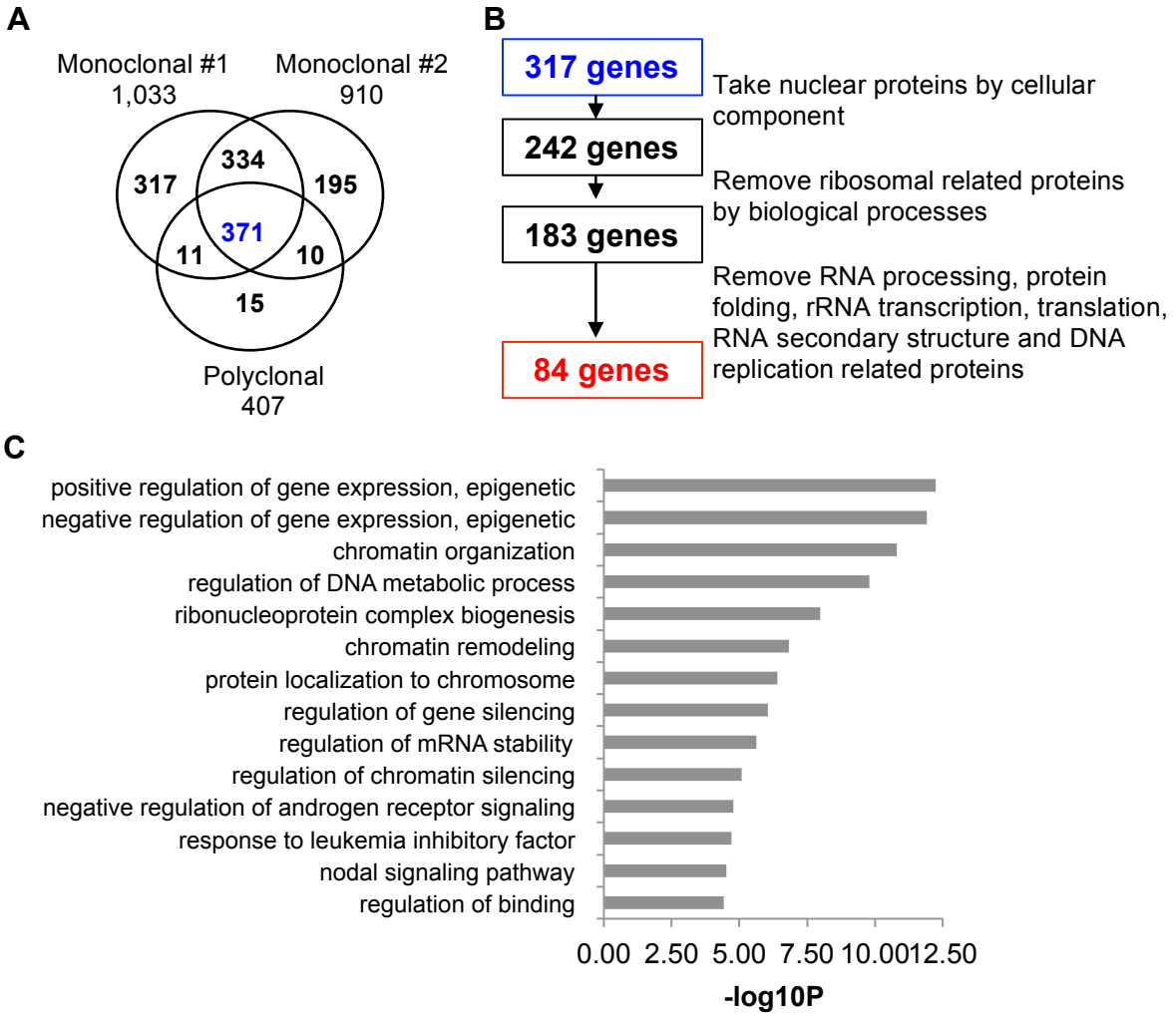

Figure 3.1 Foxh1-associated proteins are involved in epigenetic regulation of gene expression A) Venn diagram of Foxh1 associated proteins in three populations of mFoxH1-SFB transduced mESCs (E14) that are detected by mass spectrometry. B) Flowchart to select nuclear proteins by through GO analysis-biased filtering from proteins detected from all three samples of Fig3.1A. C) GO analysis using Foxh1-associated 84 nuclear protein.

Table 3.1. List of 84 nuclear proteins (Firgure 3.1B) from Foxh1 baited mass spectrometry in mESCs with its biological GO description

Input ID	Description	Biological Process (GO)
Actb	actin beta	GO:0051623 positive regulation of norepinephrine uptake
Aifm1	apoptosis inducing factor mitochondria associated 1	GO:1904045 cellular response to aldosterone
Ankrd17	ankyrin repeat domain 17	GO:1900245 positive regulation of MDA-5 signaling pathway
Atp2a2	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 2	GO:1903233 regulation of calcium ion-dependent exocytosis of neurotransmitter
Cebpz	CCAAT enhancer binding protein zeta	GO:0045944 positive regulation of transcription by RNA polymerase II
Ckap4	cytoskeleton associated protein 4	GO:0043312 neutrophil degranulation
Dnajc21	DnaJ heat shock protein family (Hsp40) member C21	GO:0006457 protein folding
Dnajc9	DnaJ heat shock protein family (Hsp40) member C9	GO:0032781 positive regulation of ATPase activity
Dnmt3l	DNA methyltransferase 3 like	GO:1905642 negative regulation of DNA methylation
Drg1	developmentally regulated GTP binding protein 1	GO:1901673 regulation of mitotic spindle assembly
Esco2	establishment of sister chromatid cohesion N-acetyltransferase 2	GO:0034421 post-translational protein acetylation
Ezh2	enhancer of zeste 2 polycomb repressive complex 2 subunit	GO:0098532 histone H3-K27 trimethylation
Fam98b	family with sequence similarity 98 member B	GO:0006388 tRNA splicing, via endonucleolytic cleavage and ligation
Foxh1	forkhead box H1	GO:1900164 nodal signaling pathway involved in determination of lateral mesoderm left/right asymmetry
Fus	FUS RNA binding protein	GO:1905168 positive regulation of double-strand break repair via homologous recombination
Fxr1	FMR1 autosomal homolog 1	GO:2000637 positive regulation of gene silencing by miRNA
G3bp1	G3BP stress granule assembly factor 1	GO:0090090 negative regulation of canonical Wnt signaling pathway
Glyr1	glyoxylate reductase 1 homolog	GO:0055114 oxidation-reduction process
Gnai2	G protein subunit alpha i2	GO:0140199 negative regulation of adenylate cyclase-activating adrenergic receptor signaling pathway involved in heart process
Gnl3	G protein nucleolar 3	GO:1904816 positive regulation of protein localization to chromosome, telomeric region
Gnl3l	G protein nucleolar 3 like	GO:1904816 positive regulation of protein localization to chromosome, telomeric region
Gpx4	glutathione peroxidase 4	GO:0110076 negative regulation of ferroptosis
Hdac1	histone deacetylase 1	GO:0061198 fungiform papilla formation
Hdlbp	high density lipoprotein binding protein	GO:0034384 high-density lipoprotein particle clearance
Hells	helicase, lymphoid specific	GO:0031508 pericentric heterochromatin assembly
Hist1h1a	histone cluster 1 H1 family member a	GO:0031936 negative regulation of chromatin silencing
Hist1h1e	histone cluster 1 H1 family member e	GO:0098532 histone H3-K27 trimethylation
Hnrnpa0	heterogeneous nuclear ribonucleoprotein A0	GO:0070935 3'-UTR-mediated mRNA stabilization
Hnrnpab	heterogeneous nuclear ribonucleoprotein A/B	GO:0001837 epithelial to mesenchymal transition
Hnrnpd	heterogeneous nuclear ribonucleoprotein D	GO:1905663 positive regulation of telomerase RNA reverse transcriptase activity
Hnrnpr	heterogeneous nuclear ribonucleoprotein R	GO:0061157 mRNA destabilization
Jarid2	jumonji and AT-rich interaction domain containing 2	GO:0034721 histone H3-K4 demethylation, trimethyl-H3-K4-specific
Lbr	lamin B receptor	GO:0030223 neutrophil differentiation
Lyar	Ly1 antibody reactive	GO:0048821 erythrocyte development
Mtdh	metadherin	GO:0031663 lipopolysaccharide-mediated signaling pathway
Mybbp1a	MYB binding protein 1a	GO:2000210 positive regulation of anoikis
Myef2	myelin expression factor 2	GO:2000815 regulation of mRNA stability involved in response to oxidative stress
Ncl	nucleolin	GO:1901838 positive regulation of transcription of nucleolar large rRNA by RNA polymerase I
Nemf	nuclear export mediator factor	GO:1990116 ribosome-associated ubiquitin-dependent protein catabolic process
Nifk	nucleolar protein interacting with the FHA domain of MKI67	GO:0009303 rRNA transcription;GO:0098781 ncRNA transcription
Nop16	NOP16 nucleolar protein	GO:0042273 ribosomal large subunit biogenesis
Nsd1	nuclear receptor binding SET domain protein 1	GO:0000414 regulation of histone H3-K36 methylation
Nvl	nuclear VCP like	GO:1904749 regulation of protein localization to nucleolus
Nxf1	nuclear RNA export factor 1	GO:0016973 poly(A)+ mRNA export from nucleus

Ogt	O-linked N-acetylglucosamine (GlcNAc) transferase	GO:0061087 positive regulation of histone H3-K27 methylation
Pfkp	phosphofructokinase, platelet	GO:0061621 canonical glycolysis;GO:0061620 glycolytic process through glucose-6-phosphate
Phb	prohibitin	GO:2000323 negative regulation of glucocorticoid receptor signaling pathway
Pisd	phosphatidylserine decarboxylase	GO:0006646 phosphatidylethanolamine biosynthetic process
Polr1c	RNA polymerase I and III subunit C	GO:0006362 transcription elongation from RNA polymerase I promoter
Polr1e	RNA polymerase I subunit E	GO:0001188 RNA polymerase I preinitiation complex assembly
Polr2a	RNA polymerase II subunit A	GO:0050434 positive regulation of viral transcription
Polr2e	RNA polymerase II subunit E	GO:0050434 positive regulation of viral transcription
Prdx1	peroxiredoxin 1	GO:0019430 removal of superoxide radicals
Prkci	protein kinase C iota	GO:0046326 positive regulation of glucose import
Rbbp7	RB binding protein 7, chromatin remodeling factor	GO:0034080 CENP-A containing nucleosome assembly
Rbm34	RNA binding motif protein 34	
Rcc2	regulator of chromosome condensation 2	GO:0072356 chromosome passenger complex localization to kinetochore
Rps19bp1	ribosomal protein S19 binding protein 1	
Rrp12	ribosomal RNA processing 12 homolog	
Ruvbl1	RuvB like AAA ATPase 1	GO:1904874 positive regulation of telomerase RNA localization to Cajal body
Ruvbl2	RuvB like AAA ATPase 2	GO:1904874 positive regulation of telomerase RNA localization to Cajal body
Senp3	SUMO specific peptidase 3	GO:0016926 protein desumoylation;
Serbp1	SERPINE1 mRNA binding protein 1	GO:0030578 PML body organization
Skp1	S-phase kinase associated protein 1	GO:0002223 stimulatory C-type lectin receptor signaling pathway
Smad2	SMAD family member 2	GO:1900224 positive regulation of nodal signaling pathway involved in determination of lateral mesoderm left/right asymmetry
Smad3	SMAD family member 3	GO:0019049 evasion or tolerance of host defenses by virus
Smarca5	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5	GO:0034080 CENP-A containing nucleosome assembly
Suz12	SUZ12 polycomb repressive complex 2 subunit	GO:0070734 histone H3-K27 methylation
Tecr	trans-2,3-enoyl-CoA reductase	GO:0035338 long-chain fatty-acyl-CoA biosynthetic process
Tet1	tet methylcytosine dioxygenase 1	GO:0090310 negative regulation of methylation-dependent chromatin silencing
Tex10	testis expressed 10	GO:0006364 rRNA processing
Tmpo	thymopoietin	GO:0006355 regulation of transcription, DNA-templated
Trim25	tripartite motif containing 25	GO:1902187 negative regulation of viral release from host cell
Trim28	tripartite motif containing 28	GO:1902187 negative regulation of viral release from host cell intermediate
Tubb4a	tubulin beta 4A class IVa	GO:0031115 negative regulation of microtubule polymerization
Tubb4b	tubulin beta 4B class IVb	GO:0097711 ciliary basal body-plasma membrane docking
Urb2	URB2 ribosome biogenesis homolog	GO:0042254 ribosome biogenesis;GO:0022613 ribonucleoprotein complex biogenesis
Vcp	valosin containing protein	GO:1903007 positive regulation of Lys63-specific deubiquitinase activity
Vdac2	voltage dependent anion channel 2	GO:0007339 binding of sperm to zona pellucida
Wdr18	WD repeat domain 18	GO:0006267 pre-replicative complex assembly involved in nuclear cell cycle DNA replication
Ybx3	Y-box binding protein 3	GO:1902219 negative regulation of intrinsic apoptotic signaling pathway in response to osmotic stress
Yme1l1	YME1 like 1 ATPase	GO:0006851 mitochondrial calcium ion transmembrane transport
Zscan4c	None	None

Table 3.2. Top-ranked Foxh1-associated proteins based on the average number of peptides detected from all three replicates from Foxh1 baited mass spectrometry

Protein	n						
NCL	75	TUBB4A	6	HIST1H1E	4	GPX4	2
MYBBP1A	68	AIFM1	5	PHB	3	JARID2**	2
FOXH1	67	POLR1E	5	HELLS	3	SENP3	2
TRIM28	33	RUVBL2	5	TET1	3	HNRNPD	2
GNL3	17	RCC2	5	VCAC2	3	DNMT3L	2
ACTB	9	POLR1C	4	OGT	3	EZH2**	2
POLR2A	9	HIST1H1A	4	SKP1	3	HDAC1	2
FXR1	8	SNSD1	4	SMAD3*	3	POLR2E	2
SMAD2*	8	SMARCA5	4	SUZ12**	3		
RUVBL1	7	GNL3L	4	FAM98B	2		

*: FOXH1 known interaction proteins

**: subunits of PRC2 complex

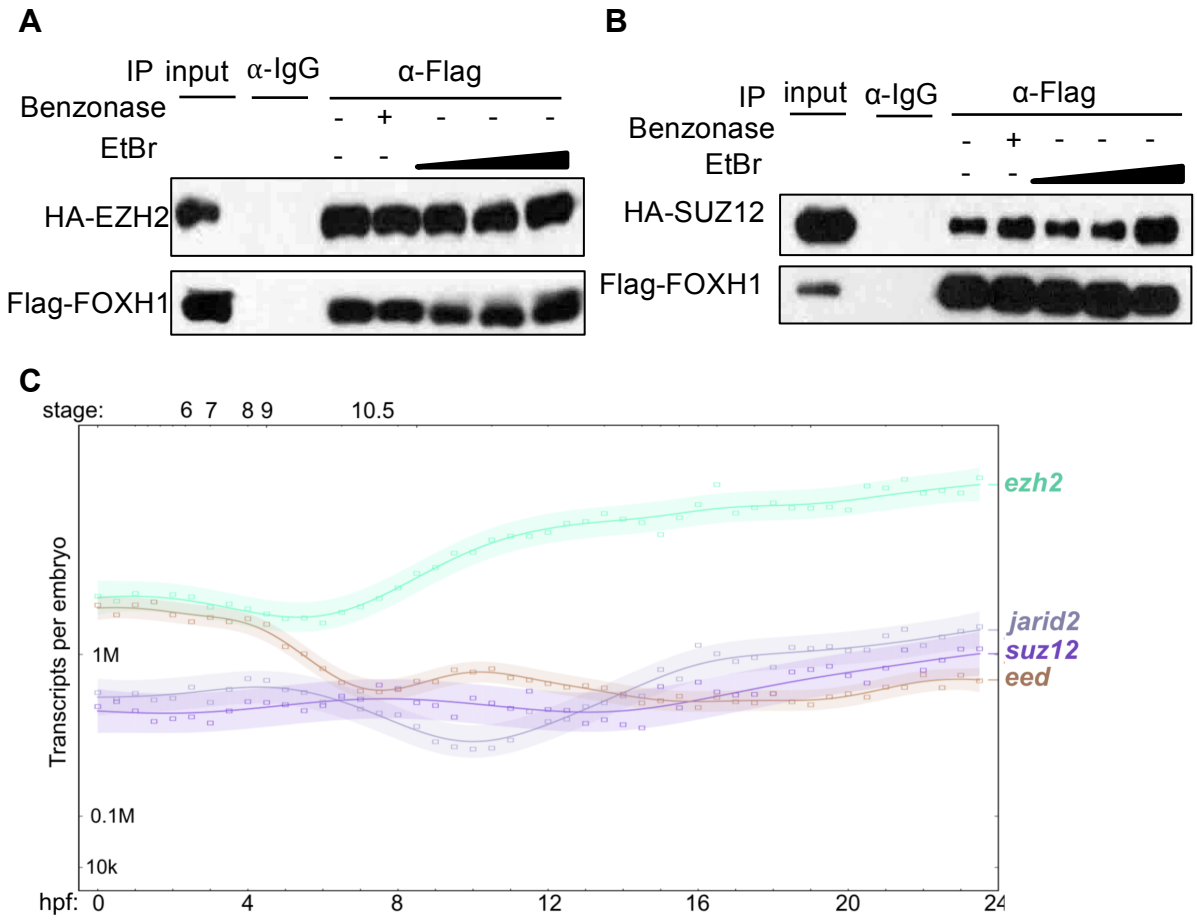

Figure 3.2 Polycom-associated proteins interact with Foxh1 IP-western blot analysis of (A) HA-Ezh2 and Flag-Foxh1 and (B) HA-Suz12 and Flag-Foxh1 in HA-PRC2 subunits and 3XFlag-Foxh1 transfected HEK293T cells. Benzonase or ethidium bromide (EtBr) was treated for the nucleic acid free condition. C) Temporal transcriptional expression level of PRC2 subunits by mRNA-seq over the course of early *Xenopus* tropicalis development (Owen et al., 2016)

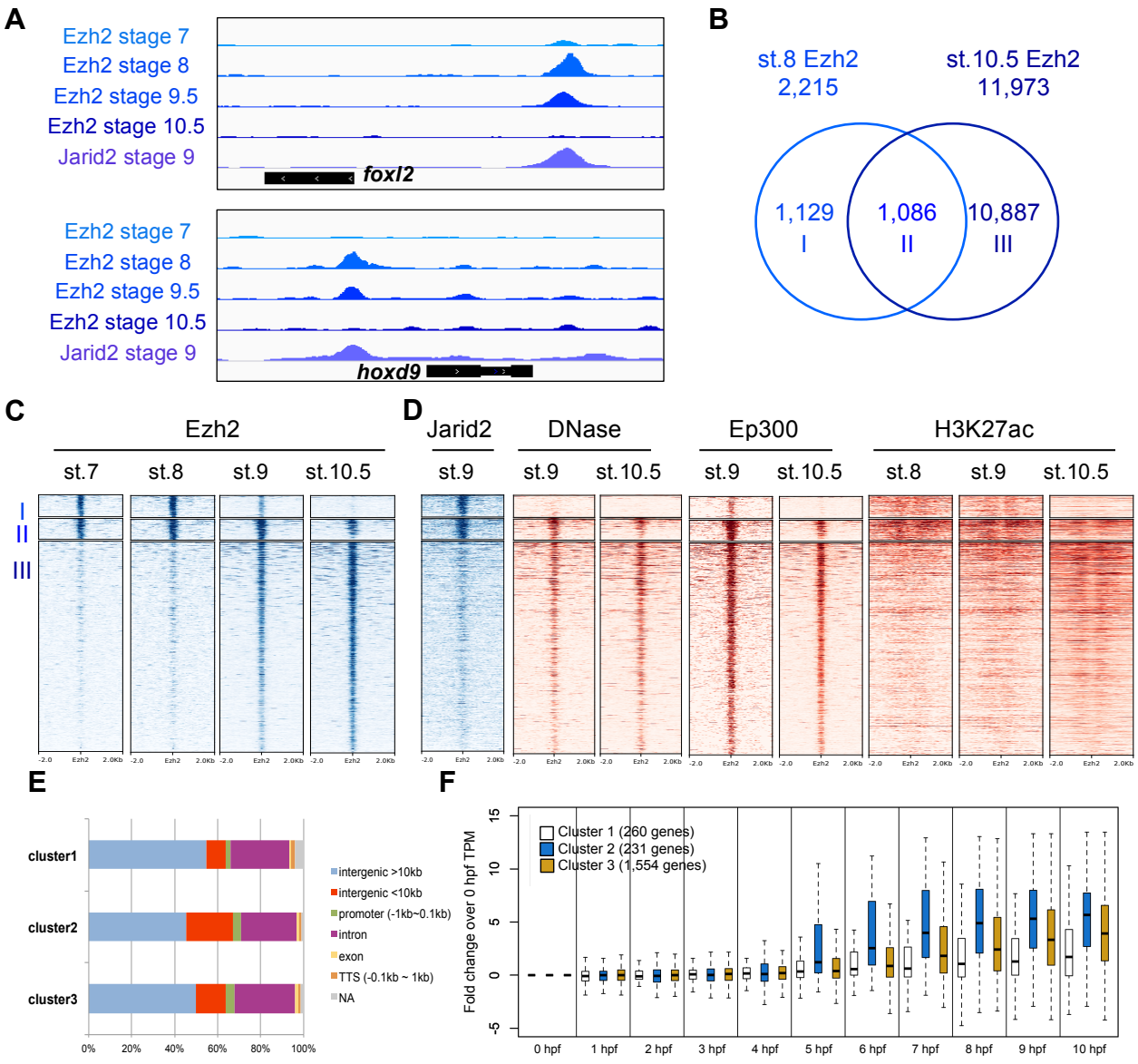

Figure 3.3 Ezh2 binding is dynamic and Ezh2 bound genes display dynamic transcriptional activity A) A genome-wide profile of the time course Ezh2 ChIP-seq and stage 9 Jarid2 ChIP-seq near zygotic genes, *foxl2* and *hoxd9*. B) The overlap of Ezh2 IDR peaks between stage 8 and stage 10.5. Each group of Ezh2 peaks is used for the clusters of heatmaps. C) Clustered heatmaps ordered by Ezh2 ChIP-seq signals on the summit of all Ezh2 peaks. D) Heatmaps of Jarid2, DNase, Ep300, and H3K27ac using the same order of peaks in Ezh2 heatmap in panel (C) before. E) Genomic distribution of Ezh2 peaks in each cluster. F) Temporal gene expression of Ezh2-bound genes in each cluster.

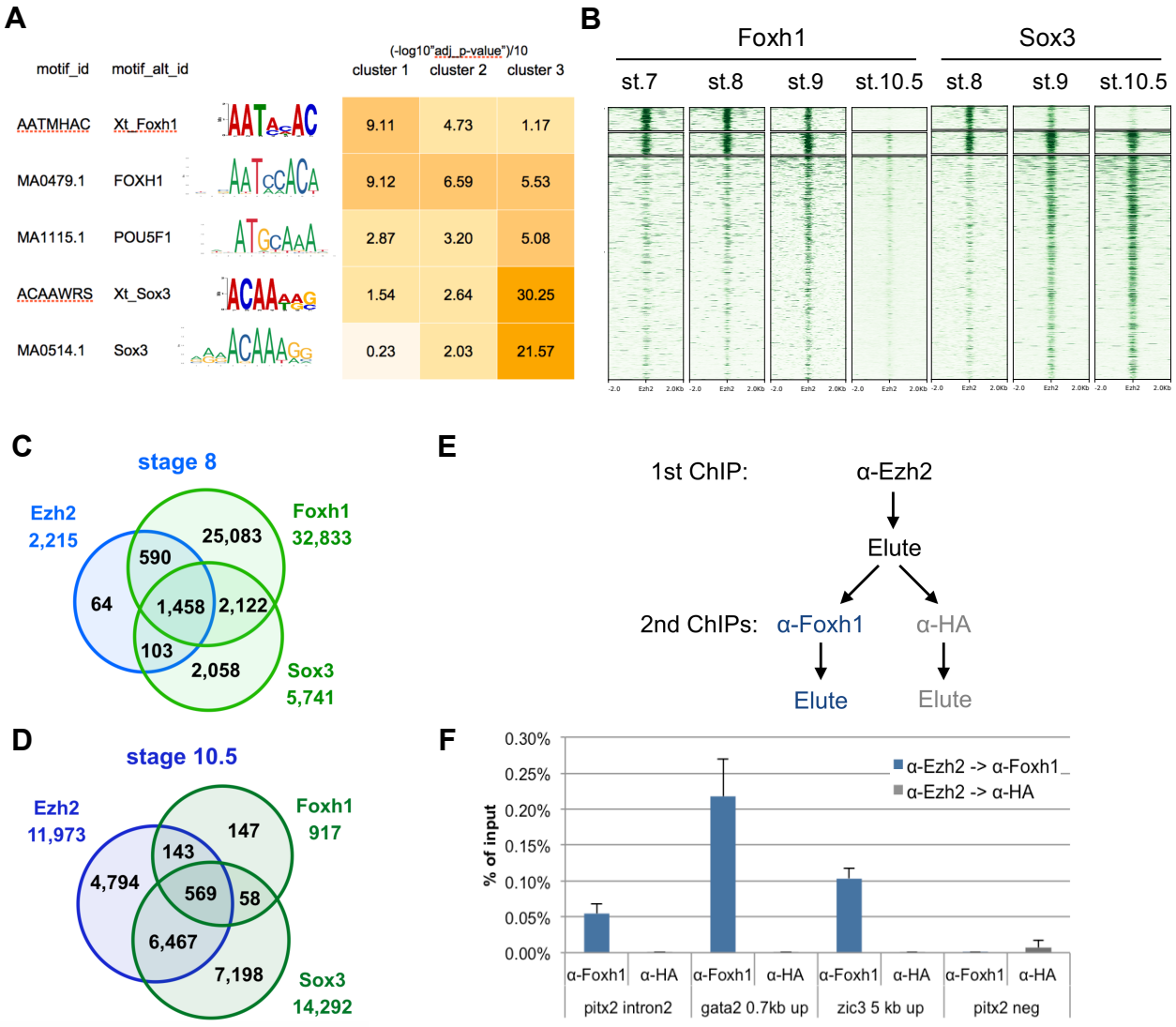

Figure 3.4 Ezh2 forms a complex with Foxh1 A) Motif analysis on sequences under the clustered Ezh2 peaks to see the TF motif enrichments. B) Heatmaps of the time course Foxh1 and Sox3 ChIP-seqs using the order of Figure 3.2C. Overlaps of Ezh2, Foxh1 and Sox3 IDR peaks at (C) stage 8 and (D) stage 10.5. E) Schematic diagram of the sequential ChIP. F) Sequential ChIP-qPCR using anti-Ezh2 followed by anti-HA or anti-Foxh1.

Table 3.3. Sequence near Foxh1 CRISPR target regions from F1 embryos

Sequence around the mutation site (from F0 female #1)

L	Y	R	E	58 G	G	T

	515aa		
GGCCGCCCTTGTACCGAGAG**GGG**GGCACCTGGAGCCCAGACAGA	WT		n=43
GGCCGCCCTTGTACCG-GAGGGGGGCACCTGGAGCCCAGACAGA	1del	20.9%	(9/43)
GGCCGCCCTTGTACC--AAGGGGGGCACCTGGAGCCCAGACAGA	2del_1sub	25.6%	(11/43)
GGCCGCCCTTGTACC-----TGGAGCACCTGGAGCCCAGACAGA	5del_2sub	27.9%	(12/43)
GGCCGCCCTTGGCCGCCCTT----GCACCTGGAGCCCAGACAGA	4del_9sub	25.6%	(11/43)

Sequence around the mutation site (from F0female #2)

GGCCGCCCTTGTACCGAGAG**GGG**GGCACCTGGAGCCCAGACAGA	WT		n=24
GGCCGCCCTTGTA----GAGGGGGGCACCTGGAGCCCAGACAGA	4del	20.9%	(10/24)
GGCCGCCCTTGTACCGAG--GGGGGCACCTGGAGCCCAGACAGA	2del	25.6%	(13/24)
GGCCGCCCTTGTACCGTCGAGGGGGGCACCTGGAGCCCAGACAGA	1del_2sub	27.9%	(1/24)

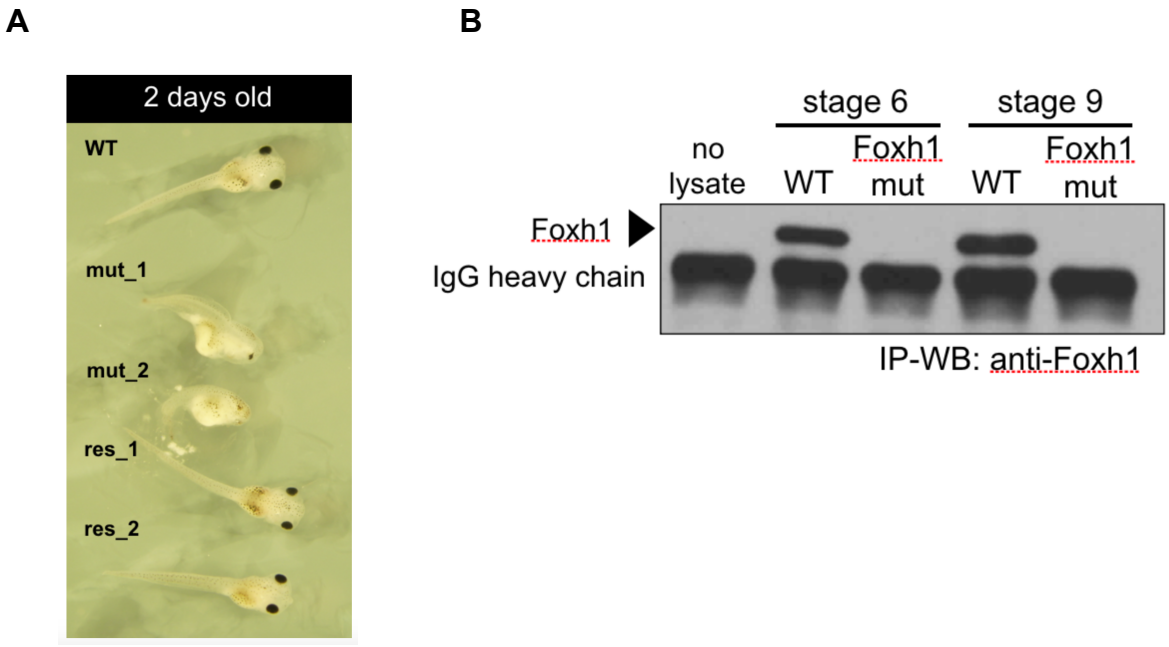

Figure 3.5 Generation of Foxh1 mutant embryos A) Phenotype of Foxh1 mutant embryos and rescued embryos by *foxh1* mRNA injection. B) IP-western blot analysis using anti-Foxh1. Arrowhead indicates Foxh1 at ~56kDa. The lower band corresponds to the IgG heavy chain from the immunoprecipitation.

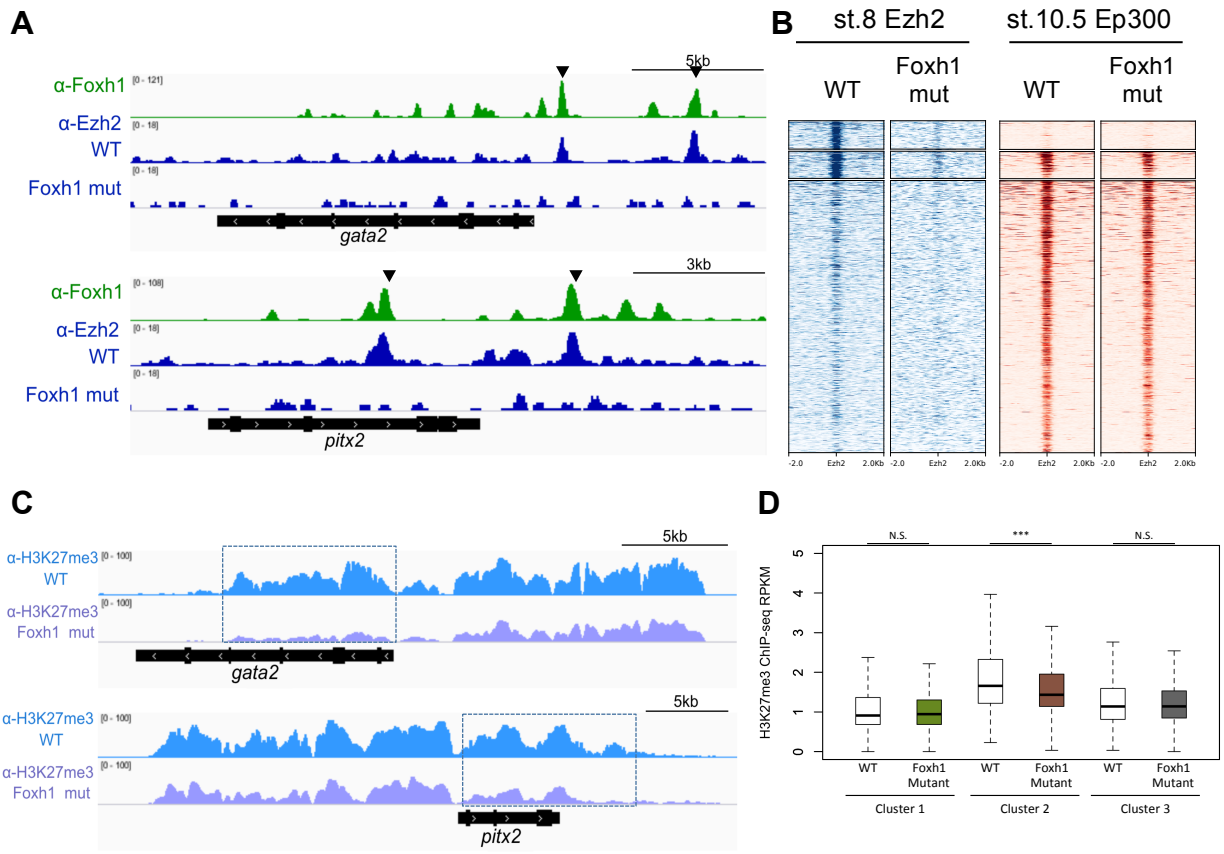

Figure 3.6 Ezh2-mediated H3K27me3 activity is Foxh1-dependent A) Genome browser view of Foxh1 peaks and Ezh2 peaks between wild-type (WT) and Foxh1 mutant embryos. Arrowheads mark the Ezh2 peaks that are diminished on Foxh1 mutant embryos. B) Comparison of stage 8 Ezh2 and stage 10.5 Ep300 ChIP signals between WT and Foxh1 mutant embryos. The peak region order is same as in Figure 3.2C. C) Genome-wide view of H3K27me3 activity between stage 10.5 WT and Foxh1 mutant embryos. The regions of reduced H3K27me3 enrichments were indicated with dotted boxes. D) Box-plot of H3K27me3 enrichment between WT and Foxh1 mutant in Ezh2 associated genes from Figure 3.2F.

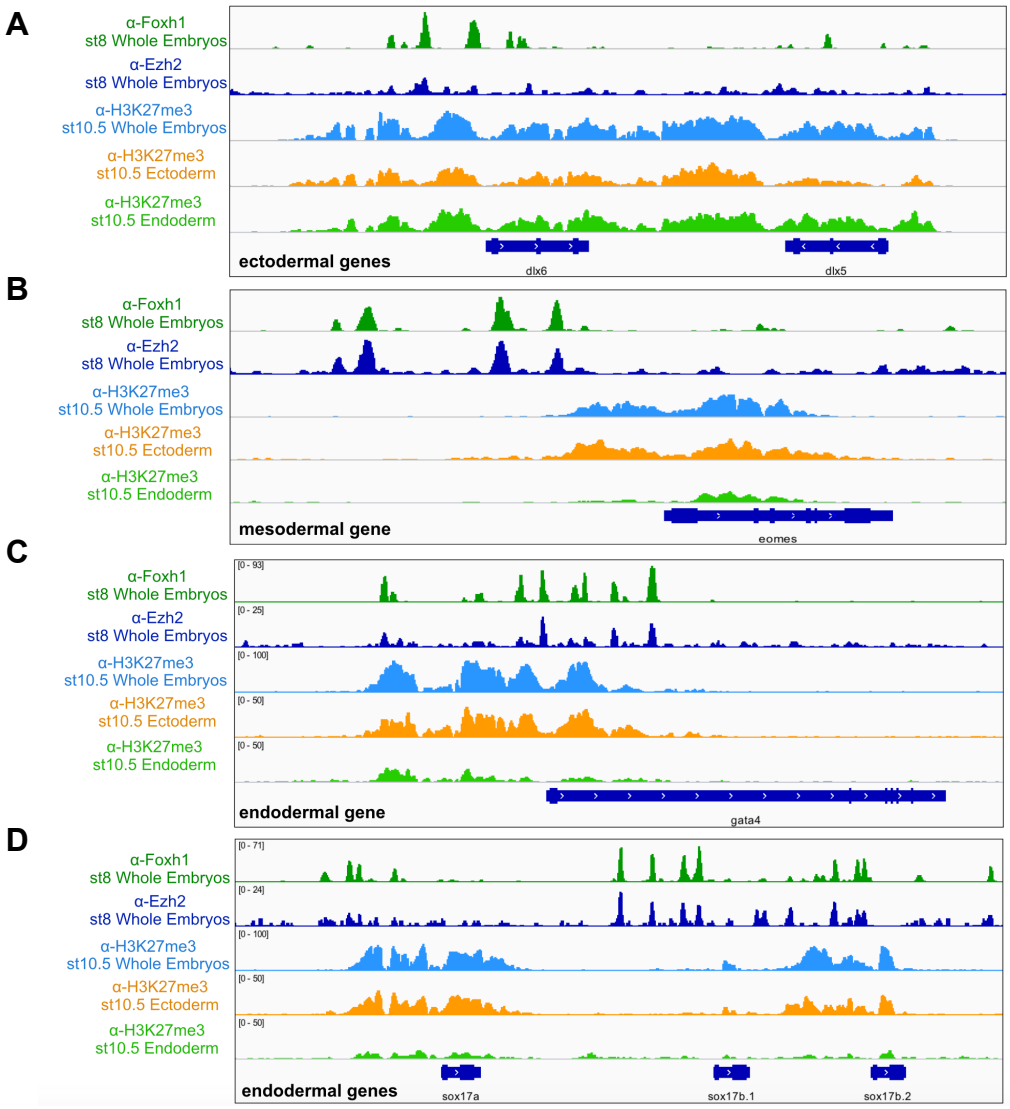

Figure 3.7 H3K27me3 activity is regionally regulated H3K27me3 deposits differently on germ layer specific developmental genes (A) ectodermal genes, *dlx5* and *dlx 6*; (B) mesodermal gene, *eomes*; endodermal genes, (C) *gata4* and (D) *sox17a, sox17b.1* and *sox17b.2*.

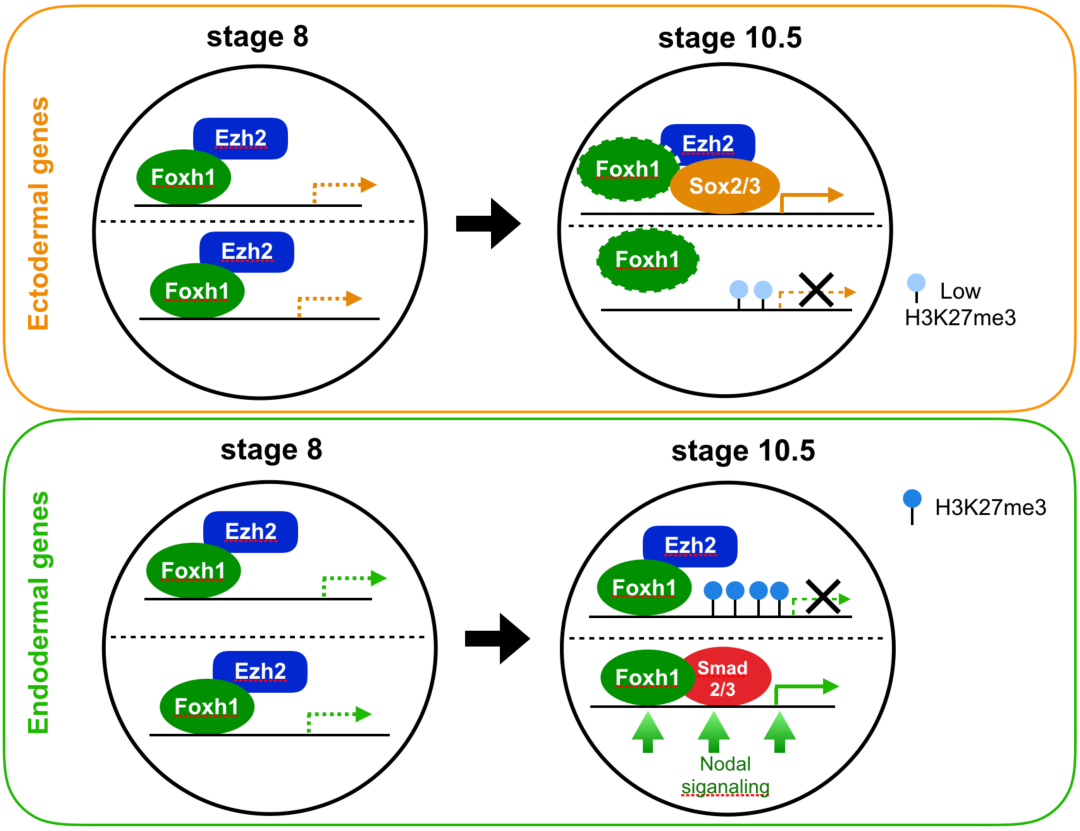

Figure 3.8 Proposed model of Foxh1-dependent regional H3K27me3 activity Foxh1 bookmarks *cis*-regulatory modules (CRMs) of germ layer-specific genes and recruits Ezh2 at early blastula stage and Ezh2 recruitment on endodermal CRMs on ectoderm germ layer deposits H3K27me3 to repress spatially inappropriate genes.

Experimental Procedures

Constructs and Viruses

Mouse Foxh1 ORF was generated from mESC E14 cDNA and subcloned into pDONOR201 vector using Gateway Technology (Invitrogen, Carlsbad, CA) as the entry clone. This entry clone was subsequently recombined into the modified lentiviral pLV-EF1a-IRES-Puro vector (Hayer et al., 2016) for the expression of C-terminal triple (S tag-Flag tag-SBP tag, or SFB) tagged fusion proteins (Wang et al., 2014).

The lentiviral mFoxh1 supernatant was generated by transient transfection of 293T cells with helper plasmids pSPAX2 and pMD2G (kindly provided by Dr. Wang, University of California, Irvine) and harvested 48 hours later. The supernatant was passed through a 0.45-um filter and used to infect mESC E14 cells with the addition of 8ug/ml polybrene.

Cell Culture and Transfection

mESC E14 cells were maintained in Glasgow minimal essential medium (G-MEM) supplemented with 10% fetal bovine serum (FBS), 2mM L-glutamine, 1mM sodium pyruvate, 1X MEM non-essential amino acid solution, 1X bME solution, and 6*10^5 U murine LIF at 37^{o}C in 5% CO2 (v/v). 293T cells were maintained in Dulbecco modified essential medium (DMEM) supplemented with 10% FBS. All culture media contained 1X penicillin streptomycin antibiotics. Plasmid transfection was performed using polyethylenimine.

Tandem Affinity Purification of SFB-tagged mFoxh1 Complexes

mESC E14 cells were infected twice with lentivirus encoding SFB-fused mFoxh1. Cells stably expressing Foxh1-SFB fusion protein were selected in 2ug/ml puromycin and confirmed by immunostaining and Western blotting. Single clonal E14 lines expressing mFoxh1-SFB fusion protein were established using serial dilution method. Affinity purification was performed as previously described (Wang et al., 2014). Eluted proteins were subjected to SDS-PAGE, excised and sent for mass spectrometry analysis (performed by Taplin Mass Spectrometry Facility, Harvard Medical School). This experiment was done using the following E14 cells: non-transduced E14 cells, mFoxh1-SFB transduced E14 cells with a heterogeneous population, and two independent single clonal mFoxh1-SFB transduced E14 cells.

Frog husbandry and embryo handling

Xenopus tropicalis frogs were raised and maintained under the University of California, Irvine Institutional Animal Care Use Committee (IACUC). Mature male and female frogs raised in the laboratory and/or purchased from NASCO (University of Virginia stock). *X. tropicali*s females were pre-primed with 10 units (U) of human chorionic gonadotropin (Chorulon HCG, Merck and Co.) 1-3 nights before eggs collection and then were primed with 100U of HCG on the day of embryo collection. Eggs were collected in 1X Marc's modified Ringers (MMR) and *in vitro* fertilized using testis macerated in 1X MMR containing 1 mg BSA/mL. Sperm suspension into eggs was diluted with 1/9X MMR to activate the sperm, and 10 minutes after sperm addition, the fertilized eggs were de-jellied using 3% cysteine in 1/9x MMR, pH 7.8. De-jellied embryos were cultured at 25^{o}C in 1/9X MMR until the desired stage according to Nieukwoop-Faber developmental table (Nieuwkoop and Faber, 1994).

Preparation and microinjection of CRISPR sgRNA design and Cas9

CRISPR-mediated mutagenesis in *Xenopus* was performed as described (Blitz et al., 2013; Nakayama et al., 2014) with modifications. Foxh1 gRNA was designed to mutagenize at the beginning of the translation site of *foxh1*. The *foxh1*-targeted sequence was 5'-GGCCGCCCTTGTACCGAGAGGGG-3′. Linearized pCasX plasmid used for the transcription of hCas9 mRNA by using the T7 mMessage mMachine kit (Ambion). The dose range injected was 3–4 ng hCas9 mRNA/embryo and 150-200 pg foxh1 gRNA/embryo. The cocktail of hCas9 mRNA and foxh1 gRNA in 4 nL of the injection volume was microinjected into four sites per embryo around the equator. During microinjection, embryos were in agarose-coated plates containing 1X MMR at room temperature and then embryos were subsequently cultured in agarose-coated plates in 1/9X MMR at 24–25^{o}C until desired stages.

Genomic PCR and Sequencing for F1 genotyping

Individual F1 embryos were obtained by *in vitro* fertilization of eggs from Foxh1 CRISPR F0 female or wild type (WT) female with sperm from a WT male. Each single embryo was transferred to 0.2-mL PCR tubes containing 100 µl of lysis buffer (50 mM Tris, pH 8.8, 1 mM EDTA, 0.5% Tween 20, 200 µg/ml proteinase K). Embryos are incubated at 56^{o}C for 2 h to overnight, followed by incubation at 95^{o}C for 10 min to inactivate proteinase K. Lysates were centrifuged for 1 min at 4^{o}C and 1 µl aliquots were used directly in 20µL PCR reactions. The locus of interest was amplified by PCR with the following primers: Forward: 5’-CCACTTGCTGAAGGTTCGTT-3’ and Reverse: 5’-AATATGGTGGCTTGGCGTAG-3’. PCR

product was purified and eluted in 20µL and then 1µL of the purified PCR product was sequenced (GeneWiz) using the same primers used for PCR.

Chromatin immunoprecipitation (ChIP) assay

ChIP using *Xenopus tropicalis* embryos was performed as previously described (Charney et al., 2017). The following antibodies were used: anti-Ezh2 (Abcam ab3748 or Abcam ab191250), anti-Foxh1 (Chiu et al., 2014), anti-HA (Abcam ab9110) anti-Sox3 (Zhang et al., 2003), and anti-H3K27me3 (Upstate/Millipore 07-449).

For sequential ChIP assay, the first round of ChIP was performed as the single round of ChIP. After the immunoprecipitation with the first antibody, chromatin was eluted in 1x TE, pH 8.0, with 10mM DTT, 500 mM NaCl and 0.1% SDS at 37°C for 30 minutes. The eluate was then diluted 10 times of the volume with RIPA buffer. The diluted eluate was incubated with the second antibody and the rest of the procedure was performed as the single round of ChIP.

Quantitative PCR after ChIP (ChIP-qPCR) was performed on Roche LightCycler 480 II using SYBR Green I master mix (Roche). The following primers were used for ChIP-qPCR and sequential ChIP-qPCR.

pitx2 intron2:

F: 5'-ATCTGCTCCCATCTCTCCAA-3'; R: 5'-CAAACAGGGCTCATTGAGGA-3'

gata2 0.7kb upstream:

F: 5'-GTCGCTCTGCTCAGCTCTTC-3'; R: 5'-CCGTTTCACAGATGTGGACT-3'

zic3 5kb upstream:

F: 5'-ggaaatggaactggggaaag-3'; R: 5'-GGGTGATCTGAGCCAAATTC-3'

pitx2 negative (20kb down):

F: 5'-TGCACCTAGGTTTGGGTAGG-3'; R: 5'-GAGGGTGGAAAGGGGTTAAG-3'

ChIP-seq libraries were generated using Nextflex ChIP-seq kit (Bioo Scientific). The quality of the libraries was measured using an Agilent Bioanalyzer 2100 and KAPA qPCR and the libraries were sequenced using an Illumina platform at the UC Irvine Genomics High Throughput Facility.

Sequence Alignment and Visualization

ChIP-seq read files were aligned to the X. tropicalis genome v9.0 (Xenbase, http://ftp.xenbase.org/pub/Genomics/JGI/Xentr9.0/) using Bowtie v1.0.0 (Langmead et al., 2009) with the following script.

```
$bowtie --sam -p 32 -m 1 Xentro9 reads-1.fastq mapped-1.sam
```

Duplicate reads were removed using Samtools (Li et al., 2009).

```
$samtools view -Sb mapped-1.sam > mapped-1.bam
$samtools sort mapped-1.bam -o mapped-1.sorted.bam
$samtools rmdup -s mapped-1.sorted.bam mapped-1.nodup.bam'
```

Bigwig (bw) files were generated for the visualization using HOMER (Heinz et al., 2010).

```
$makeTagDirectory mapped-1_TagDirectory mapped-1.nodup.bam Xentro9.fa
-single
$makeUCSCfile mapped-1_TagDirectory -bigWig Xentro9.size -o sample-
1.bw'
```

The bw file was loaded into Integrative Genomics Viewer (IGV) v2.5.0 (Robinson et al., 2011).

Peak calling and Irreproducibility Discovery Rate (IDR) Analysis

Peaks were called using MACS2 v2.1.0 (Zhang et al., 2008) with the option -p 0.001 for biological replicates to perform IDR analysis but otherwise, default options were used.

```
$macs2 callpeak -t sample-1.nodup.bed -c input_DNA.nodup.bed -f BED -g
1.16e9 -p 1e-3 -n sample-1_peaks
```

To select high-confidence peaks between two biological replicates, IDR v2.0.1 was used to identify high-confidence peaks between pseudoreplicates generated from pooled reads of biological replicates with default options (Li et al., 2011).

```
$ cat sample-1.nodup.bed sample-2.nodup.bed > pooled-sample.bed
$ cat pooled-sample.bed | shuf | split -d -l nlines - pooled-sample.pr
```

The number for nlines can get from '(cat rep0.pooled.bed | wc -l)/2'

```
$ mv pooled-sample.pr00 pooled-sample.pr1.bed
```

```
$ mv pooled-sample.pr01 pooled-sample.pr2.bed
```

MACS2 was run to get the narrowPeak files from pseudoreplicates.

```
$ idr --samples pooled-sample.pr1.narrowPeak pooled-
sample.pr2.narrowPeak --peak-list pooled-sample.narrowPeak --input-
file-type narrowPeak --rank signal.value --output-file idr_pooled-
sample –plot
```

Heatmap

DeepTools v2.0.0 (Ramírez et al., 2014) was used to generate heatmaps around peak summits or genes.

First, a bigwig file was generated from the BAM files using bamCoverage in deepTools.

```
$ bamCoverage --bam sample.nodup.bam --binSize 50 --normalizeUsingRPKM
--ignoreDuplicates -of bigwig -o sample-signal.bw
```

To generate heatmaps around peak summits, the 'computeMatrix' command with the subcommand 'reference-point' was used to generate the matrix underlying the heatmaps.

```
$ computeMatrix reference-point -R peak-summits.bed -S sample-
signal.bw --binSize 50 -b 2000 -a 2000 --missingDataAsZero --
referencePoint center --outFileName output_Matrix -p 4
```

To generate heatmaps around genes, he 'computeMatrix' command with the subcommand 'scale-regions' was used.

```
$ computeMatrix scale-regions -R genes.bed -S sample-signal.bw --
binSize 50 -m 7500 -b 10000 -a 10000 --missingDataAsZero --outFileName
output_matrix -p 4
```

Heatmap was generated from the matrix using the 'plotHeatmap' command. K-means clustering option was used to cluster the heatmap.

```
$ plotHeatmap --matrixFile output_Matrix --outFileName output.pdf --
colorMap Colors  --heatmapHeight 10 --refPointLabel peaks --xAxisLabel
'' -z Peakname --plotTitle title --kmeans # --outFileSortedRegions
output_SortedRegions
```

Motif Analysis

MEME-ChIP from the MEME Suite 5.0.5 (http://meme-suite.org/tools/meme-chip) was used for motif analysis (Bailey et al., 2009) with the default setting. The 500bp centered on the summit of a peak was generated as the input FASTA sequence using the Bedtools v2.19.1 (Quinlan and Hall, 2010).

```
$ bedtools slop -i summits.bed -g Xentro9.sizes -b 249 >
summits_500bp.bed
$ bedtools getfasta -fi Xentro9.fa -bed summits_100bp.bed -fo
summits_500bp.fa
```

The accession of published ChIP-seq datasets

Published datasets used in this research can be downloaded from NCBI GEO (https://www.ncbi.nlm.nih.gov/geo/) using the GEO accession numbers: GSE85273 for stage 8,9, and 10.5 Foxh1 ChIP-seq data (Charney et al., 2017); GSE41161 for stage 9 Jarid (van Heeringen et al., 2014); GSE67974 for stage 9 and 10.5 Ep300 (Hontelez et al., 2015); and GSE56000 for stage 8,9, and 10.5 H3K27ac (Gupta et al., 2014).

Acknowledgements

I would like to thank the following individuals who contributed to the work in this chapter: Jeff J. Zhou and Dr.Wenqi Wang performed and shared unpublished mass-spectrometry data and Jeff J. Zhou performed immuno-western using 293T cells; Jessica Cheung contributed sequential ChIP-qPCR datasets; Margaret Fish contributed to generating F0 Foxh1 CRISPR mutants; Rebecca Charney performed Foxh1 immuno-western; Kitt Paraiso assisted with the published temporal RNA-seq analysis; Ira Blitz generated the ChIP-seq libraries. I would also like to thank the University of California, Irvine (UCI) High Performance Computing Cluster (https://hpc.oit.uci.edu/) for their helpful support and the UCI Genomic High Throughout Facility for sequencing support.

CHAPTER 4

Conclusions And Discussion

My dissertation focused on the epigenetic functions of the maternal TF, Foxh1. Through the investigation of chromatin states and co-regulators bindings to DNA using Foxh1-deficient *Xenopus* embryos generated by CRISPR/Cas9 technology, I propose a critical role of Foxh1 in the recruitment of the PRC2 core subunit, Ezh2, for the spatiotemporal regulation of gene expression. This *in vivo* functional study of a maternal TF highlights how maternal TFs control chromatin state by interacting with chromatin modifiers during early embryogenesis. In this chapter, I highlight key findings from my dissertation work, discuss their broader implications for understanding epigenetic regulation in early embryogenesis, and point to critical future experiments to further elucidate the functions of maternal TFs.

The epigenetic role of maternal Foxh1 early binding

Forkhead TF, Foxh1, is the well-known TF in the Nodal signaling pathway by recruiting phosphorylated Smad2/3 on the target CRMs. We previously found that this maternal TF bookmarks mesendodermal CRMs as early as the 32-cell stage, before zygotic gene activation, and that this binding occurs before RNA pol II recruitment and the enrichment of epigenetic enhancer marks like Ep300 and H3K4me1 (Charney et al., 2017). Our findings also suggested Foxh1's role as a pioneer factor, which can endow transcriptional competence by recruiting other TFs and additional regulators to create active enhancer or by binding to nucleosomal DNA and opening up local chromatin (Zaret and Carroll, 2011; Iwafuchi-Doi and Zaret, 2014).

To investigate whether Foxh1 has a function as a pioneer factor, I initially monitored chromatin accessibility using Foxh1-deficient *Xenopus* embryos. There are two major approaches to detect chromatin accessibility: DNase-seq and ATAC-seq.

Deoxyribonuclease I (DNase I) hypersensitive sites sequencing (DNase-seq) identifies genome-wide open chromatin regions that are accessible to regulatory factors like transcription factors and are relatively nucleosome-free genomic regions. DNase I can digest nucleosome-depleted DNA, which is presumably bound by transcription factors, but DNA in nucleosomes or higher-order chromatin fibers is less accessible to this nuclease (Neph et al., 2012). I was able to adapt DNase-seq on early gastrula *Xenopus* embryos (Chapter 2), but this method was not optimal for earlier stages. The low number of cells in earlier stage embryos (~500 cells per blastula embryo compared with approximately 10,000 cells per early gastrula embryo) hampered the adaption of this method for early embryos. On the other hand, ATAC-seq requires fewer cells than DNase-seq (requiring a minimum of 50,000 cells, which is equivalent to five early gastrula embryos, are required) and was more practical than DNase-seq to do genomic work for embryos.

Transposase-Accessible Chromatin followed by sequencing (ATAC-seq) is also a simpler protocol than DNase-seq for detection of open chromatin (Buenrostro et al., 2013). While the ATAC-seq technique was adaptable to *Xenopus* gastrula and subsequent stage embryos (stage 11; Bright and Veenstra, 2019), it has not been possible to identify open chromatin regions of the earlier stage embryos (e.g., blastula embryos) due to high yolk contamination during nuclei isolation. Recent modified ATAC-seq protocols (Corces et al., 2017; Cusanovich et al., 2018) encouraged me to adopt this technique for early *Xenopus* embryos. Currently, I have improved the resolution of chromatin accessibility using fixed and gently sheared chromatin from *Xenopus* embryos (Buenrostro et al., 2013) (J.C., unpublished data). I expect this improved low-input ATAC-seq and single-cell ATAC-seq methods (Corces et al., 2017; Mezger et al., 2018) will lead to exciting breakthroughs in our

understanding of how maternal TFs coordinates chromatin accessibility during early embryogenesis.

Candidates for Foxh1-associated epigenetic regulators

Our mESC Foxh1 MS data (Table 3.1) showed Foxh1 is associated with most of the PCR2 subunits -Suz12, Jarid2, Ezh2, and RBBP7. However, in the list, there are other proteins known to participate in epigenetic regulation. These proteins include ACTB (actin beta), POLR2A (RNA polymerase II subunit A), POLR1C (RNA polymerase I and III subunit C), HIST1H1A (histone cluster 1 H1 family member a) and more - which could interact with Foxh1. GO analysis (Figure 3.1.C) of those Foxh1-associated proteins showed that most of them are related to epigenetic regulation of gene expression. This suggests Foxh1 could interact with other epigenetic regulators. Two interesting candidates of epigenetic regulators associated with Foxh1 are histone cluster 1, H1 family members, HIST1H1A and HIST1H1E. In *Xenopus*, maternally inherited H1 linker variant, H1M, persists in chromatin until its somatic variants are synthesized at MBT (Dimitrov and Wolffe, 1996; Dworkin-Rastl et al., 1994) and maternal H1 generates less stable chromatin, which helps rapid early cell divisions and to initiate transcription during MBT (Freedman and Heald, 2010). H1 incorporation on chromatin remodeling during nuclear reprogramming has been reported in frogs and flies using *Xenopus* egg extracts and *Drosophila* preblastodermic embryo extracts, respectively (Jullien et al., 2010; Satovic et al., 2018). Next, linker histone H1b has been reported to interact with a homeobox protein transcription factor, Msx1, to repress a myogenic master regulator, MyoD, that is activated upon mesoderm induction (Lee et al., 2004). This suggests a potential role of linker histone H1 in mesoderm specification.

Furthermore, another forkhead TF, Foxa, functions as a pioneer factor by displacing linker histones to keep nucleosomes accessible to retain tissue-specific enhancers in mammalian chromatin (Iwafuchi-Doi et al., 2016). Since forkhead proteins and linker histones H1 and H5 possess winged-helix DNA-binding domains (Cirillo et al., 1998) and a subset of Foxh1 bindings occur prior to zygotic Foxa1 binding in *Xenopus* (Charney et al., 2017), it will be interesting to monitor whether H1 occupation depends on Foxh1 and whether Foxh1-bookmarked mesendodermal enhancer endows accessibility on the subset of linker histone H1 which have Foxh1 binding sites.

PRC2 has been reported to methylate H1, dependent upon the contents of the PCR complex (Kuzmichev et al., 2004). Depending on the isoforms of the Eed, Ezh2-containing complexes, PRC2 directs its histone lysine methyltransferase activity toward histone H3 lysine 27 (H3K27) or histone H1b lysine26 (H1K26). PRC2 can dock onto H1K26me3 substrates. However, docking to H1K26me3 decreases the enzymatic activity of PRC2 (Xu et al., 2010). This suggests that the PRC2 complex on maternal H1K26me3 holds the enzymatic activity of PRC2 until ZGA because the emergence of newly synthesized H1 after ZGA reduces PRC2 binding on maternal H1K26me3 and at the same time, it increases the repressive enzymatic activity of PRC2 to methylate H3K27me3. Since I found that Ezh2 recruitment and PRC2 activity depend on Foxh1 binding and the onset of H3K27me3 mark, it would be interesting to investigate whether Foxh1-mediated PRC2 activity before and after ZGA is due to linker histone H1 and H1K26me3 activity.

Foxh1-dependent Ezh2 recruitment and H3K27me3 activity

Among all the Foxh1 interactors from mESC FOXH1 mass spec data, I noted that the major components of PRC2 are all consistently detected - SUZ12, EZH2, EED, and JARID2 (Table 3.2). PRC2 is a well-studied histone modifier during embryonic development with its classical role to repress target genes by depositing H3K27me3, but it has been unclear how PRC2 is recruited to mark H3K27me3 (reviewed in Margueron and Reinberg, 2011). First I was able to confirm the direct interaction of Ezh2 and Foxh1 in the *Xenopus* early blastula embryos. Second, by monitoring the loss or decrease of Ezh2 binding on Foxh1-deficient *Xenopus* embryos, I was able to observe the requirement of maternal TF bookmarking for the PRC2 recruitment during early embryogenesis. Third, I showed that both dynamic and persistent Ezh2 binding is associated with the transcriptional activity of target genes (Figure 3.3). Next, Foxh1-Ezh2 binding peaks overlap with those of maternal TF Sox3. Importantly, Ezh2 motif analysis showed high enrichment of the Sox3 motif, especially among de novo Ezh2 peaks (cluster III in Figure 3.4B). Finally, I also observed that H3K27me3 modification is significantly compromised in the absence of Foxh1 in some regions (Figure 3.6C), but overall the change is not statistically significant at the whole genome level (Figure 3.6 D). I also noted that H3K27me3 marks are enriched in endodermal genes compared with ectodermal genes, which has been previously reported (Akkers et al., 2009; van Heeringen et al., 2014).

Taken together, I propose the following mechanism (Figure 3.8). First, the maternal TF Foxh1 marks its CRMs as early as the 32-cell stage, and then Ezh2 is recruited on Foxh1-binding sites at the early blastula stage (stage 7, about 250 cells per embryo). Ezh2 binds to Foxh1-bookmarked sites and this appears to occur in all embryonic cells before ZGA. This view is consistent with the findings that *foxh1* mRNA is ubiquitously expressed in the

blastula embryos and that *ezh2* is detected in both animal and vegetal region of 8-cell embryos. At the blastula stage when Nodal signaling is activated in endodermal, Smad2/3, are recruited to Foxh1-bound CRMs, which results in the eviction of Ezh2 from the complex. On the other hand, in ectodermal cells where Nodal signaling is absent, the Ezh2-Foxh1 complex is maintained on the endodermal genes and this results in the recruitment of the PRC2 complex. Consequently, the endodermal genes in ectodermal cells will acquire repressive marks and ensure a proper lineage commitment.

The maintenance of PRC2 at target sites is required for PRC2 activity to deposit and spread H3K27me3 marks. It has been suggested PRC2 facultative subunits could have this role instead of the core subunits of PRC2 (Laugesen et al., 2016). Depending on the components of the facultative subunits, PRC2 is specified into PRC2.1 and PRC2.2. PRC2.1 contains the PCL proteins (PHF1, MTF2, or PHF19) and EPOP or PALI1, and PRC2.2 involves AEBP2 and JARID2 as the facultative subunits (Alekseyenko et al., 2014; Hauri et al., 2016; Grijzenhout et al., 2016). Our Foxh1 mESCs MS data did not list any PRC2.1 facultative subunits but showed one of the PRC2.2 facultative subunits, Jarid2. Furthermore, I showed that Jarid2 binding also overlaps with Foxh1 and Ezh2 binding in stage 9 *Xenopus* embryos (Figure3.3D). Further examination of Jarid2 binding at later stages would help to understand the role of early recruitment of PCR2, which is mediated by Foxh1 before H3K27me3 activity, and which PRC2 complex is involved in lineage specification during early embryogenesis. Also, it will be interesting to assess whether maternal and zygotic TFs have a distinct role in the recruitment and activity of PRC2 considering Ezh2 and Jarid2 binding occurs prior to H3K27me3 marking.

Combinatorial functions of maternal TFs during early embryogenesis

Motif enrichment analysis of Ezh2 binding regions (Figure 3.4A) showed Fox, Sox, motif enrichment. Jarid2 peaks are also enriched with Fox, Sox, and Pou motifs (van Heeringen et al., 2014). The coincidence of the similar motif enrichment from those two subunits of PRC2 - Ezh2 and Jarid2 - suggests that combinatorial interaction of TFs and the epigenetic landscape would be required for PRC2 recruitment. My Sox3 ChIP-seq analysis also showed a strong correlation between Sox3 and Ezh2-Foxh1 binding regions (Figure 3.4B). It will be interesting to investigate whether the co-occupancy of multiple TFs is necessary for the establishment of enhanceosomes. Co-localization of PcG components with pluripotency factors Oct4, Sox2, and Nanog was reported in ESCs (Bernstein et al., 2006; Boyer et al., 2006; Lee et al., 2006). Moreover, those Oct4, Sox2, and Nanog co-bound regions form super-enhancers (SE), which are distinguished from typical enhancers by size, transcription factor binding density and content, to control the pluripotent state (Whyte et al., 2013). In *Xenopus*, maternal TFs -Otx1, Vegt, and Foxh1 - mediated SEs are reported to control endodermal cell fate specification during ZGA (Paraiso et al., 2019). The combinatorial binding of maternal Otx1, Vegt, and Foxh1 forms *Xenopus* endodermal SEs that have stronger signals of enhancer marks - H3K4me1, Ep300, and H3K27ac - than regular enhancers (Paraiso et al., 2019). They also showed strong signals of polycomb markers - H3K27me3, Jarid2, and Ezh2 - on those *Xenopus* endodermal SEs. Therefore, it will be interesting to see whether multiple Foxh1/Sox/Pou TFs forming enhanceosomes to promote the PRC2 complex formation and subsequent H3K27me3 occurs to regulate epigenetic modification during early embryogenesis.

Single TF, Foxh1, depletion did not affect either Ep300 binding (Figure 3.6B) or the zygotic TF, Foxa1, binding in early gastrula embryos (J.C., unpublished data). There are several possible reasons for these results. First, Foxa1 binding is independent of Foxh1 even though their binding sites overlap in *Xenopus* embryos and they share similar binding motifs (Charney et al., 2017). Since Foxa is the well-known pioneer factor for hepatic specification, which is differentiated from the endodermal cells and critical for endoderm development across diverse organisms (Zaret and Carroll, 2011; Friedman and Kaestner, 2006; Ben-Tabou de-Leon, 2011), Foxa might also have its own function as a pioneer factor during germ layer specification and it could be independent of Foxh1. Second, combinatorial TF binding or the formation of super-enhancers may be required for the establishment of lineage-specific CRMs and the recruitment of zygotic TFs. Foxh1 binding sites are enriched with other motifs like Sox and Pou (Charney et al., 2017) and Foxh1 forms super-enhancers with another maternal TFs, Vegt and Otx1 (Paraiso et al., 2019). Additionally, I show that Ezh2 and Sox3 binding overlaps, and Sox3 motifs are enriched on Ezh2 peaks (Figure 3.4A-D). On the other hand, motif analysis of Ep300 peaks has a high Sox motif enrichment score (J.C., unpublished data). If multiple TFs are required compared to a single TF to recruit zygotic TF or to form active CRMs, then a single TF knockdown is not sufficient to observe its role in the combinatorial regulation of gene expression and chromatin state. Multiple TF knockdowns will help to answer this question.

Roles of maternal TFs for the spatial regulation of the epigenetic landscape

In Chapter 3, I showed the regionally defined H3K27me3 activity, specifically in the ectodermal tissue of the early *Xenopus* gastrula embryos. The ectodermal-localized PRC2

activity on endodermal genes suggests that histone modification is spatially regulated to prevent aberrant lineage commitment during germ layer specification. In addition, I showed that the binding of Ezh2, the core catalytic subunit of PRC2, to deposit H3K27me3, depends on Foxh1 binding. However, H3K27me3 marking was less observed in endodermal genes. One possible explanation of this phenomenon is that PRC2 activity could be different in each germ layer of gastrula embryos. To investigate this possibility, first the protein expression of the PRC2 subunits - Ezh2, Eed, Suz12, and Jarid2 - needs to be examined in each germ layer, even their respective mRNAs are ubiquitously expressed in early gastrula embryos (Blitz et al., 2016). Another possible explanation is that multiple TF bindings stabilize Ezh2 and PRC2 binding on the target sites. The increase in overlaps between Sox3 and Ezh2 bindings at stage 10.5 and increased enrichment of Sox motifs under stage 10.5 Ezh2 peaks (Figure 3.4A-D) supports this idea. To test this hypothesis, Ezh2 and H3K27me3 activity could be examined in each germ layer when Sox3 is knocked-down by MO injection, and then the use of double knockouts of Foxh1 and Sox3 embryos by Sox3 MO injection in Foxh1-deficient embryos will help to answer this question.

In conclusion, my dissertation work has revealed the epigenetic role of maternal Foxh1 by regulating the recruitment of chromatin modifier Ezh2 and H3K27me3 activity and proposes Foxh1-dependent Ezh2 and PRC2 complex recruitment introduce the regionally differential H3K27me3 activity during *Xenopus* germ layer specification. This work and the proposed future work discussed in this chapter will lead to exciting breakthroughs in our understanding of how maternal TFs regulate the epigenome during early embryogenesis.

REFERENCES

Akkers, R.C., van Heeringen, S.J., Jacobi, U.G., Janssen-Megens, E.M., FranCoijs, K.-J., Stunnenberg, H.G., and Veenstra, G.J.C. (2009). A Hierarchy of H3K4me3 and H3K27me3 Acquisition in Spatial Gene Regulation in Xenopus Embryos. Developmental Cell 17, 425–434.

Alekseyenko, A.A., Gorchakov, A.A., Kharchenko, P.V., and Kuroda, M.I. (2014). Reciprocal interactions of human C10orf12 and C17orf96 with PRC2 revealed by BioTAP-XL cross-linking and affinity purification. Proceedings of the National Academy of Sciences 111, 2488–2493.

Amador-Arjona, A., Cimadamore, F., Huang, C.-T., Wright, R., Lewis, S., Gage, F.H., and Terskikh, A.V. (2015). SOX2 primes the epigenetic landscape in neural precursors enabling proper gene activation during hippocampal neurogenesis. Proceedings of the National Academy of Sciences 112, E1936–E1945.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res 21, 381–395.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. Cell 125, 315–326.

Blitz, I.L., Biesinger, J., Xie, X., and Cho, K.W.Y. (2013). Biallelic genome modification in F 0Xenopus tropicalisembryos using the CRISPR/Cas system. Genesis 51, 827–834.

Blitz, I., Paraiso, K., IlyaPatrushev, Chiu, W.Y., Cho, K.Y., and Gilchrist, M.J. (2017). Developmental Biology. Developmental Biology 426, 409–417.

Blythe, S.A., and Wieschaus, E.F. (2016). Establishment and maintenance of heritable chromatin structure during early Drosophila embryogenesis. eLife 5, e20148.

Bogdanović, O., van Heeringen, S.J., and Veenstra, G.J.C. (2011). The epigenome in early vertebrate development. Genesis 50, 192–206.

Borchers, A., and Pieler, T. (2010). Programming Pluripotent Precursor Cells Derived from Xenopus Embryos to Generate Specific Tissues and Organs. Genes 1, 413–426.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature Methods 10, 1213–1218.

Cao, R., and Zhang, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Current Opinion in Genetics & Development 14, 155–164.

Cattell, M.V., Garnett, A.T., Klymkowsky, M.W., and Medeiros, D.M. (2012). A maternally established SoxB1/SoxFaxis is a conserved feature of chordate germ layer patterning. Evolution & Development 14, 104–115.

Cha, S.-W., McAdams, M., Kormish, J., Wylie, C., and Kofron, M. (2012). Foxi2 Is an Animally Localized Maternal mRNA in Xenopus, and an Activator of the Zygotic Ectoderm Activator Foxi1e. PLoS ONE 7, e41782.

Charney, R.M., Forouzmand, E., Cho, J.S., Cheung, J., Paraiso, K.D., Yasuoka, Y., Takahashi, S., Taira, M., Blitz, I.L., Xie, X., et al. (2017a). Foxh1 Occupies cis-Regulatory Modules Prior to Dynamic Transcription Factor Interactions Controlling the Mesendoderm Gene Program. Developmental Cell 40, 595–607.e4.

Chen, H., Einstein, L.C., Little, S.C., and Good, M.C. (2019).
Spatiotemporal Patterning of Zygotic Genome Activation in a Model Vertebrate Embryo. Developmental Cell 49, 852–866.e857.

Chiu, W.T., Charney Le, R., Blitz, I.L., Fish, M.B., Li, Y., Biesinger, J., Xie, X., and Cho, K.W.Y. (2014). Genome-wide view of TGF /Foxh1 regulation of the early mesendoderm program. Development 141, 4537–4547.

Cirillo, L.A., McPherson, C.E., Bossard, P., Stevens, K., Cherian, S., Shim, E.Y., Clark, K.L., Burley, S.K., and Zaret, K.S. (1998). Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. The EMBO Journal 17, 244–254.

Coleman, R.T., and Struhl, G. (2017). Causal role for inheritance of H3K27me3 in maintaining the OFF state of a DrosophilaHOX gene. Science 356, eaai8236.

Collart, C., Owens, N.D.L., Bhaw-Rosun, L., Cooper, B., De Domenico, E., Patrushev, I., Sesay, A.K., Smith, J.N., Smith, J.C., and Gilchrist, M.J. (2014). High-resolution analysis of gene activity during the Xenopusmid-blastula transition. Development 141, 1927–1939.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nature Methods 14, 959–962.

Cusanovich, D.A., Reddington, J.P., Garfield, D.A., Daza, R.M., Aghamirzaie, D., Marco-Ferreres, R., Pliner, H.A., Christiansen, L., Qiu, X., Steemers, F.J., et al. (2018). The cis-regulatory dynamics of embryonic development at single-cell resolution. Nature 555, 538–542.

Dahl, J.A., Jung, I., Aanes, H., Greggains, G.D., Manaf, A., Lerdrup, M., Li, G., Kuan, S., Bin Li, Lee, A.Y., et al. (2016). Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. Nature 537, 548–552.

Daniels, D.L., and Weis, W.I. (2005). β-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. Nature Publishing Group 12, 364–371.

Davidson, E., and Levin, M. (2005). Gene regulatory networks. Proceedings of the National Academy of Sciences 102, 4935–4935.

Davidson, E.H., and Erwin, D.H. (2006). Gene Regulatory Networks and the Evolution of Animal Body Plans. Science 311, 796–800.

De Domenico, E., Owens, N. D. L., Grant, I. M., Gomes-Faria, R., & Gilchrist, M. J. (2015). Molecular asymmetry in the 8-cell stage Xenopus tropicalis embryo described by single blastomere transcript sequencing. Developmental Biology, 408(2), 252–268. doi.org/10.1016/j.ydbio.2015.06.010

Fukuda, M., Takahashi, S., Haramoto, Y., Onuma, Y., Kim, Y.-J., Yeo, C.-Y., Ishiura, S., and Asashima, M. (2010). Zygotic VegT is required for Xenopus paraxial mesoderm formation and is regulated by Nodal signaling and Eomesodermin. Int. J. Dev. Biol. 54, 81–92.

Grijzenhout, A., Godwin, J., Koseki, H., Gdula, M.R., Szumska, D., McGouran, J.F., Bhattacharya, S., Kessler, B.M., Brockdorff, N., and Cooper, S. (2016). Functional analysis of AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ESCs. Development 143, 2716–2723.

Haines, J.E., and Eisen, M.B. (2018). Patterns of chromatin accessibility along the anterior-posterior axis in the early Drosophila embryo. PLoS Genet 14, e1007367.

Hamm, D.C., and Harrison, M.M. (2018). Regulatory principles governing the maternal-to-zygotic transition: insights from Drosophila melanogaster. Open Biology 8, 180183.

Harrison, M.M., Li, X.-Y., Kaplan, T., Botchan, M.R., and Eisen, M.B. (2011). Zelda Binding in the Early Drosophila melanogaster Embryo Marks Regions Subsequently Activated at the Maternal-to-Zygotic Transition. PLoS Genet 7, e1002266.

Hashimshony, T., Feder, M., Levin, M., Hall, B.K., and Yanai, I. (2015). Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. Nature 519, 219–222.

Hauri, S., Comoglio, F., Seimiya, M., Gerstung, M., Glatter, T., Hansen, K., Aebersold, R., Paro, R., Gstaiger, M., and Beisel, C. (2016). A High-Density Map for Navigating the Human Polycomb Complexome. CellReports 17, 583–595.

Heasman, J., Crawford, A., Goldstone, K., Garner-Hamrick, P., Gumbiner, B., McCrea, P., Kintner, C., Noro, C.Y., and Wylie, C. (1994). Overexpression of cadherins and underexpression of β-catenin inhibit dorsal mesoderm induction in early Xenopus embryos. Cell 79, 791–803.

Heasman, J. (2006). Maternal determinants of embryonic cell fate. Seminars in Cell and Developmental Biology 17, 93–98.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-Enhancers in the Control of Cell Identity and Disease. Cell 155, 934–947.

Ho, L., Ronan, J.L., Wu, J., Staahl, B.T., Chen, L., Kuo, A., Lessard, J., Nesvizhskii, A.I., Ranish, J., and Crabtree, G.R. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. Proceedings of the National Academy of Sciences 106, 5181–5186.

Hontelez, S., van Kruijsbergen, I., Georgiou, G., van Heeringen, S.J., Bogdanović, O., Lister, R., and Veenstra, G.J.C. (2015). Embryonic transcription is controlled by maternally defined chromatin state. Nature Communications 6, 10148EP–.

Horb, M.E., and Thomsen, G.H. (1997). A vegetally localized T-box transcription factor in Xenopus eggs specifies mesoderm and endoderm and is essential for embryonic mesoderm formation. Development 124, 1689.

Hurlstone, A., and Clevers, H. (2002). T-cell factors: turn-ons and turn-offs. The EMBO Journal 21, 2303–2311.

Iwafuchi-Doi, M., and Zaret, K.S. (2014). Pioneer transcription factors in cell reprogramming. Genes & Development 28, 2679–2692.

Iwafuchi-Doi, M., Donahue, G., Kakumanu, A., Watts, J.A., Mahony, S., Pugh, B.F., Lee, D., Kaestner, K.H., and Zaret, K.S. (2016). The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. Molecular Cell 62, 79–91.

Jevtić, P., and Levy, D.L. (2015). Nuclear Size Scaling during Xenopus Early Development Contributesto Midblastula Transition Timing. Current Biology 25, 45–52.

Juan, A.H., Wang, S., Ko, K.D., Zare, H., Tsai, P.-F., Feng, X., Vivanco, K.O., Ascoli, A.M., Gutierrez-Cruz, G., Krebs, J., et al. (2016). Roles of H3K27me2 and H3K27me3 Examined during Fate Specification of Embryonic Stem Cells. CellReports 17, 1369–1382.

Jukam, D., Shariati, S.A.M., and Skotheim, J.M. (2017). Zygotic Genome Activation in Vertebrates. Developmental Cell 42, 316–332.

Kiecker, C., Bates, T., and Bell, E. (2015). Molecular specification of germ layers in vertebrate embryos. Cellular and Molecular Life Sciences 73, 923–947.

Kim, J., Lee, Y., Lu, X., Song, B., Fong, K.-W., Cao, Q., Licht, J.D., Zhao, J.C., and Yu, J. (2018). Polycomb- and Methylation-Independent Roles of EZH2 as a Transcription Activator. Cell Reports 25, 2808–2820.e4.

Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the regulatory epigenome. Nat Rev Genet 1–14.

Kofron, M., Demel, T., Xanthos, J., Lohr, J., Sun, B., Sive, H., Osada, S., Wright, C., Wylie, C., and Heasman, J. (1999). Mesoderm induction in Xenopus is a zygotic event regulated by maternal VegT via TGFbeta growth factors. Development 126, 5759–5770.

Laprell, F., Finkl, K., and Müller, J. (2017). Propagation of Polycomb-repressed chromatin requires sequence-specific recruitment to DNA. Science 356, 85–88.

Laugesen, A., Højfeldt, J.W., and Helin, K. (2016). Role of the Polycomb Repressive Complex 2 (PRC2) in Transcriptional Regulation and Cancer. Cold Spring Harb Perspect Med 6, a026575.

Lee, H., Habas, R., and Abate-Shen, C. (2004). Msx1 Cooperates with Histone H1b for Inhibition of Transcription and Myogenesis. Science 304, 1675–1678.

Lee, M.T., Bonneau, A.R., Takacs, C.M., Bazzini, A.A., DiVito, K.R., Fleming, E.S., and Giraldez, A.J. (2013). Nanog, Pou5f1 and SoxB1 activatezygotic gene expression during thematernal-to-zygotic transition. Nature 503, 360–364.

Levine, M. (2010). Transcriptional Enhancers in Animal Review Development and Evolution. Current Biology 20, R754–R763.

Li, E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet 3, 662–673.

Li, X., Tran, K.M., Aziz, K.E., Sorokin, A.V., Chen, J., and Wang, W. (2016). Defining the Protein-Protein Interaction Network of the Human Protein Tyrosine Phosphatase Family. Mol Cell Proteomics 15, 3030–3044.

Liang, H.-L., Nien, C.-Y., Liu, H.-Y., Metzstein, M.M., Kirov, N., and Rushlow, C. (2008). The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. Nature 456, 400–403.

Lindeman, L.C., Andersen, I.S., Reiner, A.H., Li, N., Aanes, H., Østrup, O., Winata, C., Mathavan, S., Müller, F., Aleström, P., et al. (2011). Prepatterning of Developmental Gene Expression by Modified Histones before Zygotic Genome Activation. Developmental Cell 21, 993–1004.

Liu, G., Wang, W., Hu, S., Wang, X., and Zhang, Y. (2018). Inherited DNA methylation primes the establishment of accessible chromatin during genome activation. Genome Research 28, 998–1007.

Liu, X., Wang, C., Liu, W., Li, J., Li, C., Kou, X., Chen, J., Zhao, Y., Gao, H., Wang, H., et al. (2016). Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. Nature 537, 558–562.

Loose, M., and Patient, R. (2004). A genetic regulatory network for Xenopus mesendoderm formation. Developmental Biology 271, 467–478.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 389, 251–260.

Lustig, K.D., Kroll, K.L., Sun, E.E., and Kirschner, M.W. (1996). Expression cloning of a Xenopus T-related gene (Xombi) involved in mesodermal patterning and blastopore lip formation. Development 122, 4001–4012.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. Nature 469, 343–349.

Mezger, A., Klemm, S., Mann, I., Brower, K., Mir, A., Bostick, M., Farmer, A., Fordyce, P., Linnarsson, S., and Greenleaf, W. (2018). High-throughput chromatin accessibility profiling at single-cell resolution. Nature Communications 1–6.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448, 553–560.

Miyamoto, K., Teperek, M., Yusa, K., Allen, G.E., Bradshaw, C.R., and Gurdon, J.B. (2013). Nuclear Wave1 Is Required for Reprogramming Transcription in Oocytes and for Normal Development. Science 341, 1002–1005.

Nakayama, T., Blitz, I.L., Fish, M.B., Odeleye, A.O., Manohar, S., Cho, K.W.Y., and Grainger, R.M. (2014). Cas9-Based Genome Editing inXenopus tropicalis (Elsevier Inc.).

Nekrasov, M., Wild, B., and Müller, J. (2005). Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. EMBO Rep 6, 348–353.

Neph, S., Vierstra, J., Stergachis, A.B., Reynolds, A.P., Haugen, E., Vernot, B., Thurman, R.E., John, S., Sandstrom, R., Johnson, A.K., et al. (2012). An expansive human regulatory lexicon encoded in transcription factor footprints. Nature 488, 83–90.

Novaka, A and Dedhar,S. (1999). Signaling through β-catenin and Lef/Tcf. Cell Mol Life Sci.;56(5-6):523-37.

Osada, S.I., and Wright, C.V. (1999). Xenopus nodal-related signaling is essential for mesendodermal patterning during early embryogenesis. Development 126, 3229–3240.

Owens, N.D.L., Blitz, I.L., Lane, M.A., Patrushev, I., Overton, J.D., Gilchrist, M.J., Cho, K.W.Y., and Khokha, M.K. (2016). Measuring Absolute RNA Copy Numbers at High Temporal Resolution Reveals Transcriptome Kinetics in Development. CellReports 14, 632–647.

Paraiso, K.D., Blitz, I.L., Coley, M., Cheung, J., Sudou, N., Taira, M., and Cho, K.W.Y. (2019). Endodermal Maternal Transcription Factors Establish Super-Enhancers during Zygotic Genome Activation. CellReports 27, 2962–2977.e2965.

Paranjpe, S.S., and Veenstra, G.J.C. (2015). Biochimica et Biophysica Acta. BBA - Gene Regulatory Mechanisms 1849, 626–636.

Peterson, C.L., and Workman, J.L. (2000). Promoter targeting and chromatin remodeling by the SWI/SNF complex. Current Opinion in Genetics & Development 10, 187–192.

Pukrop, T., Gradl, D., Henningfeld, K.A., Knöchel, W., Wedlich, D., and Kühl, M. (2001). Identification of Two Regulatory Elements within the High Mobility Group Box Transcription Factor XTCF-4. Journal of Biological Chemistry 276, 8968–8978.

Reid, C.D., Steiner, A.B., Yaklichkin, S., Lu, Q., Wang, S., Hennessy, M., and Kessler, D.S. (2016). Developmental Biology. Developmental Biology 414, 34–44.

Ribeiro, L., Tobias-Santos, V., Santos, D., Antunes, F., Feltran, G., de Souza Menezes, J., Aravind, L., Venancio, T.M., and Nunes da Fonseca, R. (2017). Evolution and multiple roles of the Pancrustacea specific transcription factor zelda in insects. PLoS Genet 13, e1006868.

Roël, G., van den Broek, O., Spieker, N., Peterson-Maduro, J., and Destrée, O. (2003). Tcf-1 expression during Xenopus development. Gene Expression Patterns 3, 123–126.

Schier, A.F. (2003). Nodal Signaling in Vertebrate Development. Annu. Rev. Cell Dev. Biol. 19, 589–621.

Schuettengruber, B., Bourbon, H.-M., Di Croce, L., and Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. Cell 171, 34–57.

Schulz, K.N., Bondra, E.R., Moshe, A., Villalta, J.E., Lieb, J.D., Kaplan, T., McKay, D.J., and Harrison, M.M. (2015). Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early Drosophilaembryo. Genome Research 25, 1715–1726.

Shan, Y., Liang, Z., Xing, Q., Zhang, T., Wang, B., Tian, S., Huang, W., Zhang, Y., Yao, J., Zhu, Y., et al. (2017). PRC2 specifies ectoderm lineages and maintainspluripotency in primed but not naïve ESCs. Nature Communications 1–14.

Shen, M.M. (2007). Nodal signaling: developmental roles and regulation. Development 134, 1023–1034.

Stennard, F., Carnac, G., and Gurdon, J.B. (1996). The Xenopus T-box gene, Antipodean, encodes a vegetally localised maternal mRNA and can trigger mesoderm formation. Development 122, 4179–4188.

Sudou, N., Yamamoto, S., Ogino, H., and Taira, M. (2012). Dynamic in vivo binding of transcription factors to cis-regulatory modules of cer and gsc in the stepwise formation of the Spemann-Mangold organizer. Development 139, 1651–1661.

Sun, Y., Nien, C.-Y., Chen, K., Liu, H.-Y., Johnston, J., Zeitlinger, J., and Rushlow, C. (2015). Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophilazygotic genome activation. Genome Research 25, 1703–1714.

Koide, T., Hayata, T., and Cho, K.W.Y. (2005). Xenopus as a model system to study transcriptional regulatory networks. Proceedings of the National Academy of Sciences 102, 4943–4948.

Takahashi, K., and Yamanaka, S. A decade of transcription factor-mediated reprogramming to pluripotency. Nat Rev Mol Cell Biol 17, 183EP–.

Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A., and Narechania, A. (2003). PANTHER: A Library of Protein Families and Subfamilies Indexed by Function. Genome Research 13, 2129–2141.

Tie, F., Stratton, C.A., Kurzhals, R.L., and Harte, P.J. (2007). The N Terminus of DrosophilaESC Binds Directly to Histone H3 and Is Required for E(Z)-Dependent Trimethylation of H3 Lysine 27. Molecular and Cellular Biology 27, 2014–2026.

Tsompana, M., and Buck, M.J. (2014). Chromatin accessibility: a window into the genome. Epigenetics & Chromatin 7, 251.

van Heeringen, S.J., Akkers, R.C., van Kruijsbergen, I., Arif, M.A., Hanssen, L.L.P., Sharifi, N., and Veenstra, G.J.C. (2014). Principles of nucleation of H3K27 methylation during embryonic development. Genome Research 24, 401–410.

Vastenhouw, N.L., Zhang, Y., Woods, I.G., Imam, F., Regev, A., Liu, X.S., Rinn, J., and Schier, A.F. (2010). Chromatin signature of embryonic pluripotency is established during genome activation. Nature 464, 922–926.

Wang, W., Li, X., Huang, J., Feng, L., Dolinta, K.G., and Chen, J. (2014). Defining the Protein–Protein Interaction Network of the Human Hippo Pathway. Mol Cell Proteomics 13, 119–131.
White, J.A., and Heasman, J. (2007). Maternal control of pattern formation inXenopus laevis. J. Exp. Zool. 310B, 73–84.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell 153, 307–319.

Wilson, B.G., and Roberts, C.W.M. (2011). SWI/SNF nucleosome remodellers and cancer. Nat Rev Cancer 11, 481–492.

Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. Genes & Development 25, 2227–2241.

Zhang, B., Zheng, H., Huang, B., Li, W., Xiang, Y., Peng, X., Ming, J., Wu, X., Zhang, Y., Xu, Q., et al. (2016). Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. Nature 537, 553–557.

Zhang, C., and Klymkowsky, M.W. (2007). The Sox axis, Nodal signaling, and germ layer specification. Differentiation 75, 536–545.

Zhang, C., Basta, T., Fawcett, S.R., and Klymkowsky, M.W. (2005). SOX7 is an immediate-early target of VegT and regulates Nodal-related gene expression in Xenopus. Developmental Biology 278, 526–541.

Zhang, C., Basta, T., Hernandez-Lagunas, L., Simpson, P., Stemple, D.L., Artinger, K.B., and Klymkowsky, M.W. (2004). Repression of nodal expression by maternal B1-type SOXs regulates germ layer formation in Xenopus and zebrafish. Developmental Biology 273, 23–37.

Zhang, J., Houston, D.W., Lou King, M., Payne, C., Wylie, C., and Heasman, J. (1998). The Role of Maternal VegT in Establishing the Primary Germ Layers in Xenopus Embryos. Cell 94, 515–524.

Zhang, J., and King, M.L. (1996). Xenopus VegT RNA is localized to the vegetal cortex during oogenesis and encodes a novel T-box transcription factor involved in mesodermal patterning. Development 122, 4119–4129.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). DNase-seq to Study Chromatin Accessibility in Early Xenopus tropicalisEmbryos. Genome Biology 2019, pdb.prot098335.

Zhou, Y., Bin Zhou, Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature Communications 1–10.