

NORTHWESTERN UNIVERSITY

Quantitative Landscape of *Xenopus laevis* Transcriptome Dynamics During Lineage  
Restriction

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Biological Sciences

By

Kristin Lorette Johnson

EVANSTON, ILLINOIS

December 2022

## Abstract

### Quantitative Landscape of *Xenopus laevis* Transcriptome Dynamics During Lineage Restriction

Kristin Lorette Johnson

One of the fundamental questions in developmental biology is how a single cell gives rise to a complex organism. More specifically, how a totipotent egg divides into cells that become increasingly restricted in their potential. Development is a process of increasingly restricted cellular potential, and here I home in on the transition from pluripotent cells that are able to give rise to all cell types, to multipotent germ layers with differing and finite progenitor populations. I examine this transition from pluripotency to multipotency at high resolution at the transcriptome level in order to understand the dynamics of the transcriptome as cells exit pluripotency. Important biology is missed in our understanding of the exit from pluripotency when only the start and end points are studied. Much of the transition out of pluripotency and toward specific multipotent cell populations is driven by genes transiently expressed in a non-monotonic fashion, and thus a high resolution study is needed to understand how cells are able to exit pluripotency and how the transcriptome changes differ in cells pushed towards different germ layers.

In this thesis, I developed a highly reproducible pipeline for analyzing the *Xenopus* transcriptome at six time points in the transition from pluripotency to four multipotent cell populations, endoderm, mesoderm, and the ectodermal derivatives, epidermis and neural progenitor cells. I provide quantitative support for the neural default model by demonstrating that the path to the neural lineage is the shortest and most linear, as these cells achieve an early equilibrium in their transcriptome dynamics not seen in the other three lineages. I identify novel divergent roles for BMP signaling before and after developmental stage 10.5. It has long been known that BMP signaling, in part, drives epidermal formation. Here I show that the epidermal and neural lineages have a largely shared trajectory until stage 10.5, at which point BMP signaling is activated in the epidermal lineage only, and these two lineages diverge. Importantly, early activation of BMP signaling does not drive an early divergence between these two lineages, but rather pushes cells to a mesodermal fate, suggesting cells are not able to appropriately respond to BMP to form epidermis until the time it comes endogenously, stage 10.5. The mechanism by which BMP

signaling is held off until the appropriate time is driven, in part, by the presence of the maternally deposited transcription factor, *Dand5*. I also show that exposure to high levels of Activin leads to an early divergence in the transcriptome that is unique to endoderm cells, and I use this high resolution transcriptome data to provide temporal resolution to a previously established gene regulatory network (GRN) for the mesendoderm, as well as predict novel members of this GRN. The work in this thesis enhances our understanding of how cells progress from pluripotent to multipotent cell populations and provides valuable insight into the transcriptome at several time points during this transition to four lineages, providing a framework by which to identify previously identified key players in this transition, and to establish novel gene regulatory networks and enhance already established networks.

## Table of Contents

<b>Abstract.....</b>	<b>2</b>
<b>List of Figures.....</b>	<b>10</b>
<b>Chapter 1.....</b>	<b>13</b>
<b>General Introduction.....</b>	<b>13</b>
<b>Embryonic Development and Lineage Restriction.....</b>	<b>14</b>
Figure 1.1 Waddington's Epigenetic Landscape.....	15
Figure 1.2 Waddington's Modified Epigenetic Landscape.....	17
<b>Pluripotency and the Stem Cell State.....</b>	<b>18</b>
<b>Primary Germ Layers: An Overview.....</b>	<b>19</b>
Figure 1.3 Gastrulation of <i>Xenopus</i> Embryos.....	20
<b>Ectoderm and the Neural Default Model.....</b>	<b>21</b>
Figure 1.4 BMP Signaling is Crucial for Ectodermal Differentiation.....	23
Figure 1.5 Ectodermal Explants will form Neural Progenitors in Absence of BMP.....	24
<b>Mesendoderm.....</b>	<b>26</b>
Figure 1.6 Role of Maternal VegT in Mesendoderm Formation.....	27
Figure 1.7 Mesendoderm Gene Regulatory Network.....	29
<b>Neural Crest as the Fourth Germ Layer.....</b>	<b>30</b>
Figure 1.8 Neural Crest cells yield Ectodermal and Mesodermal Derivatives.....	31
Figure 1.9 Neural Crest GRN.....	33
<b>TGF-Beta Signaling and its Role in Development.....</b>	<b>34</b>
Figure 1.10 TGF-Beta Signaling.....	35
<b>FGF and Wnt Signaling Pathways.....</b>	<b>37</b>
Figure 1.11 Wnt and TGF-Beta Crosstalk.....	38
<b>Epigenetics in Lineage Restriction: An Overview .....</b>	<b>39</b>
Figure 1.12 Epigenetic Regulation of DNA and Histones.....	40
Figure 1.13 DNA Methylation Pattern Established and Maintained by DNMTs.....	43

<b>Role of DNA Methylation in Pluripotency Maintenance.....</b>	<b>44</b>
<b>Role of DNA Methylation in Alternative Splicing.....</b>	<b>46</b>
<b>Small Molecule Inhibitors Targeting DNMTs.....</b>	<b>46</b>
Figure 1.14 DNA Methylation Regulates Alternative Splicing.....	47
Figure 1.15 5-Aza Inhibits DNMTs.....	48
<b><i>Xenopus</i> as a Model System.....</b>	<b>49</b>
Figure 1.16 <i>Xenopus laevis</i> Development Time Frame.....	50
Figure 1.17 Animal Cap Assay.....	51
<b>Specific Questions to be Addressed in This Thesis.....</b>	<b>52</b>
<b>Chapter 2.....</b>	<b>54</b>
<b>Quantitative Analysis of Transcriptome Dynamics Provides Novel Insights into</b>	
<b>Developmental State Transitions.....</b>	<b>54</b>
<b>Introduction.....</b>	<b>55</b>
<b>Results.....</b>	<b>58</b>
Naïve Animal Pole Cells from <i>Xenopus</i> Blastula can be Programmed to any Lineage.....	59
Figure 2.1 Dynamics of Maternal Transcripts Demonstrate Assay Reproducibility.....	60
Figure 2.2 Four Independent Lineages are Generated.....	61
PCA and Time Series Analysis Reveal Novel Lineage-Specific Dynamics.....	62
Figure 2.3 Transcription Factor Expression Shows Overlap Between Lineages.....	63
Figure 2.4 PCA Reveals Path to Neural Lineage is One-Dimensional.....	65
Figure 2.5 PCA Reveals Stage 9 to 13 Distance is Shortest in Neural Lineage.....	66
Gene Expression Dynamics Provide Novel Insights into the Neural Default State.....	67
Figure 2.6 Temporal DESeq Reveals Neural Lineage Reaches Early Equilibrium.....	68
Figure 2.7 Each Lineage has both Unique and Shared Transcriptome Dynamics.....	69
Figure 2.8 Epidermal and Neural Lineages Show Early Overlapping Dynamics.....	70
Figure 2.9 Neural GO Terms are not Specific to Neural Lineage Until Stage 11.....	72

Figure 2.10 Gene Expression Dynamics of Epidermal Formation.....	73
Figure 2.11 Between Lineage DESeq Shows Epidermal and Neural Differences.....	75
Figure 2.12 Expression Dynamics of Pluripotency Genes Maintained in Neural.....	76
Robust BMP Signaling is Initiated in Explants Around the Onset of Gastrulation.....	77
Early Response to Activin and BMP4/7 Reveals Unexpected Overlap.....	77
Figure 2.13 BMP Signaling is Detectable at Epidermal-Neural Divergence Point.....	78
Figure 2.14 Gene Expression of BMP Targets.....	79
Figure 2.15 Western Blot Shows Robust Transient Activin Response.....	81
Figure 2.16 Dynamics of Transcriptional Response to Activin and BMP.....	82
Figure 2.17 Dynamics of Genes Immediately Responding to BMP.....	84
Figure 2.18 Dynamics of Genes with Later BMP Response.....	85
Figure 2.19 PCA of Endoderm and Activin and BMP Generated Mesoderm.....	86
Figure 2.20 Dorsal / Ventral Clustering.....	87
Time Series Data Provides Novel Insights into Mesendoderm GRN.....	88
Figure 2.21 Treatment with BMP4/7 Elicits Early BMP Signaling.....	89
Figure 2.22 Dynamics of Genes with Immediate Activin Response.....	90
Figure 2.23 Dynamics of Genes with Later Activin Response.....	91
Figure 2.24 DE Genes Between Mesoderm and Endoderm.....	92
Figure 2.25 Dynamics of Mesendoderm GRN.....	93
Figure 2.26 Dynamics of Mesendoderm Onset.....	95
Figure 2.27 Proposed Mesendoderm GRN Candidates.....	96
Figure 2.28 WGCNA on Mesoderm and Endoderm Stages 10-11.....	97
Early BMP Signaling Drives Ventral Mesoderm not Early Epidermal Divergence.....	98
Figure 2.29 Overlap in Temporal Dynamics of Mesoderm and Epidermis.....	100
Figure 2.30 DESeq Between Epidermal and Mesoderm Lineages.....	101
Figure 2.31 Varying Patterns of Transcriptome Response to Early BMP.....	102

Figure 2.32 Exogenous BMP at Stage 10.5 Does not Induce Mesoderm.....	103
BMP Signaling is Restrained Until Stage 10.5 by <i>dand5</i> Activity.....	104
Figure 2.33 Maternal BMP Inhibitors Explain Endogenous BMP Timing.....	105
Figure 2.34 <i>Dand5</i> MO Induces Early BMP and Mesoderm Formation.....	106
<b>Discussion.....</b>	<b>107</b>
Neural as the Default State.....	108
Surprising Overlap in Transcription Response to BMP4/7 and Activin.....	109
The Timing of BMP Signaling is Critical for Proper Lineage Segregation.....	110
<b>Materials and Methods.....</b>	<b>112</b>
Embryological Methods.....	112
RNA Isolation, cDNA Synthesis, Sequencing and qRT-PCR.....	113
Western Blot Analysis.....	113
RNA Seq Processing and Computational Analysis.....	114
Vertebrate Animals.....	116
Availability of Data and Materials.....	116
<b>Chapter 3.....</b>	<b>117</b>
<b>DNA Methylation is Required for Normal Lineage Restriction .....</b>	<b>117</b>
<b>Introduction.....</b>	<b>118</b>
<b>Results.....</b>	<b>122</b>
Inhibition of DNA Methylation Blocks Lineage Formation.....	122
Figure 3.1 5-Aza Heterogeneously Affects Pluripotency Factors.....	123
Figure 3.2 5-Aza Partially Prolongs <i>Sox3</i> and Prevents <i>EpK</i> Expression.....	124
Figure 3.3 5-Aza Reduces Cellular Competency.....	125
DNA Methylation Inhibition Does Not Significantly Change Transcriptome Until Stage 13.....	126
Figure 3.4 Hierarchical Clustering of RNA Seq TPM.....	127
Figure 3.5 DESeq Between Conditions.....	129
Figure 3.6 Comparison of 5-Aza Effects at Different Developmental Stages.....	130

Figure 3.7 GO Enrichment of Genes Higher in DMSO Caps at Stage 13.....	131
Figure 3.8 GO Enrichment of Genes Higher in 5-Aza Caps at Stage 13.....	132
5-Aza Upregulates Cell Signaling Genes and Downregulates Structural Genes.....	133
Figure 3.9 DESeq Between Developmental Stages.....	134
Figure 3.10 Comparison of Transcriptome Changes Over Time.....	135
Figure 3.11 GO Enrichment of Genes Increasing in Response to 5-Aza.....	136
Figure 3.12 GO Enrichment of Genes Increasing with DMSO Only.....	137
Figure 3.13 GO Enrichment of Genes Increasing in Both Conditions.....	138
Figure 3.14 GO Enrichment of Genes Decreasing in Response to 5-Aza.....	139
Figure 3.15 GO Enrichment of Genes Decreasing in DMSO Only.....	140
Figure 3.16 GO Enrichment of Genes Decreasing in Both Conditions.....	141
PCA Suggests Alternative Splicing Accounts for Majority of Variance.....	142
Figure 3.17 PCA of 5-Aza And Primary Lineages.....	143
Figure 3.18 Principal Components Contributing to Stage 13 Variance.....	144
Figure 3.19 GO Term Enrichment for Top Genes Loading PC1.....	145
<b>Discussion.....</b>	<b>146</b>
<b>Materials and Methods.....</b>	<b>149</b>
Embryological Methods.....	149
RNA Isolation and Sequencing.....	150
RNA Seq Processing and Computational Analysis.....	150
Vertebrate Animals.....	151
<b>Chapter 4.....</b>	<b>152</b>
<b>General Discussion.....</b>	<b>152</b>
<b>Quantitative Insights into The Dynamics of Lineage Restriction.....</b>	<b>154</b>
<b>Quantitative Backing for and Biological Insights into the Neural Default Model.....</b>	<b>156</b>
<b>Expansion of Mesendoderm GRN.....</b>	<b>157</b>
<b>Examination of BMP vs Activin Targets.....</b>	<b>158</b>

<b>Dand5 Temporally Restrains BMP Signaling.....</b>	<b>159</b>
<b>Effect of Inhibition of DNA Methylation on Lineage Specification.....</b>	<b>161</b>
<b>Role of DNA Methylation on Regulation of Intron Splicing.....</b>	<b>162</b>
<b>Concluding Remarks.....</b>	<b>163</b>
<b>Significance of the Thesis Work.....</b>	<b>164</b>
<b>References.....</b>	<b>166</b>

### List of Figures

Figure 1.1 Waddington's Epigenetic Landscape.....	15
Figure 1.2 Waddington's Modified Epigenetic Landscape.....	17
Figure 1.3 Gastrulation of <i>Xenopus</i> Embryos.....	20
Figure 1.4 BMP Signaling is Crucial for Ectodermal Differentiation.....	23
Figure 1.5 Ectodermal Explants will form Neural Progenitors in Absence of BMP.....	24
Figure 1.6 Role of Maternal VegT in Mesendoderm Formation.....	27
Figure 1.7 Mesendoderm Gene Regulatory Network.....	29
Figure 1.8 Neural Crest cells yield Ectodermal and Mesodermal Derivatives.....	31
Figure 1.9 Neural Crest GRN.....	33
Figure 1.10 TGF-Beta Signaling.....	35
Figure 1.11 Wnt and TGF-Beta Crosstalk.....	38
Figure 1.12 Epigenetic Regulation of DNA and Histones.....	40
Figure 1.13 DNA Methylation Pattern Established and Maintained by DNMTs.....	43
Figure 1.14 DNA Methylation Regulates Alternative Splicing.....	47
Figure 1.15 5-Aza Inhibits DNMTs.....	48
Figure 1.16 <i>Xenopus laevis</i> Development Time Frame.....	50
Figure 1.17 Animal Cap Assay.....	51
Figure 2.1 Dynamics of Maternal Transcripts Demonstrate Assay Reproducibility.....	60
Figure 2.2 Four Independent Lineages are Generated.....	61
Figure 2.3 Transcription Factor Expression Shows Overlap Between Lineages.....	63
Figure 2.4 PCA Reveals Path to Neural Lineage is One-Dimensional.....	65
Figure 2.5 PCA Reveals Stage 9 to 13 Distance is Shortest in Neural Lineage.....	66
Figure 2.6 Temporal DESeq Reveals Neural Lineage Reaches Early Equilibrium.....	68
Figure 2.7 Each Lineage has both Unique and Shared Transcriptome Dynamics.....	69
Figure 2.8 Epidermal and Neural Lineages Show Early Overlapping Dynamics.....	70
Figure 2.9 Neural GO Terms are not Specific to Neural Lineage Until Stage 11.....	72

Figure 2.10 Gene Expression Dynamics of Epidermal Formation.....	73
Figure 2.11 Between Lineage DESeq Shows Epidermal and Neural Differences.....	75
Figure 2.12 Expression Dynamics of Pluripotency Genes Maintained in Neural.....	76
Figure 2.13 BMP Signaling is Detectable at Epidermal-Neural Divergence Point.....	78
Figure 2.14 Gene Expression of BMP Targets.....	79
Figure 2.15 Western Blot Shows Robust Transient Activin Response.....	81
Figure 2.16 Dynamics of Transcriptional Response to Activin and BMP.....	82
Figure 2.17 Dynamics of Genes Immediately Responding to BMP.....	84
Figure 2.18 Dynamics of Genes with Later BMP Response.....	85
Figure 2.19 PCA of Endoderm and Activin and BMP Generated Mesoderm.....	86
Figure 2.20 Dorsal / Ventral Clustering.....	87
Figure 2.21 Treatment with BMP4/7 Elicits Early BMP Signaling.....	89
Figure 2.22 Dynamics of Genes with Immediate Activin Response.....	90
Figure 2.23 Dynamics of Genes with Later Activin Response.....	91
Figure 2.24 DE Genes Between Mesoderm and Endoderm.....	92
Figure 2.25 Dynamics of Mesendoderm GRN.....	93
Figure 2.26 Dynamics of Mesendoderm Onset.....	95
Figure 2.27 Proposed Mesendoderm GRN Candidates.....	96
Figure 2.28 WGCNA on Mesoderm and Endoderm Stages 10-11.....	97
Figure 2.29 Overlap in Temporal Dynamics of Mesoderm and Epidermis.....	100
Figure 2.30 DESeq Between Epidermal and Mesoderm Lineages.....	101
Figure 2.31 Varying Patterns of Transcriptome Response to Early BMP.....	102
Figure 2.32 Exogenous BMP at Stage 10.5 Does not Induce Mesoderm.....	103
Figure 2.33 Maternal BMP Inhibitors Explain Endogenous BMP Timing.....	105
Figure 2.34 Dand5 MO Induces Early BMP and Mesoderm Formation.....	106
Figure 3.1 5-Aza Heterogeneously Affects Pluripotency Factors.....	123

Figure 3.2 5-Aza Partially Prolongs Sox3 and Prevents <i>EpK</i> Expression.....	124
Figure 3.3 5-Aza Reduces Cellular Competency.....	125
Figure 3.4 Hierarchical Clustering of RNA Seq TPM.....	127
Figure 3.5 DESeq Between Conditions.....	129
Figure 3.6 Comparison of 5-Aza Effects at Different Developmental Stages.....	130
Figure 3.7 GO Enrichment of Genes Higher in DMSO Caps at Stage 13.....	131
Figure 3.8 GO Enrichment of Genes Higher in 5-Aza Caps at Stage 13.....	132
Figure 3.9 DESeq Between Developmental Stages.....	134
Figure 3.10 Comparison of Transcriptome Changes Over Time.....	135
Figure 3.11 GO Enrichment of Genes Increasing in Response to 5-Aza.....	136
Figure 3.12 GO Enrichment of Genes Increasing with DMSO Only.....	137
Figure 3.13 GO Enrichment of Genes Increasing in Both Conditions.....	138
Figure 3.14 GO Enrichment of Genes Decreasing in Response to 5-Aza.....	139
Figure 3.15 GO Enrichment of Genes Decreasing in DMSO Only.....	140
Figure 3.16 GO Enrichment of Genes Decreasing in Both Conditions.....	141
Figure 3.17 PCA of 5-Aza And Primary Lineages.....	143
Figure 3.18 Principal Components Contributing to Stage 13 Variance.....	144
Figure 3.19 GO Term Enrichment for Top Genes Loading PC1.....	145

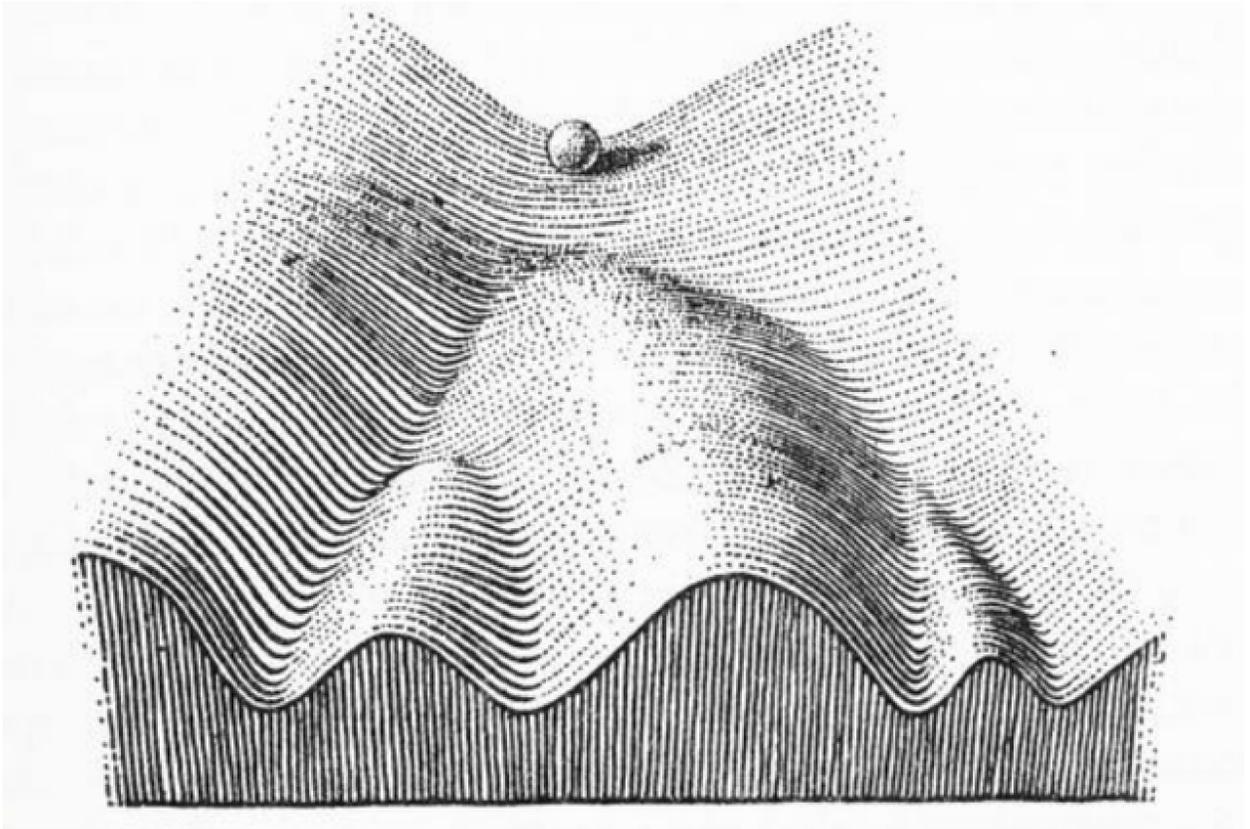
## **Chapter One**

### **General Introduction**

One of the most pressing questions in developmental biology is how a totipotent egg gives rise to a complex multicellular organism. At the onset of development, a single cell has the potential to give rise to every cell type in an entire organism, but as development progresses and cells divide, cellular potential decreases progressively and largely irreversibly. A totipotent egg first gives rise to a population of pluripotent cells, which have the ability to give rise to all embryonic cell types. The first restriction in cellular potential occurs when cells become one of three multi-potent germ layers: mesoderm, endoderm, and ectoderm. Each of these germ layers is still able to give rise to a multitude of different cell types but is restricted in their derivatives. The cells arising from these germ layers become increasingly restricted in the types of derivatives they form as development progresses and the cells become increasingly specified. Here, I characterize the progression from pluripotent stem cell to four distinct multipotent cell populations, epidermis, neural progenitor, endoderm, and mesoderm. I provide novel quantitative evidence for the neural default model, home in on a critical time period for Bone Morphogenetic Protein (BMP) signaling, identify *Dand5* as a maternally provided BMP inhibitor playing a critical role in regulating this timing, and propose novel members of the mesoderm gene regulatory network (GRN). Finally, I explore the effects of inhibition of DNA methylation on lineage formation and provide evidence for the important role of DNA methylation in regulating splicing.

### **Embryonic Development and Lineage Restriction**

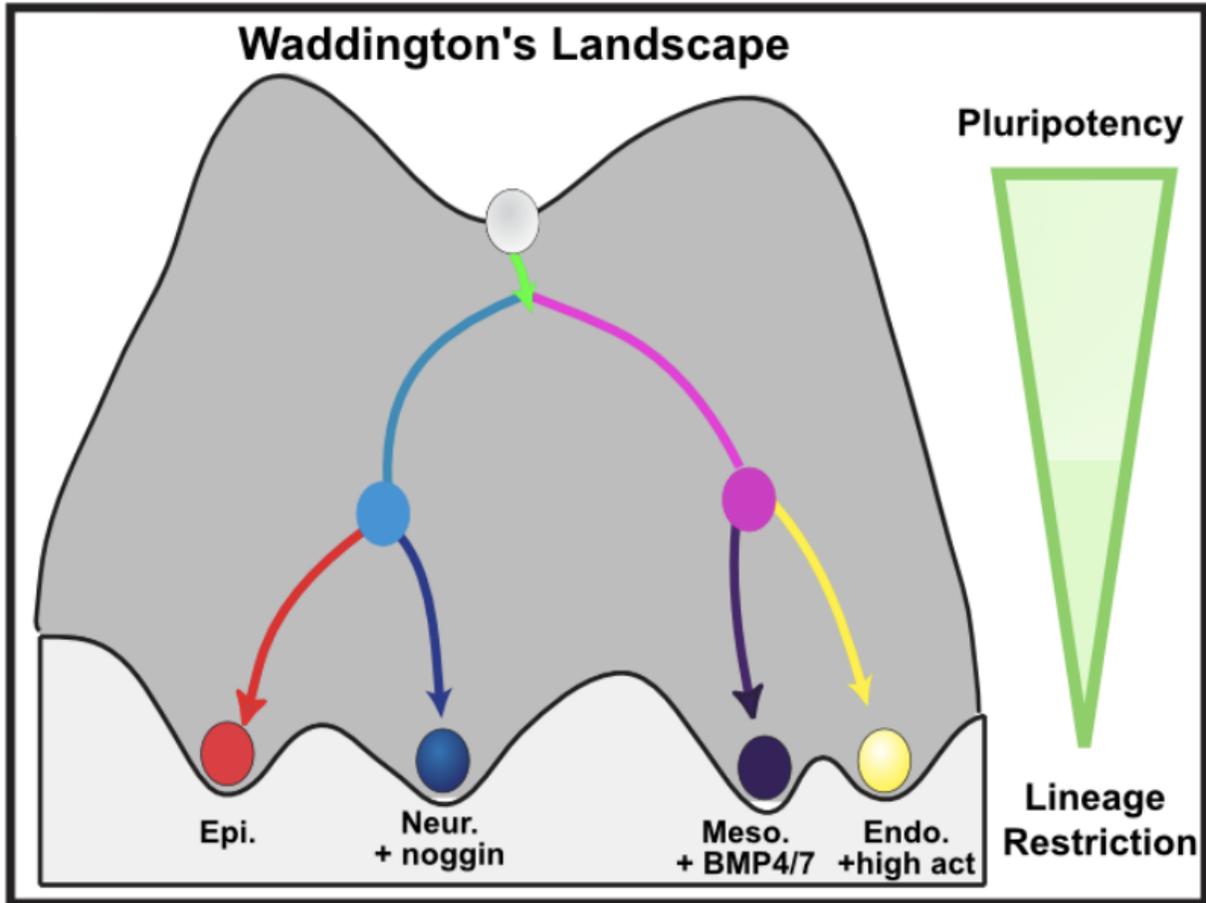
The process by which a pluripotent cell becomes increasingly restricted in its cellular potential was famously depicted by Conrad Waddington in 1957 as a topological landscape (Waddington, 1957). In his proposal, he envisioned a pluripotent cell as a marble atop a hill, with several paths the marble can take to the bottom of the hill. Which path this marble takes to the bottom of the hill will determine which cell type it will become, and while there are multiple decision points for each marble, they cannot defy gravity, and thus these cells cannot reverse course and increase their potential at any point (Figure 1.1). One notable exception to this progressive restriction is the neural crest, a cell population that maintains its stem cell characteristics late into development, which will be addressed in a later section. Though Waddington's original landscape was a simplistic and largely theoretical view of the cellular decrease in



**Figure 1.1 Waddington's Epigenetic Landscape.** Waddington's original picture of his epigenetic landscape, proposed in 1957, envisions a marble atop a hill, with a number of different decision points before commitment to a specific path towards the bottom. This model is used across a variety of fields today and is the foundation of the premise of the work in this thesis (Adapted from Waddington, 1957).

potential, he was remarkably perceptive in his proposed epigenetic landscape. Waddington first coined the term epigenetics in 1940 to encompass both the environmental and genetic factors that contribute to cellular phenotypes. Waddington described the development of an organ as a series of steps along one of many distinct epigenetic paths, which are affected by multiple genes, often in opposition to one another, with sequential states of development resulting from the balance of influence of many genes as well as the environment (Waddington 1940, Robertson 1977, Tronick and Hunter 2016). Waddington later expanded on the fluidity, or lack thereof, of these paths, naming his different cellular pathways chreods. He proposed that these paths were generally sufficiently canalized to be resistant to environmental noise and genetic variation below a certain threshold, further emphasizing that developmental changes could only be driven by gene complexes and their products, not individual genes (Waddington 1942, Robertson 1977, Goldberg et al. 2007, Baedke 2013). He described these paths as becoming increasingly canalized over time, and increasingly affected by the environmental and physiological forces interacting with genes (Waddington 1957, Tronick and Hunter 2016). Waddington suggested that different cell types were distinct from one another due to the variance of hundreds if not thousands of genes, as well as contributions from the environment (Waddington, 1968). Scientists have been contributing data to fill in the gaps of his largely theoretical model in the fields of developmental biology, cell and molecular biology and genetics ever since Waddington proposed his initial metaphor (Figure 1.2).

While Waddington deserves much credit for his groundbreaking theory on lineage restriction and embryonic development, it is important to note that he was not the first scientist to propose that a complex of genes in addition to environmental and physiological factors work in concert to drive embryonic development. Ernest Everett Just was one of the foremost embryologists in the early 1900s and had unmatched early success in fostering and studying normal embryonic development in laboratory settings because he understood the importance of the influence of nature and was able to simulate those settings in his lab. Just was also the first person to link cell surface changes and morphological changes to different stages of embryonic development and the first to identify critical forces of contraction and conduction in the cytoplasm (Crow 2008, Byrnes and Newman 2014). For these contributions, among others, and for his acknowledgement that both genetics and environment played a role in development,



**Figure 1.2 Waddington's Modified Epigenetic Landscape.** Waddington's Epigenetic Landscape provides a framework for studying lineage restriction, here I extrapolate current knowledge of germ layer formation onto the original landscape.

Just is considered to be the “Forgotten Father of Epigenetics” (Byrnes 2015). Alas, the remarkable ability of science to “forget” the contributions of women and people of color is outside the scope of this thesis.

### **Pluripotency and the Stem Cell State**

Cells atop Waddington’s landscape, pluripotent cells, maintain the remarkable, yet transient potential to become any cell type in the adult organism. Importantly, these cells have the ability not only to form derivatives of all three germ layers, ectoderm, mesoderm, and endoderm, and therefore all specified cell types, but also to self-renew for a limited time, thus retaining their potency by proliferating rather than differentiating (Young 2011). Pluripotent stem cells in the model organism for this thesis work, *Xenopus laevis*, are known as animal pole cells or animal caps. These cells are similar in potency to inner cell mass cells (ICM) in mice, and human embryonic stem cells (Snape et al. 1987, Evans and Kaufman 1981, Martin 1981, Thomson et al. 1998, Shambloott et al. 1998). This potential of pluripotent cells is stabilized and regulated by a core group of genes known as the pluripotency gene regulatory network (Li and Belmonte 2017). The key members of this pluripotency network are the Pou family member *Pou5f3.1 / Oct3/4*, *Nanog* (the functional equivalent of which is *Vent* in *Xenopus*) and *Sox2*, all of which both autoregulate and regulate one another in feed forward loops (Boyer et al. 2005) and share downstream targets that keep cells from differentiating (Loh et al. 2006). While these transcription factors have been shown to maintain pluripotency, a slightly different set of transcription factors, *Oct3/4*, *Sox2*, *Klf2* and *c-Myc*, together are able to induce pluripotency in already differentiated mouse and human somatic cells, re-endowing these cells with ability to differentiate into any of the three germ layers (Takahashi and Yamanaka 2006, Takahashi et al. 2007). Interestingly, *Nanog*, which is able to block differentiation into endoderm (Chambers et al. 2003, Hamazaki et al. 2004), mesoderm (Suzuki et al. 2006a,b), and neural (Ying et al. 2003) and thus acts as a potent pluripotency maintenance factor in vivo, is not required to reprogram differentiated human and mouse somatic cells to a pluripotent state (Niwa 2007).

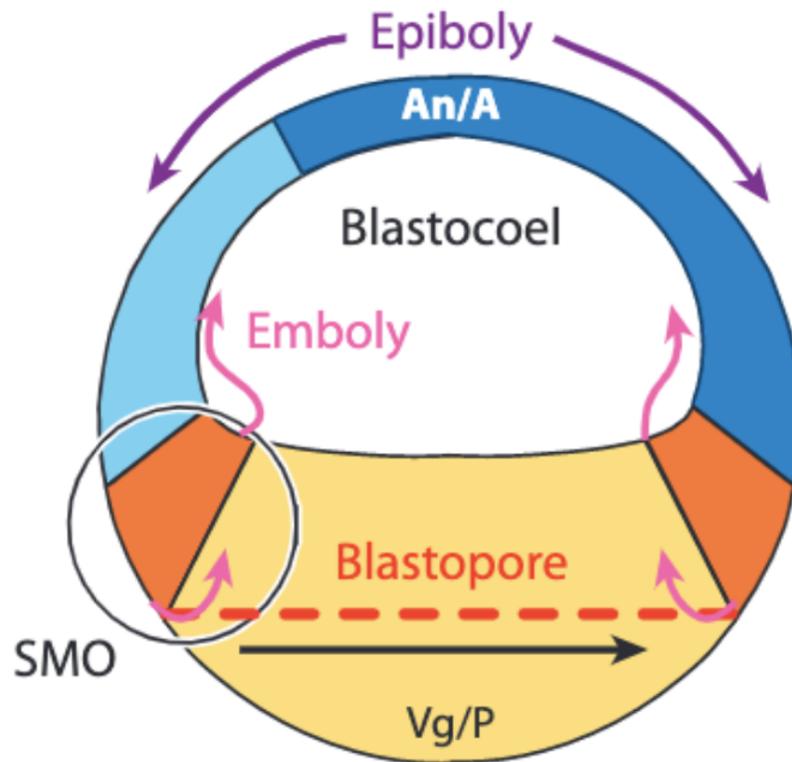
Expression of genes that allow for maintenance of and ultimately exit from the pluripotent state is largely dictated by chromatin accessibility and chromatin structure. One of the most fundamental epigenetic markers of pluripotency is histone methylation, with the trimethylation of lysine 27 of histone 3

(H3K27me3) reflecting a closed chromatin state, being indicative of transcriptional inactivity (Cao et al. 2002, Lachner et al. 2003), and trimethylation of lysine 4 of histone 3 (H3K4me3) reflecting an open chromatin state, reflective of transcriptional activity (Strahl et al. 1999, Santos-Rosa et al. 2002), discussed in more detail in a later section. Ultimately, as the chromatin structure and gene expression change, the transient pluripotency of cells is lost, and they become increasingly restricted over developmental time as they progress down Waddington's landscape.

### **Primary Germ Layers: An Overview**

One of the earliest decision points that can be projected onto Waddington's landscape is the divergence between the ectoderm and mesendoderm, and consequently the formation of the three primary germ layers; ectoderm, mesoderm and endoderm. The observation of organized cell layers in the developing embryo was first made in the 1800s by Caspar Wolff leading to the concept of these layers being germ layers, which was introduced by Christian Pander (Pander 1817, Keicker et al. 2016). Germ layer specification begins during gastrulation, a critical time in embryonic development. In fact, Lewis Wolpert famously said "It is not birth, marriage, or death, but gastrulation which is truly the most important time in your life" (Hopwood 2022). Gastrulation is characterized by emboly, epiboly, convergence, and extension, leading to the proper formation and positioning of individual germ layers (Figure 1.3). These gastrulation movements are regulated by a number of morphogenetic processes including cell adhesion, chemotaxis, chemokinesis and planar polarity and are coordinated with cell polarity to properly pattern the embryo (Keller 1991, Solnica-Krezel and Sepich 2012). While distinct morphogenetic movements occur in different vertebrate embryos, these movements establish highly similar embryos with many shared features regardless of vertebrate species (Watabe 1995). This thesis focuses on *Xenopus* gastrulation.

During emboly, mesodermal and endodermal progenitors are internalized beneath the prospective ectoderm via the blastopore in a process called involution, during which the actin cytoskeleton converts contractile force into motile force pushing these cells over the blastopore lip to form the archenteron roof and mesodermal cell stream or mesodermal mantle. The endoderm and mesoderm precursors then form a single tissue above the blastopore, marked by constriction of bottle cells on the



**Figure 1.3 Gastrulation of *Xenopus* Embryos.** Schematic of cross section of *Xenopus* gastrula stage embryo depicting both a fate map and patterns of gastrulation movement during epiboly and emboly. SMO = Spemann-Mangold Organizer (Adapted from Solnica-Krezel and Sepich 2012).

dorsal side of the embryo, near the Spemann-Mangold organizer before migrating away from the blastopore upon the epithelial to mesenchymal transition. This process is driven by the deep layer of the marginal zone, and the bottle cells anchor the superficial epithelium to the mesodermal mantle (Keller 1981, Hardin and Keller 1988, Shih and Keller 1994, Lee and Harland 2007). Epiboly is characterized by the thinning and spreading of germ layers due to interdigitation, as multiple cell layers are transformed into one larger layer as cells wedge between one another (Keller 1980). Finally, during convergence and extension, the embryo undergoes a narrowing of the mediolateral (ML) axis and elongates along the anteroposterior (AP) axis during which cells rearrange and change shape due to mediolateral intercalation behavior (MIB). These intercalations change the tissue geometry, and consequently the tissue interactions, and therefore have a significant effect on early embryonic patterning, including formation of the notochordal-somitic boundary (Keller 1978, Keller 1985, Shih and Keller 1992 a,b).

In addition to the physical cell movements and interactions that change embryo morphology during gastrulation, several signaling pathways provide both inductive and repressive signals that lead to germ layer specification, the initiation of which is driven by Spemann's Organizer (Spemann & Mangold 1924, Harland & Gerhart 1997, Niehrs 2004). Beta-catenin is dorsally enriched, initiating a signaling cascade that establishes the organizer that will coordinate gastrulation and pattern germ layers (Carnac et al. 1996) A ventral to dorsal BMP signaling gradient is also established and a cadherin-dependent cell adhesion gradient develops reverse of the BMP gradient, that helps determine cell movements during gastrulation (De Robertis & Kuroda 2004, Solnica-Krezel and Sepich 2012). These signaling pathways and cellular movements are tightly controlled both spatially and temporally within the embryo to allow for proper germ layer formation.

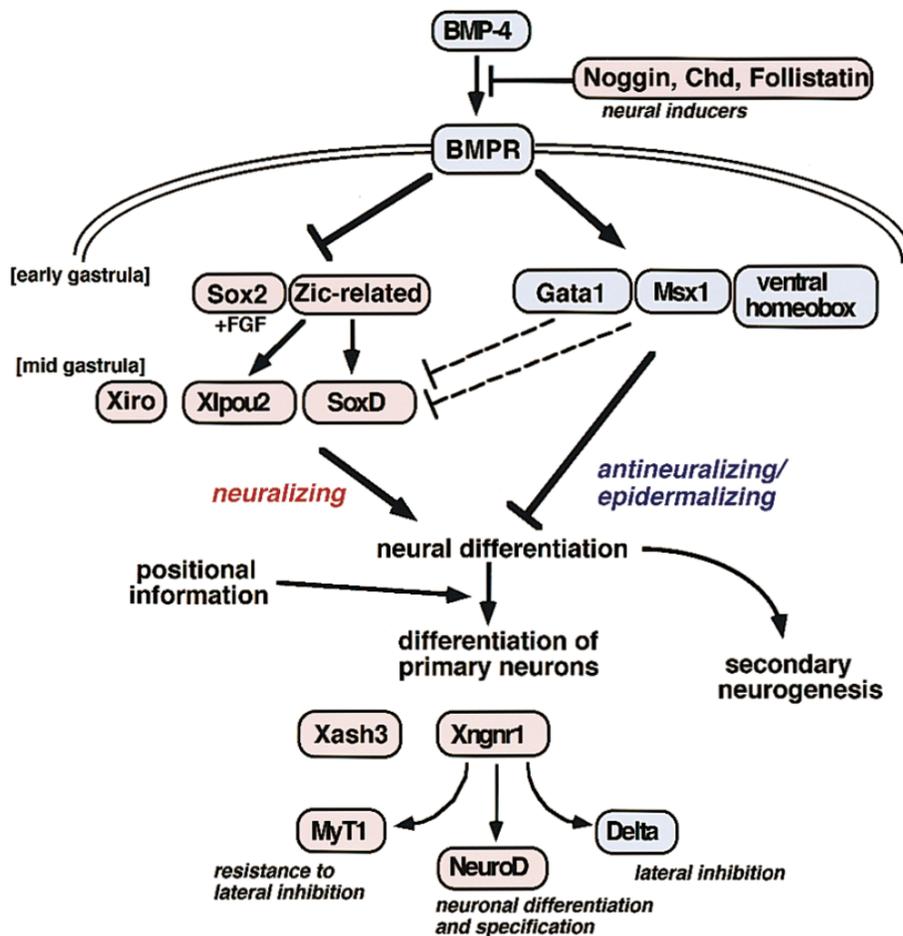
### **Ectoderm and the Neural Default Model**

Animal pole cells of the *Xenopus* embryo are fated to become ectoderm, but whether they form neural or non-neural ectoderm depends upon their position along the dorsal-ventral axis within the embryo. During gastrulation, dorsal ectoderm is induced to form neural progenitor cells. This neural induction was first discovered by Spemann and Mangold's grafting experiments in newts, in which the

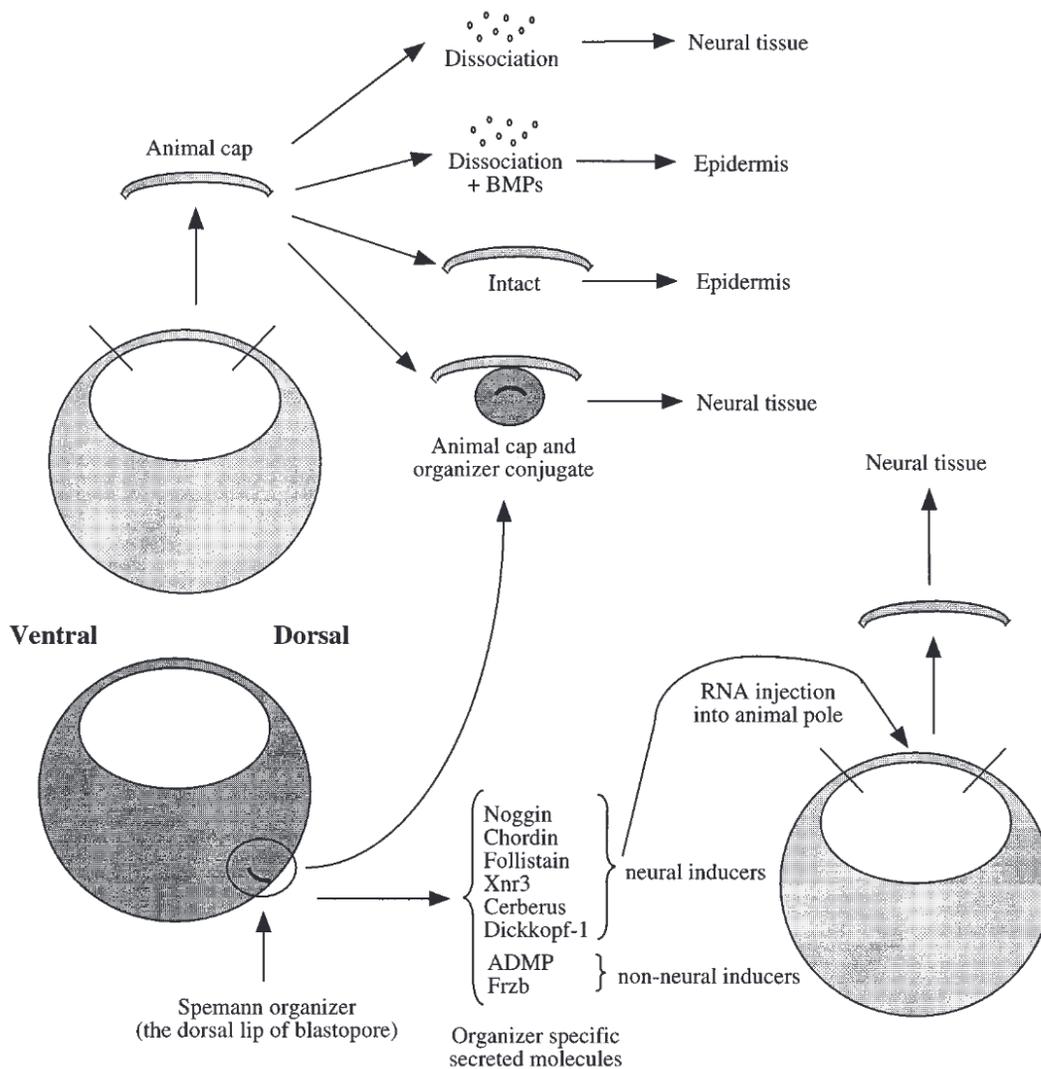
dorsal lip of the blastopore was grafted into the ventral side of a host, causing formation of secondary D-V and A-P axes as well as formation of a secondary nervous system from the host's ventral ectodermal tissue that would normally form epidermis (Spemann and Mangold 1924). These experiments implied there were neural inducing signals emanating from this tissue, and several endogenous neural inducers have since been identified. A neural inducer can be defined as a gene active at gastrula stages in the dorsal lip of an embryo that is secreted from the organizer and acts on the ectoderm non-cell autonomously to form neural tissue directly (Chang and Hemmati-Brivanlou 1998).

Noggin was first identified as a neural inducer based on its ability to induce neural tissue without mesodermal contamination when either injected as mRNA or provided in protein form (Lamb et al. 1993), and it is expressed zygotically at late blastula stages in the region that will form the organizer (Smith and Harland 1992). Follistatin has also been defined as a neural inducer given its expression at the correct time and place in the embryo and its ability to induce neural tissue absent mesoderm contamination when injected as mRNA (Hemmati-Brivanlou et al. 1994). A third neural inducer, Chordin, also induces neural tissue directly either by mRNA injection or when provided as protein and is also enriched in the organizer at the appropriate developmental stage (Sasai et al. 1995). Cerberus is able to induce neural tissue, though it also induces mesoderm, suggesting the neural induction is not direct (Bouwmeester et al. 1996).

While ectodermal cells in an embryo will form epidermis in the absence of neural inducing cues, it has been shown repeatedly that dissociated ectodermal cells from the animal pole, not exposed to the dorsal lip, will form neural, not epidermal cells, demonstrating that neural tissue could indeed be formed in the absence of the aforementioned neural inducers (Sato and Sargent 1989, Godsave and Slack 1989, Grunz and Tacke 1989). Additionally, neural rather than epidermal tissue was induced when a truncated form of the activin receptor was injected into embryos (Hemmati-Brivanlou and Melton 1994). These studies, among others, contradicted the idea that neural tissue had to be induced, and led to the formation of the neural default hypothesis, which suggests that neural is the default state of a cell and in the absence of any external signals ectodermal derivatives will form neural progenitors, not epidermis, and BMP is required to override this neural default state for epidermal formation (Hawley et al. 1995).



**Figure 1.4 BMP Signaling is Crucial for Ectodermal Differentiation.** Schematic showing how neural inducers block BMP signaling and downstream transcription factors activated in presence or absence of BMP signaling to induce epidermal or neural fates respectively (Adapted from Sasai 1998).



**Figure 1.5 Ectodermal Explants will form Neural Progenitors in Absence of BMP.** Schematic showing that upon dissection, intact animal caps will form epidermis, as will dissociated cells treated with BMP. Dissociated animal cap cells absent BMP, and animal caps cells exposed to organizer signals will form neural progenitor tissues. Neural inducing signals emanating from the dorsal lip of blastopore all serve as inhibitors of BMP signaling (Adapted from Chang and Brivanlou 1998).

In line with the idea of neural being the default state of the cell is the fact that all neural inducers are in fact inhibitors of BMP signaling. Noggin, Chordin and Follistatin all bind directly to at least one BMP ligand, thus preventing BMP signaling by preventing BMP from binding to its receptor (Zimmerman et al. 1996, Piccolo et al. 1996, Fainsod et al. 1997). This suggested that BMP signaling, which activates endogenously just prior to gastrulation in the embryo, is required to induce the epidermal fate, and in its absence the ectoderm forms neural progenitors by default. This has been validated by the fact that BMP4 is able to induce an epidermal fate in dissociated cells that would otherwise neuralize (Wilson and Hemmati-Brivanlou 1995). In fact, BMP4 is a morphogen critical for ectodermal patterning. At high doses, BMP induces epidermis, in its absence neural progenitors form, and intermediate levels lead to cement gland formation (Wilson et al. 1997). Epidermal specification is further promoted by expression of BMP targets, *Vent* genes, *Msx1* and *Gata1*, which also act to ventralize ectodermal tissue (Ault et al. 1997, Suzuki et al. 1997a, Xu et al. 1997, Sasai 1998). Also important in the transit from pluripotency to the epidermal fate is activation of the epidermal marker *Foxi1* by maternally provided Fox family member *Foxi2* (Cha et al. 2012, Mir et al. 2007). The epidermis is further specified by the activation of epidermal markers *Grhl1* by BMP signaling, which then directly activates *Epk* (Tao et al. 2005, Jonas et al 1985). On the other hand, *Zic* factors function downstream of neural inducers to promote a neural fate in ectodermal cells (Mizuseki et al. 1998a, Nakata et al. 1997, 1998). *Sox2* also plays a role in enhancing neuralization of cells, as does *Sox15*, specifically in anterior tissue (Mizuseki et al. 1998a,b). These transcription factors together play a crucial role in specifying ectodermal cells downstream of BMP signaling (Figure 1.4).

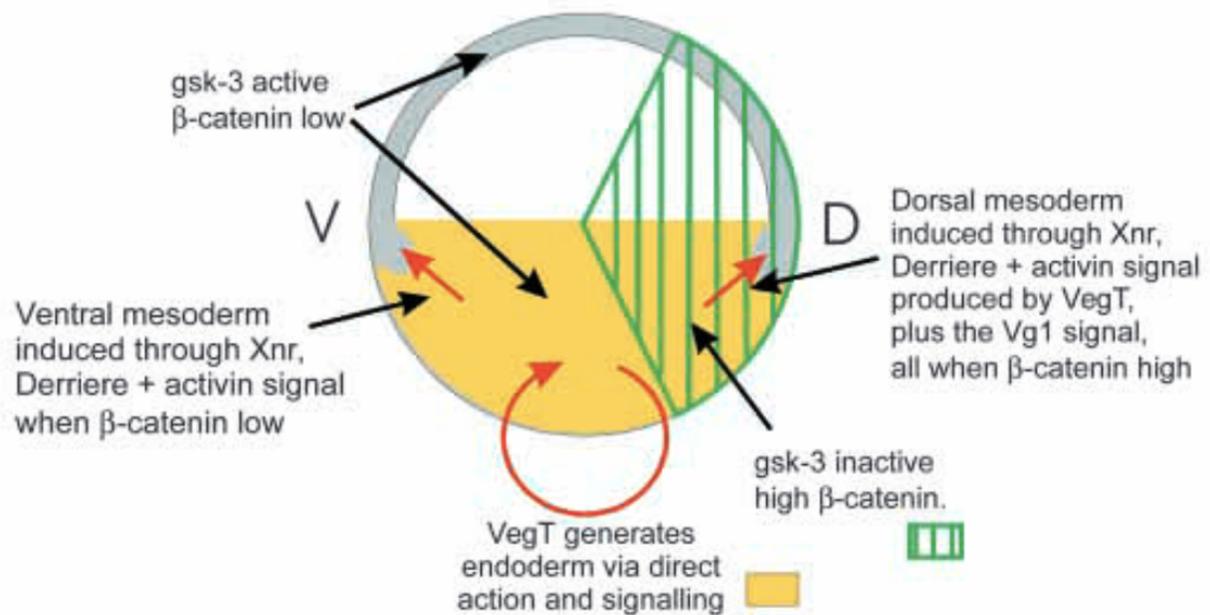
While there is significant experimental evidence in *Xenopus* to back the neural default model, in that ectodermal cells default to neural in the absence of BMP signaling, either due to dissociation of cells or the presence of BMP antagonists in intact tissue (Figure 1.5), this model is not as well supported in other model systems. In chick embryos, for example, the expression patterns of BMPs and their antagonists do not fully fit the neural default model (Streit et al. 1998), and misexpression of BMP antagonists does not induce neural progenitors, nor does ectopic expression of BMP inhibit neural plate formation in this system (Streit and Stern 1999a, Streit and Stern 1999b, Stern 2005). There is even some evidence in *Xenopus* embryos that in some cases FGF signaling is required for competence to

respond to BMP antagonists that will lead to neural induction (Linker and Stern 2004, Delaune et al. 2005, Wawersik et al. 2005), though the roles of FGF and BMP signaling in neural induction have been shown to be separable (Wills et al. 2010).

### **Mesendoderm**

On the other side of Waddington's landscape from the ectoderm are the two mesendoderm germ layers, mesoderm and endoderm. While the ectoderm arises from the animal pole cells in vertebrates, both the mesoderm and endoderm arise from cells near the vegetal margin in a mesendoderm layer, which then further differentiates into mesoderm and endoderm (Warga and Nusslein-Volhard 1999, Rodaway and Patient 2001). Evidence for bipotent mesendoderm cells has also been demonstrated in human cell lines, marked by the presence of *Gooseoid*, *E-cadherin* and *PDGF* (Tada et al. 2005). A shared mesendoderm layer of cells prior to gastrulation is evidenced specifically in *Xenopus* with grafting experiments that show tissue from the superficial marginal zone giving rise to both endoderm and mesoderm (Minsuk and Keller 1997).

While the mesoderm and endoderm germ layers differentiate early in development, they are both generated by overlapping signaling pathways and transcription factors. The maternally provided transcription factor *VegT* is critical for activating zygotic mesoderm and endoderm inducing signals (Zhang et al. 1998, Kimelman and Griffin 2000). The mechanism by which *VegT* induces mesoderm and endoderm is slightly different, as *VegT* induces endoderm both in a cell autonomous manner, as well as by activating TGF-beta signaling that contributes to endoderm formation and maintenance, whereas mesoderm is induced only by *VegT*-induced TGF-beta signaling and not in a cell-autonomous manner (Clements et al. 1999). This is supported by the fact that maternally provided *VegT* is concentrated in the vegetal pole fated to become endoderm and not the equatorial cells that will become mesoderm (Stennard et al. 1999). *Eomes* also has overlapping functions with *VegT* in early mesoderm specification (Ryan et al. 1996, Gentsch et al. 2013). TGF-beta signaling molecules that comprise the mesoderm-inducing signal include the nodal homologs *Nodal*, *Nodal1*, and *Nodal2*, *ActivinB*, and *Gdf3*, with *Gdf3* playing a role specifically in posterior mesoderm (Figure 1.6) (Sun et al. 1999, Clements et al. 1999,

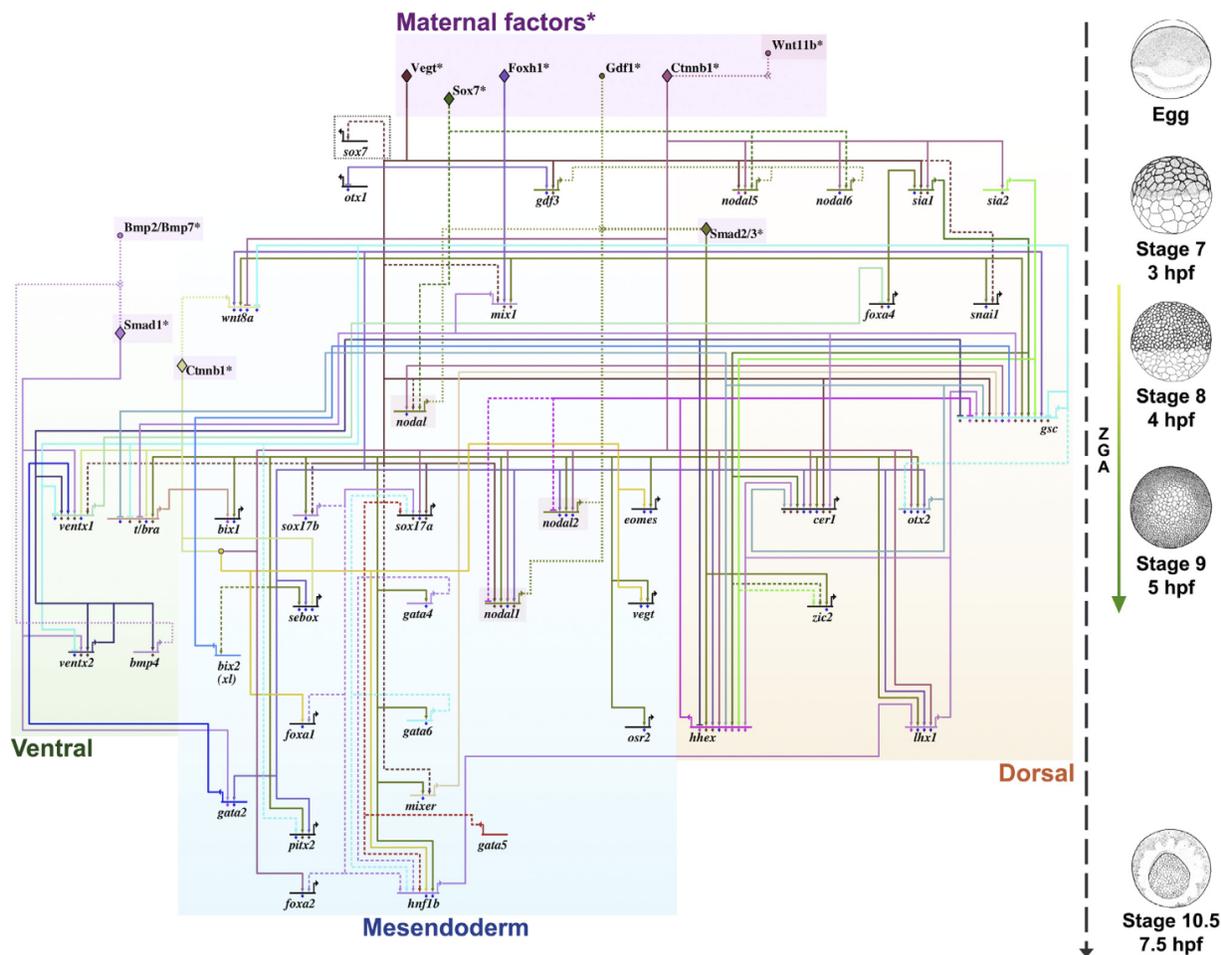


**Figure 1.6 Role of Maternal VegT in Mesendoderm Formation.** Schematic representing the cell autonomous formation of endoderm induced by VegT as well as initiation of TGF-beta signaling leading to endoderm and mesoderm formation (Adapted from Clements et al. 1999).

Kofron et al. 1999). In addition to targeting signaling molecules, maternal *VegT* also activates *Sox7*, and both maternal and zygotic *Sox7* induce *Nodal*, *Nodal1*, *Nodal2*, *Nodal5* and *Nodal6* as well as *Mixer*, *Sox17* and *Endodermin* (Zhang et al. 2005). *Gooseoid* is another critical transcription factor targeted by *VegT* and inducible synergistically by Activin, and Wnt8, leading to organizer formation (Watabe et al. 1995).

*Nodal* homologs play an important role not just in inducing mesoderm and endoderm, but also in initiation of gastrulation and in setting up Spemann's organizer. While many *Xenopus* nodal-like genes have been identified, they play nuanced roles in germ layer formation. *Nodal1*, *Nodal2* and *Nodal* are expressed in a dorsal to ventral gradient in endoderm cells and have mesoderm inducing capability. These three genes also play a role indirectly in forming neural tissue by inducing *Cerberus*, which then can inhibit *Nodal1* in a negative feedback loop, whereas *Nodal3* induces neural tissue directly (Agius et al. 2000, Osada and Wright 1999, Piccolo et al. 1999, Tavares et al. 2007, Hansen et al. 1997). *Nodal* genes then activate FGF4 further downstream which is not critical for mesoderm formation, but does play an important role in dorsal mesoderm maintenance, particularly in Brachyury regulation (Isaacs et al. 1994, Christen and Slack 1999). In addition to maternally provided transcription factors and signaling molecules, there are also several groups of zygotically activated transcription factors that are critical for mesendoderm formation, including the *Mix* family, *Foxa* family, *Gata* family and *Sox17* (Rosa 1989, Mead et al. 1996, Suri et al. 2004, Murgan et al. 2014, Chiu et al. 2014, Lemaire et al. 1998, Henry and Melton 1998, Tada et al. 1998, Periera et al. 2012, Weber et al. 2000, Afouda et al. 2005, Hudson et al. 1997).

One of the most complete ways of visualizing mesendoderm formation is through gene regulatory networks that have been compiled based on functional experiments with many individual genes. Gene regulatory networks (GRNs) are comprised primarily of transcription factors, which can generate feedback loops, and signaling molecules. These genes work to modulate when genes are turned on and off, as well as their rate of transcription, both temporally and spatially, which is critical for proper germ layer formation. The inputs and outputs of the network are regulated by cis-regulatory modules, combinations of regulatory elements and primarily enhancers, to which transcription factors bind and then recruit cofactors (Figure 1.7). From these GRNs, several network motifs have been identified, including

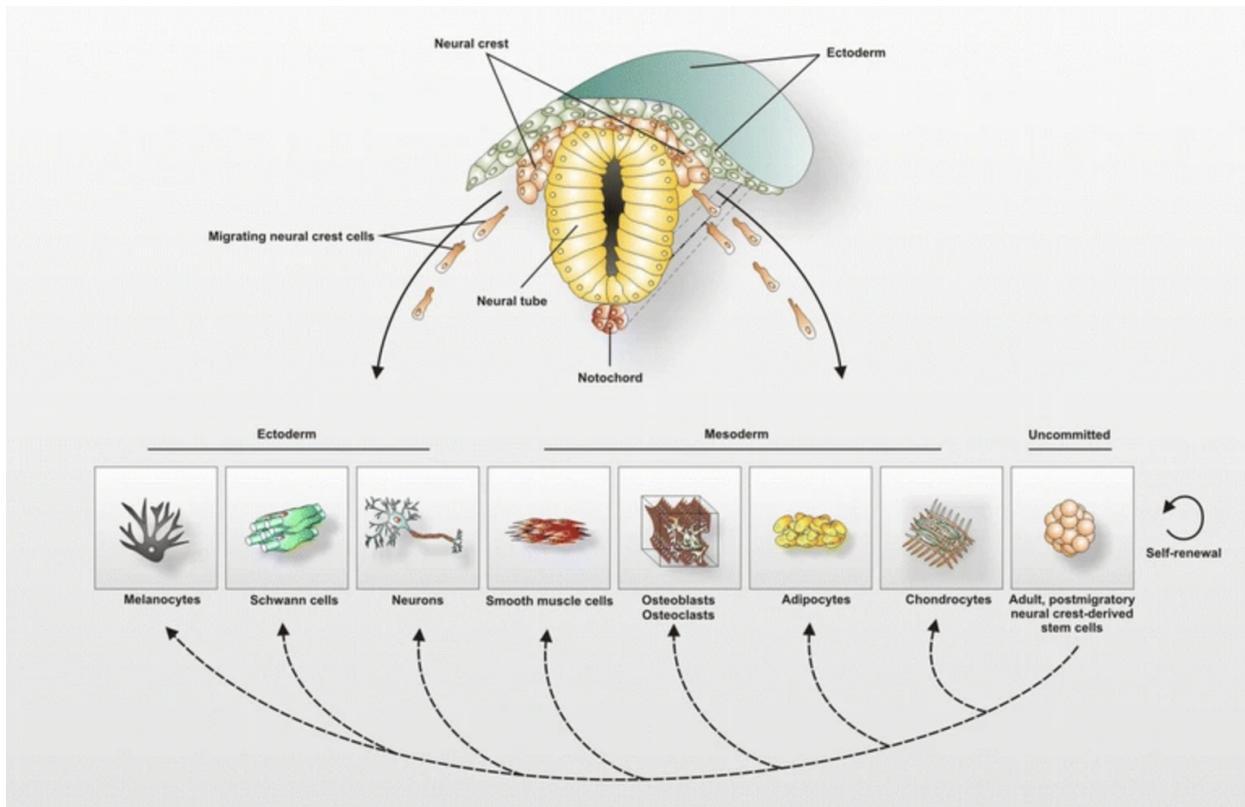


**Figure 1.7 Mesendoderm Gene Regulatory Network.** Map of known mesendoderm gene regulatory network members with their direct (solid lines) and putative (dashed lines) targets including approximate positioning on dorsal-ventral axis and time of onset by developmental stage. Diamonds represent maternal proteins and circles represent signaling ligands. (Adapted from Charney et al. 2017).

autoregulatory loops, feedback loops and feedforward loops (Charney et al. 2017). There are four genes that self-regulate during endoderm formation, *Nodal* and *Ventx2*, which both positively autoregulate, and *Wnt8* and *Gsc* that negatively autoregulate (Osada et al. 2000, Henningfield et al. 2002, Nakamura et al. 2016, Danilov et al. 1998). There are also double negative feedback loops between two genes that help control lineage boundaries, such as that between *Ventx2* and *Gsc* (Onichtchouk et al. 1996, Trindade et al. 1999) and positive feedback loops that allow for multiple genes to be expressed simultaneously in the same lineage such as between *BMP4* and *Ventx2* (Schuler-Metz 2000). Finally, feed forward loops can facilitate the transition between regulation by maternal factors to a primary wave of zygotic factors followed by a secondary wave of zygotic factors if they are positive or help define regionalization if they are negative (Charney et al. 2017). While the authenticity of the *Xenopus* mesendoderm network is high, given all the functional experiments done, the completeness is still relatively low, and large strides can be made regarding completeness with information provided by genomic sequencing on the transcript level of every gene under different conditions. Improving the completeness of GRNs will not only enhance the current understanding of embryology, but also, through comparison of GRNs, enhance the understanding of the evolution of body plans and the control of developmental processes, underscoring the importance of both genomic and functional studies (Davidson et al. 2005).

### **Neural Crest as the Fourth Germ Layer**

While most cell types become restricted to a specific cell type during the formation of primary germ layers upon gastrulation, one cell type, the neural crest, maintains multipotency late into development. These cells are unique to vertebrates and are sometimes called the “fourth germ layer” (Hall 2000). Wilhelm His first discovered the neural crest, identifying a band of cells he termed the *zwischenstrang* positioned between the neural tube and prospective epidermis (His 1868, Hall 2018). The neural crest arises from the neural plate border, in between the neural and non-neural ectoderm, which undergoes unique patterning to distinguish neural crest, as well as cranial placodes, from other ectodermal tissue (Basch et al. 2000, Pieper et al. 2012, Groves and LaBonne 2014). The neural crest is unique in its ability to form both ectodermal and mesodermal derivatives, including melanocytes,

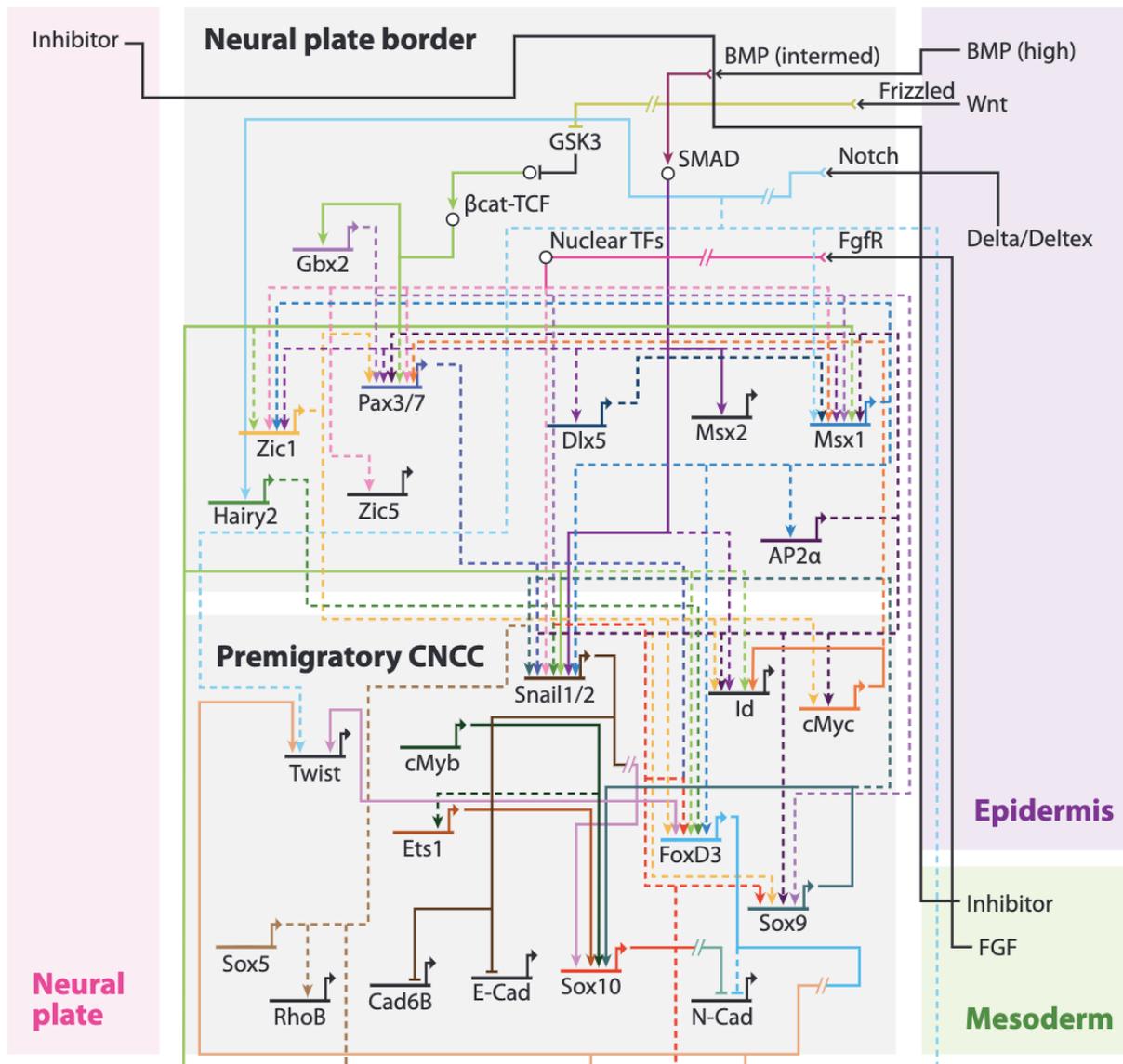


**Figure 1.8 Neural Crest Cells yield Ectodermal and Mesodermal Derivatives.** Neural crest cells are a population of cells unique to vertebrates that maintain their potency late into development and give rise to both ectodermal and mesodermal derivatives as well as maintain their ability to self-renew (Adapted from Kaltschmidt et al. 2012).

peripheral nervous system, and craniofacial structures, as well as to self-renew and is thus critical for vertebrate development (Figure 1.8) (Baroffio et al. 1988, Bronner-Fraser and Fraser 1988, Trentin et al. 2004, Morrison et al. 1999, Kaltschmidt et al. 2012). Importantly, neural crest cells are multipotent not just at the population level, but at the single cell level. The progeny of a single neural crest cell have the potential to give rise to numerous distinct derivatives, including glia, neurons, melanocytes, myofibroblasts, chondrocytes and osteocytes (Bronner-Fraser and Fraser 1989, Collazo et al. 1993, Baggiolini et al. 2015, Dupin and Sommer 2012).

Most literature in the neural crest field provides evidence for a regaining of potential upon formation of the neural plate border and neural crest. The LaBonne lab proposed an alternate model in which potency is maintained rather than regained (Buitrago-Delgado et al. 2015). This model is based upon the fact that neural crest induced animal caps are able to form endoderm, in addition to mesoderm and ectoderm derivatives. This model is also predicated on the fact that *in situ* hybridization experiments suggest expression of the neural crest marker *Foxd3* (Sasai et al. 2001) as early as blastula stages (Buitrago-Delgado et al. 2015). Multiple RNASeq datasets generated in the lab, however, suggest there is limited quantitative backing for this *Foxd3* expression. While robust at neural crest stages in *Wnt/Chrd* induced neural crest caps, *Foxd3* does not exceed 6 transcripts per million (TPM) at blastula stages, which is likely not high enough to be biologically relevant. Other single cell sequencing experiments have called into question the pluripotency maintenance model due to lack of quantitative evidence of an intermediate multipotent cell population based on single cell expression analysis, favoring instead the classical model of reacquisition of cellular potential (Briggs et al. 2018, Zalc et al. 2021). Thus, while compelling, the retention model of pluripotency is still up for debate. Indeed, just last year a third model was proposed, of cyclical fate restriction of neural crest, suggesting this process of lineage restriction may be more dynamic than initially thought, perhaps reconciling the retention and regaining models of multipotency (Kelsh et al. 2021).

Whether the multipotency and stem cell-like qualities of the neural crest are maintained, regained,

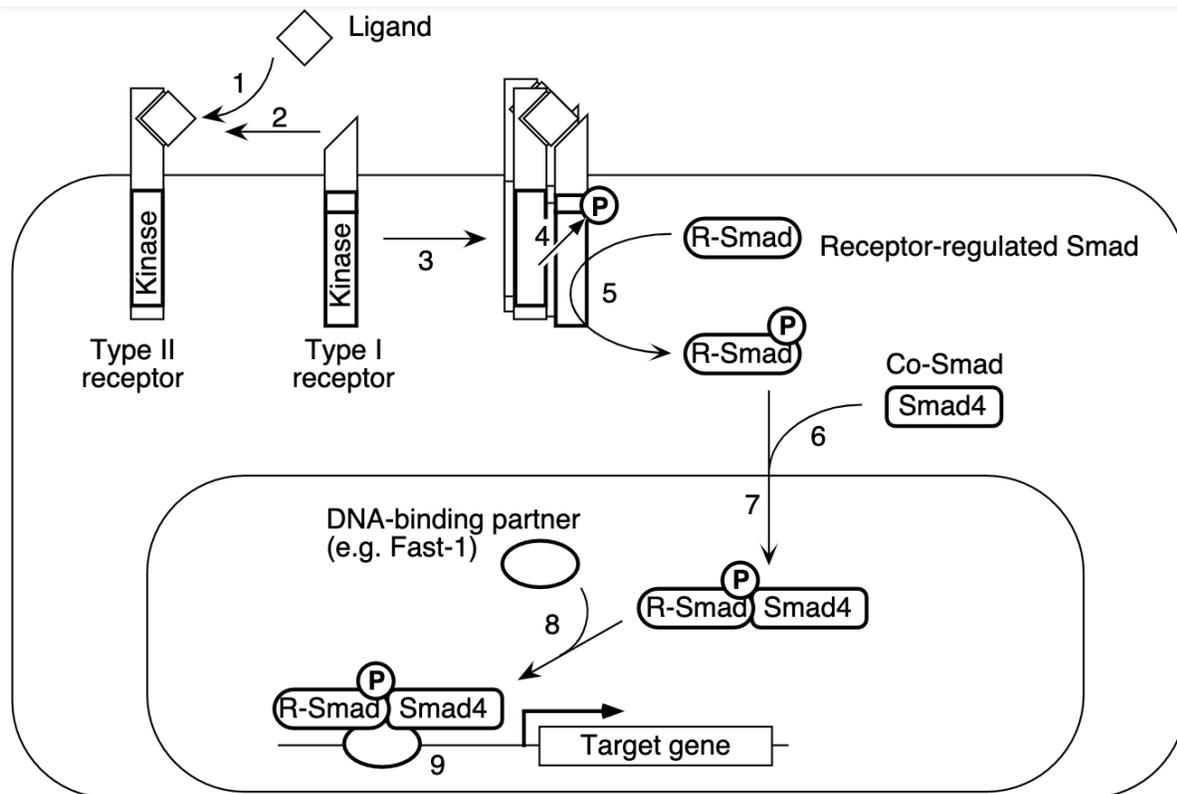


**Figure 1.9 Neural Crest GRN.** A gene regulatory network schematic for neural crest formation encompassing many functional experiments that shows the contributions of different signaling pathways and surrounding tissues that lead to the appropriate levels of signaling and transcription factors leading to neural crest formation (Adapted from Betancur et al. 2010.)

or cyclical, this potential is unique and the method by which these cells either retain or regain this multipotency is still not fully understood. A number of signaling molecules and transcription factors contribute to a gene regulatory network that regulates the induction and maintenance of neural crest cells. Intermediate BMP signaling levels are essential for neural crest specification. Moderate BMP levels, versus the high levels of BMP that specify epidermis, are achieved by inhibition of BMP by the BMP antagonists Chordin, Noggin and Follistatin. These BMP antagonists cooperate with Fibroblast Growth Factors (FGF) to induce neural crest (Monsoro-Burq et al. 2003, 2005, Geary and LaBonne 2018). Upstream of BMP, Delta-Notch signaling contributes to the formation of the neural crest and its restriction to the neural plate border (Glavic et al. 2004). Finally, the canonical Wnt signaling pathway leads to the stabilization and accumulation of beta-catenin, allowing it to interact with the Tcf/LEF family of transcription factors that upregulate downstream target genes (Saint-Jeannet et al. 1997, Wu et al. 2005). These three primary signaling pathways, BMP, FGF and Wnt, induce *Msx1*, *Dlx5*, *Gbx2*, *Pax3* and *Zic1* (LaBonne and Bronner-Fraser 1998, Feledy et al. 1999, Luo et al. 2001, Li et al. 2009, Tribulo et al. 2003, Hong and Saint-Jeannet 2007), which in turn regulate the expression of key neural crest specifier genes, *Ap2*, *Myc*, *Snail1*, *Snail2*, *Twist*, *Sox8*, *Sox9*, *Sox10*, *Foxd3* and *Id3* (Luo et al. 2003, Bellmeyer et al. 2003, Mayor et al. 1995, Hopwood et al. 1989, O'Donnell et al. 2006, Spokony et al. 2002, Aoki et al. 2003, Honoré et al. 2003, Sasai et al. 2001, Light et al. 2005, Kee and Bronner-Fraser 2005, Sauka-Spengler and Bronner-Fraser 2008). Interactions between these genes can be depicted in a gene regulatory network (Figure 1.9) and are responsible for promoting neural crest fate over neural and epidermal fates (Betancur et al. 2010).

### **TGF-Beta Signaling and its Role in Development**

TGF-beta superfamily signaling plays a critical role in establishing germ layers, with the two different branches playing different roles in each lineage. The BMP signaling branch is primarily responsible for epidermal formation, whereas the Activin / Nodal branch is primarily responsible for mesendoderm formation, with BMP playing an important role in ventralization in mesoderm. The two branches act through similar mechanisms, with the differing receptors and downstream targets leading to



**Figure 1.10 TGF-Beta Signaling.** Schematic of TGF-beta superfamily signaling demonstrating binding of ligand to type II receptor, allowing for receptor complex formation and phosphorylation of type I receptor leading to phosphorylation of receptor Smad which binds to co-Smad and is translocated into the nucleus leading the transcription of target genes (Adapted from Masagué 1998).

varying transcriptional outcomes. The general mechanics of TGF-beta signaling are binding of a ligand to Type I and Type II receptors, forming a receptor complex. This leads to the phosphorylation of the Type I receptor, which in turn phosphorylates the receptor Smad (R-Smad), which then binds to the coSmad, Smad4. The Smad complex is translocated to the nucleus where it binds to a DNA binding partner and then binds to enhancers of target genes leading to transcription (Figure 1.10) (Massagué 1998).

The ligands in TGF-beta signaling typically activate the signaling pathway as dimers, with the most potent inducer of BMP signaling being the BMP4/7 heterodimer (Aono et al. 1995, Suzuki 1997b, Nishimatsu 1998, Little and Mullins 2009). Several different dimers can activate the Activin / Nodal branch of TGF-beta signaling including between Gdf3 and Nodal, ActivinA/B, and between different Nodal proteins (Eimon and Harland 2002, Massagué 1987, Thomsen et al. 1990, Osada and Wright 1999). The manner in which these ligands bind to receptors differs slightly between the two branches of the TGF-beta superfamily. BMP ligands bind to type I and II receptors in a cooperative manner, having high affinity only to both receptors simultaneously, but not to either type I or II individually (Liu et al. 1995, Nohno et al. 1995). Activin, on the other hand, binds first to the type II receptor and then the ligand-type II receptor complex binds with type I receptors (Attisano et al. 1993, Ebner et al. 1993). For both types of receptor binding, once the ligand-receptor complex is formed the type II receptor phosphorylates the type I receptor which then phosphorylates the relevant receptor Smad proteins (Okadome et al. 1994). In the case of BMP signaling, Smads 1,5 and 8 are phosphorylated and in the case of Activin / Nodal it is Smads 2 and 3 (Thomsen 1996, Kretzschmar 1997, Graff et al. 1996, Suzuki et al. 1997c, Chen et al. 1997, Baker and Harland 1996, Macías-Silva et al. 1996). Any of these receptor Smads will bind to co-Smad Smad4 to be translocated into the nucleus (Lagna et al. 1996, Liu et al. 1996). A third group of Smads, inhibitory Smads, can disrupt TGF-beta signaling, Smad7 inhibits both Activin and BMP signaling, whereas Smad6 specifically inhibits BMP signaling (Hayashi et al. 1997, Nakao et al. 1997, Imamura et al. 1997, Hata et al. 1998).

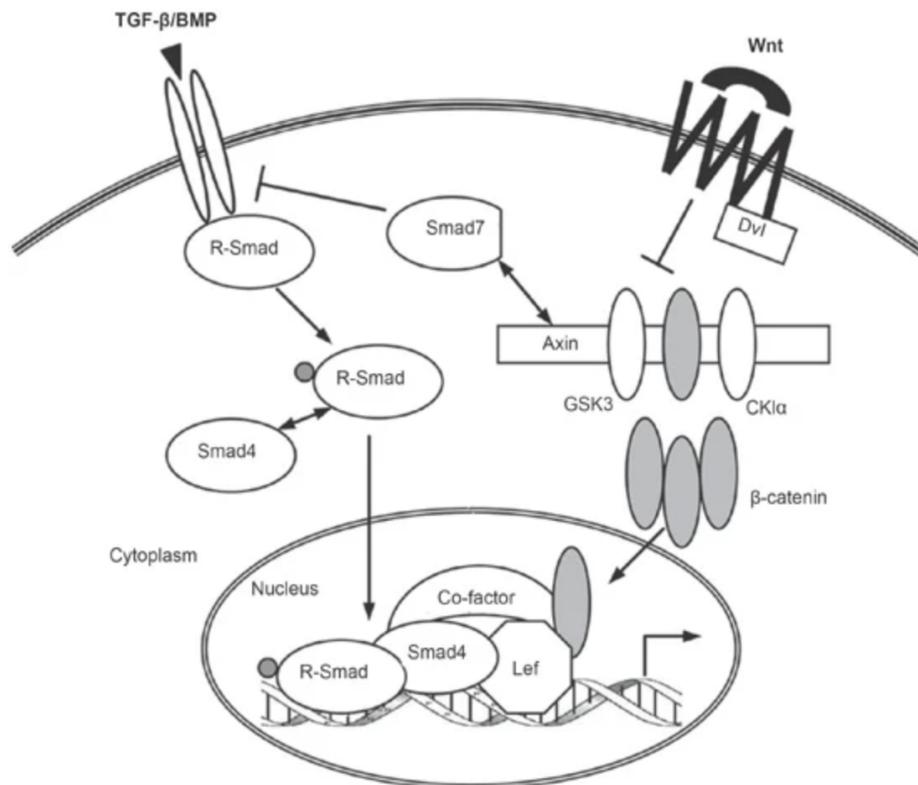
Once in the nucleus, the receptor Smad / co-Smad complex binds to DNA binding partners, Foxh1 in the case of Smad2 and Znf423 and Runx2 in the case of Smad1. Zeb2 can bind with both (Chen X et al. 1996, 1997, Watanabe and Whitman 1999, Yeo et al. 1999, Hata et al. 2000, Zhang et al. 2000,

Drissi et al. 2003, Wang et al. 2007, Verschuere et al. 1999, Blitz and Cho 1999). These complexes then go on to activate the transcription factors enumerated in previous sections. The activation of both Smad1/5/8 and Smad2/3 occur after the onset of zygotic transcription, but only Smad2/3 regulation is dependent on that transcription (Faure et al. 2000). Precise spatial and temporal regulation of both branches of the TGF-beta pathway are imperative for proper embryo patterning and germ layer formation.

### **FGF and Wnt Signaling Pathways**

While this thesis work primarily focuses on the role of TGF-beta signaling for germ layer formation, there are several other signaling pathways that interact with TGF-beta to aid in germ layer formation, patterning, and axis formation. Fibroblast Growth Factor (FGF) signaling is among the earliest signaling pathways to be activated during embryonic development and plays an important role in the maintenance of pluripotency as well as neural and neural crest induction. FGF signaling is required for proper gene expression in animal pole cells at the pluripotent state as well as for normal state transitions. Specifically, FGF mediated Map Kinase activation is required at high levels during the pluripotent state but decreases upon lineage restriction, as FGF mediated PI3 Kinase/AKT increases. MapK is required in animal cap assays for proper epidermis, mesoderm and endoderm formation and PI3K/AKT is required for proper neural and mesodermal formation (Geary and LaBonne 2018). FGF signaling has been shown to be important for formation of the neural fate, as dominant negative FGF receptors blocks neural induction by Chordin and Noggin (Launay et al. 1996) and bFGF can induce neural induction in animal caps primed with low calcium and magnesium media (Kengaku and Okamoto 1995, Lamb and Harland 1995). Ectopic FGF8 has also been shown to stimulate neural induction in whole embryos (Hardcastle et al. 2000). FGF has also been shown to play an important role in mesoderm formation. It can induce mesoderm in animal caps (Slack et al. 1987) and FGF has been shown to be required for mesodermal formation and patterning in embryos (Amaya et al. 1991) specifically through FGF mediated MapK (LaBonne et al. 1995, Umbhauer et al. 1995).

Wnt signaling is another signaling pathway critical in early embryonic development for proper germ layer formation. Wnt signaling is critical for dorsal ventral axis formation, the asymmetry for which is



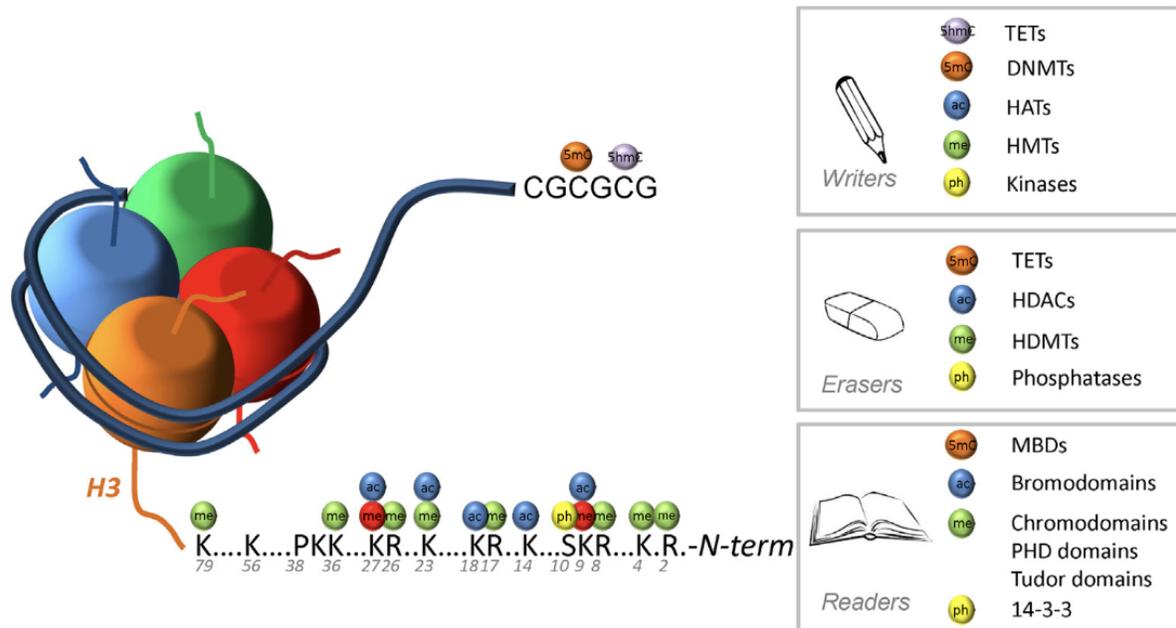
**Figure 1.11 Wnt and TGF-Beta Crosstalk.** Schematic displaying how Wnt and TGF-beta pathways crosstalk by forming a complex with Smad4 to regulate axis formation (Adapted from Guo and Wang 2009).

established upon cortical rotation of a *Xenopus* embryo after fertilization (Vincent and Gerhart 1987). During cortical rotation, maternally deposited transcripts in the vegetal pole are translocated to the dorsal side of the embryo, among which is Beta-catenin, an intracellular transducer of Wnt signaling (Elinson and Rowning 1988, Browning et al. 1997). Wnt5 and Wnt11 function as a homodimer to synergistically mediate the early dorsalizing signal through Frizzled-7 (Schroeder 1999, Sumanas 2000, Cha et al. 2008, 2009). Translocation of Beta-catenin and Dishevelled into the nucleus is indicative of active Wnt signaling and leads to the expression of dorsal specific genes (Schneider et al. 1996, Miller et al. 1999).

This maternal Wnt/Beta-catenin pathway plays a role in triggering organizer activity and inducing neural fate by initiating transcription of *Sia* and *Twin*, which repress BMP and induce BMP inhibitors (Brannon and Kimelman 1996, Carnac et al. 1996, Laurent et al. 1997, Baker et al. 1999, Wessely et al. 2001, Kuroda et al. 2004). Initiation of *Sia* and *Twin* is done synergistically with TGF-beta signaling as Beta-catenin and Lef1/Tct, downstream components of the Wnt signaling cascade, form a complex with Smad4 (Figure 1.11) (Nishita et al. 2000, Guo and Wang 2009). After initial neural induction, Wnt signaling also works cooperatively with FGF on anteroposterior neural patterning (McGrew et al. 1997). In addition to neuralizing activities, Wnt can also dorsalize mesoderm as a target of Xbra through Dishevelled, rather than the canonical Wnt pathway (Saka et al. 2000, Tada and Smith 2000). Non-canonical Wnt signaling, through the gene *Dishevelled*, plays an important role during axis formation in the convergent extension of both mesoderm and neural cells (Sokol 1996, Wallingford et al. 2000, Wallingford and Harland 2001). Wnt8 also cooperatively patterns mesoderm with BMP2/4, by which it is regulated (Hoppler and Moon 1998, Guo and Wang 2009).

### **Epigenetics in Lineage Restriction: An Overview**

In addition to the signaling pathways and transcription factors that orchestrate germ layer formation, numerous post translational modifications dictate the accessibility of DNA and thus play an important role in exiting pluripotency and lineage formation. Over developmental time the potential of cells is increasingly restricted due to progressive gene silencing while specific transcription factors are turned on to bias cells towards specific fates. This gene silencing is regulated in large part by epigenetics,



**Figure 1.12 Epigenetic Regulation of DNA and Histones.** Schematic showing methylation of DNA CpG islands as well as the locations of lysine and arginine methylation and acetylation on the histone tail, and which readers, writers, and erases regulate these different epigenetic marks (Adapted from Hu et al. 2014a).

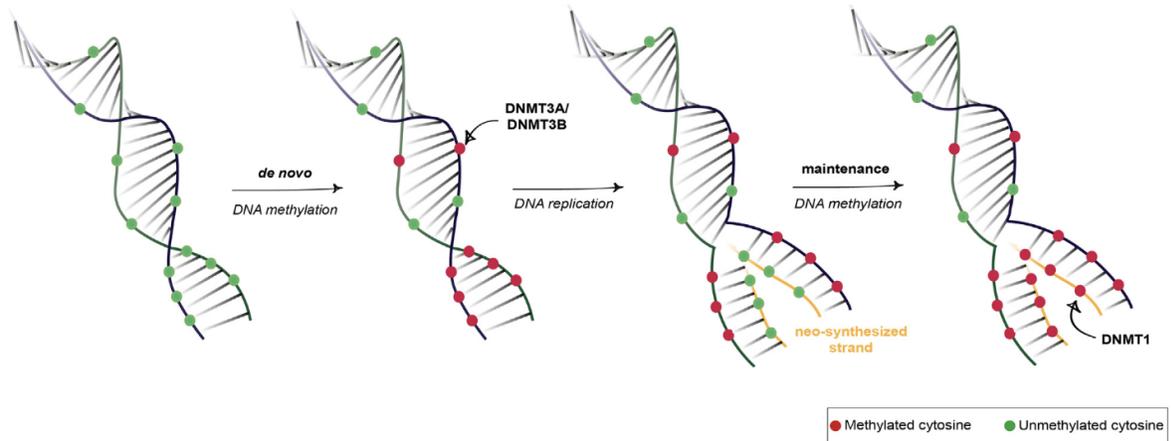
through modifications of both chromatin structure and DNA methylation (Lunyak and Rosenfeld 2008). Epigenetic regulation is a complex dynamic process by which epigenetic writers put down epigenetic marks, typically acetylation or methylation markers on DNA or histone tails. Proteins with bromodomains, chromodomains or tudor domains then bind to these epigenetic marks and serve as readers. These marks can then be removed by epigenetic erasers, typically histone deacetylases, lysine demethylases and phosphatases (Hu et al. 2014a, Falkenberg and Johnstone 2014) (Figure 1.12).

When chromatin is more loosely packed DNA is more accessible and genes are in their active state. Chromatin in pluripotent cells is globally decondensed in the form of euchromatin and becomes increasingly condensed into heterochromatin during differentiation. Embryonic stem cells are enriched for chromatin architectural proteins (Kurisaki et al. 2005). These proteins bind loosely to chromatin of embryonic stem cells and are highly dynamic, as cells undergo lineage restriction the dynamic nature of these proteins is lost (Meshorer et al. 2006). Histones are acetylated specifically on their lysine residues. Acetylation of these lysine residues neutralizes their charge and thus weakens their bond with negatively charged DNA in turn opening up histone structure and increasing DNA accessibility (Hong et al 1993). Consequently, acetylation is known to be an activating mark (Brownell et al. 1996). Histone acetylation levels are regulated by histone acetyltransferases (HATs) which add acetyl groups to lysine, and histone deacetylases (HDACs) which remove these acetyl groups, returning chromatin to its closed state. H3K27ac in particular marks active enhancer elements (Creyghton et al. 2010) and a shift from enrichment of H3K27ac on genes associated with pluripotency to those associated with lineage transcription corresponds to the onset of lineage restriction in zebrafish (Bogdanovic et al. 2012). Other acetylation marks, H3K9ac and H3K14ac are also enriched in pluripotent stem cells (Efroni et al. 2008). Finally H3K56 is enriched in human embryonic stem cells, specifically at the promoters of the key pluripotency regulators, Nanog, Oct4 and Sox2 (Xie et al. 2009). Due to the necessity of acetylation for an open chromatin state, HATs are required for the maintenance of pluripotency (Fang et al. 2014). Interestingly, HDACS, associated with gene repression, have also been shown to be required for the maintenance of pluripotency and formation of the multipotent neural crest cell population (Jamaladdin et al. 2014, Rao and Labonne 2018). Seemingly contradictorily, HDACs have also been shown to be

required for differentiation of embryonic cells and neural progenitor cells (Lee et al. 2004, Dovey et al. 2014, Hsieh 2004), with low levels of HDAC inhibitors increasing pluripotency of embryonic stem cells (Hezroni et al. 2011). HDACs likely have such a complex role in gene regulation due to the temporal and spatial shifts of acetylated histones necessary for proper embryonic development.

The nature of histone methylation is even more complicated than that of acetylation because methylation of histone can be either active or repressive depending on the position of the methyl group. H3K4Me3 is an activating mark deposited by the histone methyltransferase trithorax protein complexes, such as MLL (Nagy et al. 2002, Guenther et al. 2005, Wysocka et al. 2005). H3K4Me3 is typically found in the promoter region of genes, and its onset coincides with the zygotic genome activation and even acts to poise some genes for transcription prior to the maternal zygotic transition (Strahl et al. 1999, Santos-Rosa et al. 2002, Bernstein et al. 2005, Akkers et al. 2009, Blythe et al. 2010). H3K4Me is also crucial for inheritance of cellular memory of an active state (Ng and Gurdon 2008). Another activating mark, H3K36me3, by contrast, is deposited by the histone methyltransferase Set2 and found primarily in the coding region (Strahl et al. 2002, Pokholok et al. 2005). H3K27me3 has the opposite effect, being a transcriptional repressor, and is typically deposited by the polycomb repressor complex near the transcription start site of the same genes of H3K4Me3, added later in development as genes are repressed spatially during lineage restriction, allowing genes to rapidly decrease in expression level as cells become more specialized (Cao et al. 2002, Lee et al. 2006, Tolhuis et al. 2006, Pan et al. 2007, Akkers et al. 2009). H3K9Me2/3 is another silencing histone methylation mark deposited in a gene's promoter region by Suv39h1 and Ehmt2 and is important for proper lineage specification including differentiating between neural and neural crest cells, specifically through its removal by histone demethylase Jmjd2A (Bannister et al. 2001, Tachibana et al 2002, Schultz et al. 2002, Barski et al. 2007, Strobl-Mazula et al. 2010, Hu et al. 2014a). H3K9Me also has a strong association with transcriptionally repressive methylated DNA (Lenhertz et al. 2003, Fuks et al. 2003, Rothbart et al. 2012, Liu et al. 2013).

Importantly, these methylation and acetylation marks work together to establish chromatin accessibility. In mouse ESCs it has been shown that genes active at the pluripotent state were marked with H3KMe1 and H3K27ac, whereas poised genes needed for early development at the onset of the exit



**Figure 1.13 DNA Methylation Pattern Established and Maintained by DNMTs.** DNMT3a and 3b are responsible for de novo DNA methylation and DNMT1 maintains the DNA methylation pattern upon DNA replication (Adapted from Pechalrieu et. al 2017).

from pluripotency had only the H3KMe1 mark (Creyghton et al. 2010). Similarly, in human embryonic stem cells it was shown that active enhancers with open chromatin were marked with both H3K4Me1 and H3K27ac, whereas enhancers that are not active at pluripotent stages, but rather poised to play a role in turning genes on at the earliest stages of lineage restriction are marked with H3K4Me1 and H3K27Me3 (Rada-Iglesias et al. 2011). This same phenomenon is seen in zebrafish, with H3K27ac being enriched first in pluripotent factors, and a shift in H3K27 occupancy accompanying a shift from pluripotency to expression of lineage restricted genes as the poised state of chromatin is exited and it becomes active (Bogdanovic et al. 2012). In *Xenopus*, this signature of pluripotent enhancers H3K4Me1 and H3K27ac is established by blastula stages and likely maternally provided (Gupta et al. 2014). The fact that this process is so highly conserved emphasizes the importance of chromatin remodeling in the maintenance of pluripotency and self-renewal as well as the onset of lineage restriction.

### **Role of DNA Methylation in Pluripotency Maintenance**

Equally important to histone methylation in regulating the accessibility of DNA is the methylation of DNA itself, which leads to transcriptional repression (Keshet et al. 1986, Boyes and Bird 1991). Over the course of embryonic development, new DNA methylation marks are established in order to gradually restrict a cell's developmental potential. There are three key players in DNA methylation; DNA methyltransferase 3a (Dnmt3a), Dnmt3b, and Dnmt1, all of which catalyze the transfer of a methyl group to a cytosine residue, primarily at CpG islands (Sano and Sager 1982, Gruenbaum et al. 1982, Bestor and Ingram 1983). This methylation can simultaneously prevent binding of certain proteins with DNA while enabling other proteins to bind. Both Dnmt3a and Dnmt3b are de novo demethylases and are responsible for establishing the CpG methylation pattern initially, which is critical for regulation of specific tissue types during development (Okano et al. 1999, Liang et al. 2002, Hu et al. 2014a). De novo methyltransferases are stabilized by already methylated DNA, preventing aberrant de novo DNA methylation and promoting inheritance of methylated DNA (Sharma et al. 2011). Dnmt1 is a maintenance DNA methyltransferase, so it is recruited to hemi-methylated DNA by Uhrf1 and ensures that each newly synthesized DNA strand has the same methylation pattern as the parent strand (Figure 1.13) (Araujo et

al. 1998, Rothbart et al. 2012, Liu et al. 2013, Pechalrieu et al 2017). Methylated DNA can be demethylated passively if the methylation pattern is not maintained by Dnmt1 during DNA replication, or actively by ten eleven translocation (TET) proteins (Ito et al. 2011, Paranjpe and Veenstra 2015).

In mice, global demethylation occurs just prior to the maternal-zygotic transition (Monk et al. 1987). However, this is likely related to implantation or uterine development, as it is not seen in zebrafish or *Xenopus*, where DNA remains hypermethylated before and after the maternal-zygotic transition (Macleod et al. 1999, Veenstra and Wolffe 2001). DNA methylation has been shown in mouse ES cells to be required specifically for lineage specification, but not necessarily in the maintenance of pluripotency, self-renewal or viability (Tsumura et al. 2006). Work in zebrafish has demonstrated that interactions between Uhrf1 and Dnmt1 play a significant role in the regulation of gastrulation during embryonic development, when lineage specification begins. Uhrf1 phosphorylation of Dnmt1 is required for its stability and that crosstalk between these two proteins is essential for them both to function properly during gastrulation (Kent et al. 2016).

While Dnmt1 is broadly expressed across all tissues, it has been shown in mice that Dnmt3a and Dnmt3b have more stage and tissue specific expression. Dnmt3b is expressed ubiquitously early in development in totipotent cells, but later in development is restricted to progenitor cells in hematopoiesis, spermatogenesis and neural progenitor cells, whereas Dnmt3a is expressed after gastrulation in mesenchymal cells as well as in more specified tissues, after the progenitor stages (Watanabe et al. 2002, 2004, 2006, Feng et al. 2005). In human ESCs it has been shown that Dnmt3b is necessary to regulate the timing of neural specification and a knockdown of Dnmt3b increases expression of neural crest specifier genes including *Pax3*, *Foxd3*, *Sox10* and *Snail2* (Martins-Taylor et al. 2012). Further studies in chick show that Dnmt3b inhibits *Sox10* expression by directly methylating its promoter region, thus loss of Dnmt3b led to an excess emigration of neural crest cells from neural tube. Knockdown of both Dnmt3a and Dnmt3b has a more penetrant effect than knockdown of one of them alone, indicating they play temporally distinct roles, with Dnmt3a acting earlier than Dnmt3b (Hu et al. 2014b). It has also been demonstrated in chick that it is primarily Dnmt3a expressed in the neural crest and loss of Dnmt3a

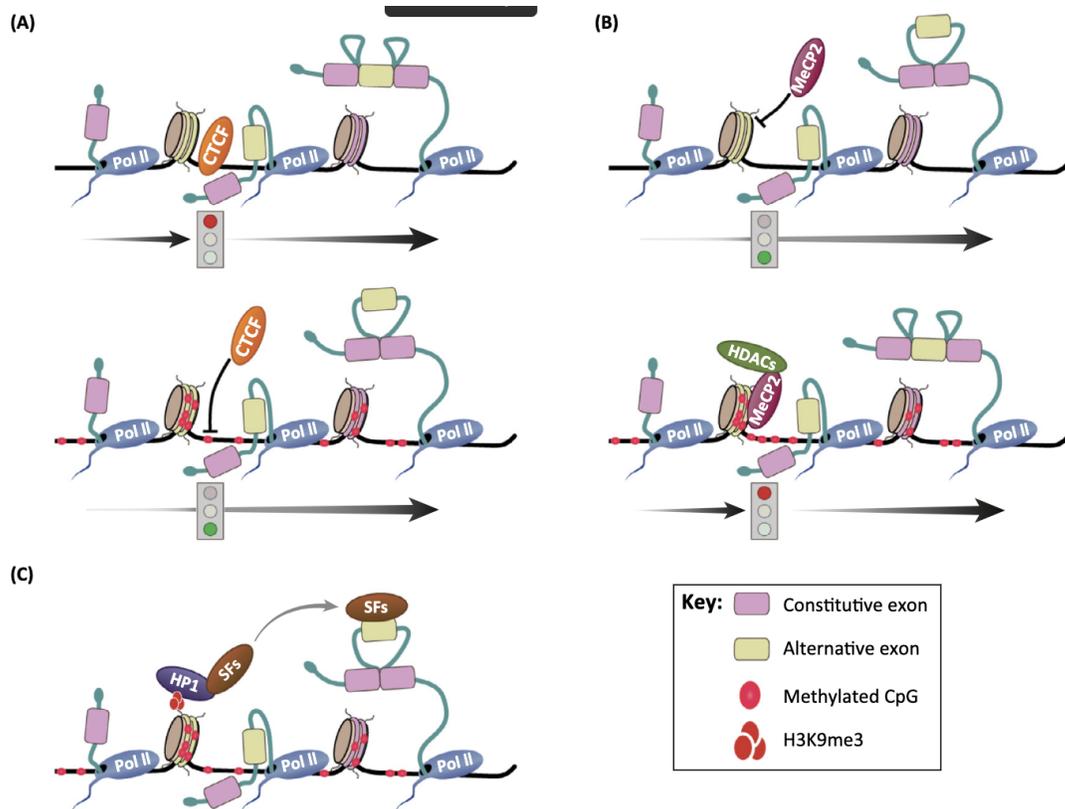
leads to a loss of some neural crest specifier genes including *Sox10*, *Snail2* and *Foxd3* and allows for expansion of neural genes, *Sox2* and *Sox3* into the neural crest (Hu et al. 2012).

### **Role of DNA Methylation in Alternative Splicing**

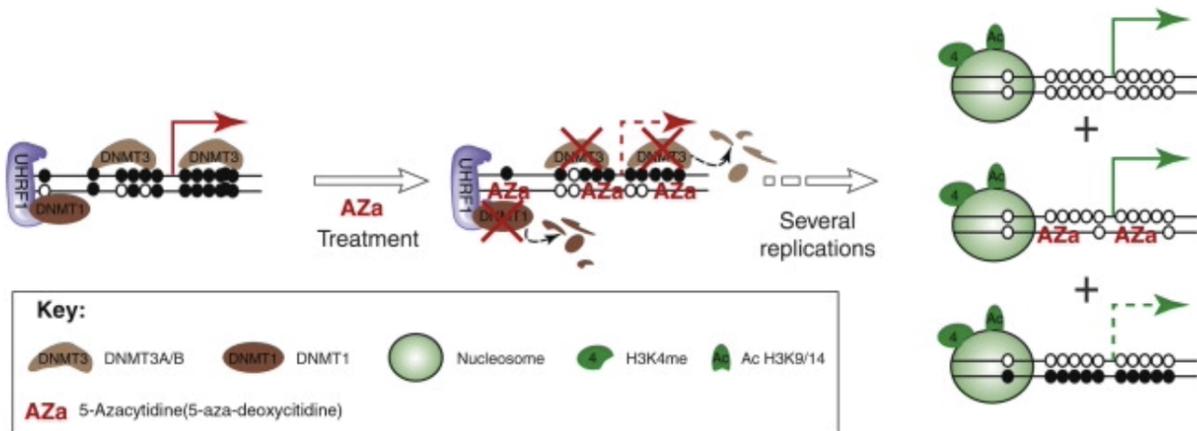
DNA methylation is primarily understood to play a role in the repression of transcription that is temporally and spatially controlled to allow for proper development and lineage formation, but it has also been shown to play an important role in regulating DNA splicing. DNA methyltransferases most often target nucleosome-bound DNA, and DNA methylation is enriched in exons compared to introns (Hodges et al. 2009, Chodavarapu et al 2010). A study in bees suggests that this enrichment of DNA methylation in exons is not merely the byproduct of nucleosome occupancy but may actually be playing a regulatory role in alternative splicing, as inhibiting Dnmt3 significantly changed intron splicing patterns, with changes in exon skipping and intron retention being directly caused by the decrease methylation (Li-Byarlay 2013). A different study in mouse ESCs strengthens the causal relationship between DNA methylation and alternative splicing as DNA methylation was switched on or off in a single gene, and adding DNA methylation increased inclusion of alternative exons (Yearim et al. 2015, Shayevitch et al. 2018, Lev Maor et al. 2015).

There are three proteins that enable the mechanism by which DNA methylation regulates alternative splicing. When the binding factor CTCF is prevented from binding by DNA methylation at its binding site, downstream Pol II pausing is released, allowing it to travel the mRNA more rapidly, leading to exon exclusion (Shukla et al. 2011) (Figure 1.14 A). DNA methylation can also recruit the protein MeCP2 which enhances exon recognition, thus reducing aberrant exclusion of alternative exons, and MeCP2 then recruits HDACs to keep histones hypoacetylated, enhancing Pol II pausing and exon inclusion (Maunakea et al. 2013) (Figure 1.14 B). Finally, DNA methylation induces H3K9Me3 which binds with the HP1 protein which then recruits splicing factors and can either enhance exon recognition when bound upstream of the exon and lower recognition when bound directly to the exon (Yearim et al. 2015). (Figure 1.14 C).

### **Small Molecule Inhibitors Targeting DNMTs**



**Figure 1.14 DNA Methylation Regulates Alternative Splicing.** (A) Schematic of DNA methylation preventing CTCF binding thus releasing Pol II pausing leading to exon exclusion. (B) Schematic of DNA recruiting MeCP2 which recruits HDACs and leads to exon inclusion. (C) Schematic of DNA methylation binding to HP1 leading to recruitment of splicing factors which has alternative effects on exon inclusion depending on where HP1 is bound (Adapted from Lev Maor et al. 2015).

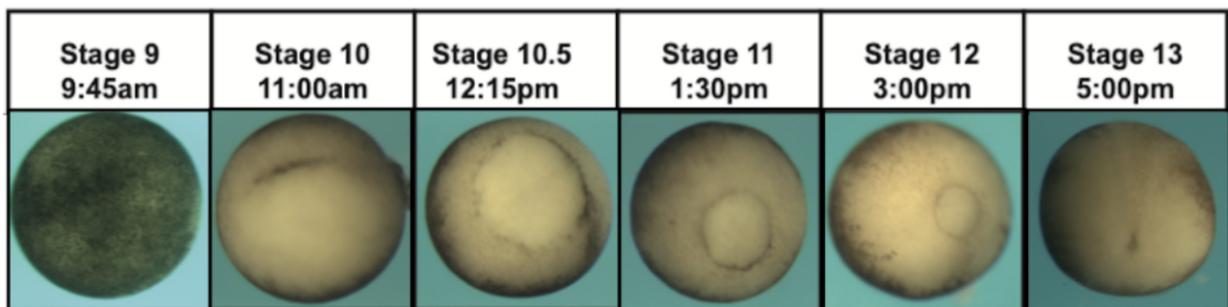


**Figure 1.15 5-Aza Inhibits DNMTs.** The pharmacological inhibitor 5-Aza-deoxycytidine inhibits both de novo (Dnmt3a/b) and maintenance (Dnmt1) Dnmts leading to fully demethylated DNA several rounds of replication after treatment (Adapted from Yang et al 2010).

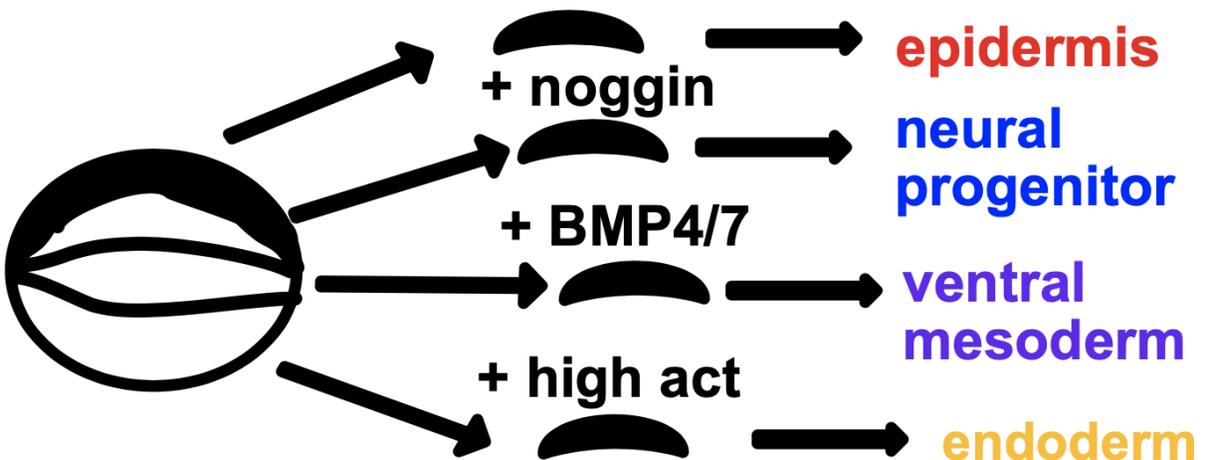
Given the significant effects DNA methylation has on both transcription and alternative splicing, many mechanisms of inhibition have been developed in order to more closely study its function. DNA methylation inhibitors are of particular interest for cancer treatment as altered DNA methylation can play a role in cancer. Tumor cells tend to be globally demethylated, leading to genomic instability, while at the same time being hypermethylated at the promoters of tumor suppressor genes (Pechalrieu et al. 2017). All three Dnmts have been shown to be overexpressed in multiple forms of cancer emphasizing the utility of a Dnmt inhibitor for cancer treatment and research (Roll et al. 2008, Yu et al. 2015). The first Dnmt inhibitors to be developed were 5-Aza-cytidine and 5-Aza-2'-deoxycytidine, which operate by similar mechanisms. Both inhibitors are incorporated into the DNA, where they irreversibly bind Dnmts, thus leading to demethylation of DNA after several rounds of replication (Creusot et al. 1982, Yang et al. 2010) (Figure 1.15). While 5-Aza-cytidine affects both DNA and RNA and must be reduced by a ribonucleotide reductase before being incorporated into the DNA 5-Aza-2'-deoxycytidine is DNA specific and is incorporated into DNA directly (Stresemann et al. 2008, Qiu et al. 2010). These drugs have both been FDA approved and shown to effectively demethylate DNA in tumor cells (Mund et al. 2005, Yang et al. 2006, Gore et al. 2006, Soriano et al. 2007). Thanks to the development of these drugs for cancer research, they are also a promising tool with which to study the role of DNA methylation in lineage specification in several model organisms.

### ***Xenopus* as a Model System**

*Xenopus laevis*, or the African clawed frog, are an excellent and widely used model for vertebrate embryology and development as well as cellular and molecular biology, genomics and toxicology. *Xenopus* have several attributes that can be used as experimental strengths for biological research. First, they can be primed with the human hormone gonadotropin to produce thousands of eggs. Importantly, these eggs can be fertilized externally, producing hundreds of synchronously developing embryos, allowing large scale experiments to be done easily. Since these embryos develop externally in a simple salt solution, they are accessible at every stage of development, which is crucial for studying the earliest stages of development. Another strength of *Xenopus* is the timing of embryonic development, both its



**Figure 1.16 *Xenopus laevis* Development Time Frame.** *Xenopus* transit from a pluripotent state (stage 9) to a lineage restricted state (stage 13), upon which primary germ layer formation has been completed in ~7 hours at 20 degrees Celsius. *Xenopus laevis* can be staged using morphological cues, the presence of bottle cells on the dorsal lip at stage 10 indicates the onset of gastrulation and the closing of yolk plug and delineation of the neural plate at stage 13 indicates that gastrulation is complete, and progenitor populations of each germ layer have been established. Animal caps achieve lineage restriction on the same time scale as if they are still in the embryo and therefore can also be staged using morphological cues of sibling embryos.



**Figure 1.17 Animal Cap Assay.** Animal caps can be directed to form any cell type in the embryo. In the absence of additional cues these cells will form epidermis. With the BMP antagonist noggin, cells will form neural progenitors. BMP4/7 heterodimer causes cells to form ventral mesoderm. High levels of activin can be used to generate endoderm

rapidity and manipulability. *Xenopus* embryos can develop healthily at a relatively wide range of temperatures, with the rate of development directly correlating to the temperature. Embryos can go from pluripotent to early lineage restricted stages in a matter of just ~7 hours at 20°C, much quicker than mice or human cell lines, allowing this developmental transition to be studied at a high resolution (Figure 1.16).

*Xenopus* also possess many features that make them great candidates for whole embryo studies. The left-right halves of the embryo are established upon the first cell cleavage, allowing one half of the embryo to be manipulated by mRNA injection, leaving the other side a perfect internal control. Furthermore, *Xenopus* cells have been extensively fate-mapped allowing for precise targeting for both overexpression studies with mRNA and knockdown studies with morpholinos.

Importantly for the work in this thesis, *Xenopus* animal poles can be dissected from the rest of the embryo at pluripotent stages, isolating a population of naive cells that can be directed to become any cell type in the embryo with the addition of the appropriate exogenous factors. (Figure 1.17). These blastula explants, or animal caps, develop in a simple high salt media, and develop and achieve lineage restriction at the same rate as if they were still in the embryos, thus their developmental stage can be confirmed by sibling embryos, even though the caps themselves lack key morphological indications of stage. Animal cap assays allow the development of specific cell types to be studied in isolation, without the influence of signals from the rest of the embryos, and consequently the effect of different manipulations on specific germ layers can also be studied. While the pseudotetraploid nature of *Xenopus* and the consequent L and S allele can complicate genomic studies, the characterization and sequencing of the *Xenopus* genome has improved significantly in recent years. The rapid early stages of development, external development and recent advancements in sequencing of the genome make *Xenopus laevis* an ideal system for studying the quantitative dynamics of lineage restriction.

### **Specific Questions to be Addressed in this Thesis**

While the pluripotent state and lineage restricted state of cells are well defined, the transition between these two levels of potential has not yet been studied at a high resolution. Additionally, it is known which signaling pathways drive cells towards each of the primary germ layers, but the timing and

dynamics of the response of the cells to these varying signals at the transcriptome level is largely unknown. The goal of this thesis is to examine at an unprecedented level the transcriptome dynamics of cells as they transit from pluripotent stem cells to four distinct multipotent germ layers.

In chapter two of this thesis, I describe the transcriptome dynamics of cells transiting to the neural, epidermal, ventral mesoderm and endoderm state. I highlight the divergence points of each of these lineages, as driven by either BMP or Nodal / Activin signaling. I demonstrate an early equilibrium of the neural lineage and a uniquely monotonic nature of the path towards the neural lineage, providing quantitative backing for the neural default model. I identify stage 10.5 as a critical time point for the effect of BMP signaling, demonstrating that cells exposed to BMP signaling before stage 10.5 will become ventral mesoderm, while cells exposed to BMP endogenously at stage 10.5 will become epidermal. I also identify *Dand5* as a maternally provided BMP inhibitor that plays an essential role in keeping BMP signaling off until the appropriate stage. Finally, I propose novel candidates for the mesendoderm gene regulatory network.

In chapter three, I examine the role of the DNA methylation on pluripotency, lineage restriction and the neural crest through *in situ* hybridization and RNA sequencing. I show via *in situ* that inhibition of DNA methyltransferases has a heterogeneous effect on pluripotency markers and impairs the ability of animal caps to undergo lineage restriction. Through GO term analysis of stage 13 RNASeq, I show an enrichment of cell signaling related GO terms in genes increased by 5-Aza and enrichment for structural GO terms in genes decreased by 5-Aza as well as a maintenance of neural and mesodermal development terms that normally turn off during normal epidermal development. Through PCA of stage 13 RNASeq I show that caps in which DNA methylation is inhibited separate from all other lineages primarily based on RNA splicing. When splicing is accounted for, caps in which DNA methylation is inhibited are most similar to the neural crest lineage.

## **Chapter Two**

### **Quantitative Analysis of Transcriptome Dynamics Provides Novel Insights into Developmental State Transitions**

During embryogenesis, the developmental potential of initially pluripotent cells becomes progressively restricted as they transit to lineage restricted states. The pluripotent cells of *Xenopus* blastula-stage embryos are an ideal system in which to study cell state transitions during developmental decision-making, as gene expression dynamics can be followed at high temporal resolution. Here we use transcriptomics to interrogate the process by which pluripotent cells transit to four different lineage-restricted states: neural progenitors, epidermis, endoderm, and ventral mesoderm, providing quantitative insights into the dynamics of Waddington's landscape. Our findings provide novel insights into why the neural progenitor state is the default lineage state for pluripotent cells and uncover novel components of lineage-specific gene regulation. These data reveal an unexpected overlap in the transcriptional responses to BMP4/7 and Activin signaling and provide mechanistic insight into how the timing of signaling inputs such as BMP are temporally controlled to ensure correct lineage decisions. Together these analyses provide quantitative insights into the logic and dynamics of developmental decision making in early embryos. They also provide valuable lineage-specific time series data following the acquisition of specific lineage states during development.

## **Introduction**

How a single cell ultimately gives rise to a patterned, complex organism is a fundamental question in biology. Embryonic development can be generalized as a process of progressive restriction of cellular potential. In vertebrates, the zygote is totipotent, but by blastula stages the three primary germ layers, ectoderm, mesoderm, and endoderm, have been specified. The fates of cells within these germ layers then become progressively restricted to single differentiated cell types characteristic of that germ layer. Conrad Waddington famously depicted this process as a topological landscape (Waddington 1957). In his model, a ball positioned at the top of the landscape represents a cell with all developmental pathways open to it. As the ball progresses down the landscape, the paths it takes dictate which lineage states will remain accessible. Waddington noted that the valleys or channels of the landscape arise from the interactions between genes and from their interactions with the cell's environment.

At blastula stages vertebrate embryos possess a transient population of pluripotent cells which, like the fertilized egg, occupy a position atop Waddington's landscape. These cells- inner cell mass cells in mammals and naïve animal pole cells in amphibians- can give rise to the derivative cell types of all three germ layers and as such can recapitulate the path to different lineage states including the relevant gene regulatory network (GRN) topology and dynamics. Studies using explants of pluripotent cells from *Xenopus* blastulae (so called "animal caps") have been central to our current understanding of the signals and transcriptional responses that direct these stem cells toward specific lineage states (Snape et al. 1987, Ariizumi and Asashima 2003, Ariizumi et al. 2009, Ariizumi et al. 2017, Satou-Kobayashi et al. 2021). Some of these signals emanate from the blastopore lip, or the Spemann-Mangold organizer, and help to direct formation and patterning of the primary germ layers (Spemann & Mangold 1924, Harland & Gerhart 1997, Niehrs 2004). Exit from the pluripotent state also coincides with the loss of expression of many maternally provided pluripotency transcripts, such as *Pou5f3.3* and *Foxi2* (Whitfield et al. 1995, Henig et al. 1998, Lef et al. 1994, Paraiso et al. 2020). As cells exit pluripotency, their potential becomes progressively restricted until their fate becomes specified and then determined.

Animal pole cells in *Xenopus* are fated to give rise to ectodermal derivatives. In situ these cells give rise to both epidermal and neural progenitor cells, as well as neural crest and cranial placodes, under the direction of signals from the organizer. Absent these signals, animal pole cells are directed by endogenous BMP signaling to become epidermis (Wilson et al. 1997). BMP2, 4, and 7 have all been shown to be potent epidermal inducers (Wilson and Hemmati-Brivanlou 1995, Suzuki et al. 1997a), and BMP4/7 heterodimers have been identified as the most physiologically relevant ligands in early embryos (Little and Mullins 2009). The binding of these ligands to type I and II BMP receptors results in phosphorylation of Smad1/5/8 and the translocation of these phospho-Smads to the nucleus together with Smad4, resulting in transcription of target genes including *Epidermal Keratin (EpK)* and *Dlx3* (Kretzschmar et al. 1997, Macías-Silva et al. 1998, Shi and Massague 2003, Jonas et al. 1985, Dirksen et al. 1994, Luo et al. 2001).

Although isolated animal pole cells will transit to an epidermal state absent additional instructions, it has been proposed that the default state of these initially pluripotent cells is a neural progenitor state

(Hemmati-Brivanlou and Melton 1994, 1997, Muñoz-Sanjuán & Brivanlou 2002). This model arose from the findings that neural “inducing” factors secreted by the organizer, including Noggin, Chordin, Follistatin and Cerberus, function as BMP antagonists (Lamb et al. 1993, Re'em-Kalma et al. 1995, Sasai et al. 1995, Piccolo et al. 1996, Piccolo et al. 1999, Iemura et al. 1998), as well as from the fact that cells neuralize upon dissociation (Grunz and Tacke 1989). The “neural default” model has not been without controversy, however. In chick embryos the expression patterns of BMPs and their antagonists do not fully fit the neural default model, misexpression of BMP antagonists does not induce neural progenitors, and ectopic expression of BMP fails to inhibit neural plate formation (Streit et al. 1998, Streit and Stern 1999a, Streit and Stern 1999b, Stern 2005). There is also evidence in *Xenopus* that FGF signaling may be required for competence to respond to BMP antagonists (Linker and Stern 2004, Delaune et al. 2005, Wawersik et al. 2005), although the roles of FGF and BMP signaling in neural induction have shown to be separable (Wills et al. 2010). By contrast, *Xenopus* animal pole explants exposed to BMP antagonists such as Noggin adopt a neural progenitor state and express neural genes including *Sox2/3* and *Otx1/2*. These transcription factors, and their homologs, play an important role in specification of the central nervous system (CNS) in vertebrates (Collignon et al. 1996, Mizuseki et al. 1998, Penzel et al. 2003, Pannese et al. 1995, Kablar et al. 1996, Andreazzoli et al. 1997, Plouhinec et al. 2017).

Explants of *Xenopus* pluripotent blastula cells have also played a central role in determining the signals that control formation of the other embryonic germ layers, mesoderm and endoderm (Asashima et al. 1990a,b, Smith et al. 1990, Henry et al. 1996). Nodal, Activin and Vg-1, ligands of the other branch of the TGF-beta family, have been shown to act as morphogens, with low levels of signaling inducing mesoderm and high levels inducing endoderm (Smith et al. 1990, Hemmati-Brivanlou & Melton 1992, Gurdon et al. 1994, McDowell et al. 1997, McDowell and Gurdon 1999, Gamer and Wright 1995, Agius et al. 2000, Dale et al. 1993, Thomsen and Melton 1993, Kessler and Melton 1995). Treatment of these cells with exogenous Activin mimics the activity of Nodal/Vg-1 (Jones et al. 1995) promoting phosphorylation of receptor Smads 2/3 (Reissmann et al. 2001, Kumar et al. 2001). High levels of Activin/Nodal signaling induce endoderm as evidenced by expression of key factors *Sox17* and *Endodermin* (Hudson et al. 1997, Sasai et al. 1996) whereas lower doses induce both dorsal and ventral mesoderm (Sokol and Melton

1992, Green et al. 1992). BMP4/7 heterodimers also possess mesoderm inducing activity, but it is limited to ventral, not dorsal mesoderm (Nishimatsu and Thomsen 1998, Hemmati-Brivanlou and Thomsen 1995, Graff et al. 1994). Much still remains to be learned about the signaling dynamics and outputs of BMP4/7 versus Activin-mediated regulation of cell state transitions during germ layer formation.

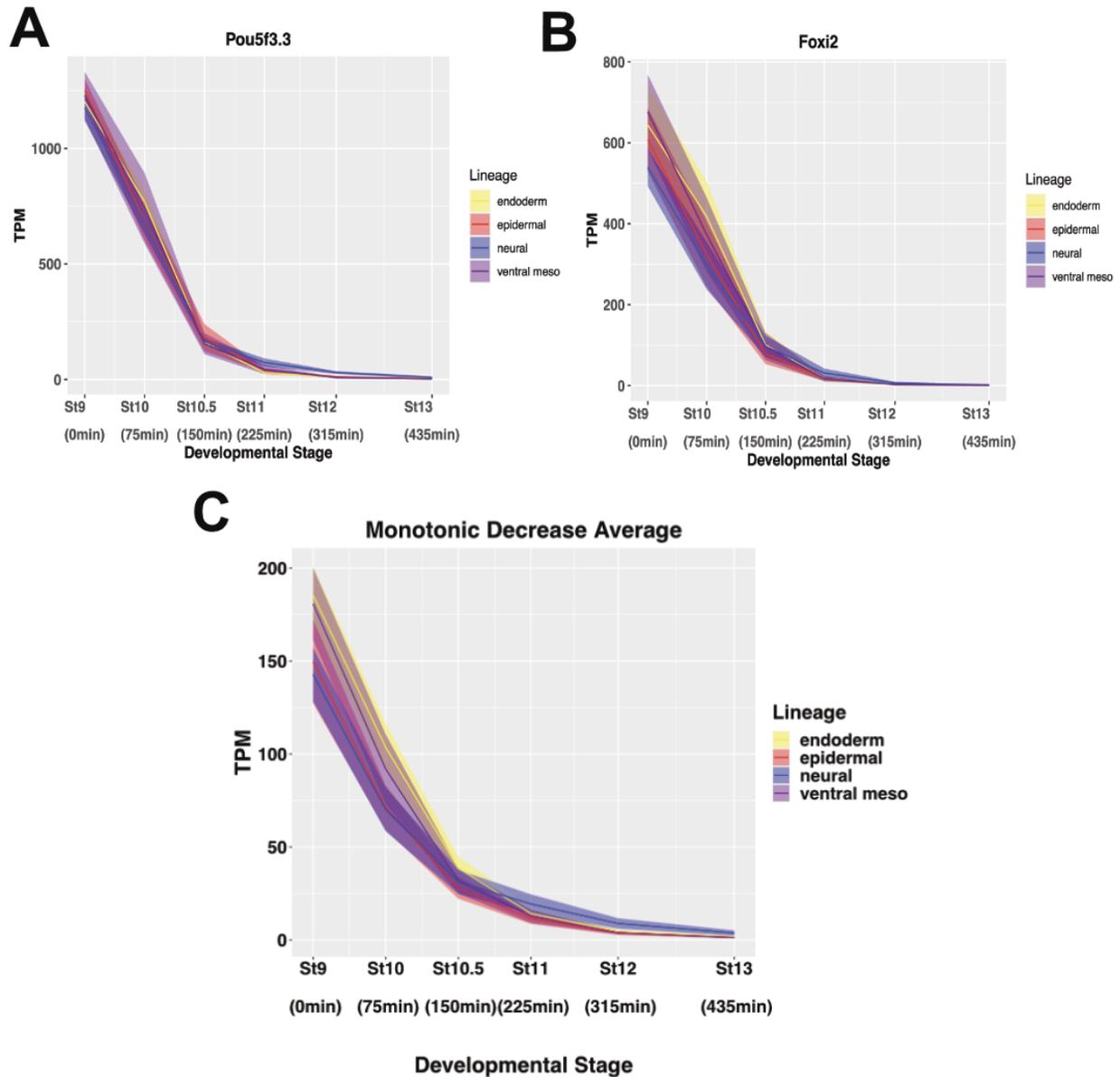
*Xenopus* animal pole explants are an ideal system to probe the dynamics of developmental decision-making during lineage restriction. These cells can be cultured in a simple salt solution and will undergo lineage restriction on the same time scale as if they had remained in vivo. Given appropriate signals, animal pole cells can be directed to any lineage progenitor state within a time frame of approximately seven hours. Analogous experiments in cultured mouse or human embryonic stem cells (ESCs) take more than a week in culture, and unlike ESCs *Xenopus* cells do not need to be artificially retained in a pluripotent state that may be distinct from the transient pluripotency that exists in vivo. The unique features of the *Xenopus* system allow initially pluripotent cells to be followed at high temporal resolution as they progress to lineage restriction, providing insights into the dynamics of developmental decision making in early vertebrate embryos. Here we develop an experimental platform and quantitative framework in which pluripotent cells explanted from blastula stage *Xenopus* embryos can be used to study the transit of these cells to four different lineage states - epidermis, neural progenitor, endoderm and ventral mesoderm - by following changes in the transcriptome at six time points during this seven-hour process. These data provide quantitative insights into the dynamics of Waddington's landscape and complement previous data sets looking at transcriptome dynamics in *Xenopus* (Owens et al. 2016, Session et al. 2016) by providing lineage specific dynamics, as well as transcriptome dynamics in response to BMP/Activin signaling. Our findings shed light on why a neural progenitor state is the default lineage state for pluripotent cells, uncover novel components of lineage specific GRNs, and provide insights into essential control of the timing of signaling inputs such as BMP for proper lineage decisions. These time-resolved data sets will serve as an important resource for future studies of developmental decision making in early vertebrate embryos.

## Results

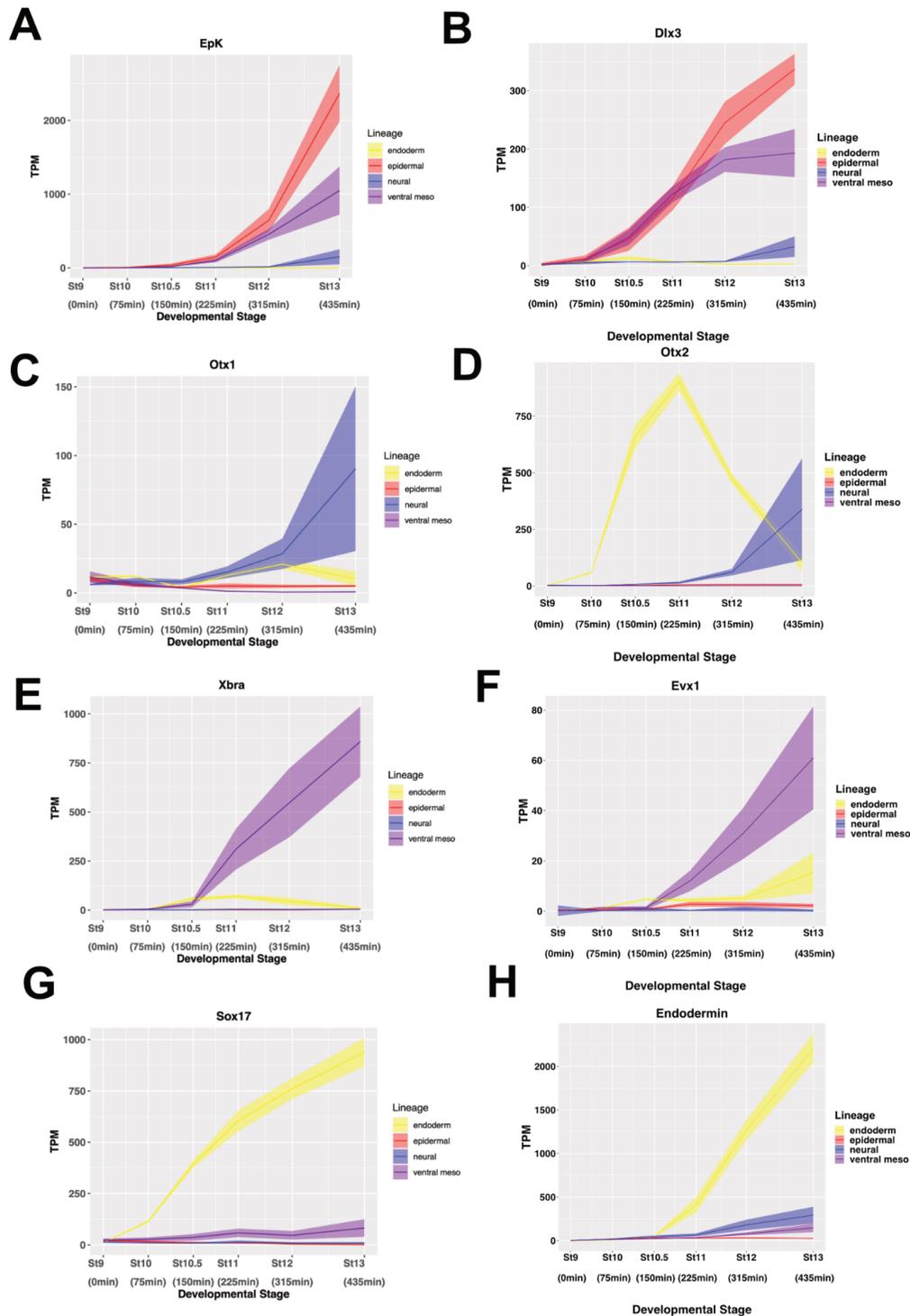
### **Naïve Animal Pole Cells from *Xenopus* Blastula can be Programmed to any Lineage**

To allow interrogation of transcriptome changes at high temporal resolution as pluripotent cells become lineage restricted, we established a highly regimented protocol for collecting synchronous populations of blastula explants across six points on the path towards lineage restriction. Late blastula stage explants (Nieuwkoop and Faber stage 9) were designated time zero and represent the pluripotent state atop Waddington's landscape. In addition, explants were collected at 75, 150, 225, 315 and 435 minutes after stage 9 (Nieuwkoop and Faber stages 10, 10.5, 11, 12, 13), confirmed by stages of sibling embryos (Nieuwkoop and Faber 1994). Stage 10, the onset of gastrulation, is marked by presence of the dorsal blastopore lip; stages 11 and 12 are mid and late gastrulae respectively. At stage 13, the neural plate stage, the developmental potential of embryonic cells has been restricted to a single lineage state, with the notable exception of neural crest cells (Prasad et al. 2012). Explants cultured without additional instructive cues transit to an epidermal state. To follow cells as they transited to an endodermal state, explants were treated with Activin (Gamer and Wright 1995). Titration experiments determined that treatment with 160ng/uL of Activin at stage 9 was the minimum concentration of Activin able to robustly induce endoderm without expression mesodermal markers, as assayed by qPCR. To direct a neural progenitor state explants were exposed Noggin (Lamb et al. 1993). 100ng/uL of Noggin was used for these experiments as it was the lowest concentration that induced *Sox3* at stage 13 and effectively blocked BMP signaling as determined by qPCR and Western blot. Treatment with BMP 4/7 heterodimers was used to induce a ventral mesoderm state (Köster et al. 1991, Nishimatsu and Thomsen 1998). 20ng/uL of BMP 4/7 was found to induce pSmad 1/5/8 at near endogenous levels, as determined by western blot, shifting the timing but not amplitude of BMP signaling. Use of BMP4/7 heterodimers to induce ventral mesoderm allows the transcriptional responses to the two different branches of TGF-beta signaling to be compared. RNA was isolated from explants at each time interval and used to generate illumina libraries for transcriptome analysis.

The transcript dynamics of maternally provided pluripotency factors *Pou5f3.3* and *Foxi2* across three biological replicates for all four lineage transitions (twelve independent experiments) demonstrates that this pipeline generated highly quantitative and highly reproducible data with minimal technical error



**Figure 2.1 Dynamics of Maternal Transcripts Demonstrate Assay Reproducibility.** (A-C) RNA Seq TPM expression over time of (A) maternally provided pluripotency marker *Pou5f3.3*, (B) maternally provided *Foxi2*, (C) average of 119 monotonically decreasing maternally provided transcripts. Graphs are sums of S+L allele. Width of lines represents SEM of three biological replicates.



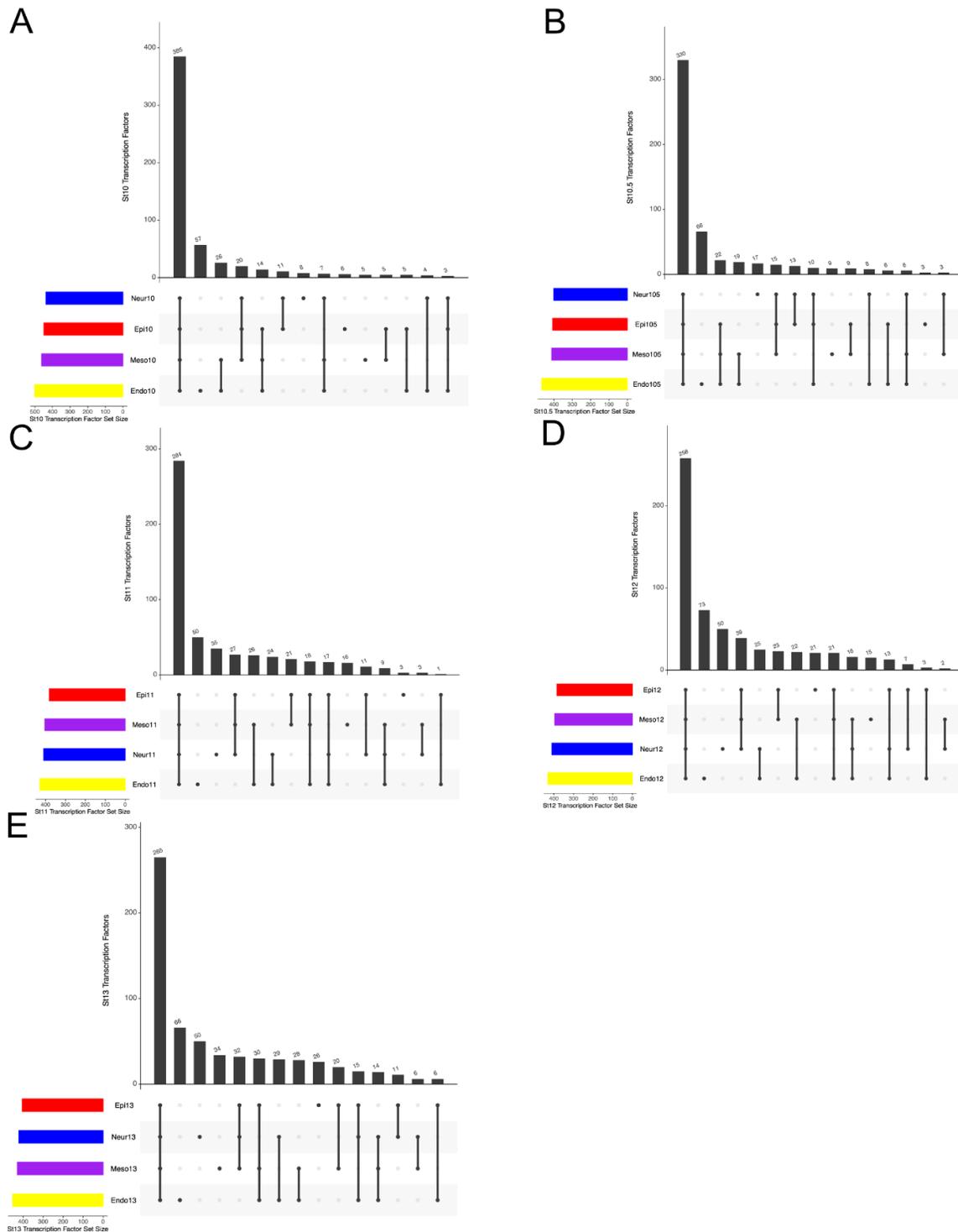
**Figure 2.2 Four Independent Lineages are Generated.** (A-H) RNA Seq TPM expression over time of (A) epidermal markers *EpK*, (B) *Dlx3*, (C) neural markers *Otx1*, (D) *Otx2*, (E) mesoderm markers *Brachyury(T)*, (F) *Evx1*, (G) endoderm markers *Sox17* (H) *Endoderm*. Graphs are sums of S+L allele. Width of lines represents SEM of three biological replicates.

(Figure 2.1 A, B). *Pou5f3.3* and *Foxi2* are representative of a class of 119 genes whose expression decreases monotonically by at least 25-fold between stage 9 and stage 13 (Figure 2.1C, Supplemental Table 1) and includes maternal transcripts that characterize the pluripotent state (Collart et al. 2014, Gentsch et al. 2019). Normalized RNA read count data, quantified as transcripts per million (TPM), confirmed generation of each of the expected lineage states, with *EpK* and *Dlx3* validating establishment of the epidermal state (Fig. 2.2 A, B), *Otx1* and *Otx2* validating transit to a neural state (Figure 2.2 C, D), *Brachyury(T)* and *Evx1* (Figure 2.2 E, F), and *Sox17* and *Endodermin* (Figure 2.2 G, H) validated the mesoderm and endoderm data sets respectively.

Global analysis of transcription factors expressed at each stage revealed that the majority are expressed in all four lineages, although their expression levels or timing may vary between them. For example, *Otx1* is unique to the endodermal lineage at stage 10 but to the neural lineage at stage 13 (Figure 2.3, Supplemental Table 2). This analysis also identified the transcription factors that at any given stage are expressed in only a single lineage and revealed that such genes are overrepresented in the endoderm lineage. We also performed weighted gene correlation network analysis (WGCNA) on transcription factors on each lineage independently with time as a continuous variable. This analysis showed that many known Activin induced transcription factors cluster together in the same module not only in the endoderm lineage, but also in the mesoderm lineage, with about half of these also clustering together in the neural lineage. Similarly, many transcription factors known to be induced by BMP cluster together in both mesoderm and epidermis, and surprisingly a small subset of these also cluster in the neural lineage, suggesting that there are subsets of transcription factors that will cluster together even in response to distinct inductive signals (Supplemental Table 3). Further study of these putative modular gene regulatory networks may provide novel insights into potential mechanisms of co-regulation of correlation factors.

### **PCA and Time Series Analysis Reveal Novel Lineage-Specific Dynamics**

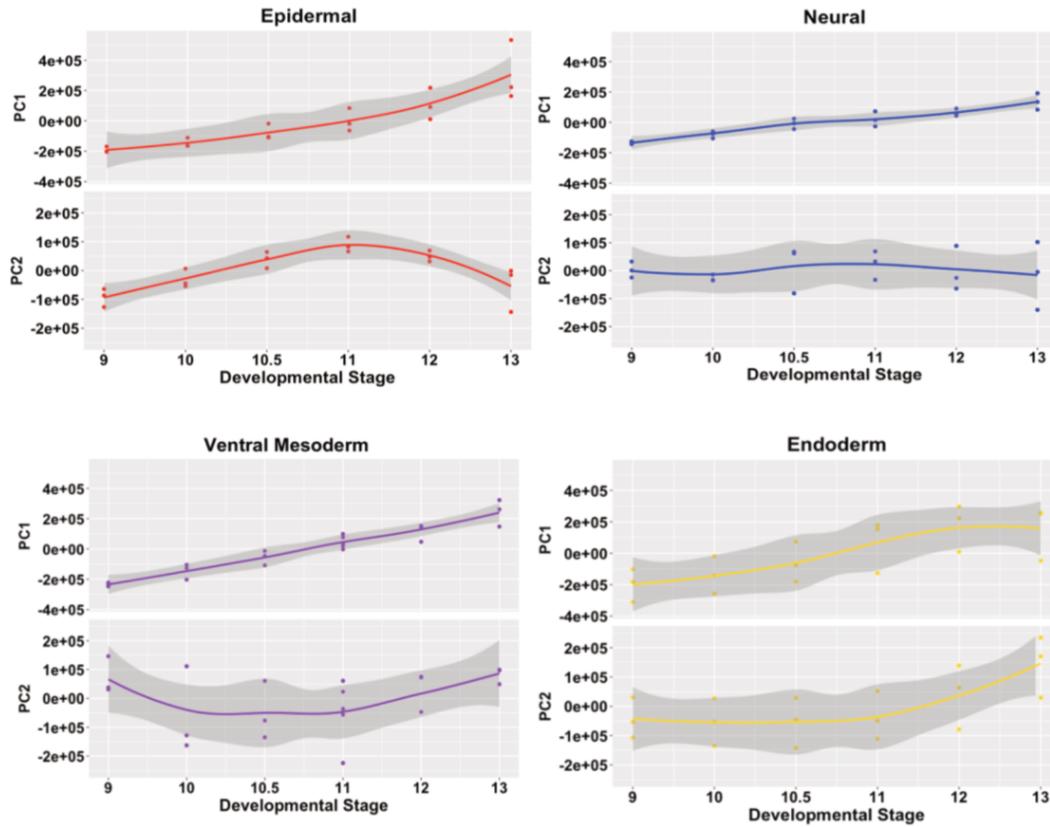
Together the transcriptomes of the four state transitions each across six time points yield 72 observations in a 45,661-dimensional gene expression space. Global insights into such high dimensional



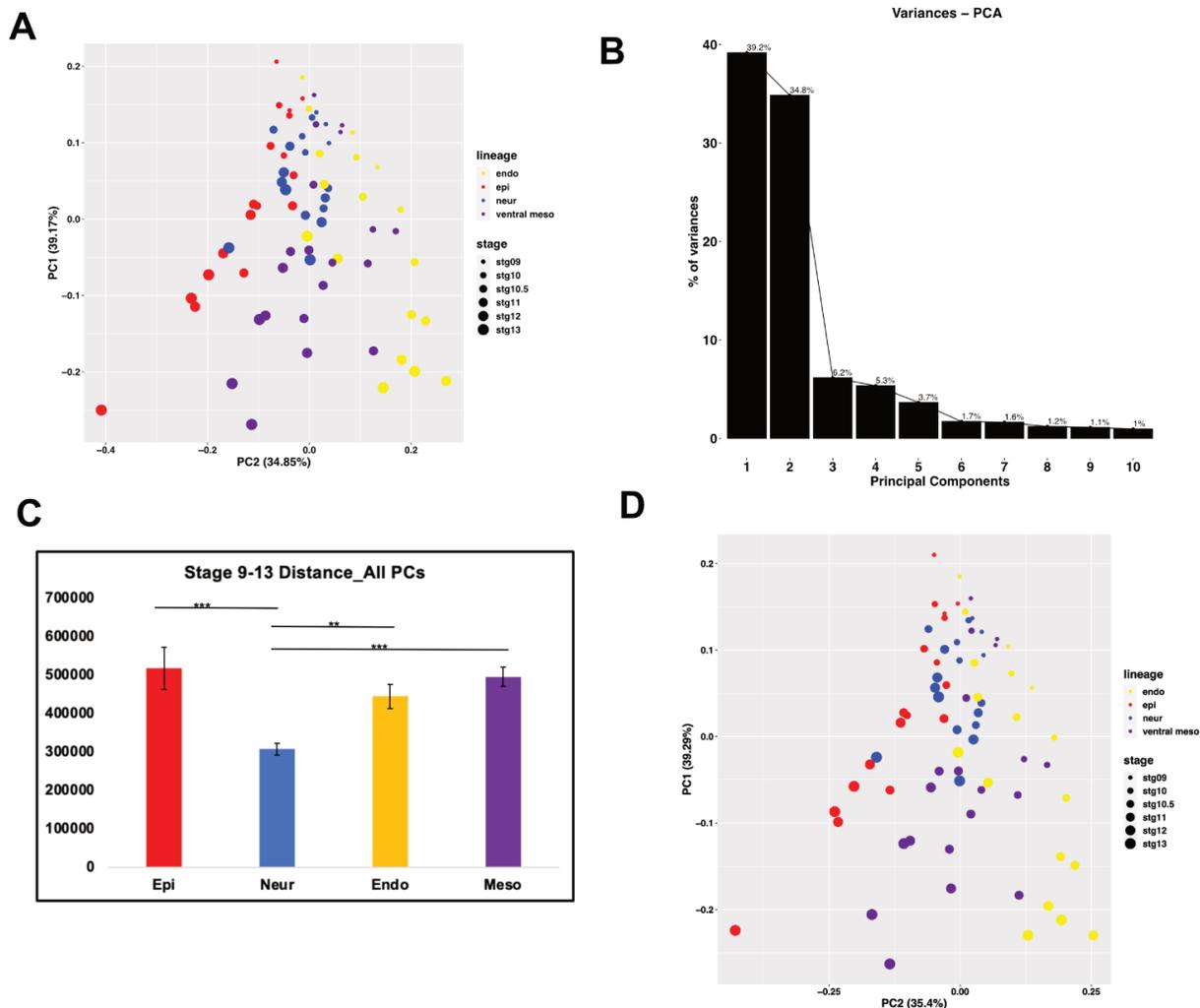
**Figure 2.3 Transcription Factors Expression Shows Overlap Between Lineages.** (A-E) UpSet plots of transcription factors expressed at a minimum of 10 TPM at (A) stage 10, (B) stage 10.5, (C) stage 11, (D) stage 12, (E) stage 13. X-axes show genes unique to each lineage and overlapping in all different combinations of lineages ordered from largest number of genes to smallest.

data require methods for dimensionality reduction. Principal Component Analysis (PCA) can provide key insights into the genes contributing most significantly to the variance between lineage states and developmental stages. We first used PCA to analyze each lineage individually, plotting the first two principal components against developmental time (Fig. 2A). For all four lineages the primary principal component (PC1) was found to be largely monotonic over time, suggesting that the majority of the gene expression variance is contributed by genes changing unidirectionally, such as pluripotency genes being turned off or lineage-specific genes being activated. Interestingly, when PC2 was plotted for each of the lineages it was found to exhibit temporal dynamics in the epidermal, ventral mesodermal and endodermal lineages that suggested genes exhibiting expression peaks at intermediate time points make a significant contribution to the variance, and potentially to these state transitions. By contrast, PC2 of the neural lineage shows no association with time, suggesting that the temporal dynamics in the neural lineage are primarily linear in time. This raises the possibility that the transition to a neural progenitor state follows a simpler trajectory than that of the other three lineages (Figure 2.4).

To gain further insights into these state transitions, PCA was carried out on all four lineages in concert. In this analysis, PC1 and PC2 together explain 75% of the variance across these data. (Figure 2.5 A, B). The distribution of these data across the PC1 axis correlates with developmental time, with the earliest (St.9) samples clustering at the top of the plot and the later samples progressively further down the y-axis (Figure 2.5 A). When examined in this context, the neural trajectory is striking as it extends a shorter distance along this axis, with the neural replicates for stages 11, 12 and 13 clustering closer to the stage 10.5 replicates for the other lineages. To quantify this, the distance between stages 9 and 13 in the full gene expression space was calculated and here too the neural lineage was found to move the shortest distance (Figure 2.5 C). Whereas PC1 correlates with developmental time, PC2 appears to distinguish the different lineage states, which after stage 10.5 show very distinct trajectories. Interestingly, endoderm and epidermis lie furthest from each other along PC2 with mesoderm lying between those states. Thus, PC2 captures the intermediate nature of the ventral mesoderm state which shares GRN features with endoderm but, like epidermis, is BMP-driven (Figure 2.5 A).



**Figure 2.4 PCA Reveals Path to Neural Lineage is One-Dimensional.** PCA for each individual lineage with the coordinates of PC1 and PC2 for each lineage plotted against developmental time.



**Figure 2.5 PCA Reveals Stage 9 to 13 Distance is Shortest in Neural Lineage.** (A) PCA performed on all four lineages simultaneously, with plot showing clustering of all lineages for PC1 vs PC2. (B) Scree plot of the variance explained by the top 10 principal components for PCA done on all lineages. (C) Distances from stage 9 to 13 for all 74 PCs for each lineage, error bars are SEM of all stage 9-13 distances for each lineage, (\*\*\*) $P < 0.005$ , (\*\*) $P < 0.01$ . (D) PCA on only zygotic transcripts with plot showing clustering for all lineages PC1 vs PC2.

As some maternally provided transcripts persist through stage 10.5 and could bias transcriptome changes, we also performed PCA on zygotic transcripts only, removing maternally provided genes (Paranjpe et al. 2013) with no zygotic transcription through stage 13 from the analysis (Figure 2.5 D). The PCA was largely unchanged using this gene set, indicating that the observed dynamics are driven primarily by zygotic transcripts.

We next used differential expression analysis (DESeq2) to gain insights into the dynamics of gene expression changes across the four state transitions. Plotting the number of genes differentially expressed between successive developmental stages revealed that the number of genes whose expression changes significantly during these state transitions is relatively modest (Figure 2.6, Supplemental Table 4). For example, between two and three thousand genes are differentially expressed between stages 9 and 10 in each lineage, which represents four to six percent of the transcriptome. Between stages 10 and 10.5 the gene expression changes in the endodermal lineage are strikingly different from those of the other lineages. The number of differentially expressed genes increases almost 60% between these stages in the endoderm, whereas there is a significant decrease in differentially expressed genes in the epidermal and neural lineages and little change in the mesoderm. While there is a gradual decrease in differentially expressed genes in the endodermal lineage between stages 10.5 and 13, this state transition continues to exhibit the most dynamic changes in gene expression compared to the other lineages. After reaching a minimum between stages 10.5 and 11 both the epidermal and ventral mesodermal lineages exhibit increasing numbers of differentially expressed genes. By stages 12-13 the endoderm and ventral mesoderm exhibit comparable gene expression dynamics. Interestingly, almost a quarter of the genes changing in the endodermal and ventral mesodermal trajectories between stages 12 and 13 are shared between these lineages, likely reflecting the overlapping landscape of the combined mesendoderm GRN (Figure 2.7 I-J). In contrast to the other three lineages, the neural trajectory exhibits very few differentially expressed genes after stage 11, providing additional evidence that this lineage reaches an early equilibrium (Figure 2.6, 2.7 G-J, Supplemental Table 5).

### **Gene Expression Dynamics Provide Novel Insights into the Neural Default State**

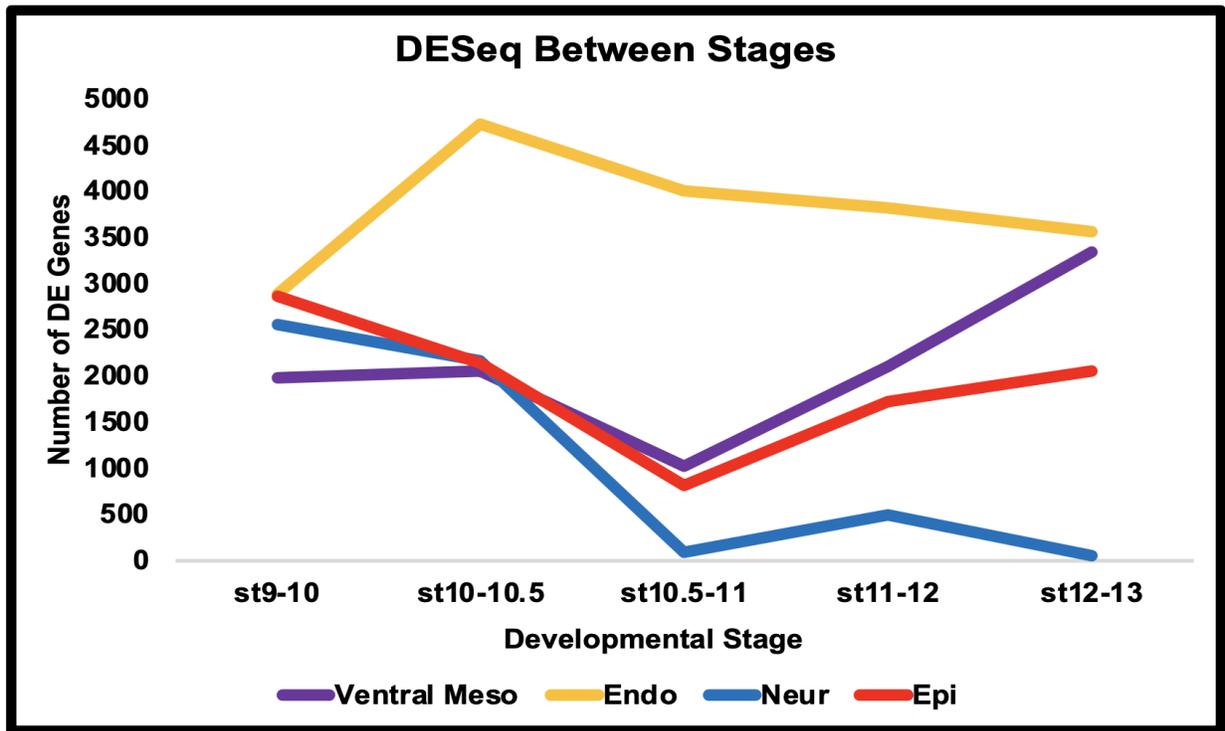
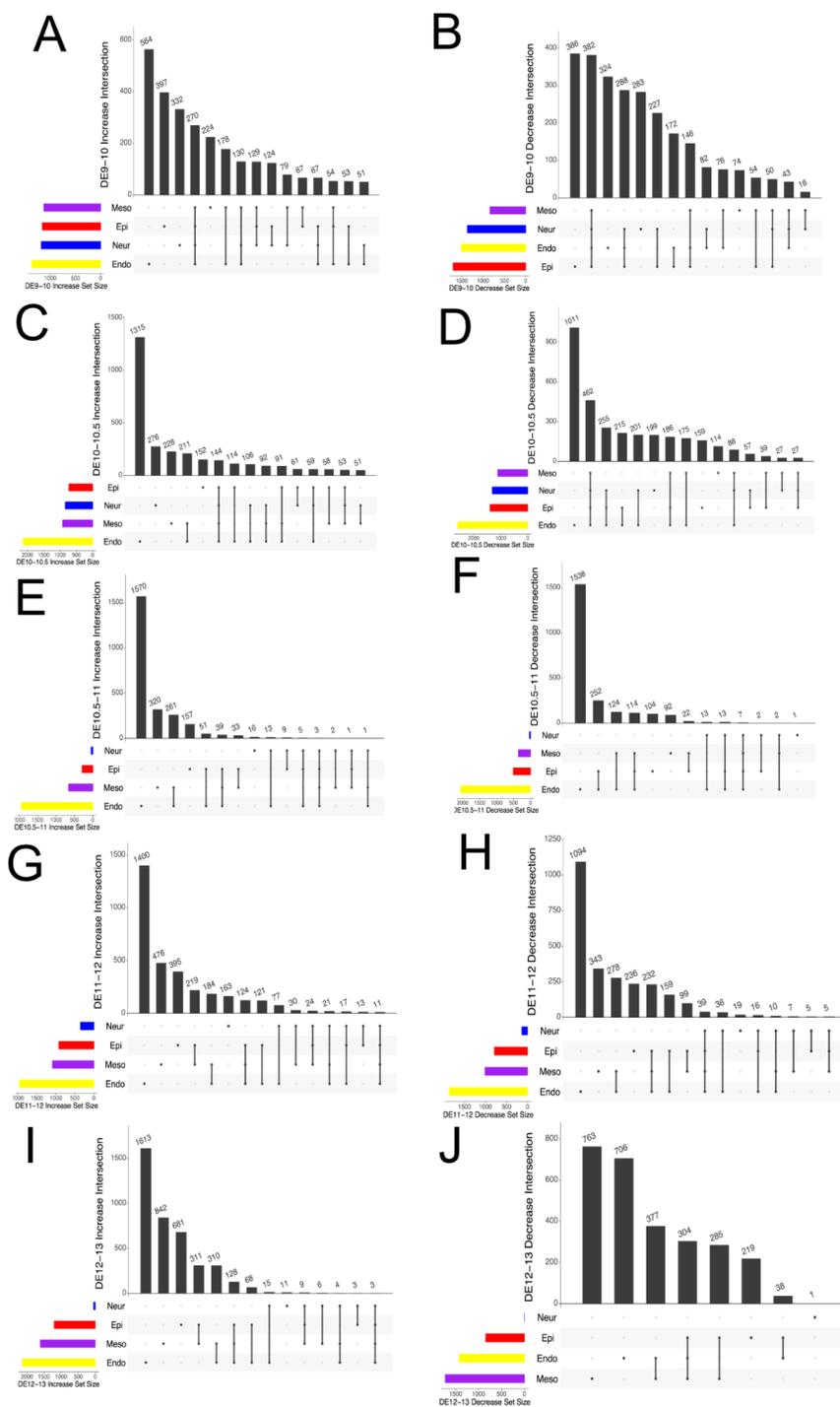
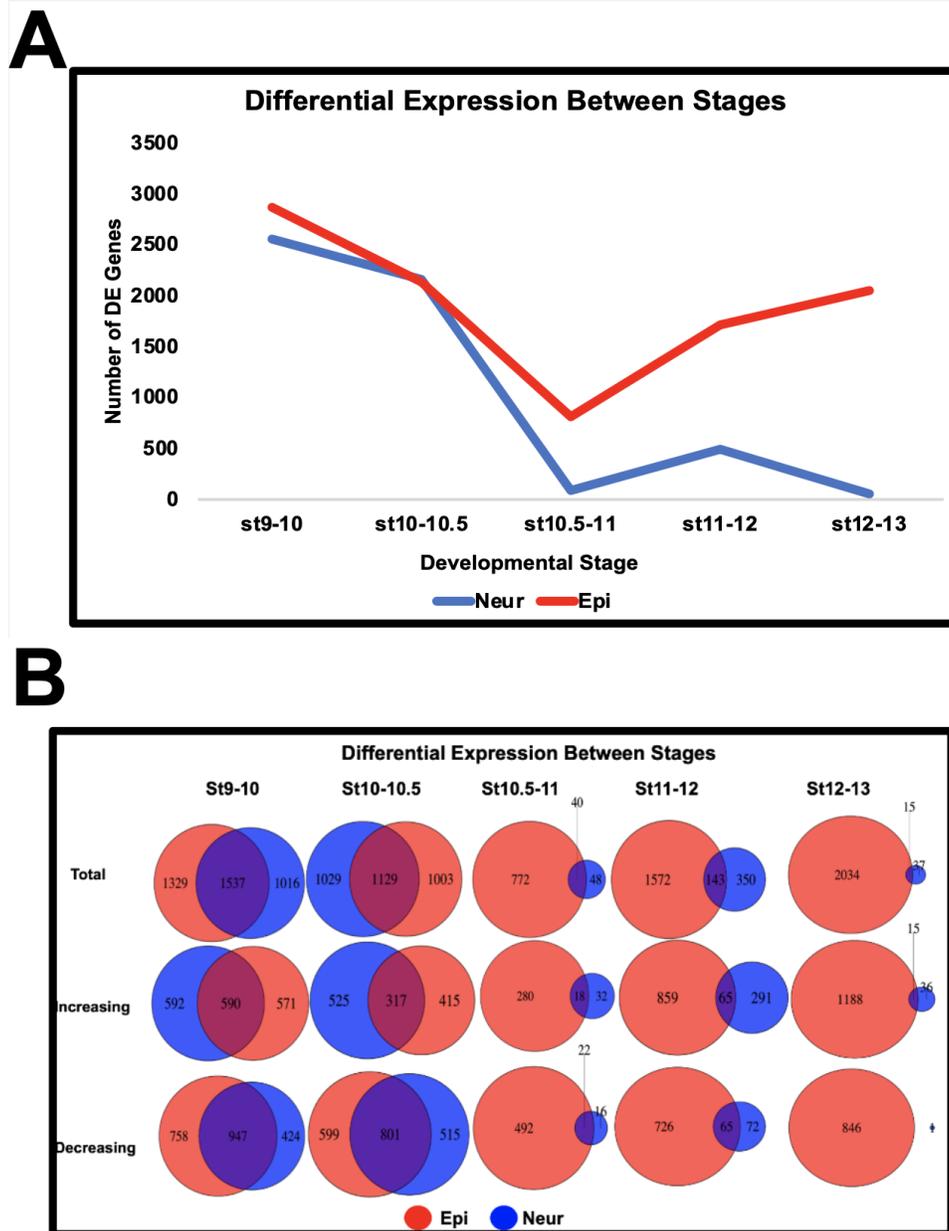


Figure 2.6 Temporal DESeq Reveals Neural Lineage Reaches Early Equilibrium. Number of genes differentially expressed between successive stages in each lineage,  $p_{adj} \leq 0.05$ , no fold change cut off.



**Figure 2.7 Each Lineage has both Unique and Shared Transcriptome Dynamics.** (A-E) UpSet plots of temporally differentially expressed genes (A) increasing between stages 9 and 10, (B) decreasing 9-10, (C) increasing 10-10.5, (D) decreasing 10-10.5, (E) increasing 10.5-11 (F) decreasing 10.5-11 (G) increasing 11-12, (H) decreasing 11-12, (I) increasing 12-13, (J) decreasing 12-13. X-axes show genes unique to each lineage and overlapping in all combinations of lineages ordered from largest number of genes to smallest.

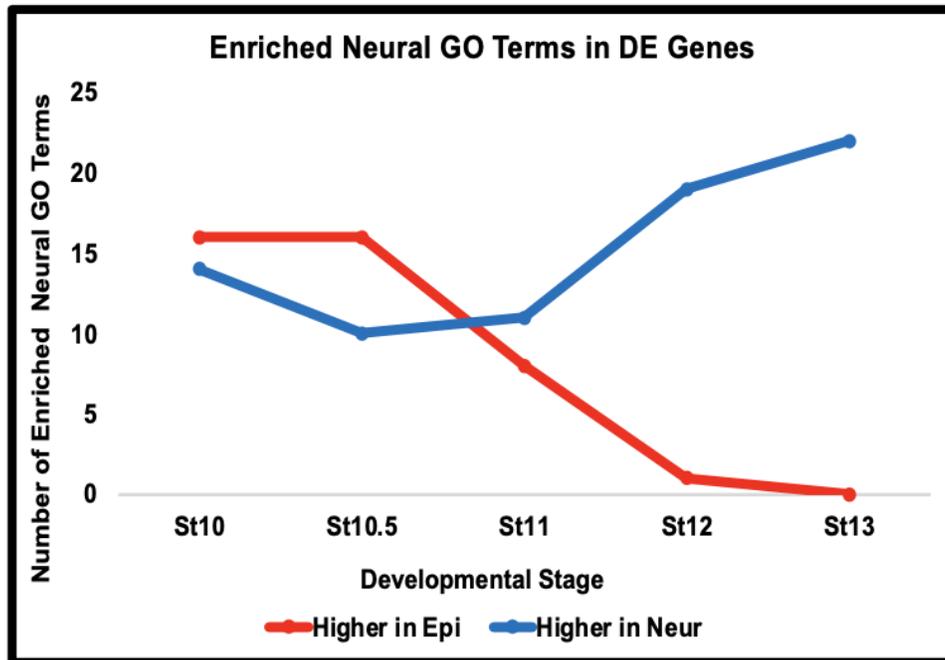


**Figure 2.8 Epidermal and Neural Lineages Show Early Overlapping Dynamics.** (A) Number of differentially expressed genes between successive developmental stages of the epidermal and neural lineage,  $p_{\text{adj}} \leq 0.05$  (B) Venn Diagrams for the total number of DE genes between stages for the epidermal and neural lineages, as well as the number of genes increasing and decreasing over time.

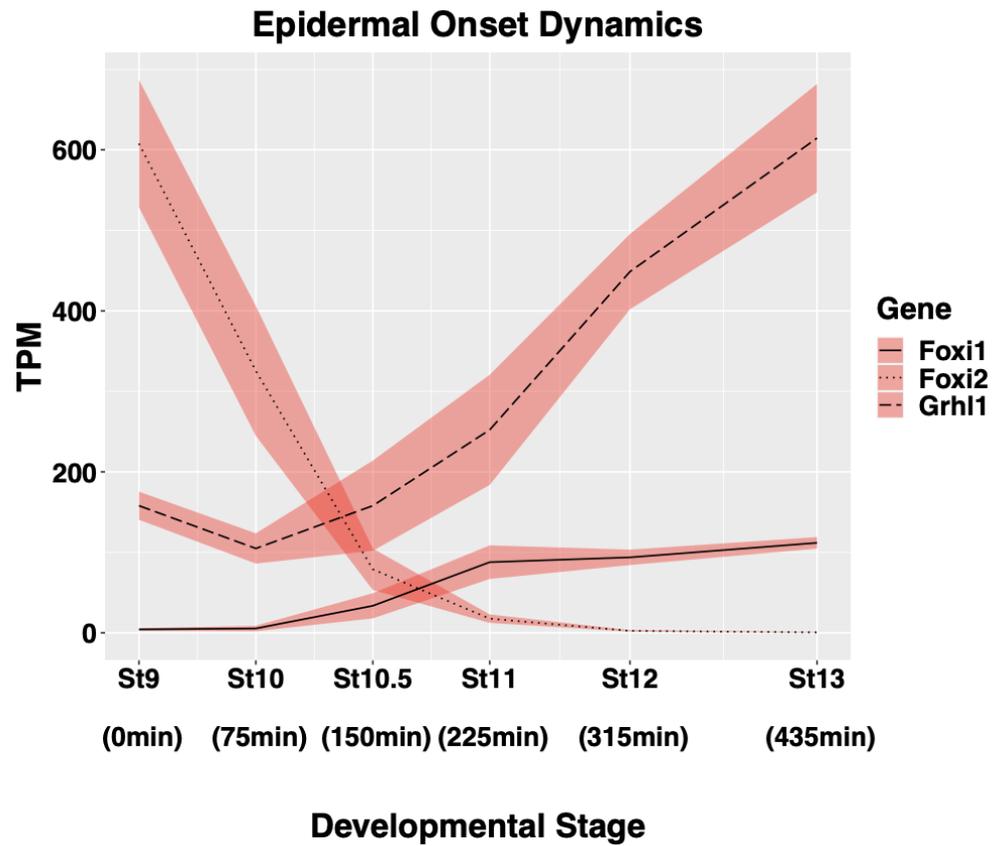
The above analyses suggested that the neural progenitor state follows a simpler trajectory than that of the other three lineages. Comparison of gene expression dynamics during transit to an epidermal versus a neural state reveals that through stage 10.5 these two lineages share a remarkably similar trajectory (Figure 2.8 A). For each, the number of genes differentially expressed between successive stages decreases, the number of genes changing is highly similar, and there is significant overlap in the genes exhibiting differential expression. Between stages 9 and 10, for example, 60% of the genes differentially expressed in the neural lineage are also differentially expressed in the epidermal lineage and that is true of 53% of genes differentially expressed between stages 10 and 10.5 (Figure 2.8 B, Supplemental Table 6). The majority of shared genes exhibit decreasing expression during these stages, in part reflecting the downregulation of pluripotency genes. Nevertheless, of the 1161 genes whose expression increases in the neural trajectory between stages 9 and 10, more than half also show increased expression in the epidermal trajectory.

Between stages 10.5 and 11, which corresponds to early gastrulation, there is a striking divergence of the epidermal and neural lineages. During these stages, gene expression dynamics largely cease in the neural lineage; fewer than 100 genes are differentially expressed between stages 10.5 and 11, and only 52 genes are differentially expressed between stages 12 and 13. By contrast, in the epidermal lineage temporal changes in gene expression begin to sharply increase and the overlap of these genes with those changing in the neural lineage is minimal (Figure 2.8 A, B). Of all genes differentially expressed between stages 12 and 13 in these two stage transitions, less than 1% are differentially expressed in both the neural and epidermal lineages, mainly because the neural lineage has ceased to change. Interestingly, the increasing gene expression dynamics that characterizes the epidermal lineage at stage 11 coincides with a loss of enrichment for neural GO terms (Figure 2.9, Supplemental Table 7). This suggests that pluripotent blastula cells possess neural-like features that begin to be lost around the onset of gastrulation as cells transit to an epidermal state but are retained and reinforced in neural progenitor cells.

The observed gene expression dynamics and GO term enrichment together point to the onset of gastrulation as a critical point on the landscape topology of early developmental decision making. To



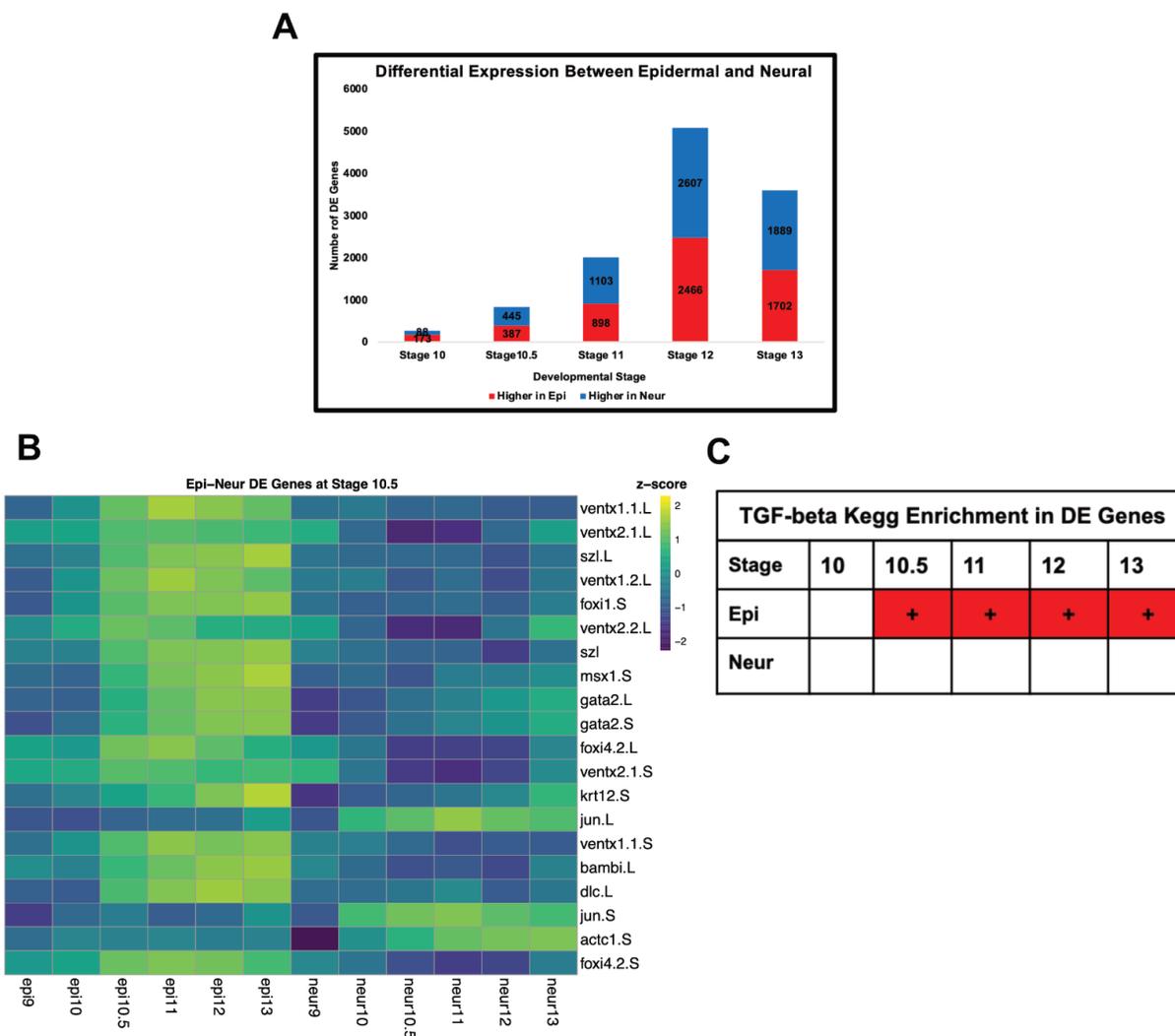
**Figure 2.9 Neural GO Terms are not Specific to Neural Lineage Until Stage 11.** Number of enriched Neural GO Terms in genes significantly higher in the neural lineage (blue) and epidermal lineage (red) at each developmental stage.



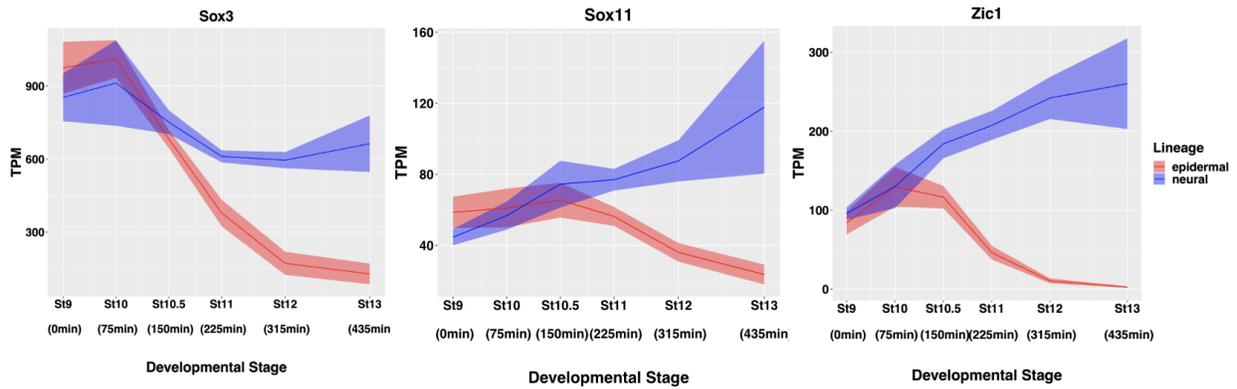
**Figure 2.10 Gene Expression Dynamics of Epidermal Formation.** TPM of *Foxi1*, *Foxi2* and *Grhl1*, revealing epidermal onset dynamics. Graph shows sum of S+L allele. Width of the line represents SEM of three biological replicates.

further explore this, we examined the expression of *Foxi1* and *Grhl1* which are key upstream components of the GRN mediating the formation of epidermis (Mir et al. 2007; Tao et al. 2005). Expression of *Foxi1* has been shown to be activated by the pluripotency factor *Foxi2* (Cha et al. 2012). Examining the expression dynamics of these three genes across the epidermal trajectory reveals a sharp increase in the expression of *Grhl1* that correlates with rapidly extinguishing expression of *Foxi2* and the onset of gastrulation (stage 10) (Figure 2.10). As early gastrulation is also when gene expression dynamics virtually cease in the neural trajectory, this suggests systems dynamics that favor a neural progenitor state over that of other lineages and a network structure that requires cells to be actively propelled toward an alternative, in this case epidermal, state.

To further examine the genes that distinguish the neural and epidermal states we analyzed the genes differentially expressed between these states at each time point on their trajectories. At stage 10 the two lineages remain strikingly similar, with only 261 genes differentially expressed (Figure 2.11 A, Supplemental Table 8). The number of differentially expressed genes increases by more than 500% between stages 10.5 and 12, driven almost entirely by gene expression dynamics in the epidermal lineage. Interestingly, 13 of the top 20 most differentially expressed genes at stage 10.5 are known BMP responsive genes and all but three, *Jun.L/S* and *Actc1.S*, are more highly expressed in the epidermal lineage (Figure 2.11 B). Consistent with this, beginning at stage 10.5 genes associated with the TGF-beta pathway in the KEGG database show enrichment in the epidermal lineage (Figure 2.11 C). As stage 10.5 represents the time when the trajectories of the neural and epidermal lineages diverge after neural reaches early equilibrium, this enrichment is consistent with a model where BMP signals actively propel cells away from a neural well, the state favored by the systems dynamics absent perturbation of the landscape, and onto the path toward an epidermal state. Together, the early equilibrium reached by the neural lineage, combined with the neural features of the pluripotent state provide new context for the neural default hypothesis. Complementing prior experimental studies, our transcriptome data suggests that neural is the default state following exit from pluripotency because of the shorter and more linear path from the pluripotent state to the neural progenitor state. This is further evidenced by the expression dynamics of genes that play important roles in both pluripotent cells and neural progenitors such as *Sox3*,



**Figure 2.11 Between Lineage DESeq Shows Epidermal and Neural Differences.** (A) Number of differentially expressed genes between lineages at each developmental stage ( $p_{adj} \leq 0.05$ ). (B) Heatmap of the top 20 DE genes by Log2FC with a minimum expression of 10TPM between the epidermal and neural lineages at stage 10.5 (C) KEGG enrichment analysis of genes differentially expressed between epidermal and neural lineage at each developmental stage. Genes significantly increased in the epidermal lineage are enriched for TGF-beta genes, as defined by KEGG database from stages 10.5-13.



**Figure 2.12 Expression Dynamics of Pluripotency Genes Maintained in Neural.** Graphs of the TPM of three pluripotency markers maintained only in the neural lineage *Sox3*, *Sox11*, and *Zic1*. Graphs are sums of S+L allele. Width of the line represents SEM of three biological replicates.

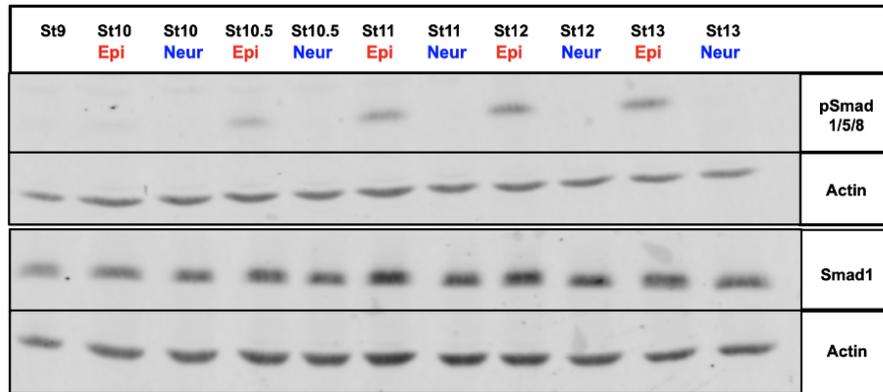
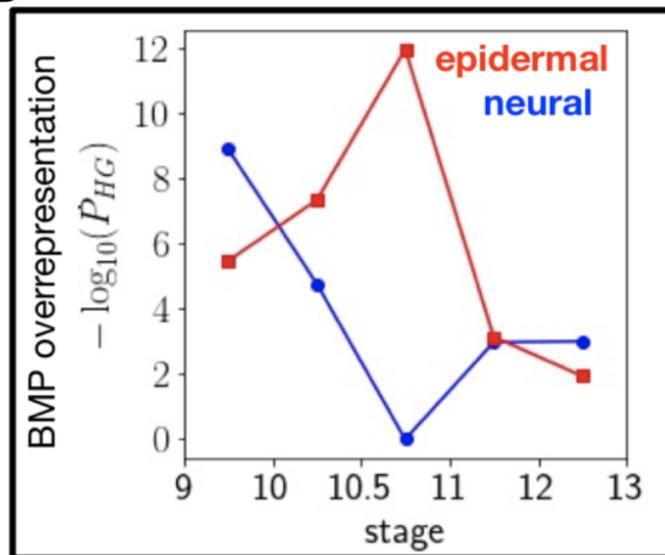
*Sox11* and *Zic1* (Penzel et al. 2003, Hyodo-Miura et al. 2002; Mizuseki et al. 1998, Nakata et al. 1998). All three of these genes retain or increase their expression in the neural lineage but are rapidly downregulated in the epidermal lineage after stage 10.5 (Figure 2.12).

### **Robust BMP Signaling is Initiated in Explants Around the Onset of Gastrulation**

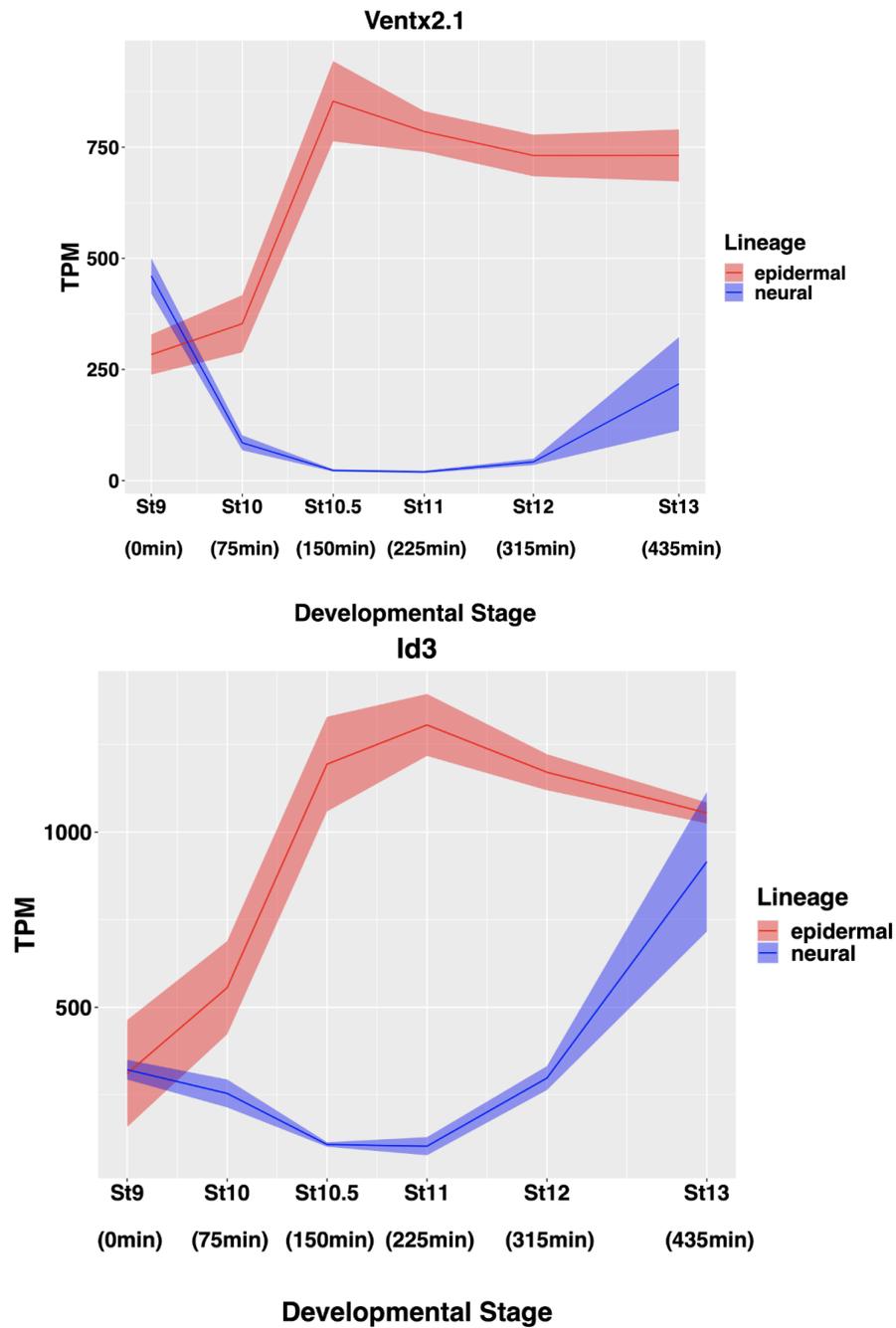
Consistent with a model where BMP signals actively propel cells away from the neural state, phosphorylation of BMP R-Smads is first detected in animal pole explants at stage 10.5 (Figure 2.13 A). Translocation of pSmad1/5/8 to the nucleus drives expression of BMP responsive genes, and its timing correlates with the divergence of the epidermal lineage from neural lineage. To gain further insights into the timing of BMP responsiveness we examined whether genes differentially expressed at successive developmental stages displayed over-representation of genes associated with BMP signaling (Kanehisa and Goto 2000) using the DESeq2 Wald test. We computed the significance at which these BMP associated genes comprised a larger fraction than would be expected by random chance via the hypergeometric  $p$ -value (Virtanen et al. 2020). The greatest divergence in overrepresentation between the epidermal and neural lineages was seen between stages 10.5 and 11, which was also the maxima for overrepresentation in the epidermal data (Figure 2.13 B). Consistent with this finding, *Ventx2.1* and *Id3*, which are both BMP targets genes, exhibit expression maxima in the epidermal lineage and minima in the neural lineages at these stages (Figure 2.14) (Onichtchouk et al. 1996, Hollnagel et al. 1999). The expression of *Id3* across these two state transitions is particularly noteworthy for its opposite intermediate non-monotonic dynamics at successive developmental stages despite comparable expression at the start and end of these lineage trajectories.

### **Early Response to Activin and BMP4/7 Reveals Unexpected Overlap**

While BMP signaling plays a central role in instructing pluripotent cells to form epidermis, it is the other branch of the TGF-beta that directs mesendoderm fates. Members of the Activin / Nodal / Vg1 / GDF1 / TGF-beta subfamily act via pSmad2/3 to activate mesodermal and endodermal target genes including *Foxh1*, *Eomes*, *Mixer*, *Tcf3* (also known as *E2a*) and *Tp53* (Chen et al. 1996, Ryan et al. 1996,

**A****B**

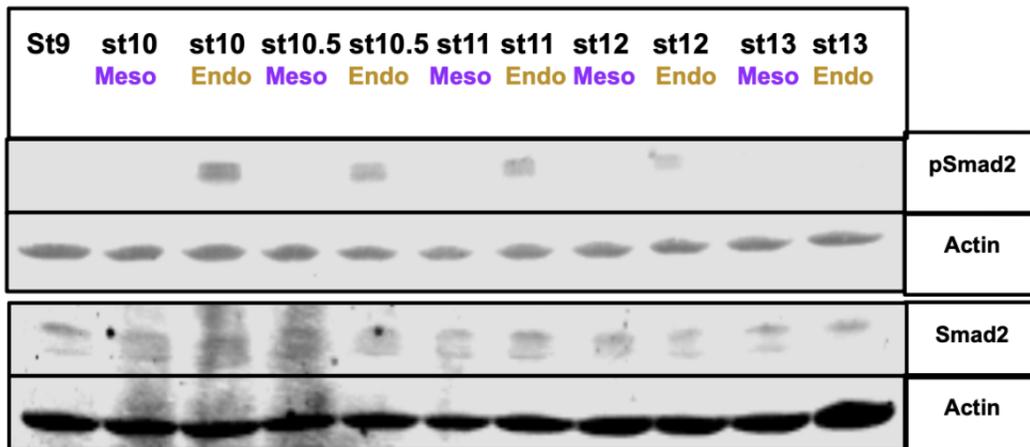
**Figure 2.13 BMP Signaling is Detectable at Epidermal-Neural Divergence Point.** (A) Western blot analysis of lysates of developing epidermal (WT) and neural (20uM K02288) explants for pSmad1/5/8 and Smad1 with Actin loading control. (B) Significance of BMP overrepresentation (hypergeometric p-value) in temporally differentially expressed genes.



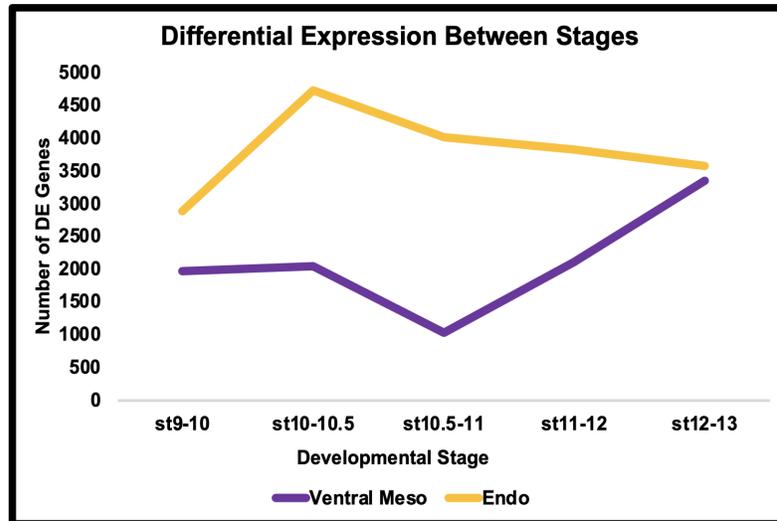
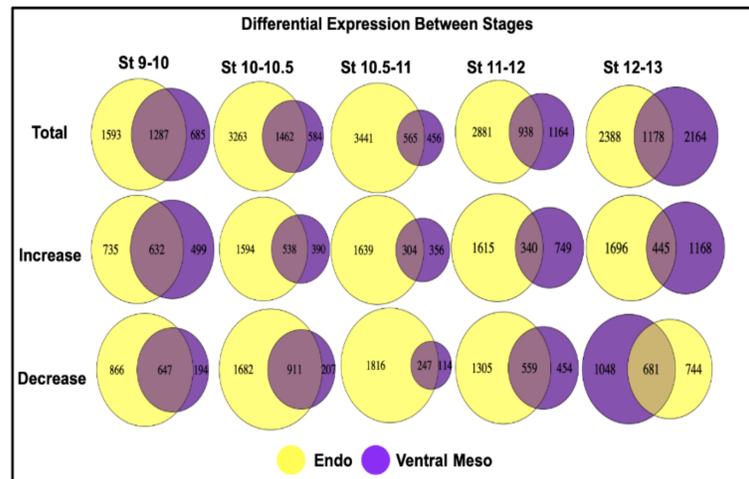
**Figure 2.14 Gene Expression of BMP Targets.** Graphs of BMP responsive genes *Ventx2.1* and *Id3* in epidermal and neural lineages. Graphs are sums of S+L allele. Width of the line represents SEM of three biological replicates.

Henry and Melton 1998, Rashbass et al 1992, Wills and Baker 2015, Cordenonsi et al. 2003). While recent work has suggested that Vg1-Nodal heterodimers mediate this process endogenously (Montague and Schier 2017), Activin has long been used for efficient mesendoderm induction in *ex-vivo* assays (Smith et al. 1990, Green et al. 1992, Hemmati-Brivanlou et al. 1992). Like Activin, BMP4/7 heterodimers have been shown to be potent mesoderm inducers at physiological concentrations. However, BMP4/7 has been reported to induce only ventral mesodermal (Suzuki 1997b, Nishimatsu and Thomsen 1998), unlike Activin and Nodal which are able to induce both ventral and dorsal mesoderm at low concentrations, as well as endoderm at high concentrations. We focused on BMP4/7-mediated mesoderm induction for this analysis because the response to Activin/Nodal is a spectrum with no clear threshold cleanly distinguishing an endodermal versus mesodermal response. An additional advantage of focusing on BMP4/7-mediated mesoderm induction is that it provides an opportunity to directly compare the signaling dynamics and transcriptional responses to the two branches of TGF-beta signaling in the same quantitative framework.

Treatment with Activin at stage 9 leads to robust signaling at stage 10 as evidenced by Western detection of phosphorylated Smad2 (p-Smad2) (Figure 2.15) and robust induction of *Sox17* beginning at stage 10 (Figure 2.2 G). Interestingly, transit to an endodermal state is distinguished from the other lineage transitions by its unique transcriptome dynamics. It is the only lineage in which there is a large increase in differentially expressed genes between stage 9 and stage 10.5, early gastrulation, after which the number of differentially expressed genes decreases (Figure 2.6, 2.16A). By contrast, treatment of stage 9 explants with BMP4/7 does not elicit an immediate increase in differentially expressed genes, distinguishing the responses to the two different arms of TGF-beta signaling between successive developmental stages (Figure 2.16A). Instead, the number of genes that are differentially expressed between stages 9-10 and 10-10.5 following BMP4/7 treatment remains fairly constant, before decreasing between stages 10.5 and 11. Intriguingly, however, the genes that are differentially expressed between these early stages in response to Activin or BMP4/7 show significant overlap. For example, approximately 48% of genes differentially expressed between stages 9 and 10 in response to BMP4/7 are also differentially expressed between those stages in response to Activin, as are 71% of genes differentially



**Figure 2.15 Western Blot Shows Robust Transient Activin Response.** (A) Western blot analysis of lysates of developing mesoderm (20ng/uL BMP4/7) and endoderm (160ng/uL Activin) explants for pSmad2 and Smad2 with Actin loading control.

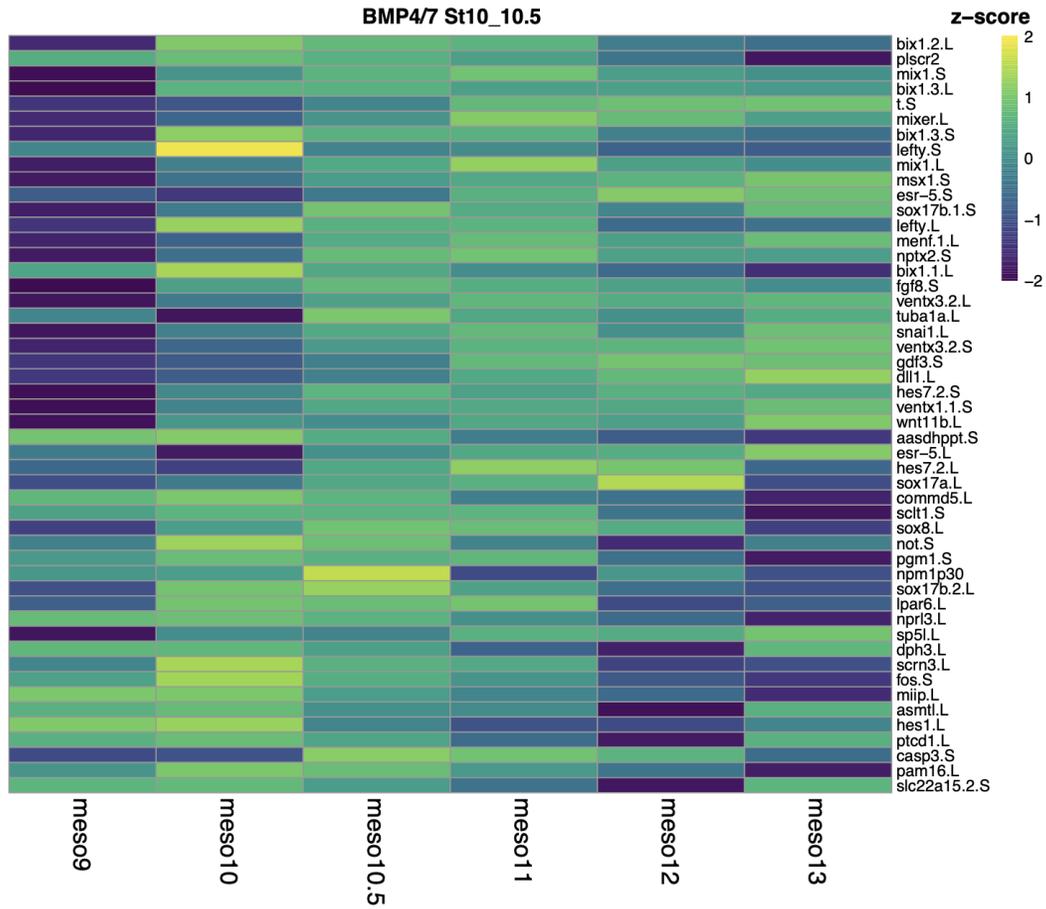
**A****B**

**Figure 2.16 Dynamics of Transcriptional Response to Activin and BMP.** (A) Number of differentially expressed genes between successive developmental stages of the endoderm and ventral mesoderm lineage,  $p_{\text{adj}} \leq 0.05$  (B) Venn Diagrams for the total number of DE genes between stages for the epidermal and neural lineages, as well as the number of genes increasing and decreasing over time.

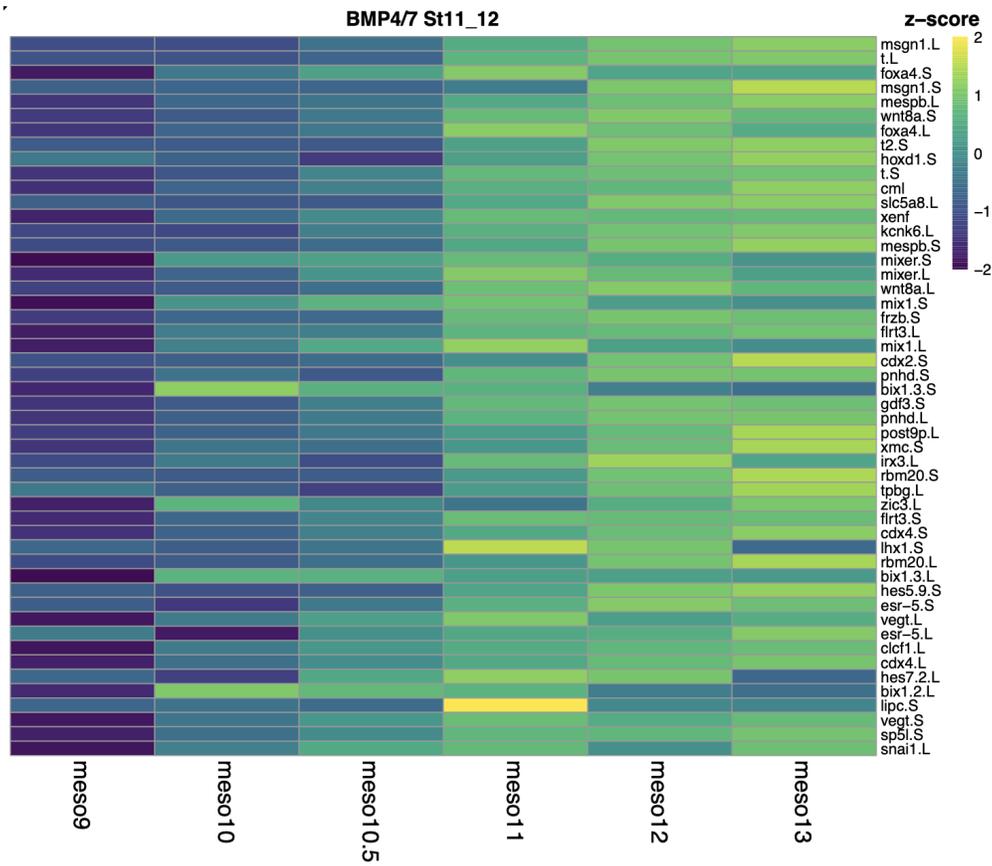
expressed in response to BMP4/7 between stages 10 and 10.5 (Figure 2.16B, Supplemental Table 9). These findings were unexpected as Smad1/5/8 and Smad2/3 generally regulate distinct target genes (Massagué and Wotton 2000, Wardle and Smith 2006).

To further explore the unexpected overlap in transcriptional responses to the two different classes of TGF-beta ligands we used DESeq2 to examine the genes that exhibit the largest log2 fold change compared to untreated explants in response to BMP4/7 treatment. Figure 2.17 shows the top 50 genes exhibiting the largest expression increase at stages 10 or 10.5 in response to BMP4/7 treatment relative to untreated explants. This gene set, which captures the immediate response to this ligand, includes previously characterized BMP target genes such as *Msx1* and *Ventx1* (Suzuki et al. 1997c, Rastegar et al. 1999) as well as the pan-mesodermal gene *Brachyury(T)* (Smith et al. 1991). Unexpectedly, it also includes a number of dorsal mesoderm/endoderm genes that have been characterized as targets of Activin/Nodal signaling including *Bix1*, *Mix1*, *Mixer* and *Sox17* (Tada et al. 1998, Rosa 1989, Chen et al. 1996, Hudson et al. 1997). Notably, activation of these genes occurs absent activation of pSmad2/3 (Figure 2.15). Using z-score scaling to visualize the expression dynamics of early responding genes across the time series revealed that genes generally associated with Activin/Nodal expression, including *Bix1* and *Mix1*, displayed non-monotonic dynamics with expression peaks at intermediate stages, whereas the expression of ventral and pan-mesodermal genes increased monotonically (Figure 2.17).

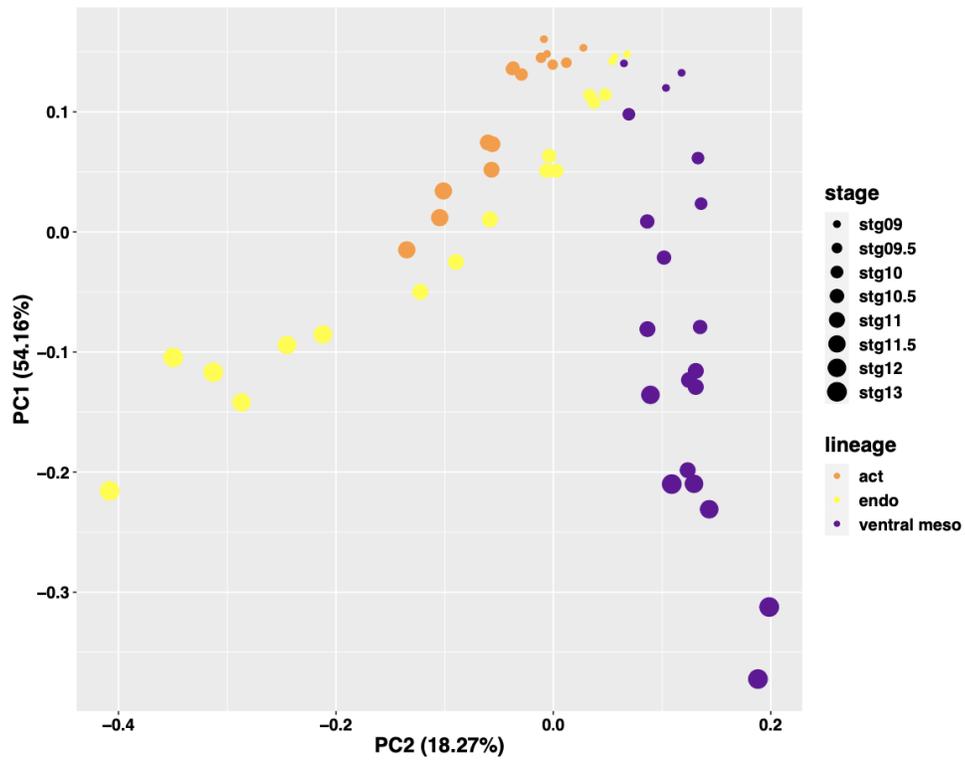
After reaching a minima between stages 10.5 and 11, the number of genes displaying dynamic expression changes in response to BMP4/7 greatly increased (Figure 2.16A). The genes exhibiting the largest log2 fold change at stages 11 or 12 were therefore similarly examined. This gene set is more enriched for ventral mesoderm associated genes than the initially responding genes, suggesting that the Activin-like response to BMP4/7 is transitory and that ventral mesoderm character is stabilized secondarily (Figure 2.18). This is consistent with a role for BMP signaling in actively ventralizing mesoderm and other tissues (Schmidt et al. 1995). We compared our endoderm and ventral mesoderm trajectories to a recently published Activin-induced mesoderm time series (Satou-Kobayashi et al. 2021) and found that the Activin treated samples cluster more closely relative to BMP4/7 treated samples (Figure 2.19). To ask if this is due to a greater enrichment of dorsal genes, we used a previously curated



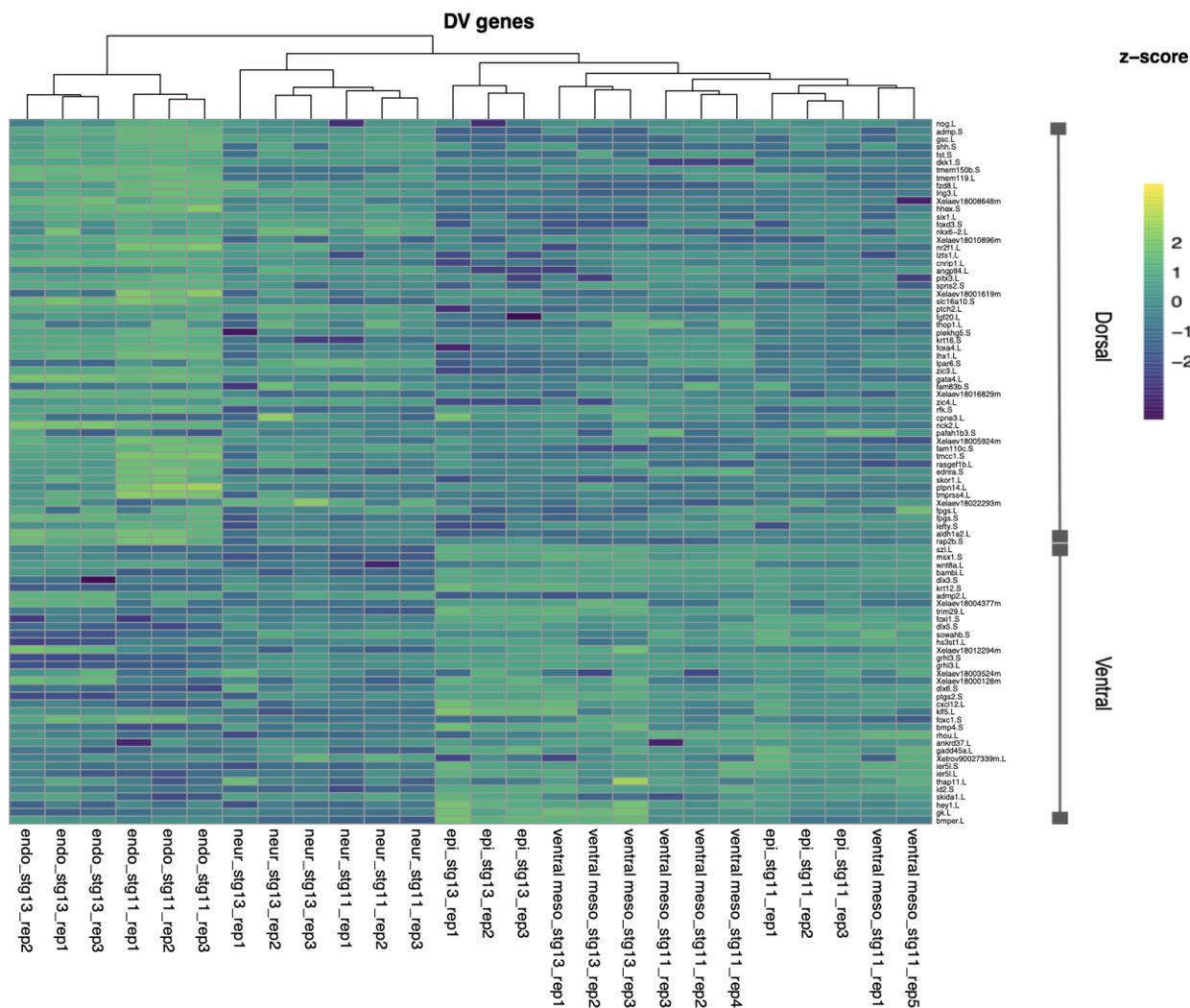
**Figure 2.17 Dynamics of Genes Immediately Responding to BMP.** Heatmaps of the top 50 genes differentially increased in response to BMP4/7 at stage 10 and/or 10.5 with expressions greater than 10 TPM.



**Figure 2.18 Dynamics of Genes with Later BMP Response.** Heatmaps of the top 50 genes differentially increased in response to BMP4/7 at stage 11 and/or 12, with expression greater than 10 TPM.



**Figure 2.19 PCA of Endoderm and Activin and BMP Generated Mesoderm.** PCA of published Activin induced mesoderm data (Satou-Kobayashi et al. 2021) with my ventral mesoderm and endoderm data shows that activin induced mesoderm clusters with my activin induced endoderm rather than BMP4/7 induced mesoderm.



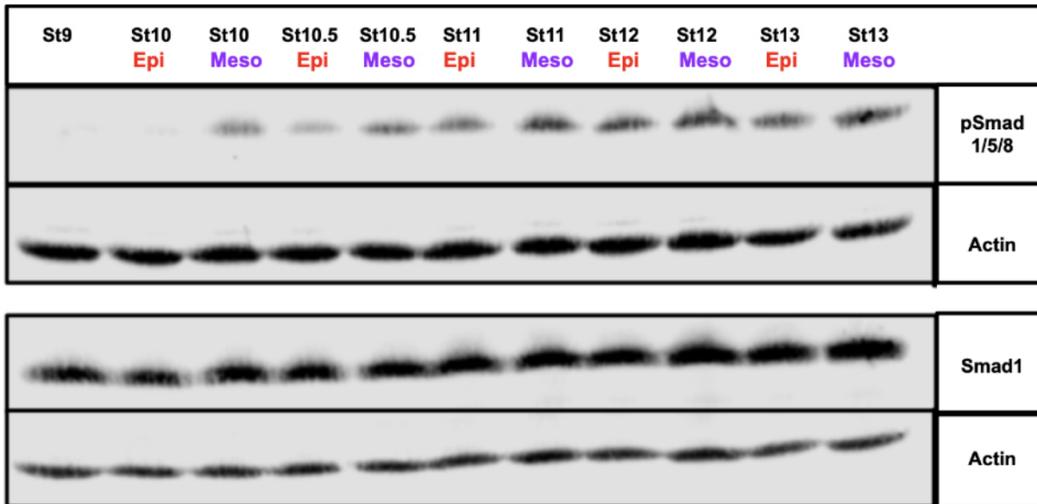
**Figure 2.20 Dorsal / Ventral Clustering.** Hierarchical clustering of Dorsal/Ventral genes (Ding et al. 2017) shows that endoderm and ventral mesoderm cluster on opposite sides of the heat map, suggesting BMP4/7 effectively ventralizes the mesoderm.

set of dorsally and ventrally induced genes (Ding et al. 2017). 31.0% of dorsally enriched genes were upregulated in the Activin-induced mesoderm samples as compared to 11.7% of ventrally enriched genes. By contrast, BMP4/7 treatment led to upregulation of 18.7% of dorsally enriched genes and 12.6% of ventrally enriched genes. Hierarchical clustering of stage 11 and 13 samples z-scored for dorsal and ventral genes identified by Ding et al. further highlights the more ventral character of BMP4/7-induced mesoderm (Figure 2.20).

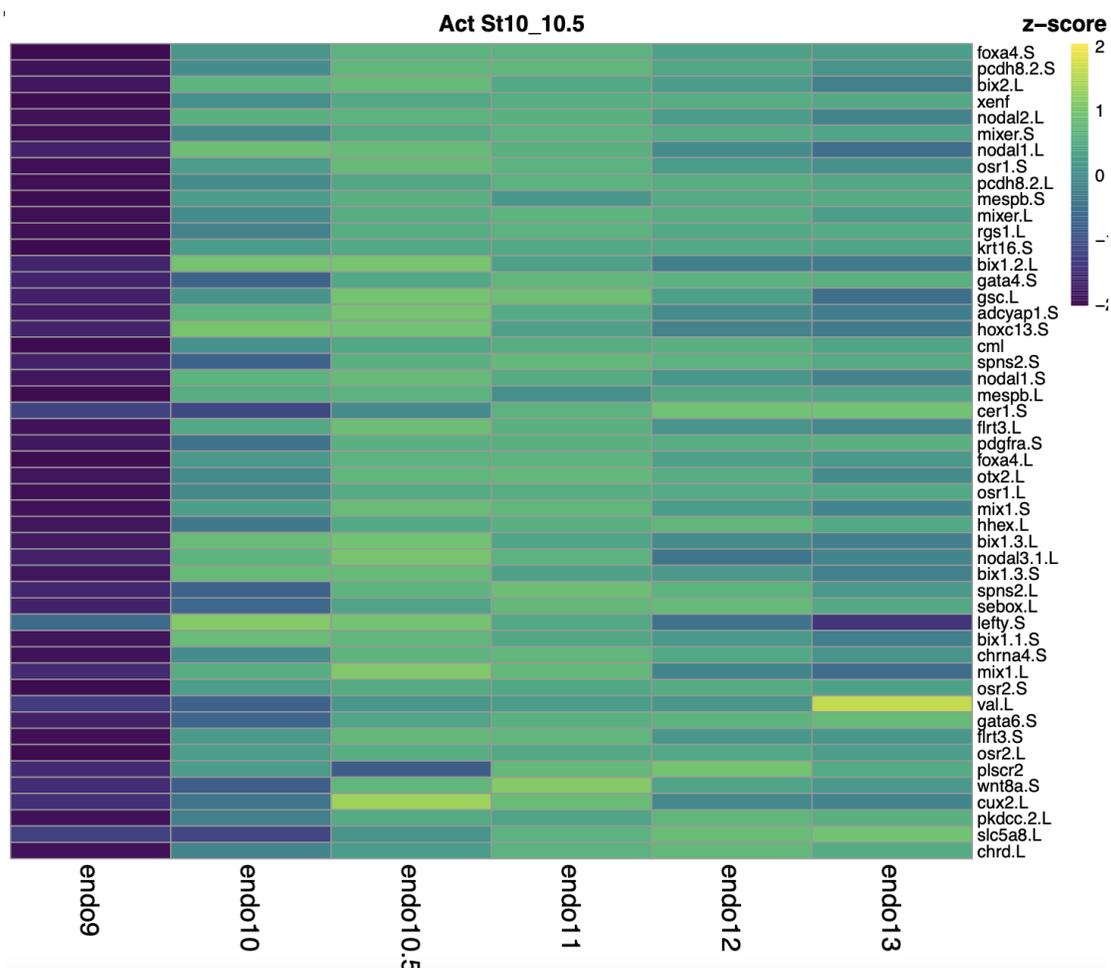
While treatment with both Activin and BMP4/7 was initiated at stage 9, Activin-mediated phosphorylation of Smad2 was transient (Figure 2.15), whereas BMP4/7-mediated phosphorylation of Smad1/5/8 persisted through stage 13 (Figure 2.21), likely contributing to the distinct gene expression dynamics in these two lineages. For example, the genes exhibiting the largest log<sub>2</sub> fold change compared to untreated explants at stages 10, 10.5 or 11, 12 in response to Activin treatment are enriched for those whose maximal expression occurs at intermediate stages of the lineage trajectory (Figure 2.22, 2.23). Importantly, while the genes activated as an early response to Activin or BMP4/7 show significant overlap (Figure 2.16B, Supplemental Table 9), the genes activated by these two classes of TGF-beta ligands nevertheless show significant differential expression with respect to one another (Figure 2.24 A, Supplemental Table 8). Over 1000 genes are differentially expressed between these trajectories as early as stage 10.5, and the number of genes differentially expressed between these lineages continues to increase over developmental time. Interestingly, analysis of KEGG pathway enrichment in these differentially expressed genes reveals that genes that are significantly higher in the endoderm lineage show enrichment for the TGF-beta pathway at stages 10 and 10.5, whereas genes that are significantly higher in the ventral mesoderm lineage are enriched for the TGF-beta pathway at stages 11 and 12 (Figure 2.24 B).

### **Time Series Data Provides Novel Insights into Mesendoderm GRN**

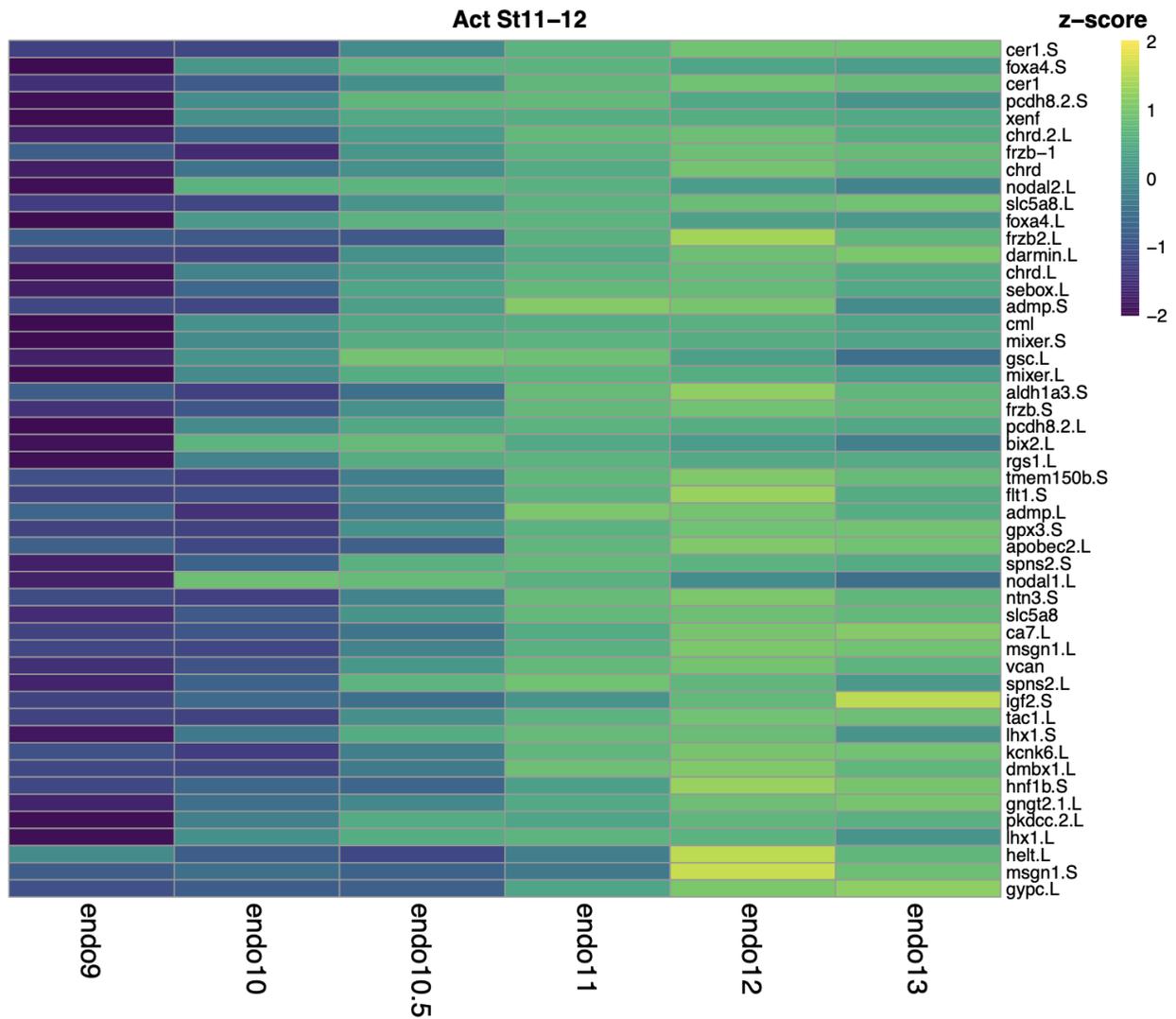
The mesendoderm gene regulatory network (GRN) has been extensively studied and has yielded a significant “parts list” of genes that make significant contributions to the formation of these lineages (Charney et al. 2017, Jansen et al 2022), however the ordering of the GRN components has lacked the



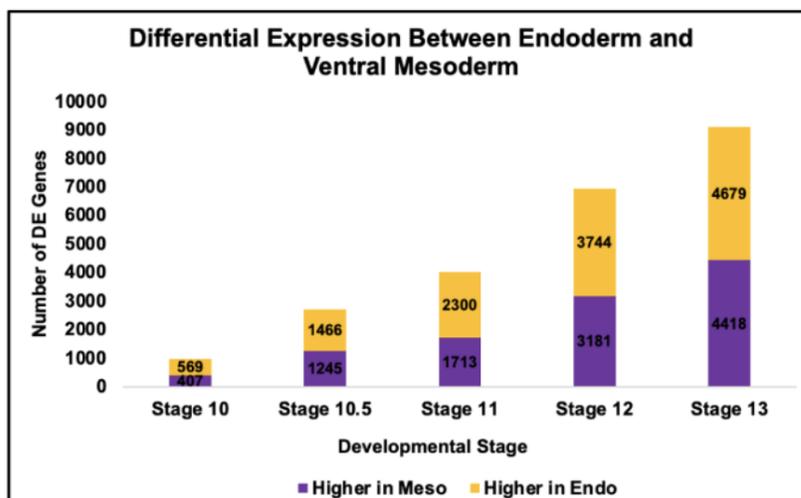
**Figure 2.21 Treatment with BMP4/7 Elicits Early BMP Signaling.** Western Blot Analysis of lysates for developing epidermal (WT) and ventral mesoderm (BMP4/7 20ng/uL) explants for pSmad1/5/8 and Smad1 with Actin loading control.



**Figure 2.22 Dynamics of Genes with Immediate Activin Response.** Heatmaps of the top 50 genes differentially increased in response to Activin at stage 11 and/or 12 with expression greater than 10 TPM.

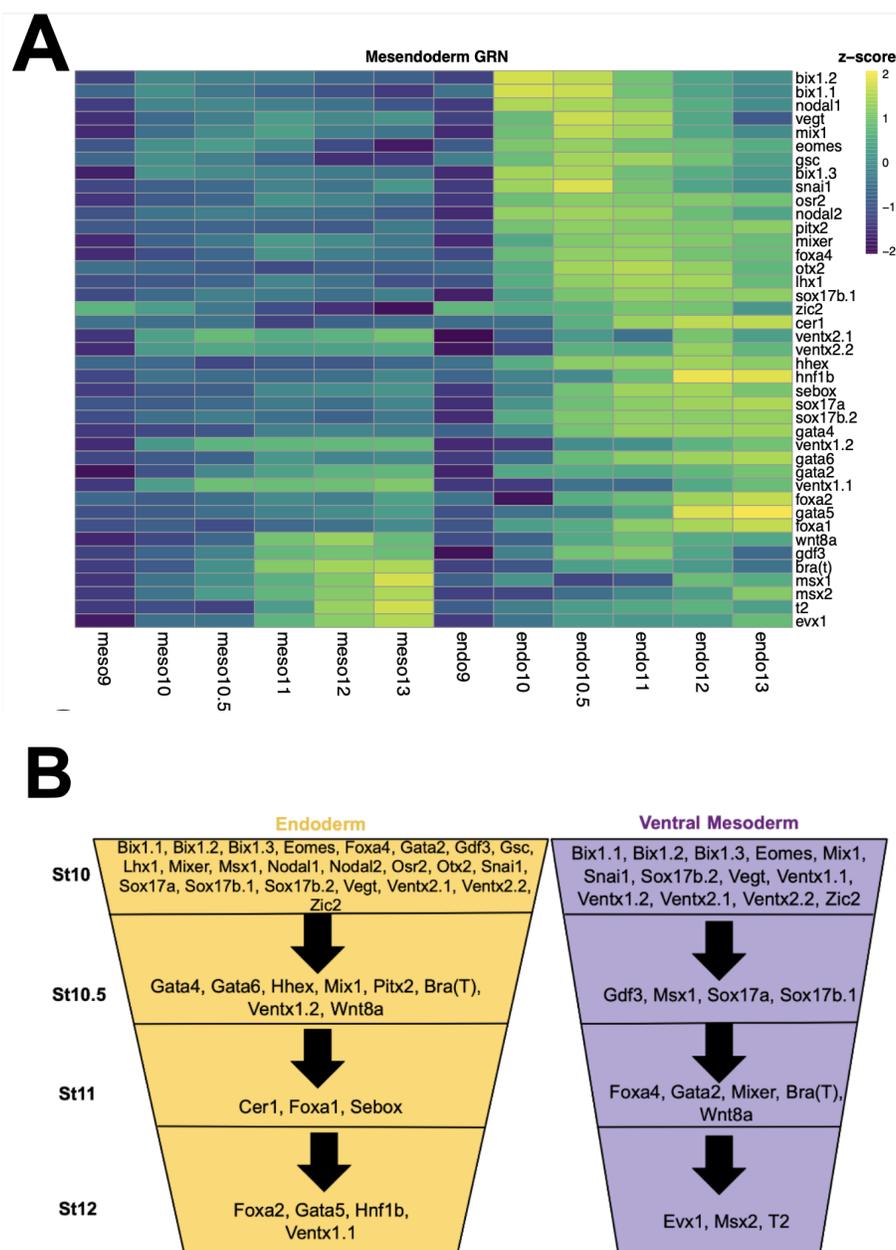


**Figure 2.23 Dynamics of Genes with Later Activin Response.** Heatmaps of the top 50 genes differentially increased in response to Activin at stage 11 and/or 12 with expression greater than 10 TPM.

**A****B**

TGF-beta Kegg Enrichment in DE Genes:					
Stage	10	10.5	11	12	13
Endo	+	+			
Meso			+	+	

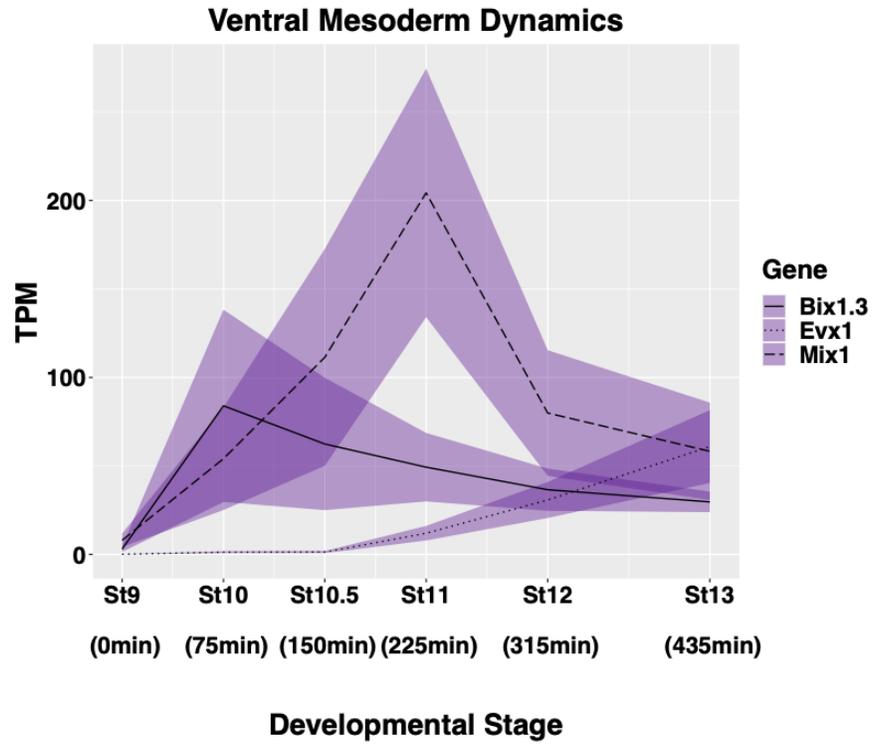
**Figure 2.24 DE Genes Between Mesoderm and Endoderm.** (A) Number of differentially expressed genes between the endoderm and ventral mesoderm lineages at each developmental stage ( $p_{adj} \leq 0.05$ ). (B) KEGG enrichment analysis of genes differentially expressed between ventral mesoderm and endoderm lineage at each developmental stage. Genes significantly increased in the endoderm lineage are enriched for TGF-beta genes, as defined by KEGG database from stages 10-10.5 and genes significantly higher in the ventral mesoderm lineage are enriched for TGF-beta genes for stages 11-12.



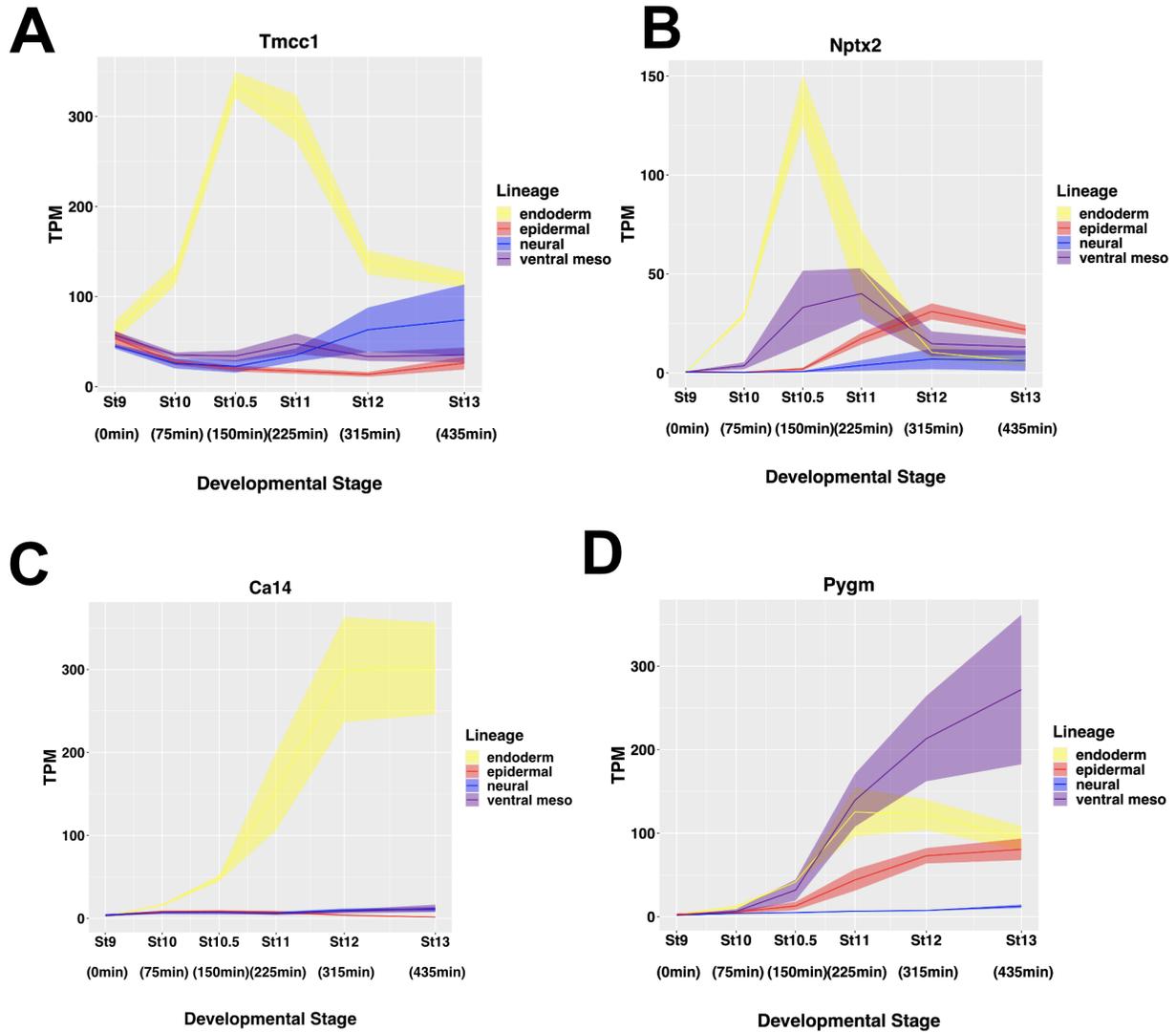
**Figure 2.25 Dynamics of Mesendoderm GRN.** (A) Heatmap of genes in the published Mesendoderm GRN across time in the endoderm and ventral mesoderm lineage with colors representing z-score of TPM expression in both lineages (Charney et al. 2017). (B) Schematic of the timing of genes from published Mesendoderm GRN in both the endoderm and ventral mesoderm lineage, as defined by expression of at least 30 TPM in the L and S allele combined for the average of three biological replicates. temporal resolution that our time series data can provide. Accordingly, we examined the expression of

forty-one validated mesendoderm GRN components across both the endoderm and ventral mesoderm lineage trajectories (Figure 2.25 A). Interestingly, many of these genes display non-monotonic expression, with their expression peaking at early time points before decreasing. These dynamics are particularly striking in the endoderm and demonstrate that many of these GRN components respond to Activin rapidly and robustly, but transiently. Indeed, expression of twenty-one of these genes is induced in the endoderm by stage 10, which is 75 minutes after ligand exposure (Figure 2.25 A, B). Interestingly, sixteen of these genes are also activated by BMP4/7 by stage 10, albeit less robustly, indicating that they are immediate responders to both classes of TGF-beta ligands. The Activin-induced expression of several GRN components, including Bix and Nodal family genes, *Vegt*, *Eomes*, *Mix1* and *Snai1*, peaks by stage 10.5 and then declines. By contrast, a second group of GRN factors, including *Sox17* and Gata2/6-related genes, *Osr2* and *Pitx2* exhibit sustained expression through stage 13. A third group of GRN components, including *Hnf1b*, *Foxa2* and *Gata5*, are not robustly turned on until mid- to late-gastrula stages (Figure 2.25 A, B). The expression of *Ventx* genes, known for their strong ventralizing activity, in response to high levels of Activin signaling was unexpected, and correlates with the downregulation of Nodal and Bix family factors. Surprisingly, *Ventx2.1* and *Ventx2.2*, which are among the first genes to respond robustly to BMP4/7, were induced as or more strongly by Activin, albeit with different temporal dynamics. Subsequently, GRN factors more closely associated with mesoderm, including *Evx1*, *Mix1*, and *Bix1.3* distinguish the BMP4/7 response from the Activin response. As with the Activin response, these BMP4/7 responding genes show distinct patterns of temporal dynamics (Figure 2.26).

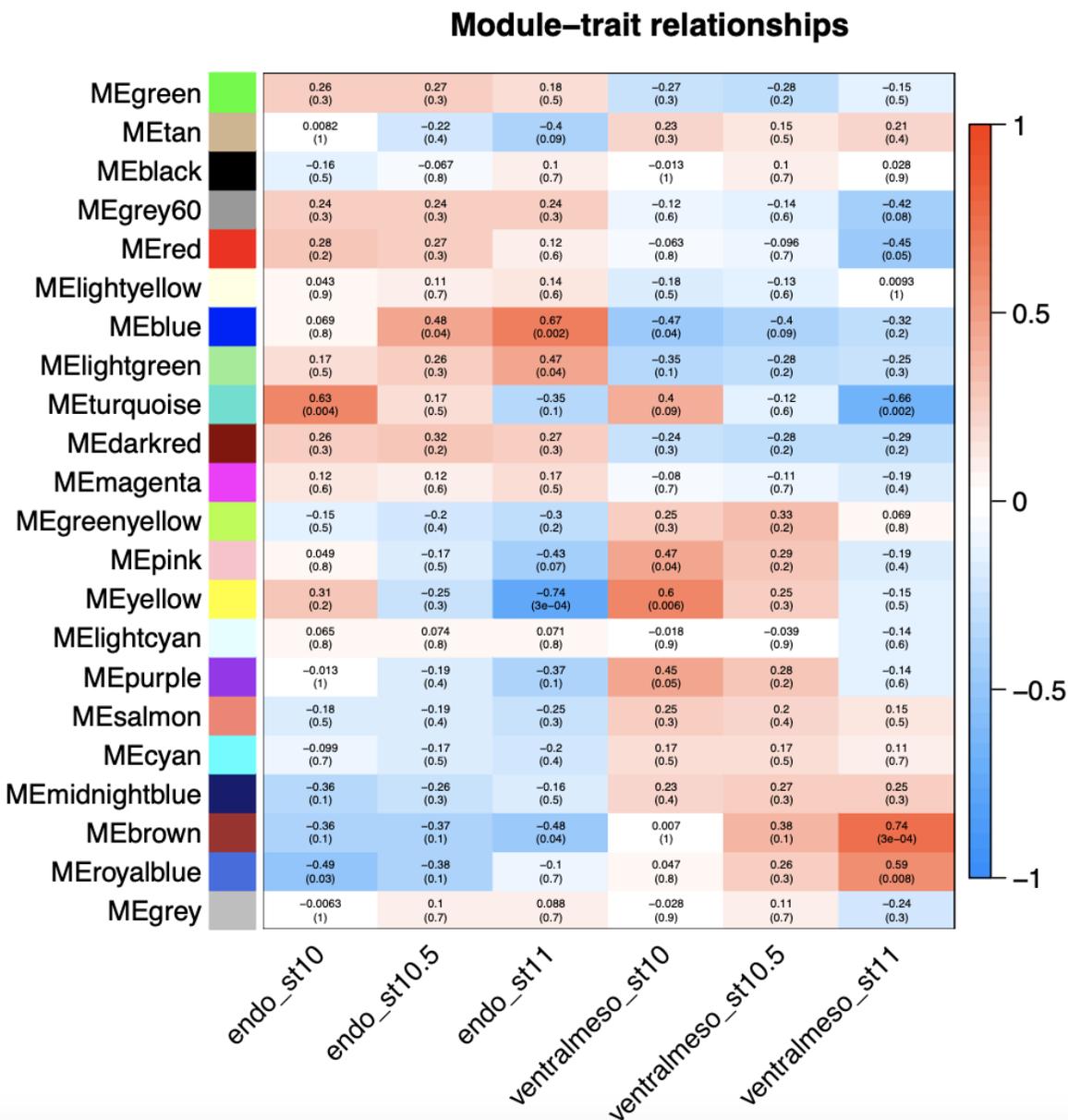
Given the distinct dynamics displayed by known mesendoderm GRN factors, we tested for differential linear and quadratic dynamics in the BMP4/7 and Activin responses using limma analysis (Law et al. 2014). This analysis identified genes displaying expression dynamics that position them as candidates for novel components the mesendoderm GRN. For example, one noted pattern was genes expressed rapidly and robustly in response to Activin but not BMP4/7, and downregulated after an early peak in expression. This pattern is exemplified by *Tmcc1* which is expressed non-monotonically in the endoderm trajectory (Figure 2.27 A). A second pattern described genes that respond to both Activin and



**Figure 2.26 Dynamics of Mesoderm Onset.** TPM of *Bix1.3*, *Evx1*, and *Mix1* revealing mesoderm onset dynamics. Graphs are sums of S+L allele. Width of the line represents SEM of three biological replicates.



**Figure 2.27 Proposed Mesendoderm GRN Candidates.** TPM of genes proposed as novel mesendoderm GRN members based on DESeq2 and Limma analysis (A) *Tmcc1* (B) *Nptx2* (C) *Ca14* (D) *Pygm*. Graphs are sums of S+L allele. Width of the line represents SEM of three biological replicates.



**Figure 2.28 WGCNA on Mesoderm and Endoderm Stages 10-11.** WGCNA on stages 10, 10.5 and 11 in the mesoderm and endoderm lineages identifies 22 gene modules, the blue module demonstrating increasing correlation over time to the endoderm lineage and the brown module demonstrating increasing correlation over time to the mesoderm lineage.

BMP4/7 with transient non-monotonic expression, such as *Nptx2* (Figure 2.27 B). A third pattern that emerged from this analysis was genes displaying a monotonic increase in expression only to Activin, such as *Ca14*, suggesting a role in the endoderm lineage specifically (Figure 2.27 C). Similarly, genes with monotonic and sustained expression in response to BMP4/7, such as *Pygm* may be novel mesoderm regulatory factors (Figure 2.27 D). While these genes are largely unstudied in mesendoderm formation, published transcriptome data sets provide further support for their involvement in mesendoderm formation. For example, *Ca14*, *Pygm* and *Tmcc1* were among genes upregulated in stage 12 animal caps in response to *Wnt* and *Nodal2* (Ding et al. 2018) and *Ca14*, *Tmcc1*, and *Nptx2* were upregulated in stage 11 embryos in response to somatic cell nuclear transfer from an endoderm cell (Hörmanseder et al. 2017). Similarly, *Pygm* has been identified as a target of *Myod* (McQueen and Pownall 2017). Thus, using *limma* analysis as an unbiased approach for detecting genes that share expression pattern dynamics allows identification of potential new members of developmental GRNs using our data sets.

To further examine the transition from a pluripotent state to endoderm vs. mesoderm we performed WGCNA on these lineages using data from stages 10, 10.5 and 11 when these state transitions display distinct dynamics. Two notable gene clusters emerged from WGCNA (blue and brown modules, Figure 2.28). These modules revealed correlations that increased over developmental time in either the endoderm (blue) or mesoderm (brown). Of the forty-one established mesendoderm GRN members, thirty were included in one of these two modules. Moreover, three of the novel GRN candidates identified by *limma* analysis were also found in one of these two clusters. The fourth of these was contained in a different module that also contained mesendoderm genes but had lower correlation values. Identification by WGCNA analysis of a cohort of genes significantly correlated with known mesendoderm GRN members provides candidate genes that may also play important roles in these two lineages (Supplemental Table 10).

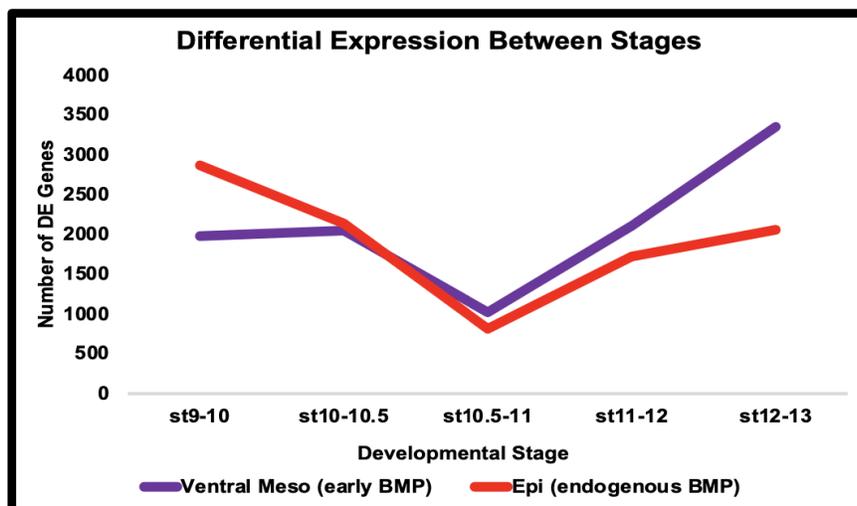
### **Early BMP Signaling Drives Ventral Mesoderm rather than Early Epidermal Divergence**

As discussed above, endogenous BMP4/7 signaling within animal pole cells will direct these cells to give rise to epidermis in the absence of BMP inhibitors, which in vivo are secreted by the organizer.

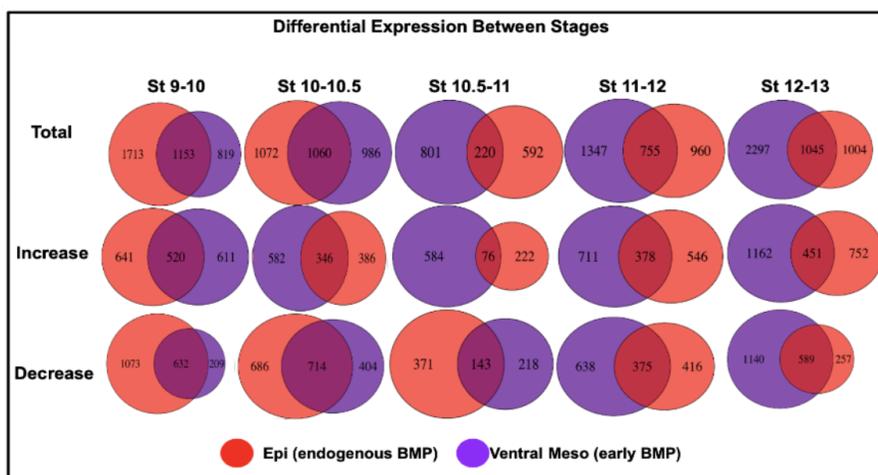
BMP signaling is detectable in these cells by stage 10.5, as evidenced by detection of pSmad1/5/8 (Figure 2.13). When explants are treated with exogenous BMP4/7 at stage 9, pSmad1/5/8 is detected at stage 10 at levels comparable to those seen at stage 10.5 in untreated explants (Figure 2.21). This allows a quantitative comparison of the transcriptional responses to the same signal when presented with shifted developmental timing – a tilting of the landscape topology. One predicted outcome of such a heterochronic shift might have been an accelerated transit to the epidermal state rather than formation of ventral mesoderm. To examine how shifting the timing of BMP activity alters the transcriptional response we first compared the transcriptome dynamics. In this context it is interesting to note that premature BMP signaling actually dampened early gene expression changes; there is an ~32% reduction in the number of genes whose expression significantly changes between stages 9 and 10 (Figure 2.29 A). This is driven almost entirely by a reduction in the number of genes whose expression decreases during those initial stages (Figure 2.29 B Supplemental Table 11). Between stages 10 and 11, the transcriptome dynamics are comparable in the two conditions (Figure 2.29 A, B). However, the number of genes displaying temporal differential expression that are shared by these two trajectories but not by the neural and endodermal state transitions (making them a general response) is remarkably low, ranging from 97 between stages 10 and 10.5, to 55 from stages 10.5 to 11 (Figure 2.7 C-F, Supplemental Table 5).

We next used DESeq2 to compare the genes differentially expressed in response to early BMP signaling. At stage 10 there are only 332 genes differentially expressed between these two conditions. The number of differentially expressed genes remains relatively low until stage 11, when it begins to increase over time (Figure 2.30 A, Supplemental Table 8). However, beginning at stage 10 the genes whose expression is higher in response to early BMP are enriched for mesoderm GO terms (Figure 2.30 B, Supplemental Table 12). Because expression of *Grhl1* is a key early driver of transit to the epidermal state (Tao et al. 2005), we examined its response to early BMP signaling. Strikingly, its expression fails to be robustly activated under this condition and the differential response is seen as early as stage 10.5 (Figure 2.31 A). Given this surprising finding we examined the expression of *Ventx2.1*, which is a direct target of BMP signaling (Onichtchouk et al. 1996). Here there was a shift in transcription dynamics that reflected the earlier onset of BMP signaling; *Ventx2.1* transcripts reach levels at stage 10 that would not

A

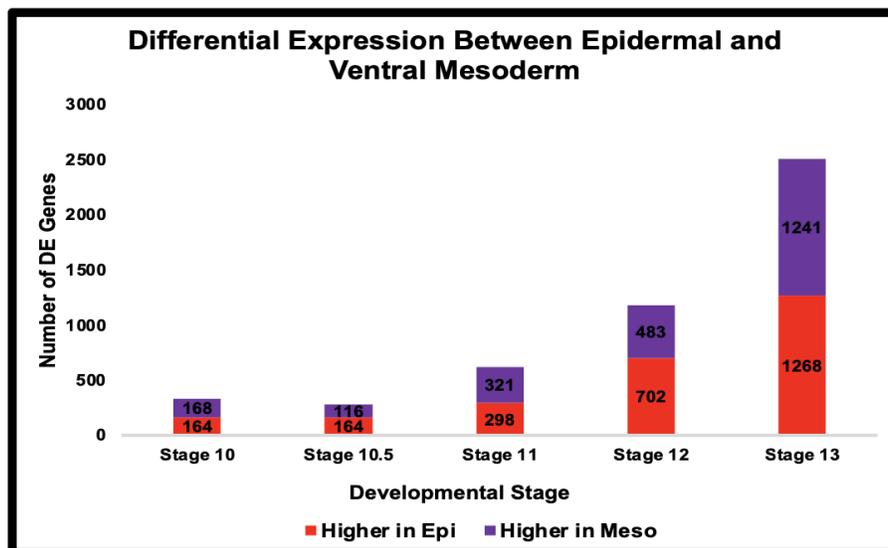


B

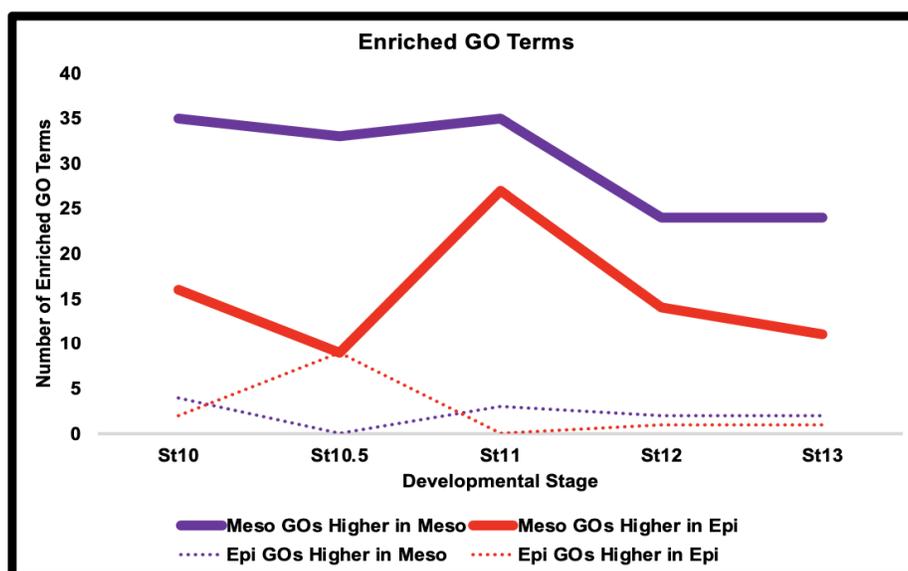


**Figure 2.29 Overlap in Temporal Dynamics of Mesoderm and Epidermis.** (A) Number of differentially expressed genes between successive developmental stages of the epidermal and ventral mesoderm lineage,  $p_{\text{adj}} \leq 0.05$  (B) Venn Diagrams for the total number of DE genes between stages for the epidermal and ventral mesoderm lineages, as well as the number of genes increasing and decreasing over time.

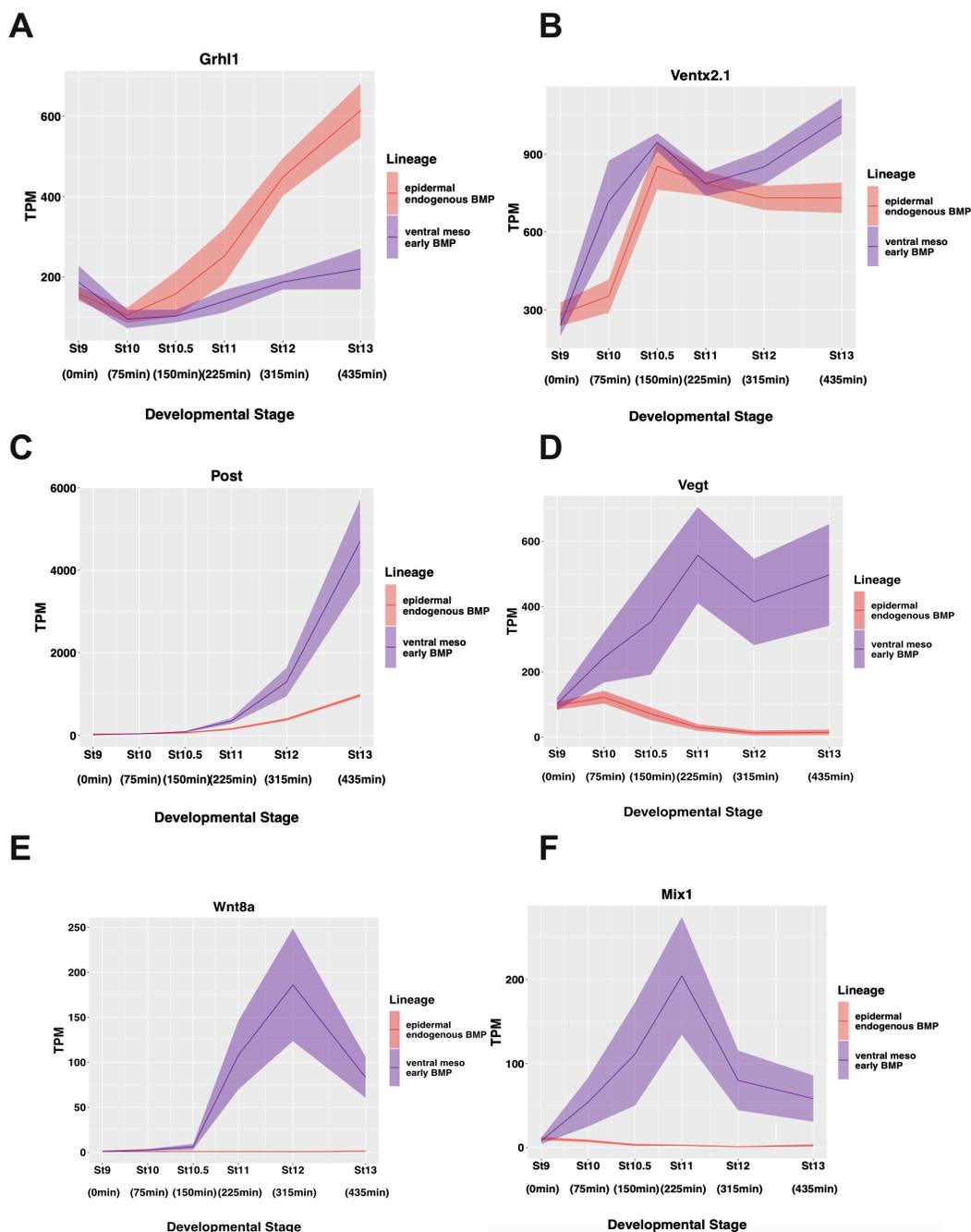
A



B

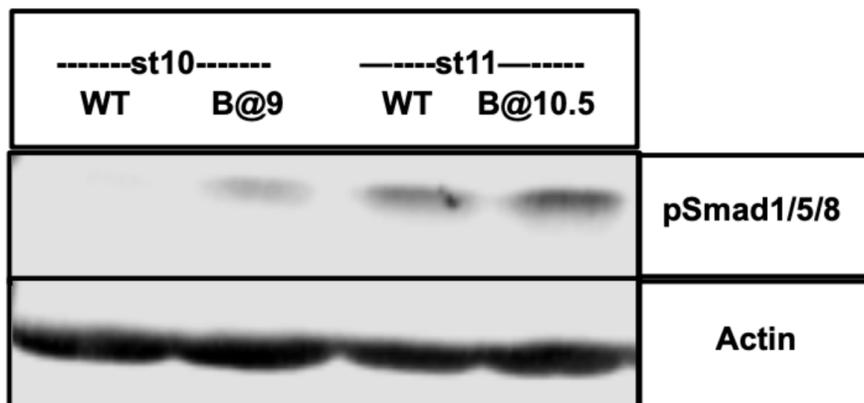


**Figure 2.30 DESeq Between Epidermal and Mesoderm Lineages.** (A) Number of differentially expressed genes between the epidermal and ventral mesoderm lineages at each developmental stage ( $p_{\text{adj}} \leq 0.05$ ). (B) Number of enriched Mesoderm GO Terms (solid line) and Epidermis GO Terms (dashed line) in genes significantly higher in the ventral mesoderm lineage (purple) and epidermal lineage (red) at each developmental stage.

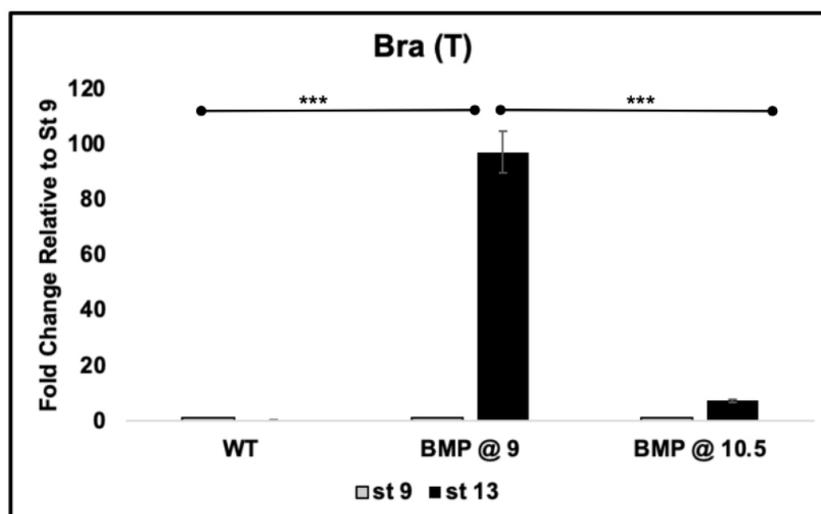


**Figure 2.31 Varying Patterns of Transcriptome Response to Early BMP.** TPM of genes representative of different expression dynamics in response to early BMP (A) *Grhl1* (B) *Ventx2.1* (C) *Post* (D) *Vegt* (E) *Wnt8a* (F) *Mix1*. Graphs are sums of S+L allele. Width of the line represents SEM of three biological replicates.

A



B

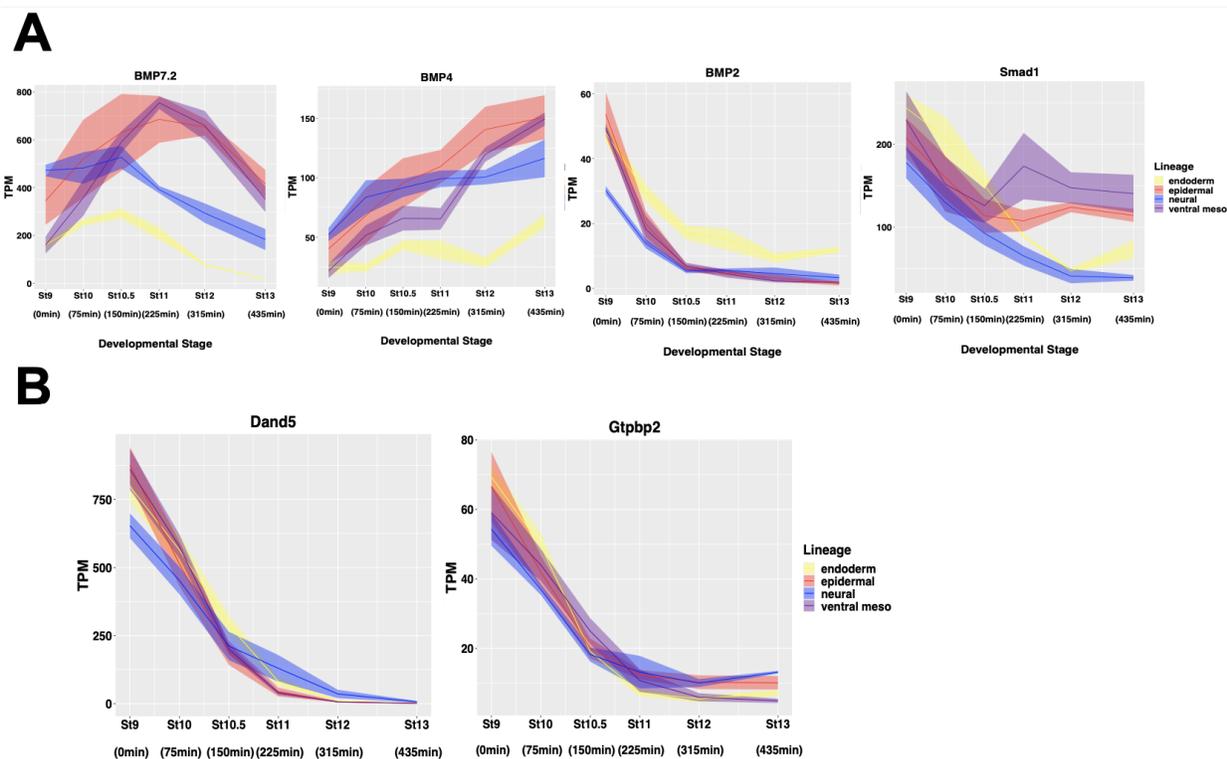


**Figure 2.32 Exogenous BMP at Stage 10.5 Does not Induce Mesoderm.** (A) Western Blot Analysis of lysates for epidermal (WT) and BMP4/7 treated at stage 9 (BMP4/7 20ng/uL) explants collected at stage 10 and epidermal (WT) and BMP4/7 treated at stage 10.5 (BMP4/7 20ng/uL) explants collected at stage 11 for pSmad1/5/8 with Actin loading control (B) qRT-PCR of animal pole explants examining the fold change from stage 9 to 13 of expression of mesodermal marker *Bra(T)* for epidermis(WT), treated with BMP4/7(20ng/uL) at stage 9 and treated with BMP4/7 (20ng/uL) at stage 10.5 (\*\* $P < 0.005$ ).

be achieved until stage 10.5 in response to endogenous BMP signals (Figure 2.31 B). Importantly, this demonstrates that BMP signaling is able to immediately elicit changes in gene expression when activated prematurely, and that this drives a heterochronic response in expression of some BMP target genes. Another BMP responsive gene, *Post*, which plays a role in conferring posterior/ventral attributes to both ectoderm and mesoderm (Sato and Sargent 1991) did not show a premature onset of expression, but instead displayed a significant increase in its amplitude of expression (Figure 2.31 C). Most striking, however, was the activation of genes categorized as Activin/Nodal target genes including *Vegt*, *Wnt8a*, and *Mix1* (Figure 2.31 D-F) (Watanabe and Whitman 1999, Christian et al. 1991, Rosa 1989). Importantly, activation of these genes in response to early BMP is not due to the inappropriate activation of pSmad2/3 (Figure 2.15). We also confirmed that it is the early timing and not the exogenous nature of BMP4/7 exposure that drives ventral mesoderm formation. While treatment of stage 10.5 explants with BMP4/7 leads to increased pSmad1/5/8 (Figure 2.32 A), it does not lead to expression of mesodermal genes (Figure 2.32 B).

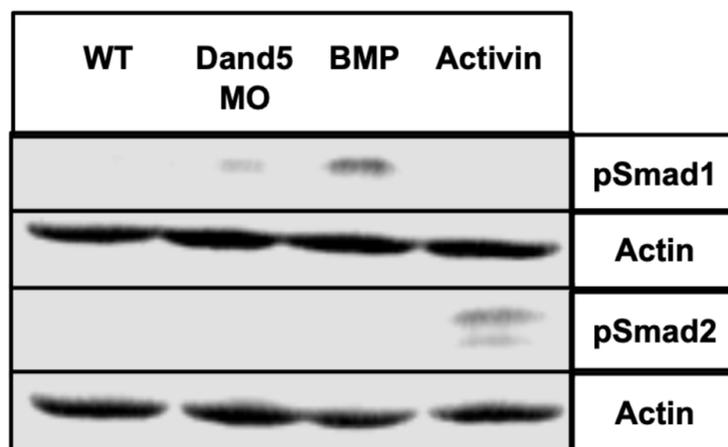
### **BMP Signaling is Restrained Until Stage 10.5 by Dand5 Activity**

The striking finding that shifting the timing of BMP signaling leads to activation of Activin/Nodal-responsive genes, and the formation of mesoderm instead of epidermis, indicates that it is essential to tightly control when cells receive this signal endogenously. *BMP2,4* and *7* ligands, as well as the receptor *Smad1* are expressed at stage 9 and 10 (Figure 2.33 A) and cells are clearly competent to respond to early BMP signals as evidenced by the shifted activation of *Ventx2.1* (Figure 2.31 D). Despite this, however, pSmad1/5/8 is not robustly detected in control explants until stage 10.5. We therefore investigated how BMP signaling is restrained in animal pole cells until the onset of gastrulation. We asked if there were BMP signaling antagonists expressed in blastula animal pole cells that were downregulated with dynamics consistent with the observed timing of pSmad1/5/8 accumulation. We identified two maternally provided BMP antagonists, *Dand5* and *Gtpbp2* (Bell et al. 2003, Vonica et al 2007, Bates et al. 2013, Reich and Weinstein 2019, Kirmizitas et al. 2014), that are expressed at stage 9 but downregulated to significantly lower levels by stage 10.5, when BMP activity is observed (Figure 2.33 B). *Dand5*, in

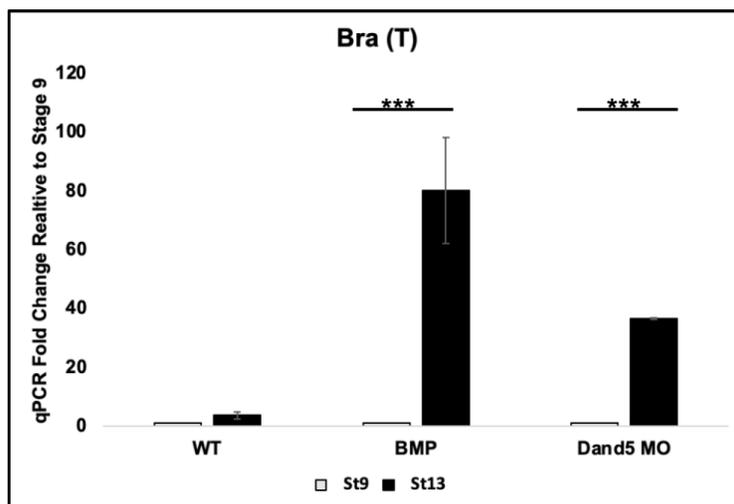


**Figure 2.33 Maternal BMP Inhibitors Explain Endogenous BMP Timing.** TPM of genes involved in BMP signaling (A) BMP heterodimer ligands *BMP7.2*, *BMP4*, *BMP2* and primary BMP target *Smad*, *Smad1* (B) maternally provided BMP antagonists *Dand5* and *Gtpbp2*. Graphs are sums of S+L allele. Width of the line represents SEM of three biological replicates.

A



B



**Figure 2.34 Dand5 MO Induces Early BMP and Mesoderm Formation.** (A) Western Blot Analysis of lysates for stage 10 epidermal (WT), *Dand5* MO injected (40pmol/embryo), ventral mesoderm (BMP4/7 20ng/uL), and endoderm (160ng/uL Activin) explants for pSmad1/5/8 and pSmad2 with Actin loading control (B) qRT-PCR of animal pole explants examining the fold change from stage 9 to 13 of expression of mesodermal marker *Bra(T)* for epidermis(WT), ventral mesoderm(BMP4/7 20ng/uL) and *Dand5* MO (300pmo/embryo) injected(\*\**P*<0.005).

particular, is robustly expressed at stages 9 and 10 in animal pole cells. To determine if *Dand5* plays a role in preventing premature BMP signaling we used a translation blocking morpholino to deplete it from early blastulae. We found that *Dand5*-depletion led to premature phosphorylation of Smad1/5/8, indicative of early BMP signaling (Figure 2.34 A), as well as expression of mesodermal marker *Bra (T)* at stage 13 (Figure 2.34 B). This suggests that a key role for *Dand5* at these stages is preventing premature BMP signaling that would generate a mesodermal rather than ectodermal state transition. As *Dand5* depletion does not increase pSmad1/5/8 or *Brachyury* levels to the same degree as BMP4/7 treatment, other BMP antagonists including *Gtpbp2*, likely cooperate in temporally constraining BMP signaling.

## Discussion

How progenitor cells decide their fate is a fundamental question central to all of developmental biology. While in many cases the inductive cues that drive these decisions have been identified, less is understood about how the timing of these signals is controlled and the dynamics of the transcriptional circuitry they activate. A related and fascinating question is how a large and complex set of transcriptional responses is canalized into a discrete set of lineage trajectories. Decades of research in *Xenopus* and other systems has shed important light not only on the signaling pathways driving lineage formation in early embryos, but also on many of the key transcriptional targets of these signals. Combined with gain and loss of function studies for individual factors, this work has allowed the construction of putative gene regulatory networks (GRNs) depicting how different lineage states are adopted. A powerful strength of the *Xenopus* system is the ability to easily isolate pluripotent cells from blastula embryos and culture them in simple saline with or without added inductive cues, and that these cells become lineage restricted over a period of approximately seven hours. This allows the transcriptional responses to inductive cues to be quantified with high temporal resolution, enabling the dynamics of individual genes as well as entire lineages to be followed. We used the pipeline we established here to study the transit of initially pluripotent cells to four distinct lineage states using bulk transcriptomics. However, this pipeline can be built upon to layer on additional analyses including ATAC to follow changes in chromatin accessibility and ChIP-seq to follow changes in epigenetic marks dynamically during lineage restriction.

### **Neural as the Default State.**

Our analysis of four different state transitions lends unexpected support for the neural default model and provides novel insights into *why* neural is the default state for pluripotent cells. Two types of analysis immediately distinguished the neural lineage from the other three state transitions. First, temporal DESeq reveals that this lineage reaches a steady state in gene expression dynamics by stage 10.5; after this time point the number of genes exhibiting differential expression is quite small (Figure 2.6). This is in marked contrast to the epidermal and mesodermal state transitions which become increasingly dynamic after stage 10.5. Similarly, principal component analysis reveals that the neural state lies closer to the pluripotent state than the other three lineages (Figure 2.5 A). This is true for both the PC1 and PC2 axes, which together explain 75% of the variation in gene expression. Since developmental times correlate along the PC1 manifold while state identities correlate along the PC2 manifold, this confirms that the time taken to transit from pluripotent to neural state is shorter than for the other lineage trajectories. Measuring the distance between stages 9 and 13 across all 74 principal components also indicates that the neural lineage is most positively correlated with pluripotency (Figure 2.5 C). Thus, neural progenitor cells occupy a position closest in state space to pluripotent cells relative to the other lineages.

The neural default model has not been without controversy. Studies in avian embryos have suggested that BMP inhibition may not be sufficient for transit to a neural progenitor state (Streit et al. 2000, Linker and Stern 2004). While work in mESCs and hESCs provided additional support for the neural default model (Smukler et al. 2006, Vallier et al. 2004), these cells do not necessarily recapitulate the *in vivo* state of inner cell mass cells. Our findings thus provide important validation of this model. Interestingly, when the neural lineage trajectory is compared to that of the epidermal trajectory, they are highly correlated through stage 10.5, early gastrulation. This is true with respect to their dynamics, as evidenced by temporal DESeq (Figure 2.8), and also supported by the very small number of genes that exhibit differential expression between these states until stage 10.5 (Figure 2.11 A). The onset of gastrulation can thus be considered a point in time when a group of equipotent cells (neural/epidermal) diverge and either continue changing state to become epidermal or do not continue changing state and

become neural. Indeed, individual cells at blastula stages and early gastrulation have been found to display multilineage gene expression (Briggs et al. 2018).

Significantly, stage 10.5 is when robust BMP signaling is first detected in animal pole cells, as evidenced by pSmad1/5/8 detection (Figure 2.13 A), and this correlates with when the trajectory of the epidermal lineage diverges from that of neural (Figure 2.8 A). It is also the time when we observe a sharp increase in expression of *Grhl1*, a key upstream component of the epidermal GRN (Tao et al. 2005), significant down-regulation of the pluripotency factor *Foxi2* (Cha et al. 2012) (Figure 2.10) and a loss of enrichment for neural GO terms in the epidermal lineage (Figure 2.9). Neural features are retained and enhanced in the Noggin-treated explants, as exemplified by the expression dynamics of transcription factors *Sox3*, *Sox11* and *Zic1* (Figure 2.12). By contrast the BMP target genes *Ventx2.1* and *Id3* exhibit expression maxima in the epidermal lineage and minima in the neural lineages around stage 10.5 (Figure 2.14). Given the distinct non-monotonic dynamics of *Id3* in the epidermal and neural lineages it is tempting to speculate that this inhibitory bHLH factor may be playing a role in suppressing the function of neuralizing factors in the prospective epidermis thus helping to canalize this state transition.

### **Surprising Overlap in Transcription Response to BMP4/7 and Activin**

Our data sets allow direct comparison of the transcriptome changes driven by the two different branches of the TGF-beta superfamily, BMP and Activin. A striking feature of the Activin-driven endoderm trajectory is its distinct dynamics, characterized by a large increase in differentially expressed genes between stage 9 and stage 10.5, early gastrulation (Figure 2.16A). This is not seen following treatment with BMP4/7. As expected, the characterized members of the mesendoderm GRN are induced in response to Activin and many of these respond rapidly and robustly, but transiently. Indeed, twenty-one of forty-one GRN factors examined are induced by stage 10, only 75 minutes after ligand exposure (Figure 2.25 A). pSmad2/3 is also robustly detected at this time, but is no longer detected by stage 13 indicating that the signaling response to Activin is transient (Figure 2.15). Distinct signaling dynamics for this branch of TGF-beta signaling may contribute to the markedly different transcriptome dynamics observed in the endodermal lineage as compared to the BMP-driven epidermal and mesodermal lineages (Figure 2.6).

Among the genes that respond immediately to Activin are several members of the Bix/Mix family of transcription factors (Pereira et al. 2012). Somewhat unexpectedly, genes generally associated with ventral fates, including *Ventx2.1*, *Ventx2.2* and *Wnt8*, were also induced by high levels of Activin. Induction of these factors occurred at later points in the trajectory, and their expression correlates with the downregulation of Bix/Mix family genes and other transiently responding factors, including *Eomes*, *Nodal*, and *Vegt*. Going forward it will be of interest to determine if these ventralizing factors play a direct role in downregulating the expression of the endodermal factors that are expressed only transiently. By contrast, *Sox17* responds immediately to Activin, and its expression increases linearly through stage 13. *Endodermin* also has a linear response to induction, although the increase in its expression does not commence until state 10.5, possibly reflecting a role for *Sox17* in its activation. The distinct dynamics of early responding endoderm genes allowed the identification of putative new members of the endoderm GRN using *limma* analysis (Figure 2.27).

Among the most surprising findings emerging from these studies was the activation by BMP4/7 of genes that are generally characterized as Activin/Nodal targets. Indeed, among the earliest responses to BMP4/7 were Mix/Bix family genes, and similar to their response to Activin, their induction was transient (Fig. 5C). Other unexpected responding genes included *Sox17*, *Eomes* and *Gsc*. As Smad2/3 phosphorylation was not observed in response to BMP4/7, this suggests that the BMP R-Smads are capable of activating expression of these Activin/Nodal targets given a permissive cellular context. In this respect it is worth noting that *Mix1.1* was previously identified in a screen for BMP4 responsive genes (Meade et al. 1996), supporting our current findings. It is intriguing that the genes exhibiting immediate responsiveness to BMP4/7 are dorsal mesendoderm factors, whereas pan and ventral mesoderm genes, including *Bra (T)*, *Wnt8*, *Post*, *Msx1* and *Evx1*, are turned on later in the trajectory. This suggests that the initial response to BMP signaling is “dorsal” as it is for Activin, and that more “ventral” attributes are a secondary response.

**The Timing of BMP Signaling is Critical for Proper Lineage Segregation.**

The level of Bmp4/7 signaling utilized in these experiments was selected to match the level of pSmad1/5/8 levels present in untreated explants at stage 10.5 (Figure 2.21). This allows comparison of the response to the same signal and amplitude but with shifted developmental timing. Receiving the same level of BMP signaling at a slightly earlier time point might have been expected to accelerate transit to the epidermal state. Indeed, the shifted onset of *Ventx2.1* expression is consistent with an accelerated response (Figure 2.31B). However, explants also respond to BMP4/7 exposure at stage 9 by inducing expression of mesendodermal factors, including *Vegt* and *Mix1* (Figure 2.31 D,F), and by suppressing the endogenous BMP-mediated increase in expression of the epidermal regulatory factor *Grhl1* (Figure 2.31 A). Thus, exposure to BMP4/7 at stage 9 elicits a fundamentally different transcriptional response than does exposure at stage 10.5. This was confirmed by treating explants with exogenous BMP4/7 at stage 10.5, which fails to elicit a mesendoderm response (Figure 2.32).

Together these findings indicate that it is critical to control the timing at which initially pluripotent cells are able to respond to endogenous BMP signaling. While we first detect pSmad1/5/8 at stage 10.5, it is likely that low levels of signaling are initiated by stage 10, as that is when increased expression of *Ventx2.1* and *Id3* is observed (Figure 2.14). Thus, cells undergo a fundamental change in competence between stages 9 and 10. Interestingly, expression of BMP inhibitors such as Noggin, Chordin, Follistatin, and Cerberus commences in the presumptive organizer region at late blastula stages (Wesseley et al. 2001), indicating that blocking BMP signaling in the marginal zone is also critical at these stages. This raised the question of how BMP signaling is restrained in blastula animal pole cells such that a mesendoderm response is prevented and ectodermal competence is established.

Using our data sets we identified two maternally provided BMP antagonists, *Dand5* and *Gtpbp2* that are expressed at stage 9 but are significantly downregulated by stage 10.5 (Figure 2.33 B). As *Dand5* displayed significantly higher levels of expression, we examined the consequences of depleting it from initially pluripotent explants. Morpholino-mediated depletion of *Dand5* depletion resulted in premature phosphorylation of Smad1/5/8 and expression of *Bra (T)* at stage 13 (Figure 2.34), indicating that a key role for *Dand5* at these stages is preventing premature BMP signaling that would generate a mesodermal rather than ectodermal state transition. Interestingly, a role for *Dand5* in controlling the

spatial response to Activin/Nodal signaling had previously been suggested, restricting this response to the mesodermal mantle (Bell et al. 2003, Bates et al. 2013, Reich and Weinstein 2019). Our findings suggest a second centrally important role for this TGF-beta antagonist in the temporal control of BMP signaling as animal pole cells exit from pluripotency. Moreover, the data sets described here will facilitate future studies into the temporal control of transcriptional responses to inductive cues across multiple embryonic lineages.

### **Materials and Methods:**

**Embryological Methods.** Wild-type *Xenopus laevis* embryos were obtained using standard methods from a daily 2pm fertilization from a single frog and placed into a 14C incubator at 2:45pm until 8:30am. Ectodermal explants were manually dissected at early blastula (stage 8-9) from embryos cultured in 1X Marc's Modified Ringer's Solution (MMR) [0.1 M NaCl, 2mM KCl, 1mM MgSO<sub>4</sub>, 2 mM CaCl<sub>2</sub>, 5mM HEPES (pH 7.8), 0.1 mM EDTA] from 8:30am-9:45am and then placed in a 20C incubator until 5pm. Groups of 12-15 explants were collected at 9:45 am, 11am, 12:15pm, 1:30pm, 3pm and 5pm using sister embryos to confirm approximate stages of 9,10, 10.5, 11, 12, and 13 based on Nieuwkoop and Faber (1994) staging. Explants for the neural progenitor lineage were generated using recombinant Noggin protein (R&D Systems) at a final concentration of 100ng/mL in low calcium magnesium media supplemented with 0.1% bovine serum albumin (BSA) as a carrier for sequencing experiments, or using 20uM K02288 (Sigma) in 0.1X MMR for Westerns. The Noggin dose utilized was the minimum amount required to induce the neural progenitors Sox3 and Sox2 at stage 13, as determined by qPCR and *in situ* hybridization. The K02288 dose utilized was the minimum amount required to effectively block BMP signaling through stage 13, as determined by pSmad158 western blot. Endoderm lineage explants were generated using recombinant Activin protein (R&D Systems) at a final concentration of 160 ng/mL in 1X MMR supplemented with 0.1% BSA. This was the lowest dose that induced endoderm without mesoderm, as determined by qPCR with Bra(T) and Sox17 primers. Mesoderm lineage explants were generated using recombinant BMP4/7 heterodimer protein (R&D Systems) at a final concentration of 20g/mL in 1X MMR supplemented with 0.1% BSA. This concentration of BMP4/7 was utilized as it

induced levels of pSmad158 at stage 10 that were comparable to levels present at stage 10.5 in untreated caps. For morpholino experiments, a previously validated translation-blocking Dand5 morpholino (Vonica and Brivanlou 2007, Maerker et al. 2021) (Gene Tools, Sequence: 'CTGGTGGCCTGGAACAACAGCATGT') was injected in 4 cells at the eight-cell stage for a total of 40pmol per embryo.

**RNA Isolation, cDNA Synthesis, Sequencing, and qRT-PCR.** RNA was isolated from blastula explants (12-15 explants) using Trizol (Life Technologies) followed by LiCl precipitation. 1  $\mu$ G of purified RNA was used as a template for synthesizing cDNA using a High Capacity Reverse Transcription Kit (Life Technologies) Quantitative (q) RT-PCR was performed using SYBR Premix ExTaq 11 (Takara Bio) and detected using the Bio-Rad CFX96 Connect system. *Brachyury* primers used were Fwd: GAA GCG AAT GTT TCC AGT TC and Rev: ACA TAC TTC CAG CGG TGG TT. Expression was normalized to *Ornithine Decarboxylase (ODC)* ODC primers used were Fwd: TGA AAA CAT GGG TGC CTA CA and Rev: TGC CAG TGT GGT CTT GAC AT. The fold change was calculated relative to stage nine samples from the same time course experiment. The results show the mean of three independent biological replicates, with error bars depicting the SEM. An unpaired, two-tailed *t*-test was utilized to determine significance. 500 ng of RNA was used for library prep with TruSeq mRNA library prep kit (Illumina) and sequenced using Next Seq 500 Sequencing (epi,neur, endo) or HiSeq4000(meso).

**Western Blot Analysis.** Blastula explants (20 explants/sample) were collected for specified stage / lineage and lysed in TNE phospho-lysis buffer [50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 0.5 mM EDTA, and 0.5% Triton X-100, 2mM Sodium Orthovanadate, 20mM Sodium Fluoride, 10mM B-Glycerophosphate, 1 MM Sodium Molybdate dihydrate] supplemented with protease inhibitors [Aprotinin, Leupeptin and phenylmethylsulfonyl fluoride (PMSF)] and a PhosStop phosphatase inhibitor and a complete Mini tablet (Roche). SDS-PAGE and western blot analyses were used to detect proteins and modifications using the following antibodies: anti-phospho SMAD2 (Ser465/467, Sigma, 1:500), Smad2 Polyclonal Antibody (Life Technologies 1:500) Anti-phospho Smad1/Smad5/Smad8 (Ser463/465, Sigma,

1:1000), Smad1 (D59D7, Cell Signal, 1:1000), Actin (A2066, Sigma-Aldrich, 1:5000). For chemiluminescence-based detection, horseradish peroxidase (HRP)-conjugated rabbit secondary antibodies were used (Vector Laboratories, 1:20,000). Results shown are representative of at least three independent experiments and are cropped only to minimize empty space.

**RNA Seq Processing and Computational Analysis.** Read quality was evaluated using FastQC (Andrews 2010). Mapping to *X. laevis* v9.2 genome downloaded from xenbase was performed using RSEM to get TPM values (Li and Dewey 2011). Alignment to the *X.laevis* v9.2 genome was performed using STAR 2.6.0 to get raw counts using standard parameters (Dobin et al. 2013). Computational Analysis of RNA Sequencing Data was performed using published R Packages. TPM data is an average of three biological replicates of combined data from S and L alleles for each gene, the width of each line represents SEM and graphs were plotted using ggplot2 (Wickham 2016). A minimum raw read count of 15 was determined computationally using the voom function of the limma package. Based on the plot of the mean variance trend, filtering out genes with expression below 15 provided the best balance between filtering out lowly expressed genes without losing genes relevant to the transcriptome dynamics, thus all analyses are performed on genes with a minimum read count of 15 in any lineage at any stage (Law et al. 2014).

Differential expression analysis was done between successive stages for each lineage and between pairs of lineages at corresponding states using DESeq2 with significance defined as  $p_{\text{adj}} \leq 0.05$  (Love et al. 2014). No fold change cut-off was used for examining temporal DESeq, in order to capture the greatest amount of dynamic change in the transcriptome over time. Overlapping DESeq genes were visualized using VennDiagram and UpsetR (Chen and Boutros 2011, Conway et al 2017). GO and KEGG enrichment was calculated using the GOSep R package (Young et al. 2012). Neural GO terms were identified based on their relevance to neural-related processes including, neuron, brain and nervous system development as well as axon formation and synapse firing. Mesoderm GO terms were identified based on their relevance to mesoderm-related processes including mesoderm formation, gastrulation and somitogenesis. DE genes for heatmaps were determined by ranking the genes DE between mesoderm

and epidermis, and between endoderm and epidermis separately based on Log2 fold change and only genes with a minimum normalized expression of 10 TPM at the relevant stages were plotted using the pheatmap package (Kolde 2015), expression was depicted using z-scores. All lists of DE genes with padj and Log2FC values are provided as supplementary tables with transcription factors bolded, Gene lists from all Venn Diagrams and UpSet plots are also provided as supplementary tables. Hierarchical clustering was done using the pheatmap package, cluster rows=TRUE, cluster columns = TRUE.

Principal Component Analysis was done with the prcomp function in the default stats package, and visualized using ggfortify (Tang and Horikoshi 2016). We assessed whether the observed patterns were statistically significant by testing whether the differences between samples of different types was significantly greater than the differences between replicates. We calculated pairwise distances between all pairs of samples in the PCA space and tested whether the distances for between- and within-type pairs differed using a nonparametric KS test. Overall, the between-type distances were far larger than the within-type distances ( $p=6.27e-03$ ). We also performed this same analysis within each pair of lineages (eg, testing whether the endo-epi distances were larger than the endo-endo and epi-epi distances) and found this pattern to be consistent though low numbers of sample pairs reduced the power of the KS test. The top 2 PCs were determined as the most significant based on the elbow of the PC plot and thus the only two that were plotted. Distances from stage 9 to 13 were calculated using the dist() function. Statistical significance of differences in distance were calculated using the Wilcoxon rank sum, wilcox.test() in R.

Pattern dynamics were determined for linear, quadratic and cubic patterns with limma analysis using the limma voom R package (Law et al. 2014). Potential mesendoderm GRN candidates were selected by examining the 10 genes with most similar differential quadratic and linear dynamics between epidermal and ventral mesoderm. Of these genes, those with minimum expression of 30 TPM, dynamics unique only to the endoderm and/or ventral mesoderm lineages and not already defined as mesendoderm GRN members were identified as possible novel GRN members. Genes graphed were those we were able to corroborate with other genomic studies.

WGCNA was performed using the WGCNA R package (Langfelder and Horvath 2008, 2012). Genes with a TPM >15 in endoderm or mesoderm at stage 10, 10.5 or 11 were included in the analysis. Analysis was run on Log2 of the TPM. Power was determined using the sft function and TOMType = "signed", minModuleSize = 30 and mergeCutHeight = 0.25.

**Vertebrate Animals.** All animal procedures were approved by Northwestern University's Institutional Animal Care and Use Committee and are in accordance with the National Institutes of Health's Guide for the care and use of Laboratory Animals.

**Availability of Data and Materials.** The datasets generated and analyzed during the current study are available in the NCBI repository and are accessible through GEO Series accession number GSE198598. Supplemental Tables can be accessed at DOI: 10.1186/s12864-022-08953-3 / PMID: 36273135.

## **Chapter Three**

### **DNA Methylation is Required for Normal Lineage Restriction**

The process by which pluripotent cells give rise to lineage restricted cells is governed by transcription factors, signaling molecules and epigenetic modifiers. DNA methylation is a key epigenetic modification in which methyl groups are added to the cytosine residues of DNA, particularly in CpG rich regions, by DNA methyltransferases (DNMTs). This limits the accessibility of DNA and thus prevents transcription factors from binding. DNA methylation is thus a key regulatory element in the lineage restriction process as it helps to spatially and temporally turn genes on and off throughout the course of embryonic development. Here I investigate the role of DNA methylation in the lineage restriction process using the DNMT inhibitor 5-Aza-2'-deoxycytidine (5-Aza) to block DNA methylation starting at the two cell stage. I show that inhibition of DNA methylation has a stochastic effect on pluripotency genes and causes a partial retention of the pluripotency factor *Sox3* in stage 13 animal caps and allows for only partial restriction to the epidermal fate and a loss of competency to form the mesoderm and endoderm fate. RNA Sequencing shows that blocking DNA methylation has a significant effect on RNA splicing during the lineage restriction process and that stage 13 5-Aza treated caps are most similar to the neural crest lineage, and not the epidermal lineage to which animal caps typically default.

## **Introduction**

Early embryonic development is characterized by the progressive restriction of stem cells as their potential is limited and cells respond to spatial and temporal cues to undergo lineage restrictions. While the specifics of this process are governed and defined primarily by transcription factors and signaling molecules, epigenetics also play a crucial role in ensuring the appropriate DNA is accessible for transcription in order for this process to go smoothly. The epigenome is regulated by writers, readers, and erasers which add, interpret, and remove epigenetic marks upon both histone and DNA, enabling the appropriate genes to be transcribed at the correct time and place (Hu et al. 2014a).

DNA methylation causes transcriptional repression by the addition of a methyl group to the cytosine residue of DNA, primarily at CG repeats (Gruenbaum et al. 1982, Bestor and Ingram 1983, Keshet et al. 1986, Boyes and Bird 1991). The methylation of DNA is initiated by the de novo methyltransferases Dnmt3A and Dnmt3B and maintained by Dnmt1 upon DNA replication (Okano et al.

1999, Liang et al. 2002, Araujo et al. 1998). In *Xenopus*, DNA is hypermethylated before and after the maternal-zygotic transition (Macleod et al. 1999, Veenstra and Wolffe 2001). DNA methylation has been shown to be required for lineage specification in mouse ESCs and Dnmt1 is required for proper regulation of gastrulation in zebrafish (Tsumura et al. 2006, Kent et al. 2016). Dnmt3A and Dnmt3B have nuanced tissue specific expressions. Dnmt3B is necessary for proper neural progenitor specification (Martins-Taylor et al. 2012) and Dnmt3A plays an important role in neural crest specification (Hu et al. 2012).

DNA can also be demethylated, either passively through lack of maintenance or actively through ten eleven translocation (TET) enzymes (Ito et al. 2011, Paranjpe and Veenstra 2015). TET enzymes oxidize methyl groups to hydroxymethyl groups by converting 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), which is then converted to 5-formylcytosine (5fC) and finally 5-carboxylcytosine(5caC). Both 5fC and 5caC can be targeted for excision by thymine DNA glycosylase and the DNA is repaired to its unmethylated state through base excision repair (Tahiliani et al. 2009, He et al. 2011, Ito et al. 2011, Weber et al. 2016, Wu and Zhang 2017). DNMTs and TETs both operate to maintain appropriate levels of DNA methylation.

Various histone acetylation and methylation marks work in concert to turn off pluripotency genes and turn on the appropriate lineage restriction genes. Loosely packed chromatin enhances DNA accessibility, thus acetylation acts as an activating mark, since the acetyl group neutralizes the charge on lysines upon which it is deposited, most commonly lysines 9 and 27, weakening their bond with DNA (Hong et al 1993, Brownell et al. 1996, Efroni et al. 2008, Creighton et al. 2010). Histone methylation, on the other hand, can be either activating, in the case of H3K4Me<sub>3</sub>, H3K36me<sub>3</sub> or repressive, as with H3K27me<sub>3</sub> and H3K9Me<sub>2/3</sub> (Strahl et al. 1999, 2002, Cao et al. 2002, Schultz et al. 2002). These methylation and acetylation marks work in concert with one another, as is the case with H3KMe<sub>1</sub> and H3K27ac, which together mark active genes, while H3KMe<sub>1</sub> alone marks poised genes (Creighton et al. 2010).

Importantly, histone methylation has a critical role in regulating the methylation of DNA itself. The ATRX-DNMT3-DNMT3L (ADD) domain of Dnmt3A specifically recognizes H3K4 absent the activating methyl marks, thus initiating de novo methylation there (Ooi et al. 2007, Otani et al. 2009, Gu et al. 2015,

Du et al. 2015). H3K436Me3 is primarily methylated on gene bodies, and Dnmt3b preferentially binds to these regions, potentially as a means of compensating for the increased 5-hmc found at transcribed gene bodies (Morselli et al. 2015, Baubec et al. 2015, Pastor et al. 2011). Dnmt3A, on the other hand, is recruited to intergenic regions by H3K36Me2/3 (Weinberg et al. 2019, Bröhm et al. 2022). Finally, the silencing histone mark H3K9Me3 is read directly by Dnmt1, leading to H3K9 and DNA methylation of the same marks, enhancing the stability of gene repression (Lehnertz et al. 2003, Ren et al. 2020).

In addition to its well established role in supporting the regulation of lineage restriction, increasing evidence is mounting that DNA methylation also plays a role in regulating exon splicing. DNA methylation is enriched in exons compared to introns (Hodges et al. 2009, Chodavarapu et al 2010). Several studies suggest that this enrichment of DNA methylation in exons is not merely a byproduct of nucleosome occupancy but may actually be playing a regulatory role in alternative splicing, by regulating exon skipping, intron retention and inclusion of alternative exons (Li-Byarlay 2013, Yearim et al. 2015, Shayevitch et al. 2018, Lev Maor et al. 2015). DNA methylation is able to regulate alternative splicing by preventing CTCF binding which releases Pol II pausing, leading to exon exclusion (Shukla et al. 2011). DNA methylation can also enhance exon recognition and inclusion through recruitment of the protein MeCP2 (Maunakea et al. 2013). Finally, DNA methylation induced H3K9Me3 can bind with the HP1 protein which then recruits splicing factors and can either enhance exon recognition when bound upstream of the exon and lower recognition when bound directly to the exon (Yearim et al. 2015).

Multiple pharmacological inhibitors of DNA methylation have been shown to effectively block DNA methylation and have been established as potential therapeutics in cancer treatment (Mund et al. 2005, Yang et al. 2006, Gore et al. 2006, Soriano et al. 2007). 5-Aza-2'-deoxycytidine is one of the most effective of these drugs as it is incorporated into DNA directly and can irreversibly bind Dnmts, leading to complete demethylation after several rounds of replication (Creusot et al. 1982, Stresemann et al. 2008, Yang et al. 2010). These inhibitors also have efficacy in model organisms and thus can be used to temporally control DNA methylation and make them exciting tools to probe the necessity of DNA methylation at different developmental stages and for specific lineages and developmental processes.

Also playing a role in cell fate decisions is the stochastic nature of fate transitions of cells from pluripotent to more lineage restricted fates. The process of differentiation of ES cells indicates that gene expression programs are cell autonomous, as they operate independently but reproducibly in each individual cell, leading to heterogeneity across a population. Single cell studies have shown that hematopoietic stem cells with multilineage potential were able to simultaneously express genes from each of their potential fates, particularly genes that play a role in balancing self-renewal and differentiation in populations of cells undergoing fate changes. These data suggest that heterogeneous gene expression patterns at the single cell level may be characteristic of populations of cells undergoing fate decisions and that transitions may be discontinuous (Laslo et al 2006, Moris et al. 2016). It has been shown that in ES cells, as opposed to tissue specific stem cells or lineage restricted cells, lineage specific markers often have chromatin states associated with both active and repressive states, as indicated by H3K4 methylation and H3K427 trimethylation of those promoters in the same cell, suggesting that they are primed for expression but restricted by chromatin modifications (Azuara et al. 2006). This indicates histone methylation may play an important role in this variability, but the role of DNA methylation in this stochasticity is yet unexplored.

Here I show via *in situ* hybridization that inhibiting DNA methylation upon the initial cell cleavage has a stochastic effect on core pluripotency factors that is partially reversible, perhaps amplifying inherent stochasticity of gene expression. I then utilize the animal cap assay to study the effect of inhibition of DNA methylation on lineage restriction. *In situ* hybridization reveals only a partial restriction to the epidermal fate, while at the same time a partial maintenance of pluripotency marker Sox3. Additionally, caps are unable to respond to inductive cues to form the mesoderm and endoderm lineages. I then use RNA sequencing to compare 5-Aza treated caps to caps directed to the epidermal, neural, mesoderm, endoderm, and neural crest lineage at stage 13. I show that the most significant effect 5-Aza has on animal caps, based on GO Term analysis, is to alter their exon splicing. When this exon splicing is accounted for in principal component analysis, 5-Aza treated caps occupy a space closest to the neural crest fate, suggesting these caps have a transcriptome suggestive of multipotency, but lack the ability to actually form any of these potential lineages. This work enhances the understanding of the role of DNA

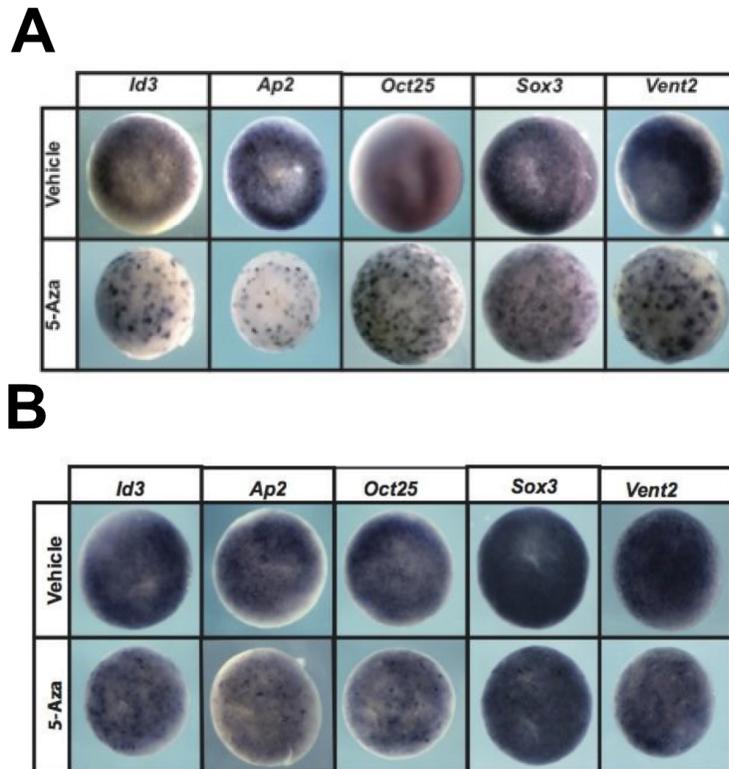
methylation in exon-splicing and its requirement for normal lineage restriction in *Xenopus laevis* animal caps.

## Results

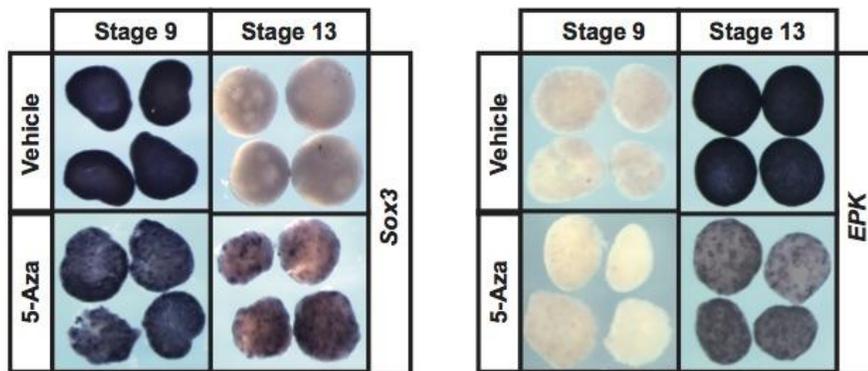
### Inhibition of DNA Methylation Blocks Lineage Formation

DNA Methylation can be temporally controlled using the pharmacological inhibitor 5-Aza-2'-deoxycytidine (5-Aza). This inhibitor acts by incorporating 5-Aza nucleosides into the DNA during replication, which targets Dnmts for proteasomal degradation, leading to completely unmethylated DNA after several rounds of replication (Yang, 2010). Treatment of *Xenopus* embryos with this inhibitor at stage 2, upon initial cell cleavage, led to stochastic expression of a set of core pluripotency factors, *Id3*, *Ap2*, *Oct25*, *Sox3* and *Vent2* at late blastula stages (Zhang et al. 1995, Winning et al. 1991, Cao et al. 2004, Penzel et al. 1997, Ladher et al. 1996). Under normal conditions, these show homogeneous gene expression in the animal pole via *in situ* hybridization (Figure 3.1 A). This phenotype is partially reversible, as when embryos are treated with same 25 uM concentration of 5-Aza at stage 2, but removed from the drug at stage 5 (16 cells) only a minor perturbation in gene expression is seen, with levels of pluripotency factors largely restored to wildtype levels. These data suggest that DNA methylation may be required for the establishment of pluripotency and thus a possible requirement for the competency of these animal pole cells and the normal developmental transition to different cell states.

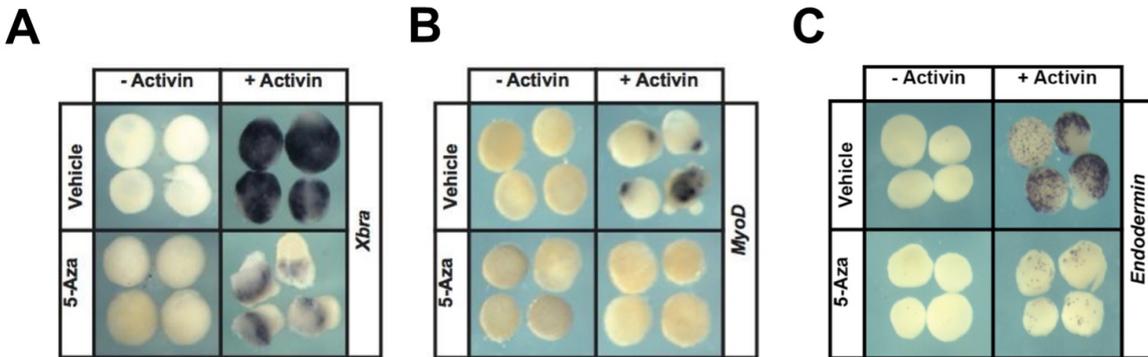
In order to investigate cell state transitions, I employed the *Xenopus* animal cap assay. In the absence of any additional cues, dissected *Xenopus* blastula explants (animal caps) will form epidermis. I first investigated the ability of these 5-Aza treated animal caps to undergo normal lineage restriction by using *in situ* hybridization to determine if these caps form epidermis by stage 13, the stage at which lineage restriction begins. Caps treated with 5-Aza showed a partial transition to the epidermal state. Animal caps showed heterogeneous expression of both the epidermal marker *EpK* and the pluripotency marker *Sox3* (Figure 3.2). This suggests that DNA methylation is required for a full exit from pluripotency and transition to the epidermal state. However, the retention of the pluripotency marker *Sox3* late into development suggested a possible prolonging of cellular competency.



**Figure 3.1 5-Aza Heterogeneously Affects Pluripotency Factors.** (A) Treatment of embryos from stage two (initial cell cleavage) through stage 9 (late blastula stages) with 25  $\mu$ M 5-Aza leads to stochastic expression of normally homogeneous pluripotency factors *Id3*, *Ap2*, *Oct25*, *Sox3*, and *Vent2* in stage 9 blastula embryos. (B) This phenotype is partially reversible. When embryos are treated with 25  $\mu$ M 5-Aza at stage 2, and the drug is washed out at stage 5 (16 cell stage) there is only a mild perturbation in gene expression, with these pluripotency factors largely restored to wild type levels.



**Figure 3.2 5-Aza Partially Prolongs Sox3 and Prevents EpK Expression.** Treatment of embryos from stage two (initial cell cleavage) through stage 9 (late blastula stages) with 25  $\mu$ m 5-Aza followed by treatment of dissected animal caps from stage 9 to stage 13 (onset of lineage restriction) leads to partial retention of pluripotency marker Sox3 and stochastic expression of the epidermal marker, EpK in animal caps at stage 13.

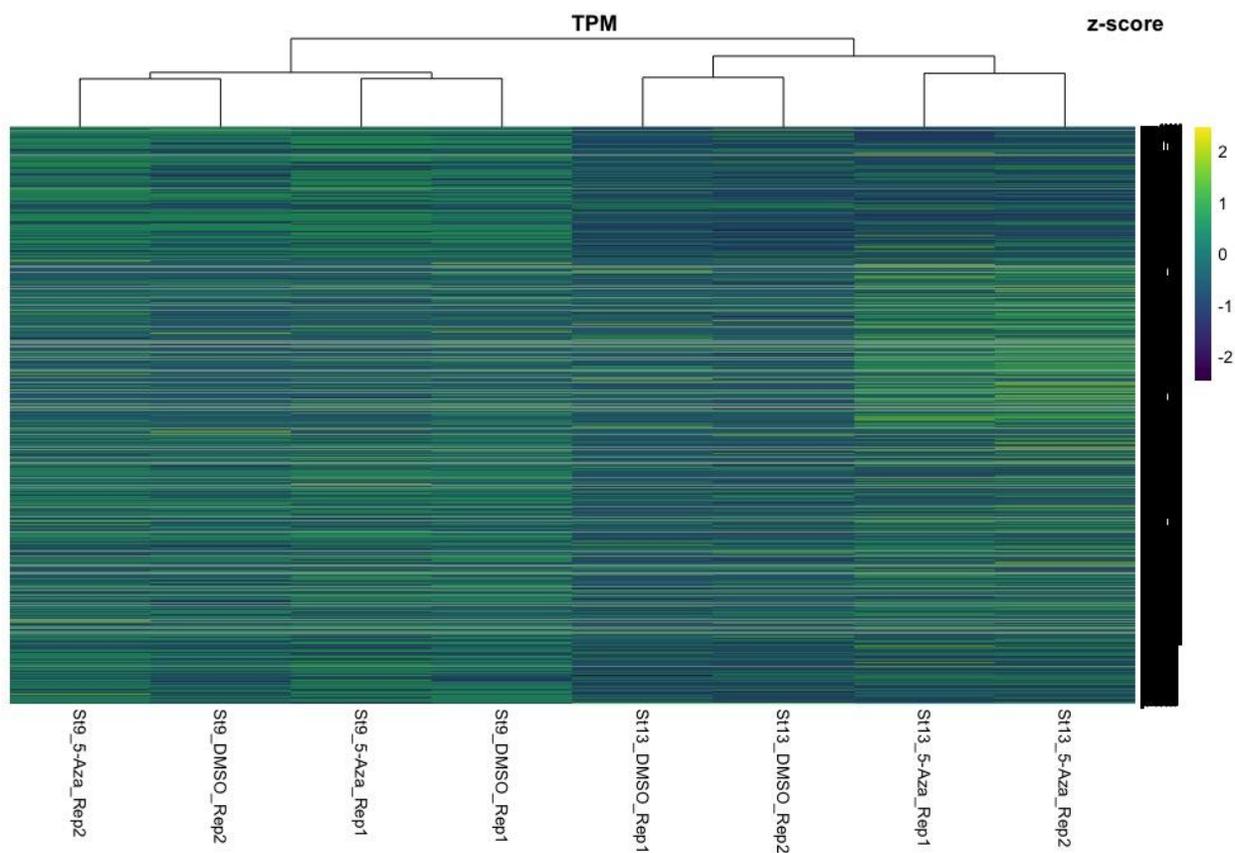


**Figure 3.3 5-Aza Reduces Cellular Competency.** Treatment of animal caps at stage 9 (late blastula stages) with 25  $\mu$ m 5-Aza leads to a loss in competency of blastula stage animal caps. These caps are unable to form mesoderm in response to low levels of Activin, as evidenced by the lack of expression of mesodermal markers *XBra* and *myoD*. These caps are also unable to form endoderm in response to high levels of Activin, shown by the lack of expression of endoderm markers *endodermin* and *Sox17*.

The heterogeneous expression of pluripotency markers in 5-Aza treated embryos indicated a decrease in overall pluripotent markers due to the stochastic expression of these genes in both embryos and animal caps, which might suggest a decrease in competency, however the prolonged expression of these genes into neurula stages suggested a possible mechanism for prolonged competency. To investigate the competency question, I took further advantage of the *Xenopus* animal cap assay. Animal caps can be directed to form any cell type in the embryo when exposed to the appropriate signaling molecules. To determine if blocking DNA methylation affects the ability of cells to form other lineages, I treated animal caps dissected from wild type embryos with 25  $\mu$ M 5-Aza upon dissection at stage 9 in order to determine how blocking DNA methylation affects the competency of animal caps in the absence of any residual effects from blocking DNA methylation at stage 2. When I treated animal caps with 25  $\mu$ M 5-Aza simultaneously with low levels of Activin, which is used to induce mesoderm, I saw a severe reduction of the mesodermal marker *Brachyury* (*XBra*) at stage 11.5 (Figure 3.3 A) and a complete loss of the mesodermal marker *Myod* at stage 18 (Figure 3.3 B). Similarly, upon simultaneous treatment of animal caps with 5-Aza and high levels of Activin to induce endoderm, I saw a nearly complete loss of the endodermal marker *Endodermin* (Figure 3.3 C). These results reflect a severe reduction in competency of animal caps in which DNA methylation is blocked, shown by the inability of caps treated with 5-Aza to form either the mesoderm or endoderm lineages in response to Activin.

### **DNA Methylation Inhibition Does Not Significantly Change Transcriptome Until Stage 13**

In my experiments, *in situ* hybridization shows prolonged expression of the pluripotency marker *Sox3*, but animal caps show a lack of competency. This suggests these pluripotency genes are not expressed at high enough levels to maintain the potential in cells signature of the pluripotent state. Since I was not seeing normal lineage restriction, characteristic of wildtype animal caps, I wanted to quantitatively assess changes to the transcriptome resulting from blocking DNA methylation by 5-Aza treatment using RNASeq. I treated whole embryos with 25  $\mu$ M 5-Aza or the equivalent volume of the vehicle DMSO at stage 2, cut animal caps at stage 9 and treated the dissected animal caps again with 25  $\mu$ M 5-Aza or DMSO at stage 9. I then collected 25 animal caps per condition at stage 9 and 13, performed the



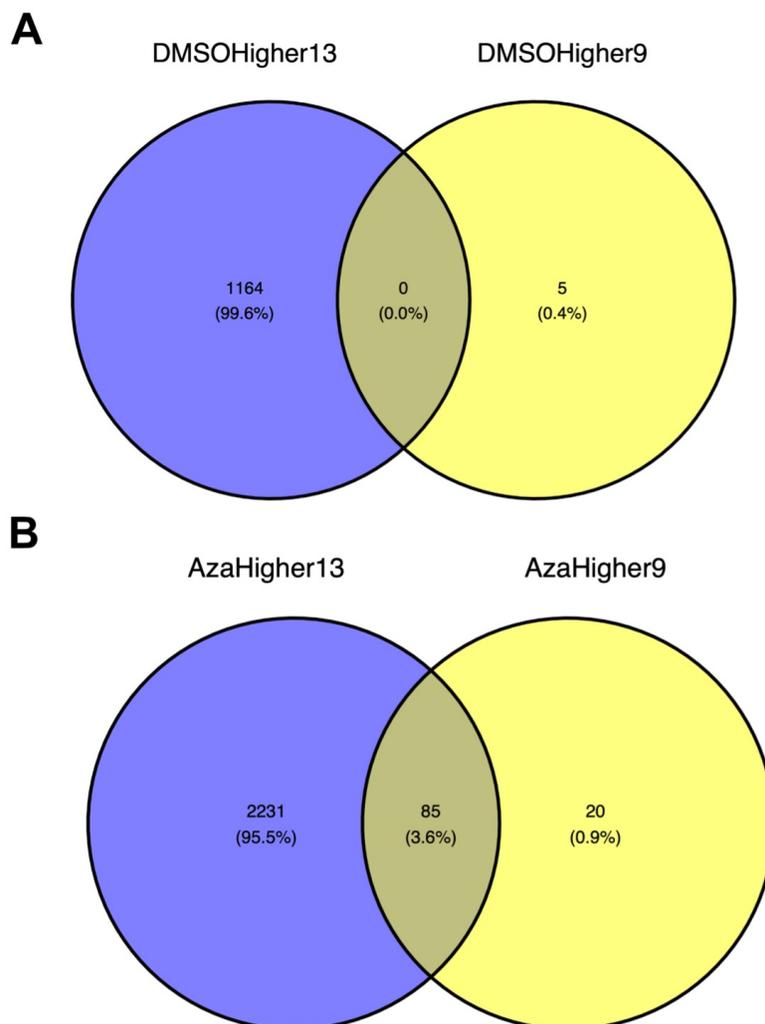
**Figure 3.4 Hierarchical Clustering of RNA Seq TPM.** Heat Map with hierarchical clustering of TPM of each sample shows that stage 9 replicates cluster by day, indicating there is not a significant difference between DMSO and 5-Aza treatments at stage 9. By stage 13, samples cluster by treatment, not day, indicating 5-Aza is causing biological differences that rise above sample to sample biological variability.

experiment in duplicate and used RNA sequencing to assess the transcriptional response to 5-Aza at stages 9 and 13.

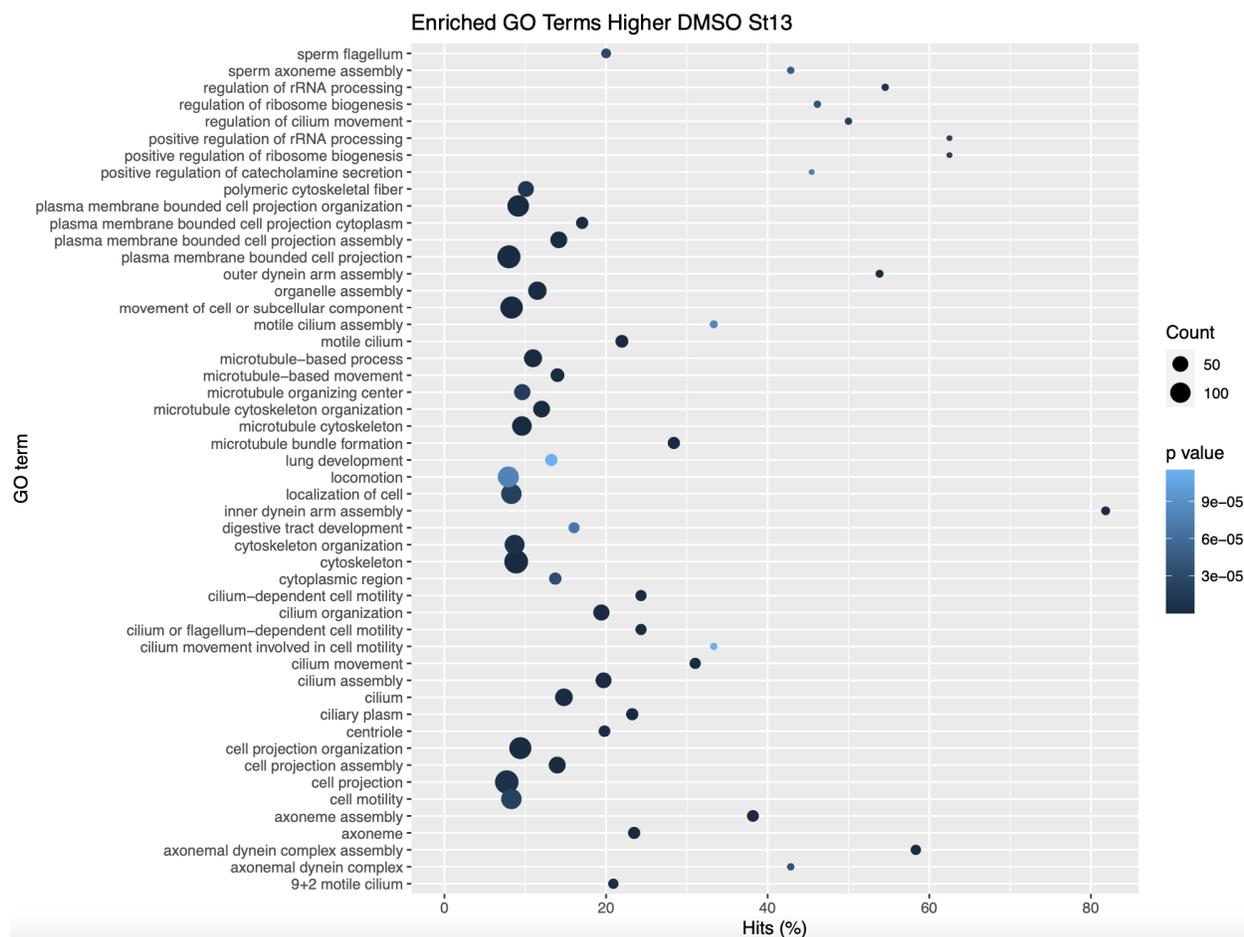
After aligning our data to the *Xenopus* genome to quantify gene expression in both raw counts and transcripts per million (TPM), I first assessed the sequencing results using hierarchical clustering of the TPM data (Figure 3.4). Clustering shows that, despite differences seen in *in situ* experiments, the samples cluster by day rather than condition for stage 9, suggesting that 5-Aza is not having a greater impact on the transcriptome than inherent frog to frog biological variability. By stage 13, however, samples cluster by condition, rather than by replicate, indicating the impact of loss of DNA methylation from 5-Aza treatment is greater than inherent sample to sample biological variability at this stage. Given hierarchical clustering revealed 5-Aza was significantly impacting the transcriptome by stage 13, I used DESeq2 to assess which genes were significantly differentially expressed both between conditions at each developmental stage, as well as over time in each condition in order to glean insight into the changes occurring in the transcriptome as a result of 5-Aza treatment. A p value of  $<.05$  and Log2FC of  $>1.5$  were minimum criteria for a gene to be considered significantly differentially expressed.

I first assessed changes in the transcriptome between conditions at each developmental stage. As expected based on hierarchical clustering, there were very few differentially expressed genes between the DMSO and 5-Aza treated genes at stage 9. Only 110 genes were differentially expressed between DMSO and 5-Aza at stage 9, and of those 105 were upregulated upon 5-Aza treatment (Figure 3.5 A, B). DESeq at stage 13 showed 3480 genes significantly differentially expressed between the two conditions, with two-thirds of those genes being upregulated by 5-Aza and one-third down regulated (Figure 3.5 C, D). I next wanted to determine which changes were a general response to 5-Aza seen at both stage 9 and 13, and which genes were significantly changed in response to 5-Aza at just one developmental stage, with stage 13 being of particular interest. Venn Diagrams of the differentially expressed genes reveal that, of the genes higher in the DMSO control group, none are downregulated by 5-Aza at both stage 9 and 13 and only 5 total are downregulated at stage 9, whereas 1164 are downregulated at stage 13 (Figure 3.6 A). Venn Diagrams of the genes upregulated by 5-Aza shows that again the vast majority of genes (2231) upregulated by 5-Aza are only significantly higher at stage 13, with 85 genes upregulated

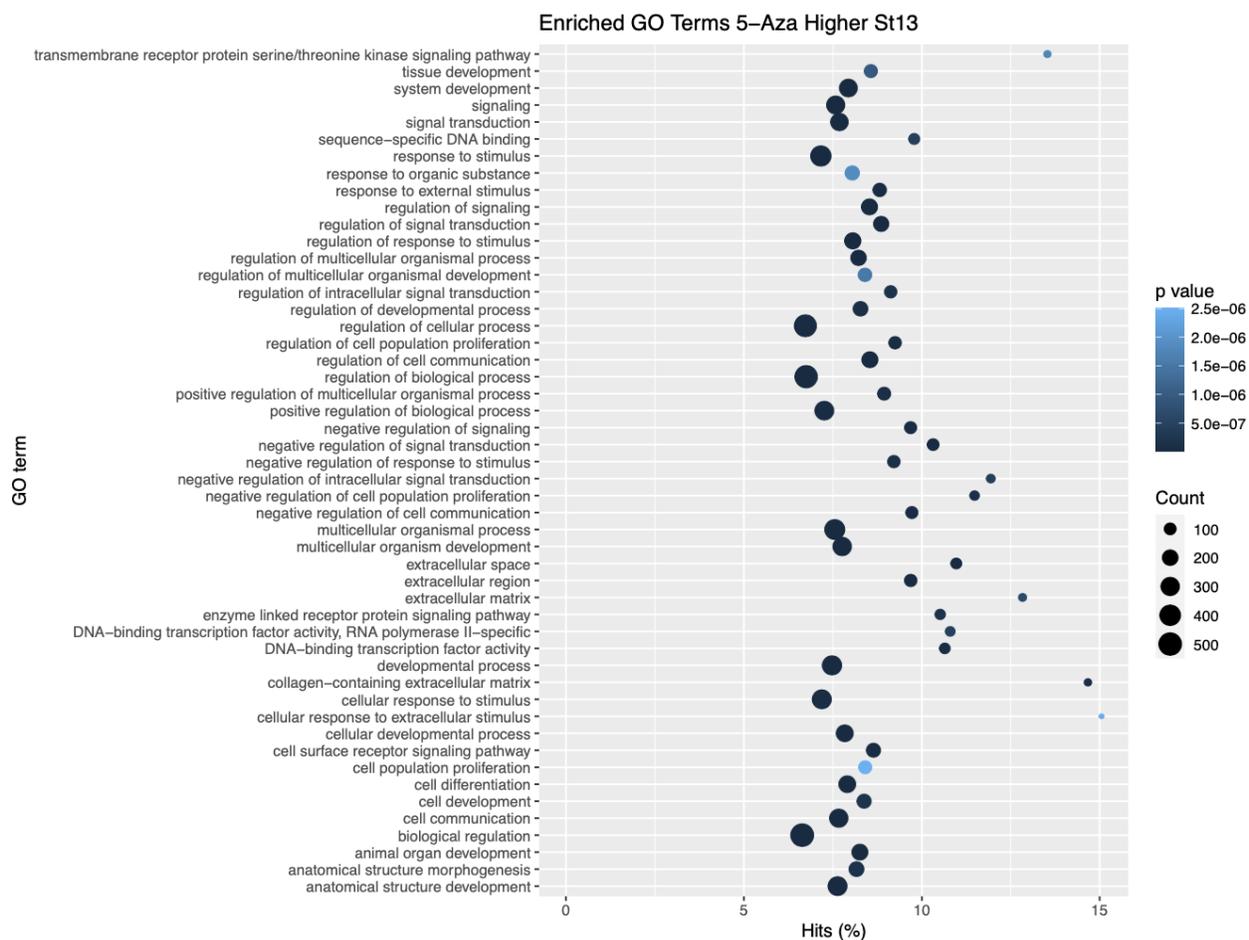




**Figure 3.6 Comparison of 5-Aza Effects at Different Developmental Stages.** Venn Diagrams of Differentially expressed genes with a Log<sub>2</sub>FC cut off of 1.5 between DMSO and 5-Aza at stages 9 and 13. (A) Genes higher in the DMSO control group, which are downregulated by 5-Aza. There are no genes downregulated by 5-Aza at both stage 9 and 13 and only 5 total downregulated at stage 9, whereas 1164 are downregulated at stage 13. (B) Genes upregulated by 5-Aza shows that again the vast majority of genes (2231) upregulated by 5-Aza are only significantly higher at stage 13, with 85 genes upregulated at both stages and only 20 genes uniquely upregulated at stage 9.



**Figure 3.7 GO Enrichment of Genes Higher in DMSO Caps at Stage 13.** Top 50 enriched GO Terms from GO Enrichment analysis of genes down regulated by 5-Aza at stage 13.



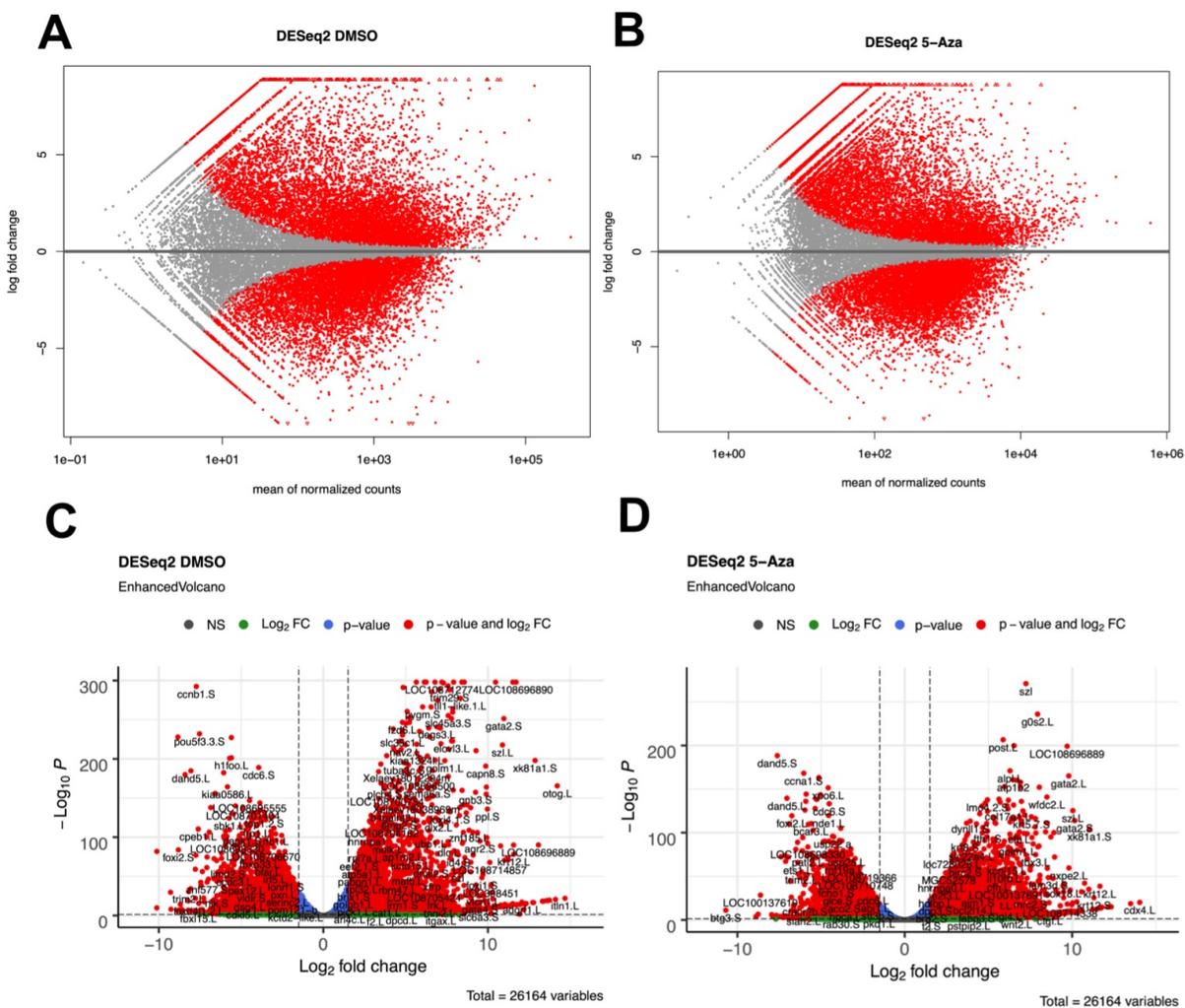
**Figure 3.8 GO Enrichment of Genes Higher in 5-Aza Caps at Stage 13.** Top 50 enriched GO terms from GO Enrichment analysis of genes upregulated by 5-Aza at stage 13.

at both stages and only 20 genes uniquely upregulated at stage 9 (Figure 3.6 B).

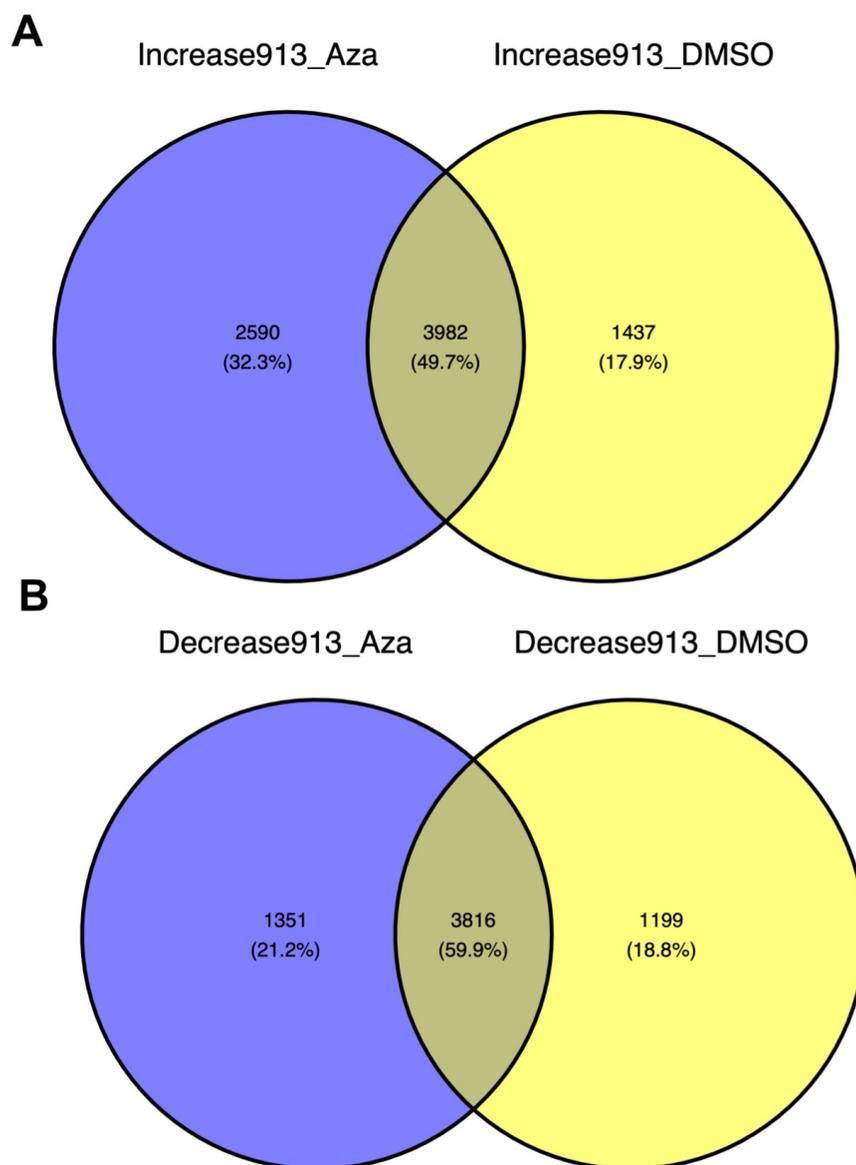
### **5-Aza Upregulates Cell Signaling Genes and Downregulates Structural Genes**

In an effort to see if there are any shared characteristics of the genes in each of these categories I performed gene ontology (GO) enrichment on each of these groups, which reveals biological processes significantly overrepresented in these DE Genes. Since very few genes are significantly differentially expressed between the two conditions and hierarchical clustering suggests these differences do not outweigh biological noise, I have focused the GO term analysis on genes that are only differentially expressed between these conditions at stage 13. GO enrichment of genes downregulated by 5-Aza shows enrichment for structural genes, such as genes required for cytoskeleton and microtubule based processes as well as genes required for locomotion, cilia development and other cellular motility processes (Figure 3.7). Genes upregulated by 5-Aza, however, show enrichment for processes related to cell signaling, communication and response to stimuli (Figure 3.8). This GO term analysis suggests that 5-Aza may be negatively affecting the structural integrity of cells and cellular motility, while enhancing cell signaling and communication, possibly because these cells are not being directed towards a single cell fate, but instead are responding to many signaling pathways at low levels simultaneously.

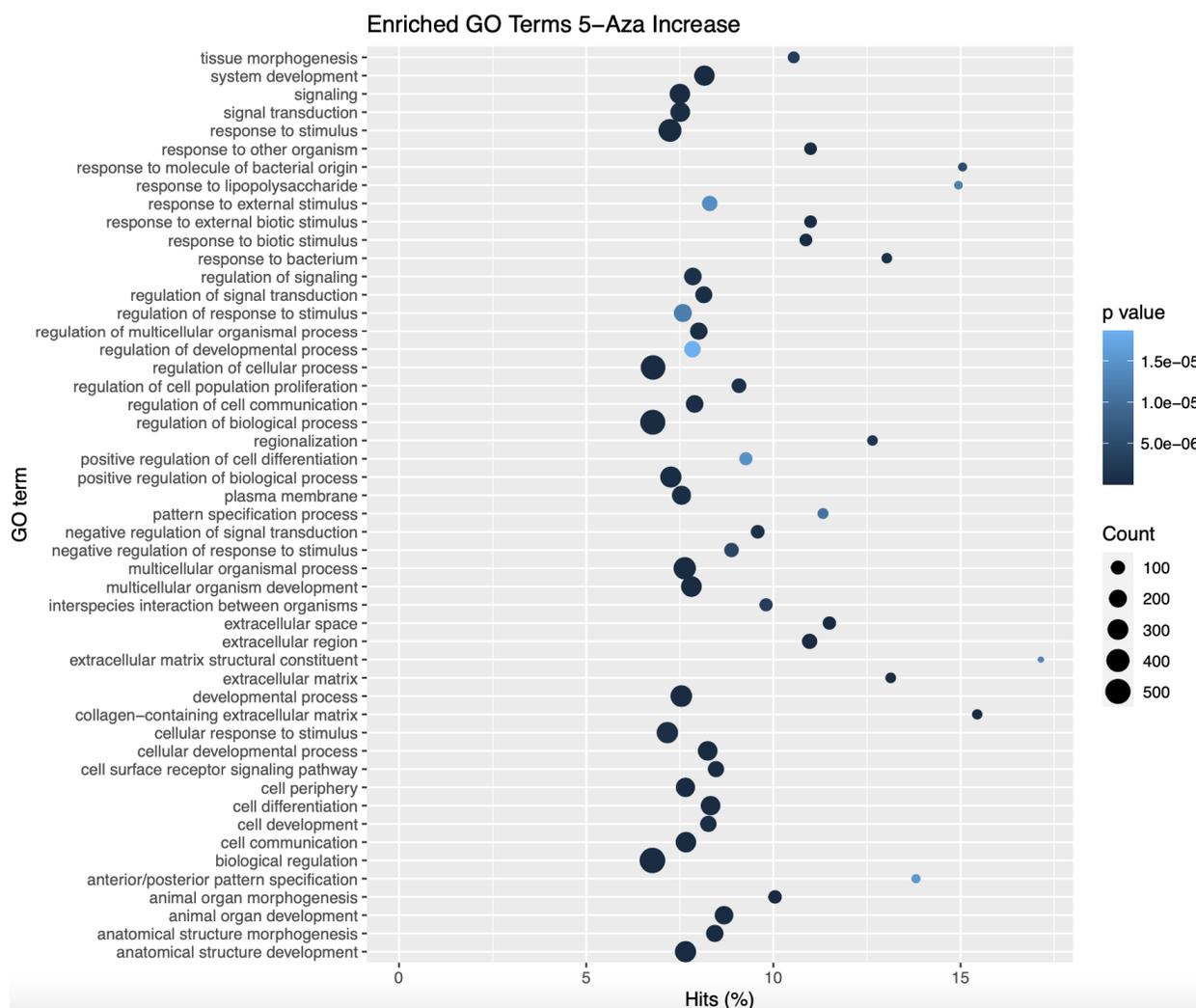
In addition to examining how genes differ between conditions, I was also interested in how treatment with 5-Aza affected how the transcriptome changes over time. I did this by performing DESeq between stage 9 and 13 for each condition separately. Volcano and MA plots reveal that while relatively equal numbers of genes are up and downregulated over time in the control DMSO condition, a disproportionate number of genes are upregulated over time in the 5-Aza treated caps (Figure 3.9). I then asked which genes were changing in the same direction over time in both conditions, and which genes were being up or downregulated over time in only one condition. These data are depicted in Venn Diagrams which show a 50% overlap in genes that are upregulated and a 60% overlap in genes that are downregulated. Of genes upregulated over time, 32% are upregulated only by 5-Aza and 18% are upregulated only in the control condition. The split was more even amongst downregulated genes, with 21% downregulated only upon 5-Aza treatment and 19% downregulated only in the control condition



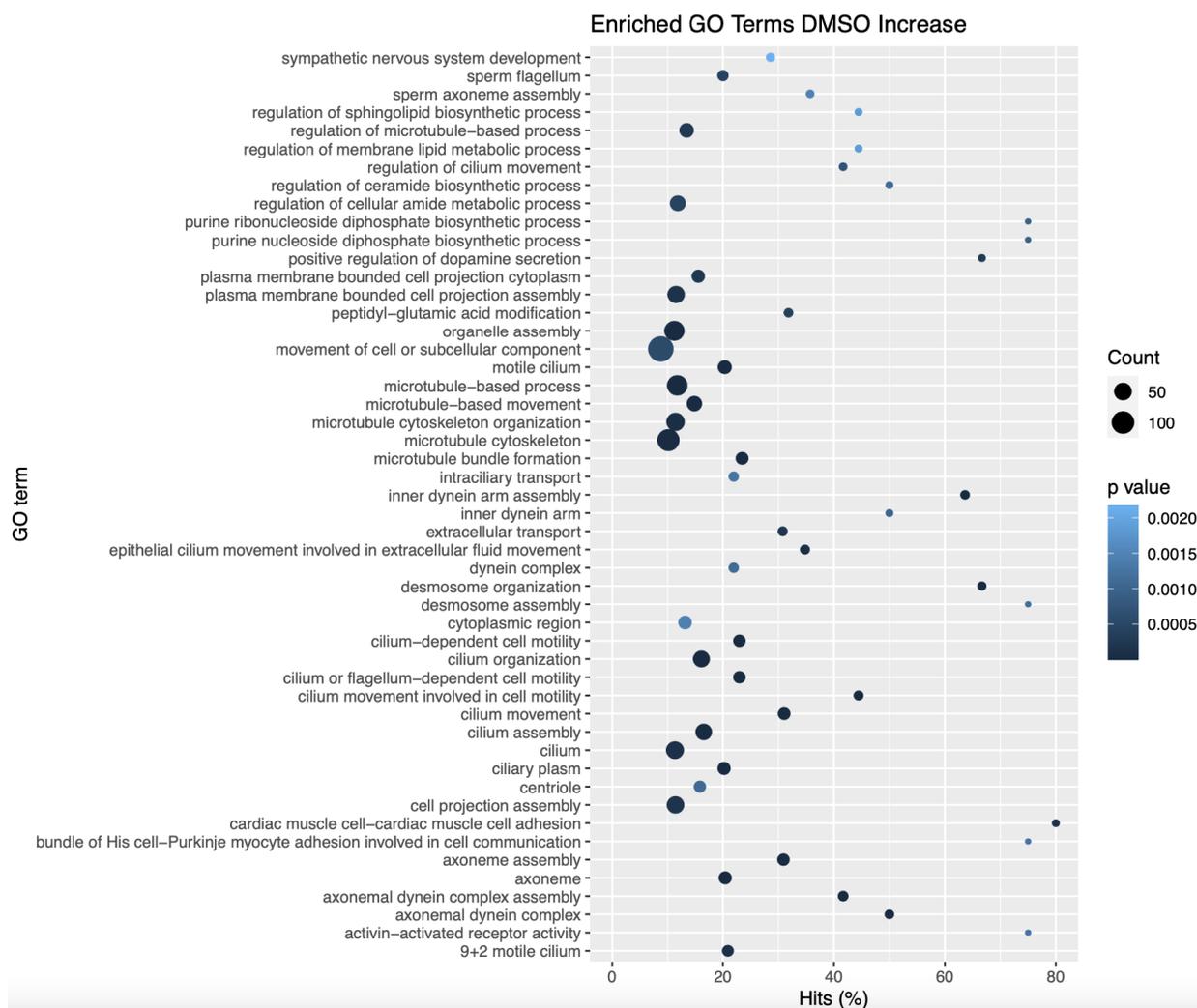
**Figure 3.9 DESeq Between Developmental Stages.** MA Plots (A,B) and Volcano Plots (C,D) of changes over time (stage 9 to 13) in the DMSO control (A,C) and 5-Aza treated (B,D) animal caps.



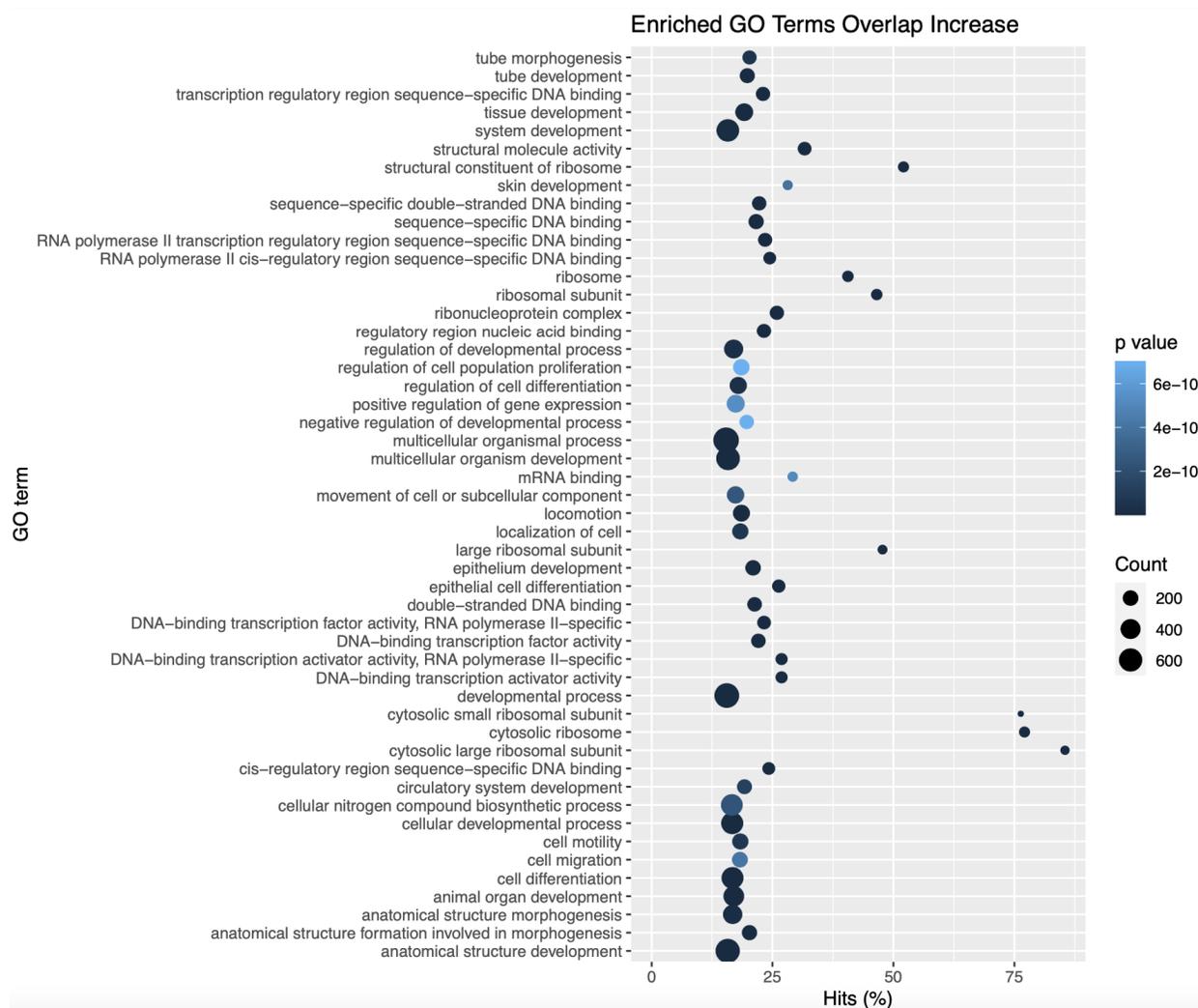
**Figure 3.10 Comparison of Transcriptome Changes Over Time.** Venn Diagrams of Differentially expressed genes with a Log<sub>2</sub>FC cut off of 1.5 between stage 9 and 13 in DMSO and 5-Aza treated caps. (A) Half the genes increasing over time are shared in both conditions, one third are increased only by 5-Aza and less than one fifth are increased only by DMSO (B) Nearly 60% of decreased genes are shared between conditions while approximately one fifth of the decreased genes are unique to each condition.



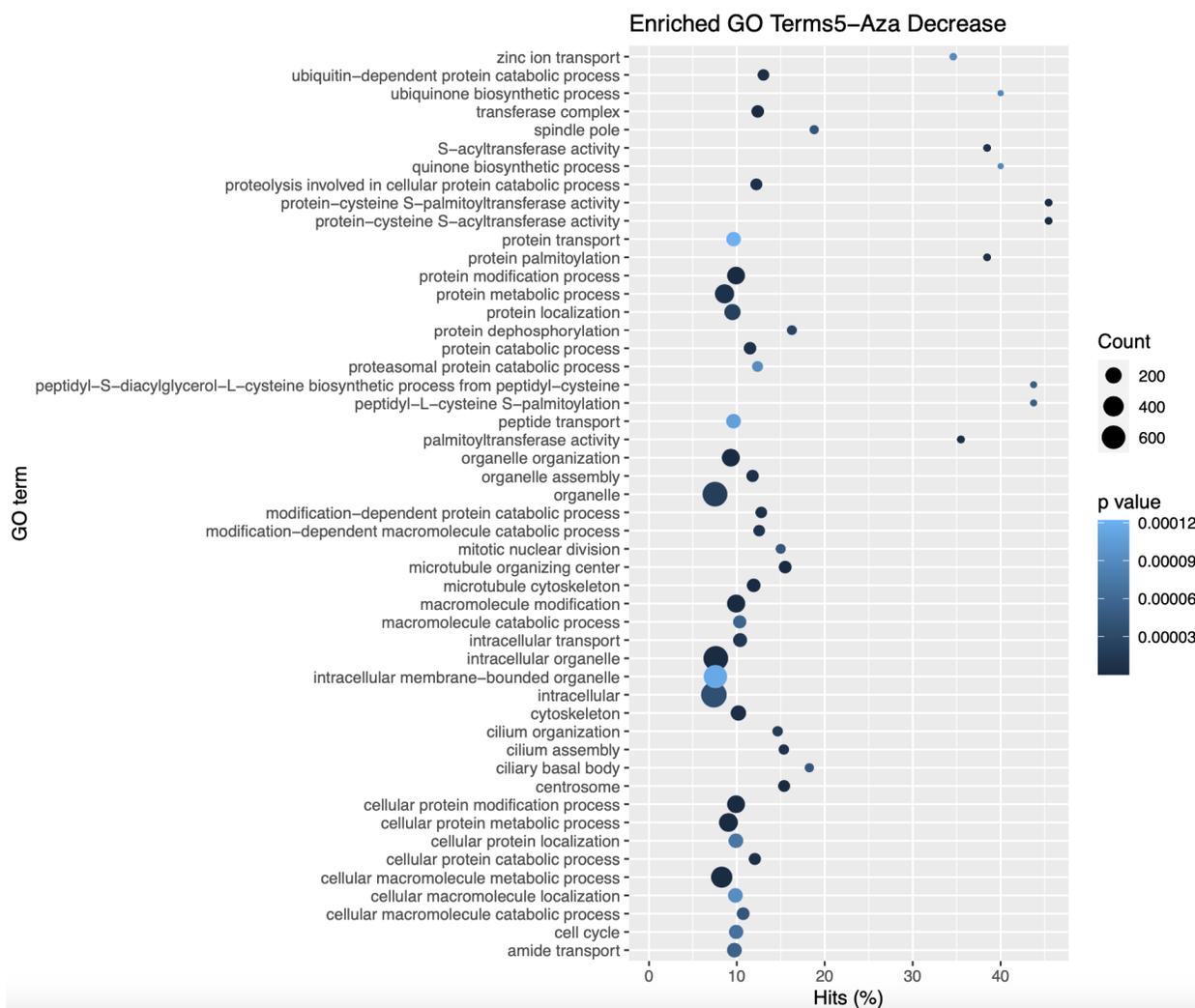
**Figure 3.11 GO Enrichment of Genes Increasing in Response to 5-Aza.** Top 50 enriched GO Terms from GO Enrichment analysis of genes upregulated over time only upon 5-Aza treatment.



**Figure 3.12 GO Enrichment of Genes Increasing with DMSO Only.** Top 50 enriched GO Terms from GO Enrichment analysis of genes upregulated over time only upon DMSO control treatment.



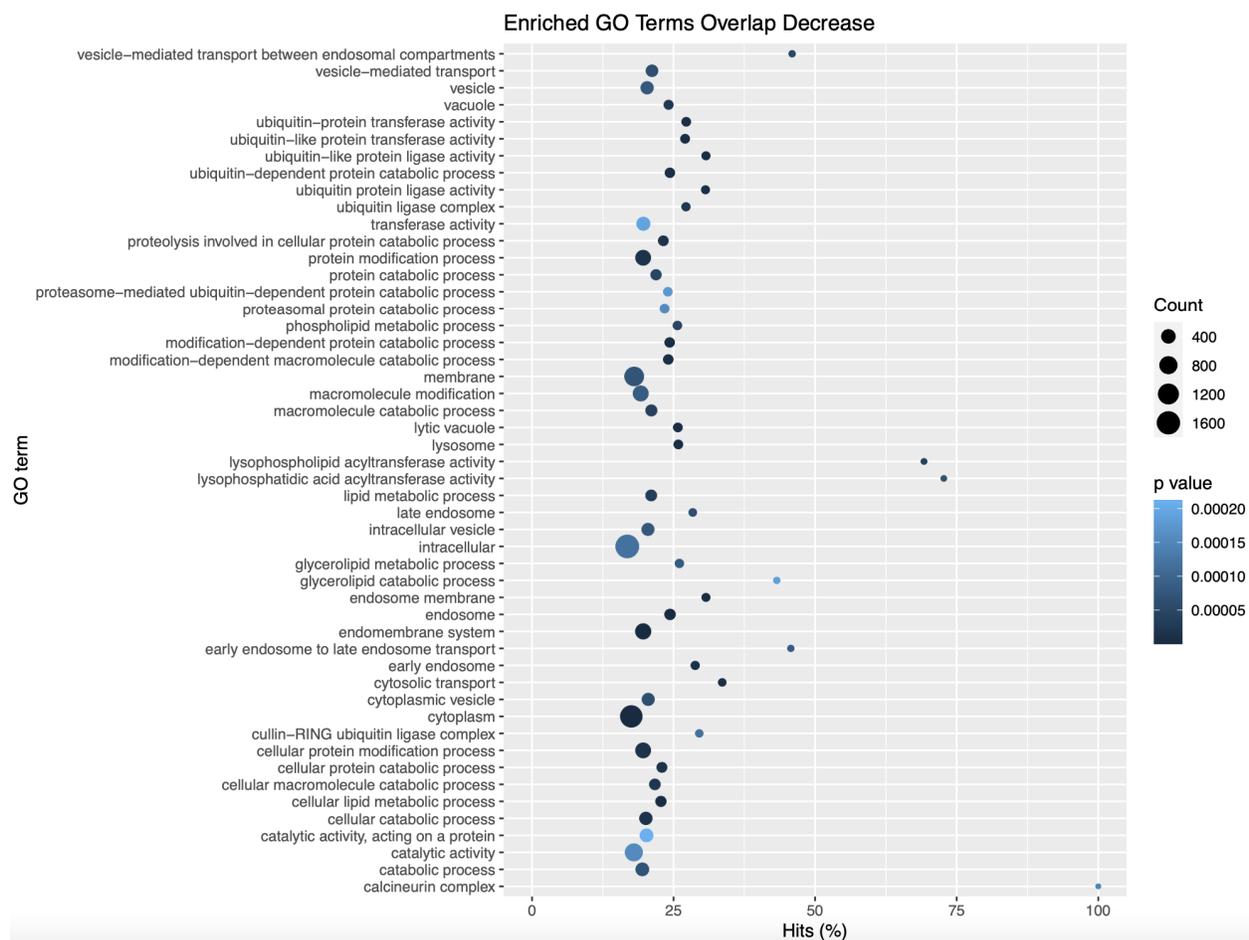
**Figure 3.13 GO Enrichment of Genes Increasing in Both Conditions.** Top 50 enriched GO Terms from GO Enrichment analysis of genes upregulated over time in both the DMSO and 5-Aza treated conditions.



**Figure 3.14 GO Enrichment of Genes Decreasing in Response to 5-Aza.** Top 50 enriched GO Terms from GO Enrichment analysis of genes downregulated over time only upon 5-Aza treatment.



**Figure 3.15 GO Enrichment of Genes Decreasing in DMSO Only.** Top 50 enriched GO Terms from GO Enrichment analysis of genes downregulated over time only upon control DMSO treatment.



**Figure 3.16 GO Enrichment of Genes Decreasing in Both Conditions.** Top 50 enriched GO Terms from GO Enrichment analysis of genes upregulated over time in both the DMSO and 5-Aza treated conditions.

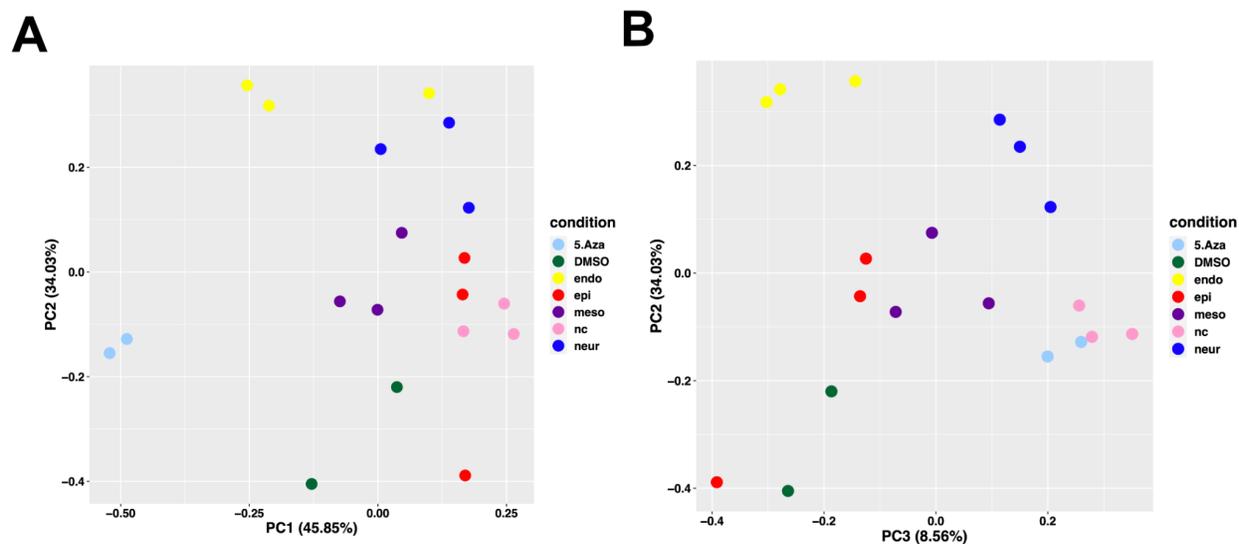
(Figure 3.10).

I again performed GO Term enrichment analysis on each of these categories to see how 5-Aza is affecting transcriptional changes over time and what biological processes are overrepresented in each category. Similar to the GO terms of genes upregulated by 5-Aza at stage 13, the GO terms of genes changing significantly over time only in response to 5-Aza are dominated by terms related to cellular signaling, communication and response to external stimuli (Figure 3.11). GO terms of genes that were upregulated only in the control condition overlapped significantly with GO terms of genes higher in DMSO at stage 13, dominated by cellular structure and motility GO terms, with the addition of terms related to metabolic and biosynthetic processes (Figure 3.12). Surprisingly, GO terms of genes whose upregulation was not affected by 5-Aza treatment was dominated by terms related to DNA binding, despite the fact that DNA methylation was inhibited, which directly impacts DNA binding. This suggests that while inhibiting DNA methylation may not affect DNA binding ability, it does bias where binding is able to occur. Other GO terms of genes unaffected by 5-Aza treatment include cell proliferation and differentiation as well as organismal development (Figure 3.13).

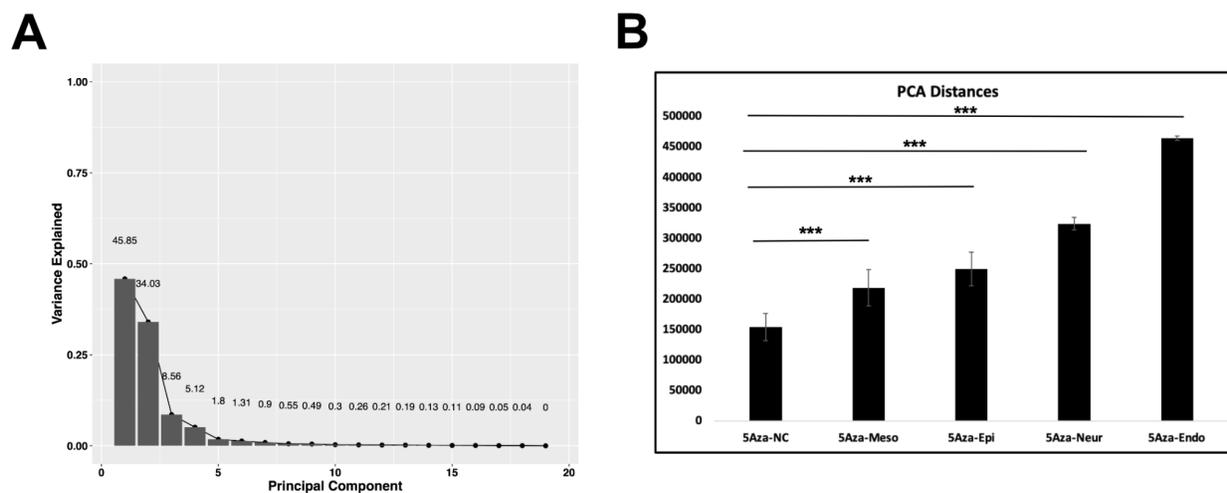
I also investigated how 5-Aza affected which genes were downregulated over time using GO Term analysis. Genes that were downregulated over time only in response to 5-Aza were dominated by terms relating to post translational protein modifications, including phosphorylation and ubiquitination, as well as protein catabolic and metabolic processes (Figure 3.14). By contrast, GO terms of genes that normally decrease over time, but are maintained in response to 5-Aza include terms related to embryo patterning, as well as neural and mesodermal formation (Figure 3.15). GO terms of genes downregulated over time regardless of 5-Aza treatment, similar to those down regulated only by 5-Aza, are predominantly related to protein modification, metabolism, and catabolism (Figure 3.16).

### **PCA Suggests Alternative Splicing Accounts for Majority of Variance**

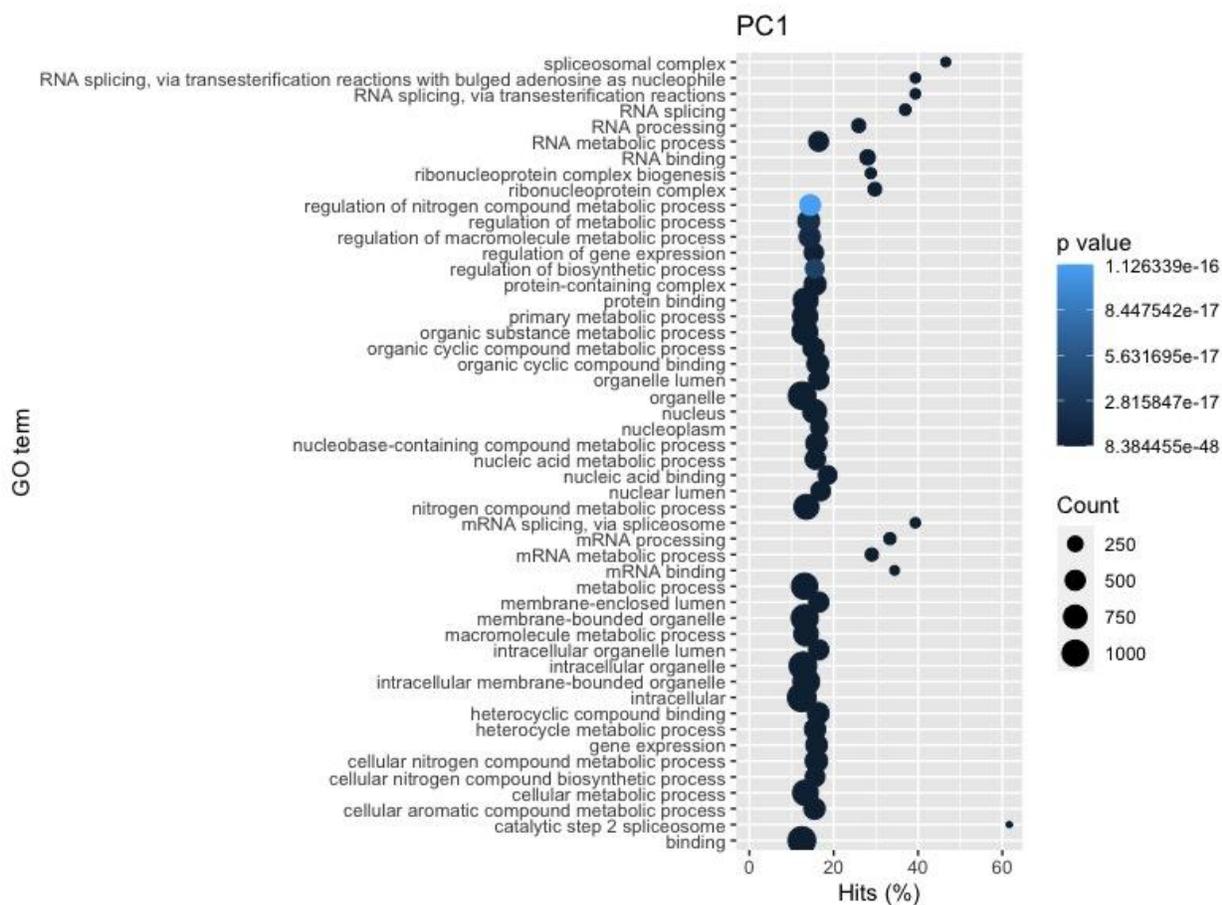
These analyses show that 5-Aza treatment was having a significant effect on over 10% of the genome with non-negligible expression. They also showed that, of the genes whose expression was affected by 5-Aza, there was enrichment of cell signaling related GO terms in genes increased by 5-Aza



**Figure 3.17 PCA of 5-Aza and Primary Lineages.** Principal Component Analysis of stage 13 data of the 5-Aza and DMSO treated animal caps, as well as epidermal, neural, mesoderm, endoderm and neural crest lineages. (A) Plot of PC1 and PC2 shows that the 5-Aza samples are markedly different from all other lineages along PC1 axis (B) Plot of PC2 and PC3 shows 5-Aza treated embryos clustering with the neural crest lineage.



**Figure 3.18 Principal Components Contributing to Stage 13 Variance.** (A) Scree plot of all 19 principal components in the PCA of all lineages shows the variance explained by each one. (C) Distance calculated between lineages of principal components 2-19 shows 5-Aza treated caps are significantly closer to the neural crest lineage than any other lineage.



**Figure 3.19 GO Term Enrichment for Top Genes Loading PC1.** GO Term Analysis performed on all genes with a loading  $>.001$  on PC1 (1856 genes) reveals that the genes contributing most significantly to this principal component are dominated by genes playing a role in RNA splicing.

as well as a maintenance of neural and mesodermal development terms that turn off during normal epidermal development. Because I have stage 13 transcriptome data for epidermis in the absence of DMSO, neural progenitors, mesoderm, endoderm, and neural crest, I wanted to see if 5-Aza treatment was not only biasing cells away from the epidermal lineage, but also towards other lineages. I assessed this using principal component analysis (PCA) run on stage 13 of all five of these lineages simultaneously with the 5-Aza and DMSO control treated embryos. This analysis showed two principal components with significant contributions to the variance. The first principal component (PC), accounting for 45.85% of the variance differentiates between 5-Aza treated samples and all other samples (Figure 3.17A). The GO terms of the genes with the most heavily weighted loadings in this component are dominated by RNA splicing related terms, likely due to the effect of DNA methylation on this process (Lev Maor et al. 2015, Yearim et al. 2015, Shayevitch et al. 2018). While this PC is important for differentiating between samples with and without DNA methylation, it is so dominant that it obfuscates the comparison between lineages. When I plotted the second two principal components, I can see that the 5-Aza treated caps cluster very closely with the neural crest lineage, whereas the DMSO control is closest to epidermis (Figure 3.17 B). In order to confirm that the 5-Aza samples were indeed closest to the neural crest lineage across PC space, I calculated the distance between conditions across all principal components except PC1, which together explain 54.15% of the variance (Figure 3.18 A). This calculation confirms that the 5-Aza treated embryos occupy a PC space that is significantly closer to the neural crest lineage than to any other lineage (Figure 3.18 B). The fact that 5-Aza treated embryos are most similar to the neural crest lineage suggests that preventing DNA from being methylated may be holding cells in a primed or multipotent state, while preventing cells from actually undergoing lineage restriction as evidenced by their lack of competency to form any of the primary germ layers.

## **Discussion**

It is well established that DNA methylation is a key regulator of DNA accessibility, and thus in controlling when and where different genes are expressed. However, the role of DNA methylation varies greatly amongst different species. For example, mammals undergo genome wide demethylation before

the maternal zygotic transition and implantation, but *Xenopus* and zebrafish maintain a hypermethylated state of DNA before and after the maternal zygotic transition (Monk et al. 1987, Macleod et al. 1999, Veenstra and Wolffe 2001). DNA is methylated by two de novo methyltransferases, Dnmt3A and 3B, and maintained by Dnmt1. Dnmt1 has been shown to be critical for the regulation of gastrulation in zebrafish (Kent et al. 2016) and Dnmt3A and 3B have been shown to play temporally distinct roles in neural crest formation in chicks (Hu et al. 2014b). Despite the evidence for the function of DNMTs in specific stages of development and specific tissues, the effect of global demethylation in early developmental stages in *Xenopus* had not been characterized.

*In situ* hybridization showed a heterogeneous effect of inhibition of DNA methylation on pluripotency markers at blastula stages, and while qualitatively these data seem to reflect an overall loss of expression of these genes, this is not reflected in quantitative analysis of the RNASeq data, as none of these core pluripotency factors were differentially expressed between 5-Aza and DMSO treated embryos. In fact, the number of genes differentially expressed between these two conditions at the blastula stage was negligible; less than .5% of the genes transcribed at this stage. Hierarchical clustering demonstrates that between sample variation was greater than between condition variation. The fact that this marked phenotype seen via *in situ* hybridization is not reflected quantitatively in the RNASeq data is surprising. However, it is possible that inhibition of DNA methylation is amplifying biological stochasticity, and that while there is a loss of transcription of pluripotency markers in some cells, there is an increase in others, leading to an averaging effect in the bulk RNASeq data.

It would be interesting to perform these sequencing experiments at the single cell level to be able to measure the cell to cell variability in DMSO versus 5-Aza treated caps to determine if it is in fact amplified by 5-Aza. It is also, of course, imperative to perform Bisulfite sequencing or to have another form of a direct readout in the effect that 5-Aza is having on methylation in order to ensure that these results are not the result of off-target effects. If loss of methylation is confirmed it would also be interesting to perform not just single cell sequencing but single cell spatial sequencing to see if the same 'salt and pepper' pattern seen *in situ* is reflected in the spatial sequencing data. A flattened, unhealed animal cap,

with optimization, would likely lend itself well to such analysis. This could also mitigate unwanted effects on the transcriptome of the non-physiological conditions of dissociation.

While differences at blastula stages were minimal, inhibition of DNA methylation with 5-Aza did have a marked effect on transcriptome at stage 13, the onset of lineage restriction. GO Term analysis for genes differentially decreased at stage 13 by 5-Aza is enriched for structural terms, suggesting 5-Aza treatment may be affecting structural integrity of cells, which did appear to be the case with caps treated with high levels of this inhibitor, though not at the 25uM dose used for sequencing. This could be a direct role of loss of methylation or could be an off target effect of toxicity. There is, however, no enrichment for apoptotic GO terms, suggesting that the loss of gene expression seen in some cells via *in situ* is not due to cell death. Interestingly, genes differentially increased by 5-Aza are enriched for cell signaling terms. This is again seemingly contradictory to the phenotype seen via *in situ* hybridization in which animal caps have an impaired ability to form epidermal, mesodermal, or neural lineages suggesting that loss of DNA methylation leads to a loss of cellular competency and a loss in the ability to respond to signaling cues. However, there is at least partial adoption of all these lineages, as well as a partial maintenance of *Sox3* expression, which may suggest a phenotype similar to that seen with inhibition of HDAC in which cells express genes of many lineages at low levels (Rao and LaBonne 2018). These caps may therefore be responding to many signaling pathways at low levels, leading to an enrichment in signaling genes, even though the response to any single pathway is not robust enough to lead to adoption of these lineages. This is further supported by the maintenance of mesodermal and neural GO terms upon 5-Aza treatment, which are usually lost over time, suggesting possible low level responses to signaling pathways driving these lineages.

This inability to fully form any lineage demonstrated via *in situ* led us to inquire which lineage 5-Aza treated caps were most similar to at stage 13. PCA reveals that caps absent methylated DNA cluster closest to neural crest animal caps. Given caps treated with 5-Aza seem able to partially adopt multiple lineages but not commit to any suggests that they may be in a state resembling multipotency, where given the right stimulus they could still commit to multiple fates. If this is the case it makes sense that these caps would be most similar to neural crest cells, which also maintain their multipotency at these

stages and are able to give rise to derivatives from multiple germ layers. It is interesting, however, given the maintenance of the shared pluripotency / neural gene *Sox3*, that these caps weren't more similar to the neural lineage. This could be due to the fact that these caps are responding at low levels to BMP signaling and partially forming epidermis, while the neural lineage is defined by the absence of BMP signaling.

The most interesting result to come out of PCA, however, was that nearly half the variance can be explained by PC 1, on which 5-Aza treated caps are completely isolated. GO term enrichment analysis of the genes most significantly loading this component is dominated by splicing terms, suggesting that most of the variance between caps with and without DNA methylation is regulation of splicing, not the transcription of pluripotent or lineage specific genes. There is evidence that DNA methylation regulates the rate of Pol II pausing as well as exon recognition and alternative exon inclusion. These sequencing results suggest that splicing regulation might actually be a primary role of DNA methylation, rather than a role secondary to regulating transcription factor binding. Given the interest in DNA methylation inhibitors as potential cancer treatments, it would be interesting to investigate the use of moderating DNA methylation in other diseases known to be associated with splice variants.

## **Materials and Methods**

**Embryological Methods.** Wild-type *Xenopus laevis* embryos were obtained using standard methods and staged according to Nieuwkoop and Faber (1994). For animal cap assays, ectodermal explants were manually dissected at early blastula (stage 8-9) from embryos treated with either 25uM 5-Aza or the equivalent volume of the inhibitor vehicle DMSO as a control. Embryos were allowed to develop in .1X Marc's Modified Ringer's Solution (MMR) [0.1 M NaCl, 2mM KCl, 1mM MgSO<sub>4</sub>, 2 mM CaCl<sub>2</sub>, 5mM HEPES (pH 7.8), 0.1 mM EDTA]. Once dissected, animal caps were cultured in 1X MMR with either 25uM 5-Aza or the equivalent volume of DMSO. For *in situ* experiments, caps or embryos were collected at the denoted stage and fixed for one hour in 1xMEM [100mM MOPS (pH 7.4), 2mM EDTA, 1mM MgSO<sub>4</sub>] + 4% formaldehyde then dehydrated in 100% ethanol followed by 100% methanol. *In situ*

hybridization was performed using digoxigenin-labeled RNA probes and developed using BM Purple substrate (Roche) (LaBonne and Bronner-Fraser, 1998) For RNA sequencing experiments, groups of 25 explants were collected either at stage 9, after being allowed to heal for one hour, or at stage 13 using sister embryos to confirm staging and snap frozen in 500uL in Trizol in liquid nitrogen.

**RNA Isolation and Sequencing.** RNA was isolated from blastula explants (25 explants) using Trizol (Life Technologies) followed by LiCl precipitation. 500 ng of RNA was used for library prep with TruSeq mRNA library prep kit (Illumina) and sequenced using Next Seq 500 Sequencing.

**RNA Seq Processing and Computational Analysis.** Read quality was evaluated using FastQC (Andrews 2010). Mapping to *X. laevis* v9.2 genome downloaded from xenbase was performed using RSEM to get TPM values (Li and Dewey 2011). Alignment to the *X.laevis* v9.2 genome was performed using STAR 2.6.0 to get raw counts using standard parameters (Dobin et al. 2013). Computational Analysis of RNA Sequencing Data was performed using published R Packages. A minimum raw read count of 15 was determined computationally using the voom function of the limma package. Based on the plot of the mean variance trend, filtering out genes with expression below 15 provided the best balance between filtering out lowly expressed genes without losing genes relevant to the transcriptome dynamics, thus all analyses are performed on genes with a minimum read count of 15 in any lineage at any stage (Law et al 2014). TPM values for these genes were plotted using the pheatmap package (Kolde 2015). Hierarchical clustering was done using the pheatmap package, cluster rows=TRUE, cluster columns = TRUE.

Differential expression analysis was done between conditions at each developmental stage and between stages in each condition using DESeq2 with significance defined as  $p_{adj} \leq 0.05$  and  $\text{Log}_2\text{FC} > 1.5$  (Love et al, 2014). Overlapping DESeq genes were visualized using VennDiagram (Chen and Boutros 2011). GO and KEGG enrichment was calculated using the GOSeq R package (Young et al. 2012). Input

genes for GO enrichment were the categories of overlapping or unique differentially expressed genes denoted in the Venn Diagram plots.

Principal Component Analysis was done on stage 13 samples on the 5-Aza experiments as well as the stage 13 counts from lineage restriction data in chapter two with the `prcomp` function in the default stats package and visualized using `ggfortify` (Tang and Horikoshi 2016). Distances between lineages were calculated using the `dist()` function across PCs 2-6, PC The top 6 PCs are the only PCs that account for greater than 1% of the variance and PC1 was eliminated because it was disproportionately affected by drug treatment . Statistical significance of differences in distance were calculated using the Wilcoxon rank sum, `wilcox.test()` in R, significance defined by a p value < .05.

**Vertebrate Animals.** All animal procedures were approved by Northwestern University's Institutional Animal Care and Use Committee and are in accordance with the National Institutes of Health's Guide for the care and use of Laboratory Animals.

## **Chapter Four**

### **General Discussion**

Key to embryonic development is the progressive restriction of cellular potential. A totipotent fertilized egg divides into pluripotent stem cells that are further restricted during gastrulation to three multipotent germ layers; ectoderm, mesoderm and endoderm, which each give rise to unique unipotent derivatives. The developmental period of gastrulation is imperative for proper body plan formation, as this is when embryos begin to undergo significant morphological changes, a process sensitive to both environmental and genetic conditions. This process of lineage restriction has been famously, albeit crudely, depicted by Conrad Waddington in 1957 as a topological landscape (Waddington 1957). While Waddington's model was largely theoretical, he was accurate in his prediction that these developmental processes were heavily influenced by both environment and genetics, and that changes to a cell's pathway are driven by large cohorts of genes, not individual genes (Waddington 1940, 1942, 1968, Robertson 1977, Baedke 2013, Tronick and Hunter 2016). To answer the questions regarding which genes are involved in the formation of each of these lineages, *Xenopus* have been used as a tool for decades to conduct classic developmental biology experiments in which individual candidate genes have been examined through overexpression and knockdown studies to determine their necessity and sufficiency for the development of various lineages and regulation of developmental processes. These experiments have helped to flesh out this complex landscape gene by gene and have led to the construction of gene regulatory networks.

Another question arising from the idea of a landscape with multiple paths to different lineages is whether one of these paths and lineages is the default. Again, *Xenopus* have been one of the more powerful model organisms to investigate this question, with evidence suggesting that neural is the default state of a cell. The evidence for this is two fold. First, it has been shown that molecules that induce a neural fate, including Noggin, Chordin and Follistatin, are not signaling activators, but rather inhibitors of BMP signaling (Zimmerman et al. 1996, Piccolo et al. 1996 and Fainsod et al. 1997). Second, upon dissociation, when cells are no longer signaling intercellularly, animal pole cells express neural rather than the normal epidermal markers showing that indeed neural cells can form absent neural "inducers" (Sato and Sargent 1989, Godsave and Slack 1989, Grunz and Tacke 1989). This model, however, is not without controversy as ectopic expression of BMP does not inhibit neural plate formation in chicks, nor

does misexpression of its antagonists induce neural progenitors (Streit and Stern 1999a, Streit and Stern 1999b, Stern 2005). Furthermore, the neural default model lacks quantitative backing.

Finally, investigation of the epigenetic portion of Waddington's epigenetic landscape has been expanded and more specifically defined to not just include environmental stimuli, but the specific influences of epigenetic readers, writers, and erasers, leading to the question of how these modifiers affect the lineage restriction process. The accessibility of DNA needs to be regulated both spatially and temporally to allow for proper lineage and cell type formation, which is regulated largely at the epigenetic level. It has been shown that both DNA and histone methylation, as well as histone acetylation all play an important role in regulating the accessibility of DNA, and thus the ability for transcription factors to bind to initiate transcription and gene expression. Histone acetylation is an activating mark (Brownell et al. 1996), DNA methylation is a repressive mark (Boyes and Bird 1991), and the role of histone methylation is context dependent (Strahl et al. 1999, 2002, Cao et al. 2002, Schultz et al. 2002).

Biological tools, ever evolving, have expanded rapidly in the last couple decades to become increasingly quantitative and powerful in their ability to assay the entire genome at once. Next Gen Sequencing allows for quantitative assessment of RNA transcripts, protein binding and chromatin accessibility among other things. The findings presented here provide novel quantitative insight into the dynamic transcriptome changes undergone by each lineage from pluripotent to early lineage restriction stages. Specifically, I provide quantitative support for the neural default model, posit novel putative members of the mesendoderm GRN and demonstrate the effects on the transcriptome of shifting the timing of BMP signaling and provide insights into how the transcriptome changes in response to loss of DNA methylation.

### **Quantitative Insights into The Dynamics of Lineage Restriction**

Classically in developmental biology, individual genes have been probed to determine if they are necessary and sufficient to perform certain functions or form different lineages. These experiments have been incredibly informative and instrumental to formation of gene regulatory networks and our current understanding of the lineage restriction process. However, paramount to understanding the role of

different genes in embryonic development is understanding their temporal and spatial dynamics. In this thesis work I have presented a quantitative, high resolution temporal analysis of the transcriptome in transit from the pluripotent state to four distinct lineages. Temporal transcriptome analysis has been performed on both *Xenopus laevis* and *Xenopus tropicalis* previously (Owens et al. 2016, Session et al. 2016), and while this expression data has proved a valuable tool for the *Xenopus* community, these are both whole embryo studies, from which lineage specific spatial information cannot be extrapolated. These studies are thus unable to distinguish between genes expressed at low to moderate levels in all cell types from genes that are expressed at high levels in a tissue specific manner. Having lineage specific information is key to understanding the role of genes that have distinct dynamics in different lineages, such as *Otx2* which plays an important role in the endoderm lineage during gastrulation and a separate and distinct role at neurula stages in the neural lineage. Furthermore, these studies have low statistical rigor, as they have only two replicates with very large statistical error in many genes. Nevertheless, these data sets have been a valuable tool for directing studies of specific genes, and the four data sets presented in this thesis work provide valuable and novel lineage specific insight into *Xenopus* transcriptome dynamics.

Single cell transcriptome analysis has also been performed in *Xenopus* (Briggs et al. 2018) which better provides spatial information than the aforementioned whole embryo studies. However, use of harsh buffers to dissociate whole embryos inherently changes the nature of the cells being analyzed, albeit slightly, and our experiments are performed in more physiological conditions. While our experiments are performed in animal caps, and thus *in vitro*, they differ from typical ES cell work, in which the stem cell state is artificially maintained, in that they are performed in pluripotent cells dissected from a whole embryo, fated to become multiple cell types. Transcriptome dynamics in animal caps can therefore be more naturally extrapolated to give an idea of spatial dynamics in the developing embryo than from ES cells maintained in culture. While the data sets generated for this thesis work provide important information about lineage formation in developing embryos it would be most informative to gather transcriptome data from cells in a developing embryo rather than snapshots of the transcriptome from pooled RNA as is done here. This has been done successfully in zebrafish with genome editing of

synthetic target arrays for lineage tracing (GESTALT) (McKenna et al. 2016). This technique relies on CRISPR injected barcodes and can be used to infer lineage restriction patterns based on inheritance and accumulation of mutations. This technique presents some technical challenges in *Xenopus* because CRISPR is not as well optimized in frogs as in fish, but this type of analysis in developing embryos would be an exciting next step in further elucidating patterns of lineage formation.

### **Quantitative Backing and Biological Insight into the Neural Default Model**

While it has been shown through multiple techniques that neural is the default state of cells in *Xenopus*, the question remains as to why this cell type is the default over all others. The prevailing explanation for the default nature of neural cells is that these cells are induced by inhibitory molecules, specifically inhibitors of BMP signaling. It is a compelling, yet somewhat sparse explanation, that a cell that forms in the absence of endogenous signaling would be the default state. In this thesis I've taken a quantitative approach to address the neural default hypothesis. I've shown through multiple quantitative methods that the path to the neural lineage is shortest and most linear in terms of transcriptome dynamics. PCA demonstrates that across all lineages, samples cluster along principal component 2 based on developmental stage. While endoderm, mesoderm and epidermis extend the entirety of the PC2 axis, the neural lineage does not extend beyond stage 10.5 in this axis. Temporal DESeq corroborates this idea of an early equilibrium for the neural lineage, as it is the only lineage in which a negligible number of genes change expression levels after stage 10.5. Indeed, examination of individual genes such as *Sox3*, *Sox11* and *Zic1* illustrate these linear dynamics unique to the neural lineage. Furthermore, GO Term Analysis shows that both the epidermal and neural lineage are enriched for many neural GO Terms before stage 10.5, suggesting the pluripotent gene enrichment signature overlaps heavily with the neural signature. Together these results convey that not only is neural the default because it can be formed in the absence of signaling, but it is the most similar to the pluripotent state, and thus requires fewer changes in the transcriptome to achieve. In light of the neural crest pluripotency maintenance model proposed by the LaBonne lab, it would be interesting to perform time course transcriptome analysis on the neural crest in order to see if a similarly one-dimensional path is observed.

If neural is truly the default state of a cell, as work in the literature and this thesis suggest, this begs the question of how this might affect cellular competency. It is generally understood that pluripotent cells remain competent to respond to signals directing them towards different cell fates up until the onset of gastrulation at stage 10. If the neural state is closer to the pluripotent state, then one might hypothesize that cells can be redirected away from the neural state and towards a different cell state later into development because it would not require the same reversal of the transcriptome that would be needed to direct cells away from a lineage like endoderm that undergoes significant and immediate transcriptome changes in response to Activin. It would be interesting to use the animal caps system to investigate in a systematic way how each lineage varies in competency. One hypothesis is that the stage at which cells are competent to respond to lineage inducing signals is not malleable. However, it is also possible that cells can be redirected away from neural at a later stage than other lineages, or that there is prolonged competency between ectodermal derivatives, or endodermal derivatives, but that cells cannot be redirected away from an ectodermal fate and toward an endodermal fate and vice versa. This could be addressed by providing animal caps with the signaling molecules sufficient to induce one lineage at stage 9 and then removing caps from this signal and providing a second signal towards a second lineage at different time points and examining changes in the transcriptome. While there are technical challenges to keeping caps open without changing the physiological conditions of the experiment, these can likely be overcome using a physical mechanism to keep caps open rather than chemical.

### **Expansion of Mesendoderm GRN**

Numerous functional studies have contributed to the composition of a robust mesendoderm gene regulatory network (Charney et al. 2017). This network consists of proven and putative gene connections. In this thesis I assess this gene regulatory network with quantitative analysis tools. First, by assessing the dynamics of the entire transcriptome during both mesoderm and endoderm formation I was able to provide a more precise order in which genes in this network turn on, based on a minimum TPM threshold. Second, I used multiple quantitative methods to identify novel putative targets for the mesendoderm GRN. Given the endoderm lineage has a robust and transient response to activin, it is uniquely

dominated by nonmonotonic genes. I used limma analysis to assess differential quadratic dynamics of genes in the endoderm and mesoderm lineages to identify two genes, *Tmcc1* and *Nptx2*, with patterns similar to mesendoderm GRN members that are likely playing a role in formation of, but not defining these lineages, as well as differential linear dynamics to identify two genes, *Ca14* and *Pygm*, increasing over time in these lineages that may play a role in defining these lineages.

While evidence from other RNA Sequencing studies corroborates a potential role for these four genes identified through limma analysis, they have not yet been functionally assessed. RNA Seq data suggests these four candidates are uniquely expressed in endoderm and/or mesoderm at these developmental stages. It would be interesting to characterize their expression in whole embryos with *in situ* hybridization to determine that they are indeed robustly and uniquely expressed in these lineages. It would also be interesting to test their necessity for proper formation of these lineages through morpholino knockdown experiments in both animal caps and whole embryos, as well as to use mRNA injections to see if overexpression leads to expansion of these lineages at the expense of others in whole embryos.

We corroborated these four genes as novel putative mesendoderm GRN members using WGCNA and found that indeed all four of these genes cluster in the same modules as the majority of established mesendoderm GRN members. Given WGCNA generated only three modules with moderately high correlation to the endoderm and mesoderm lineages, inclusive of known mesendoderm GRN members, it would be interesting to examine the expression levels of these relatively few genes in order to identify others with substantial expression unique to the mesendoderm lineages, and to corroborate the genes in these clusters with enriched genes in other mesoderm or endoderm RNASeq studies in order to possibly identify further mesendoderm GRN candidates that could be functionally assessed.

### **Examination of BMP versus Activin Targets**

Mesoderm induction experiments are typically done in *Xenopus* animal caps using Activin. Indeed, a transcriptome analysis of Activin induced mesoderm caps was recently published (Satou-Kobayashi et al. 2021). However, BMP4/7 heterodimers have also been shown to be potent ventral mesoderm inducers (Nishimatsu and Thomsen 1998). By using BMP4/7 instead of Activin to induce

mesoderm in my experiments I am able to not only examine transcriptome dynamics of formation of each of these lineages, but also compare the transcriptome response of cells to both branches of TGF-beta signaling. This comparison revealed surprising insights into the nature of the transcriptome response to early physiological levels of BMP. Adding physiological levels of BMP to animal caps a stage earlier than BMP signaling is detectable endogenously led to early expression of BMP target genes, as predicted. Surprisingly, however, early BMP induced not only BMP targets, but also Activin/Nodal target genes, including zygotic *Vegt* and *Eomes*. These genes were turned on in the absence of Nodal signaling, as shown by the phosphorylation of Smad1 absent Smad2 phosphorylation. Not all Nodal targets, however, show elevated transcription in response to early BMP signaling, for example *Otx2* and *Osr2* are not transcribed at biologically relevant levels in the mesoderm data set. This leads to the questions of how and why BMP signaling is inducing a subset of Nodal target genes.

This question could be best addressed by layering the RNASeq performed here with Chromatin Immunoprecipitation Sequencing (ChIPSeq) using the same experimental regimen. Specifically, Smad1 and Smad2 ChIPSeq could be performed to assess the DNA binding of the receptor Smads for the two branches of the TGF-beta signaling pathway by comparing peak enrichment at mesendoderm GRN members for both of these proteins. A logical hypothesis would be that genes transcribed in response to both Activin and Early BMP are enriched for both Smad1 and Smad2 binding sites in their promoter regions, whereas those only activated by Activin are enriched for only Smad2 binding sites. This hypothesis could be assessed using ChIP qPCR to preliminarily determine Smad1 and Smad2 binding at the promoters of Activin/Nodal target genes turned on by BMP signaling versus those that are not. However, since we have identified many putative mesendoderm GRN members through limma analysis and WGCNA it would be valuable to do sequencing experiments in order to probe Smad1 and Smad2 motif enrichment in a more unbiased way in this entire subset of genes, especially if preliminary qPCR experiments do demonstrate a difference in Smad1 binding between these two subpopulations of the mesendoderm GRN.

### **Dand5 Temporally Restrains BMP Signaling**

The markedly different transcriptome response to BMP upon a heterochronic shift implies that temporal regulation of BMP signaling is imperative for normal embryonic patterning, and thus the embryo must have a robust mechanism to temporally restrain activation of signaling. My data shows that BMP ligands and receptors are abundantly transcribed as early as stage 9, and yet BMP signaling is not detectable endogenously until stage 10.5. I showed through morpholino knockdown that *Dand5* is partially responsible for temporally restraining BMP signaling. However, even in the absence of *Dand5*, *Brachyury* is not expressed at levels equivalent to those seen with BMP4/7 mesoderm induction, which suggests that other proteins may be playing a role in regulating BMP timing. The fact that loss of maternally provided genes coincides with the onset of endogenous BMP signaling makes this set of genes obvious candidates. I identified a second known maternally provided BMP inhibitor, *Gtpbp2*, and it would be interesting to also use morpholinos to see if knockdown of *Gtpbp2* alone can induce mesoderm formation and to what extent. Given its low expression levels in comparison with *Dand5*, I would hypothesize that mesoderm induction with *Gtpbp2* would be minimal, but that double morpholino experiments might lead to induction of mesoderm at levels comparable to BMP4/7 induction. If double *Gtpbp2* and *Dand5* morpholino knockdown still leads to markedly lower mesoderm induction than BMP4/7, the list of 119 monotonically decreasing maternally provided genes that I curated could provide other candidate genes whose role in regulating BMP timing could be investigated.

Since we have some insight into how BMP signaling is regulated temporally, an interesting next question is why BMP signaling must be so tightly regulated and why the same cells respond to form an entirely different lineage when BMP signaling is introduced just one developmental stage earlier than normal. While investigation of the transcriptome provides answers to how cells respond to early BMP and how BMP signaling is restricted until stage 10.5, insights into why BMP timing is so tightly regulated are more likely to come from assessing chromatin accessibility than further probing the transcriptome. In order to address this question, Assay for Transposase-Accessible Chromatin Sequencing (ATACSeq) could be layered onto the RNASeq data sets using the same experimental setup. In general, it would be interesting to have lineage specific information about chromatin accessibility during early embryonic development that ATACSeq could provide. More specifically, it would be interesting to investigate the

accessibility of Smad1 binding partners. One possible hypothesis as to why early BMP drives cells towards mesoderm rather than an early adoption of the epidermal fate is that mesoderm genes have higher chromatin accessibility than epidermal genes before stage 10.5. This could be investigated by determining if there is ATAC peak enrichment at stages 9 and 10 in mesoderm genes at these stages and if enrichment shifts to epidermal genes at stage 10.5. Another hypothesis could be that altering BMP signaling alters chromatin accessibility which could be assessed by comparing ATAC peaks across lineages at the same stage and at shifted stages (ie: stage 10 mesoderm to stage 10.5 epidermis). Finally, I have identified a subset of 60 genes whose transcription is downregulated by noggin and upregulated by early BMP that are not already established BMP targets. These genes could be examined to see if there are ATAC peaks in the promoter and enhancer regions and if there are smad1 motifs. ATACSeq and the previously described Smad1 ChIPSeq could be used to potentially identify novel Smad1 targets amongst these genes.

Finally, it would be an exciting advance in the understanding of the timing of signaling regulation to couple a lineage tracing technique similar to GESTALT to signaling in order to precisely track when cells are first exposed to different signals and layer that with lineage tracing in order to see how timing of cell signal exposure affects cell fate at a single cell level in a developing embryo rather than an in vitro population of cells directed towards specific fates. This is likely a technique better suited for zebrafish than *Xenopus* studies but could nevertheless be a potential tool to build off the transcriptome data analyzed in this thesis.

### **Effect of Inhibition of DNA Methylation on Lineage Specification**

DNA Methylation is known to be a repressive epigenetic mark that prevents binding of transcription factors. As the transcription of genes is a dynamic process, methylation of DNA is thus also dynamically shifting so it was not obvious what effect global inhibition of DNA methylation with 5-Aza would have on lineage formation. I demonstrated via *in situ* hybridization that inhibition of DNA methylation has a stochastic effect on expression of transcription factors at the pluripotent stage. However, hierarchical clustering of RNA Sequencing data shows that effects on the transcriptome of 5-

Aza at stage 9 are not greater than sample to sample variability. Via stage 13 RNASeq I also showed that genes downregulated by 5-Aza are enriched for structural genes and genes upregulated by 5-Aza are enriched for cell signaling and that 5-Aza treated caps cluster closest to the neural crest lineage. First and foremost, this dataset would be significantly bolstered by layering the transcriptome data with bisulfite sequencing data in order to have a direct readout of the effect of 5-Aza on DNA methylation in order to rule out the possibility that the effects of 5-Aza on the transcriptome are due predominantly to off target effects. GO enrichment does, however, alleviate the main concern that phenotypes are not due to cell death as there is no enrichment for apoptotic genes in the 5-Aza samples.

Given a heterogeneous effect of 5-Aza on pluripotency factors is seen via *in situ* hybridization at stage nine, one hypothesis would be that inhibition of DNA methylation amplifies inherent stochasticity at pluripotent stages. Since very few genes are significantly differentially expressed in the RNASeq data at this stage, it is possible there is an averaging effect occurring in the population sequencing data in which gene expression is increased in some cells while decreased in others, thus no overall difference in transcript levels is reflected in the data. It would be interesting to perform single cell RNA Sequencing on stage 9 5-Aza treated versus control animal caps in order to gain insight into the inherent stochasticity at this stage and to test the hypothesis that loss of DNA methylation amplifies this stochasticity. It would be particularly interesting to do spatial single cell sequencing in order to see the pattern of heterogeneity.

Because 5-Aza treated animal caps cluster with epidermal caps at stage 9, but neural crest at stage 13 it would also be interesting to perform a time course of epidermal, neural crest and 5-Aza treated caps with the same experimental regimen in order to determine at what stage loss of DNA methylation causes caps to diverge away from their normal epidermal trajectory. A logical hypothesis would be that this divergence takes place around stage 10.5 when caps begin to commit to specific fates, as shown by my lineage specific RNA Sequencing data.

### **Role of DNA Methylation on Regulation of Intron Splicing**

In addition to its role in regulation of transcription, DNA methylation also plays a regulatory role in intron splicing through recruitment of splicing factors onto alternative exons and through control of Pol II

elongation rate (Lev Maor et al. 2015). This is generally considered to be a secondary role of DNA methylation. Strikingly, when PCA is performed on stage 13 RNASeq data from the epidermal, neural, mesoderm, endoderm, and neural crest lineages as well as stage 13 5-Aza treated animal caps, all 5 lineages with their methylation intact cluster together along the first principal component, which accounts for nearly half of the variance in the data set (45.85%), while 5-Aza clusters by itself. GO Term analysis of this principal component reveals that it is dominated by genes playing a role in splicing regulation which suggests that the most significant variation between caps with and without DNA methylation is in alternative splicing, not transcription. Recently, an algorithm called Find Rare Splicing Events in RNA-Seq (FRASER) has been developed to detect alternative splicing and intron retention in RNASeq data independent of genome annotation (Mertes et al. 2021). It would be interesting to apply this algorithm to the 5-Aza RNASeq data compared to the DMSO control to more directly compare splice variants between the two data sets. If elevated levels of aberrant splicing activities are identified, as would be expected based on GO Term analysis, it would be interesting to also apply this method to disease diagnostics in diseases associated with aberrant DNA methylation. DNA methylation inhibitors are effective and widely available due to the association of DNA methylation with tumor cells, thus making inhibition of DNA methylation a promising tool for cancer treatment. Further investigation into the role of DNA methylation on aberrant splicing, and the role of these splicing events on cancer in other diseases could help to further tune cancer therapeutics designed to target DNA methylation.

### **Concluding Remarks**

In my thesis work I have provided a high resolution quantitative analysis of the transit from the pluripotent state to four distinct lineages. I have provided quantitative backing for the long established neural default model and shown that the path to the neural fate is shorter and more one dimensional than to other lineages, thus providing a biological explanation for why neural is indeed the default state of cells. Through multiple types of pattern and network analysis, I have identified several novel putative members of the mesendoderm GRN, as well as provided more precise timing of the onset of previously established mesendoderm GRN members. I have also demonstrated a previously uncharacterized overlap in the

response to two distinct branches of TGF-beta signaling: BMP and Activin/Nodal. My data reveals that the timing of endogenous onset of BMP signaling is precisely controlled in part by Dand5, and that heterochronic shifts in physiological levels of BMP signaling lead to formation of mesoderm rather than epidermal formation. Finally, I have provided further support for the role of DNA methylation in splicing regulation.

The analysis of sequencing datasets presented in this thesis answers many questions regarding how the transcriptome changes over time in different lineages and how these changes are regulated. These data, in turn, lead to many more questions that could be addressed by functional experiments and/or layering this transcriptome data with ChIPSeq and ATACSeq. For example, with the default nature of the neural lineage and its shared characteristics with the pluripotent state, is there thus a prolonged competency to be redirected away from this state? Are these putative mesendoderm GRN candidates expressed specifically in the mesendoderm in whole embryos and are they required for mesendoderm formation? What other proteins play a role in temporally regulating BMP signaling beyond Dand5? Why does early BMP lead to mesoderm formation and not early epidermal formation? Is this a result of changes in chromatin accessibility of mesoderm vs epidermal genes early in development? Do Activin/Nodal targets that are also activated by early BMP signaling have Smad1 in addition to Smad2 binding sites in their promoter regions? These are all exciting directions to pursue in order to follow up on this thesis work and continue to flesh out the picture of lineage restriction in early embryonic development.

### **Significance of this Thesis Work**

Having a detailed understanding of the transcriptome dynamics specific to each cell type is critical to understanding early embryonic patterning and development. My thesis work provides, for the first time, high resolution transcriptome dynamics for the formation of four distinct lineages; epidermis, neural, mesoderm and endoderm. These data provide quantitative context to the neural default model, allows for prediction of novel mesendoderm GRN members and demonstrates a need for precise regulation of BMP timing and evidence for the mechanism by which BMP timing is controlled prior to gastrulation. With these

high resolution datasets I have performed quantitative analyses to provide novel biological insights into the formation of the primary germ layers. Importantly, these datasets are a valuable resource for the *Xenopus* and developmental biology communities that provide not only insights into early embryonic development, but also a jumping off point from which to explore some of the hypotheses proposed in this discussion chapter by performing both broad next-generation sequencing techniques such as ChIPSeq and ATACSeq as well as functional studies on thus far unstudied individual genes with interesting dynamics.

## References

- Afouda BA, Ciau-Uitz A, Patient R. GATA4, 5 and 6 mediate TGFbeta maintenance of endodermal gene expression in *Xenopus* embryos. *Development*. 2005 Feb;132(4):763-74. doi: 10.1242/dev.01647. Epub 2005 Jan 19.
- Agius E, Oelgeschläger M, Wessely O, Kemp C, De Robertis EM. Endodermal Nodal-related signals and mesoderm induction in *Xenopus*. *Development*. 2000 Mar;127(6):1173-83. doi: 10.1242/dev.127.6.1173.
- Akkers RC, van Heeringen SJ, Jacobi UG, Janssen-Megens EM, François KJ, Stunnenberg HG, Veenstra GJ. A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in *Xenopus* embryos. *Dev Cell*. 2009 Sep;17(3):425-34. doi: 10.1016/j.devcel.2009.08.005.
- Amaya E, Musci TJ, Kirschner MW. Expression of a dominant negative mutant of the FGF receptor disrupts mesoderm formation in *Xenopus* embryos. *Cell*. 1991 Jul 26;66(2):257-70. doi: 10.1016/0092-8674(91)90616-7.
- Andreazzoli M, Pannese M, Boncinelli E. Activating and repressing signals in head development: the role of *Xotx1* and *Xotx2*. *Development*. 1997 May;124(9):1733-43. doi: 10.1242/dev.124.9.1733.
- Andrews S. FastQC: a quality control tool for high throughput sequence data.
- Aoki Y, Saint-Germain N, Gyda M, Magner-Fink E, Lee YH, Credidio C, Saint-Jeannet JP. *Sox10* regulates the development of neural crest-derived melanocytes in *Xenopus*. *Dev Biol*. 2003 Jul 1;259(1):19-33. doi: 10.1016/s0012-1606(03)00161-1.
- Aono A, Hazama M, Notoya K, Taketomi S, Yamasaki H, Tsukuda R, Sasaki S, Fujisawa Y. Potent ectopic bone-inducing activity of bone morphogenetic protein-4/7 heterodimer. *Biochem Biophys Res Commun*. 1995 May 25;210(3):670-7. doi: 10.1006/bbrc.1995.1712.
- Araujo FD, Knox JD, Szyf M, Price GB, Zannis-Hadjopoulos M. Concurrent replication and methylation at mammalian origins of replication. *Mol Cell Biol*. 1998 Jun;18(6):3475-82. doi: 10.1128/MCB.18.6.3475. Erratum in: *Mol Cell Biol* 1999 Jun;19(6):4546.
- Ariizumi T, Asashima M. In vitro induction systems for analyses of amphibian organogenesis and body patterning. *Int J Dev Biol*. 2001;45(1):273-9.
- Ariizumi T, Takahashi S, Chan TC, Ito Y, Michiue T, Asashima M. Isolation and differentiation of *Xenopus* animal cap cells. *Curr Protoc Stem Cell Biol*. 2009 Apr;Chapter 1:Unit 1D.5. doi: 10.1002/9780470151808.sc01d05s9.
- Ariizumi T, Michiue T, Asashima M. In Vitro Induction of *Xenopus* Embryonic Organs Using Animal Cap Cells. *Cold Spring Harb Protoc*. 2017 Dec 1;2017(12):pdb.prot097410. doi: 10.1101/pdb.prot097410.
- Asashima M, Nakano H, Shimada K, Kinoshita K, Ishii K, Shibai H, Ueno N. Mesodermal induction in early amphibian embryos by activin A (erythroid differentiation factor). *Roux Arch Dev Biol*. 1990 Mar;198(6):330-335. doi: 10.1007/BF00383771.
- Asashima M, Nakano H, Uchiyama H, Davids M, Plessow S, Loppnow-Blinde B, Hoppe P, Dau H, Tiedemann H. The vegetalizing factor belongs to a family of mesoderm-inducing proteins related to erythroid differentiation factor. *Naturwissenschaften*. 1990 Aug;77(8):389-91. doi: 10.1007/BF01135742.

Attisano L, Cárcamo J, Ventura F, Weis FM, Massagué J, Wrana JL. Identification of human activin and TGF beta type I receptors that form heteromeric kinase complexes with type II receptors. *Cell*. 1993 Nov 19;75(4):671-80. doi: 10.1016/0092-8674(93)90488-c.

Ault KT, Xu RH, Kung HF, Jamrich M. The homeobox gene PV.1 mediates specification of the prospective neural ectoderm in *Xenopus* embryos. *Dev Biol*. 1997 Dec 1;192(1):162-71. doi: 10.1006/dbio.1997.8737.

Azuara V, Perry P, Sauer S, Spivakov M, Jørgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merckenschlager M, Fisher AG. Chromatin signatures of pluripotent cell lines. *Nat Cell Biol*. 2006 May;8(5):532-8. doi: 10.1038/ncb1403. Epub 2006 Mar 29.

Baedke J. The epigenetic landscape in the course of time: Conrad Hal Waddington's methodological impact on the life sciences. *Stud Hist Philos Biol Biomed Sci*. 2013 Dec;44(4 Pt B):756-73. doi: 10.1016/j.shpsc.2013.06.001.

Baggiolini A, Varum S, Mateos JM, Bettosini D, John N, Bonalli M, Ziegler U, Dimou L, Clevers H, Furrer R, Sommer L. Premigratory and migratory neural crest cells are multipotent in vivo. *Cell Stem Cell*. 2015 Mar 5;16(3):314-22. doi: 10.1016/j.stem.2015.02.017.

Baker JC, Harland RM. A novel mesoderm inducer, *Madr2*, functions in the activin signal transduction pathway. *Genes Dev*. 1996 Aug 1;10(15):1880-9. doi: 10.1101/gad.10.15.1880.

Baker JC, Beddington RS, Harland RM. Wnt signaling in *Xenopus* embryos inhibits *bmp4* expression and activates neural development. *Genes Dev*. 1999 Dec 1;13(23):3149-59. doi: 10.1101/gad.13.23.3149.

Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*. 2001 Mar 1;410(6824):120-4. doi: 10.1038/35065138.

Baroffio A, Dupin E, Le Douarin NM. Clone-forming ability and differentiation potential of migratory neural crest cells. *Proc Natl Acad Sci U S A*. 1988 Jul;85(14):5325-9. doi: 10.1073/pnas.85.14.5325.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007 May 18;129(4):823-37. doi: 10.1016/j.cell.2007.05.009.

Basch ML, Selleck MA, Bronner-Fraser M. Timing and competence of neural crest formation. *Dev Neurosci*. 2000;22(3):217-27. doi: 10.1159/000017444.

Bates TJ, Vonica A, Heasman J, Brivanlou AH, Bell E. *Coco* regulates dorsoventral specification of germ layers via inhibition of TGF $\beta$  signalling. *Development*. 2013 Oct;140(20):4177-81. doi: 10.1242/dev.095521.

Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, Akalin A, Schübeler D. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature*. 2015 Apr 9;520(7546):243-7. doi: 10.1038/nature14176. Epub 2015 Jan 21.

Bell E, Muñoz-Sanjuán I, Altmann CR, Vonica A, Brivanlou AH. Cell fate specification and competence by *Coco*, a maternal BMP, TGFbeta and Wnt inhibitor. *Development*. 2003 Apr;130(7):1381-9. doi: 10.1242/dev.00344.

Bellmeyer A, Krase J, Lindgren J, LaBonne C. The protooncogene *c-myc* is an essential regulator of neural crest formation in *xenopus*. *Dev Cell*. 2003 Jun;4(6):827-39. doi: 10.1016/s1534-5807(03)00160-6.

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ 3rd, Gingeras TR, Schreiber SL, Lander ES. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*. 2005 Jan 28;120(2):169-81. doi: 10.1016/j.cell.2005.01.001.

Bestor TH, Ingram VM. Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc Natl Acad Sci U S A*. 1983 Sep;80(18):5559-63. doi: 10.1073/pnas.80.18.5559.

Betancur P, Bronner-Fraser M, Sauka-Spengler T. Assembling neural crest regulatory circuits into a gene regulatory network. *Annu Rev Cell Dev Biol*. 2010;26:581-603. doi: 10.1146/annurev.cellbio.042308.113245.

Blitz IL, Cho KW. Finding partners: how BMPs select their targets. *Dev Dyn*. 2009 Jun;238(6):1321-31. doi: 10.1002/dvdy.21984.

Blythe SA, Cha SW, Tadjuidje E, Heasman J, Klein PS. beta-Catenin primes organizer gene expression by recruiting a histone H3 arginine 8 methyltransferase, Prmt2. *Dev Cell*. 2010 Aug 17;19(2):220-31. doi: 10.1016/j.devcel.2010.07.007.

Bogdanovic O, Fernandez-Miñán A, Tena JJ, de la Calle-Mustienes E, Hidalgo C, van Kruysbergen I, van Heeringen SJ, Veenstra GJ, Gómez-Skarmeta JL. Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis. *Genome Res*. 2012 Oct;22(10):2043-53. doi: 10.1101/gr.134833.111. Epub 2012 May 16.

Bouwmeester T, Kim S, Sasai Y, Lu B, De Robertis EM. Cerberus is a head-inducing secreted factor expressed in the anterior endoderm of Spemann's organizer. *Nature*. 1996 Aug 15;382(6592):595-601. doi: 10.1038/382595a0.

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*. 2005 Sep 23;122(6):947-56. doi: 10.1016/j.cell.2005.08.020.

Boyes J, Bird A. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell*. 1991 Mar 22;64(6):1123-34. doi: 10.1016/0092-8674(91)90267-3.

Brannon, M, and D Kimelman. "Activation of Siamese by the Wnt pathway." *Developmental biology* vol. 180,1 (1996): 344-7. doi:10.1006/dbio.1996.0306

Briggs JA, Weinreb C, Wagner DE, Megason S, Peshkin L, Kirschner MW, Klein AM. The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science*. 2018 Jun 1;360(6392):eaar5780. doi: 10.1126/science.aar5780.

Bröhm A, Schoch T, Dukatz M, Graf N, Dorscht F, Mantai E, Adam S, Bashtrykov P, Jeltsch A. Methylation of recombinant mononucleosomes by DNMT3A demonstrates efficient linker DNA methylation and a role of H3K36me3. *Commun Biol*. 2022 Mar 2;5(1):192. doi: 10.1038/s42003-022-03119-z.

Bronner-Fraser M, Fraser SE. Cell lineage analysis reveals multipotency of some avian neural crest cells. *Nature*. 1988 Sep 8;335(6186):161-4. doi: 10.1038/335161a0.

Bronner-Fraser M, Fraser S. Developmental potential of avian trunk neural crest cells in situ. *Neuron*. 1989 Dec;3(6):755-66. doi: 10.1016/0896-6273(89)90244-4.

- Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*. 1996 Mar 22;84(6):843-51. doi: 10.1016/s0092-8674(00)81063-6.
- Buitrago-Delgado E, Nordin K, Rao A, Geary L, LaBonne C. NEURODEVELOPMENT. Shared regulatory programs suggest retention of blastula-stage potential in neural crest cells. *Science*. 2015 Jun 19;348(6241):1332-5. doi: 10.1126/science.aaa3655.
- Byrnes WM, Newman SA. Ernest Everett Just: Egg and embryo as excitable systems. *J Exp Zool B Mol Dev Evol*. 2014 Jun;322(4):191-201. doi: 10.1002/jez.b.22567.
- Byrnes WM. The forgotten father of epigenetics. *Am. Sci*. 2015 Mar 1;103(2):106-9.
- Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science*. 2002 Nov 1;298(5595):1039-43. doi: 10.1126/science.1076997.
- Cao Y, Knöchel S, Donow C, Miethe J, Kaufmann E, Knöchel W. The POU factor Oct-25 regulates the Xvent-2B gene and counteracts terminal differentiation in *Xenopus* embryos. *J Biol Chem*. 2004 Oct 15;279(42):43735-43. doi: 10.1074/jbc.M407544200.
- Carnac G, Kodjabachian L, Gurdon JB, Lemaire P. The homeobox gene *Siamois* is a target of the Wnt dorsalisation pathway and triggers organiser activity in the absence of mesoderm. *Development*. 1996 Oct;122(10):3055-65. doi: 10.1242/dev.122.10.3055.
- Cha SW, Tadjuidje E, Tao Q, Wylie C, Heasman J. Wnt5a and Wnt11 interact in a maternal Dkk1-regulated fashion to activate both canonical and non-canonical signaling in *Xenopus* axis formation. *Development*. 2008 Nov;135(22):3719-29. doi: 10.1242/dev.029025.
- Cha SW, Tadjuidje E, White J, Wells J, Mayhew C, Wylie C, Heasman J. Wnt11/5a complex formation caused by tyrosine sulfation increases canonical signaling activity. *Curr Biol*. 2009 Sep 29;19(18):1573-80. doi: 10.1016/j.cub.2009.07.062.
- Cha SW, McAdams M, Kormish J, Wylie C, Kofron M. Foxi2 is an animally localized maternal mRNA in *Xenopus*, and an activator of the zygotic ectoderm activator Foxi1e. *PLoS One*. 2012;7(7):e41782. doi: 10.1371/journal.pone.0041782. Epub 2012 Jul 27.
- Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell*. 2003 May 30;113(5):643-55. doi: 10.1016/s0092-8674(03)00392-1.
- Chang C, Hemmati-Brivanlou A. Cell fate determination in embryonic ectoderm. *J Neurobiol*. 1998 Aug;36(2):128-51.
- Charney RM, Paraiso KD, Blitz IL, Cho KWY. A gene regulatory program controlling early *Xenopus* mesendoderm formation: Network conservation and motifs. *Semin Cell Dev Biol*. 2017 Jun;66:12-24. doi: 10.1016/j.semcdb.2017.03.003.
- Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics*. 2011 Jan 26;12:35. doi: 10.1186/1471-2105-12-35.

Chen X, Rubock MJ, Whitman M. A transcriptional partner for MAD proteins in TGF-beta signalling. *Nature*. 1996 Oct 24;383(6602):691-6. doi: 10.1038/383691a0. Erratum in: *Nature* 1996 Dec 19-26;384(6610):648.

Chen X, Weisberg E, Fridmacher V, Watanabe M, Naco G, Whitman M. Smad4 and FAST-1 in the assembly of activin-responsive factor. *Nature*. 1997 Sep 4;389(6646):85-9. doi: 10.1038/38008.

Chen Y, Bhushan A, Vale W. Smad8 mediates the signaling of the ALK-2 [corrected] receptor serine kinase. *Proc Natl Acad Sci U S A*. 1997 Nov 25;94(24):12938-43. doi: 10.1073/pnas.94.24.12938.

Chiu WT, Charney Le R, Blitz IL, Fish MB, Li Y, Biesinger J, Xie X, Cho KW. Genome-wide view of TGF $\beta$ /Foxh1 regulation of the early mesendoderm program. *Development*. 2014 Dec;141(23):4537-47. doi: 10.1242/dev.107227. Epub 2014 Oct 30.

Chodavarapu RK, Feng S, Bernatavichute YV, Chen PY, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ, Casero D, Bernal M, Huijser P, Clark AT, Krämer U, Merchant SS, Zhang X, Jacobsen SE, Pellegrini M. Relationship between nucleosome positioning and DNA methylation. *Nature*. 2010 Jul 15;466(7304):388-92. doi: 10.1038/nature09147. Epub 2010 May 30.

Christian JL, McMahon JA, McMahon AP, Moon RT. Xwnt-8, a *Xenopus* Wnt-1/int-1-related gene responsive to mesoderm-inducing growth factors, may play a role in ventral mesodermal patterning during embryogenesis. *Development*. 1991 Apr;111(4):1045-55. doi: 10.1242/dev.111.4.1045.

Christen B, Slack JM. Spatial response to fibroblast growth factor signalling in *Xenopus* embryos. *Development*. 1999 Jan;126(1):119-25. doi: 10.1242/dev.126.1.119.

Clements D, Friday RV, Woodland HR. Mode of action of VegT in mesoderm and endoderm formation. *Development*. 1999 Nov;126(21):4903-11. doi: 10.1242/dev.126.21.4903.

Collart C, Owens ND, Bhaw-Rosun L, Cooper B, De Domenico E, Patrushev I, Sesay AK, Smith JN, Smith JC, Gilchrist MJ. High-resolution analysis of gene activity during the *Xenopus* mid-blastula transition. *Development*. 2014 May;141(9):1927-39. doi: 10.1242/dev.102012.

Collazo A, Bronner-Fraser M, Fraser SE. Vital dye labelling of *Xenopus laevis* trunk neural crest reveals multipotency and novel pathways of migration. *Development*. 1993 Jun;118(2):363-76. doi: 10.1242/dev.118.2.363.

Collignon J, Sockanathan S, Hacker A, Cohen-Tannoudji M, Norris D, Rastan S, Stevanovic M, Goodfellow PN, Lovell-Badge R. A comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2. *Development*. 1996 Feb;122(2):509-20. doi: 10.1242/dev.122.2.509.

Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*. 2017 Sep 15;33(18):2938-2940. doi: 10.1093/bioinformatics/btx364.

Cordenonsi M, Dupont S, Maretto S, Insinga A, Imbriano C, Piccolo S. Links between tumor suppressors: p53 is required for TGF-beta gene responses by cooperating with Smads. *Cell*. 2003 May 2;113(3):301-14. doi: 10.1016/s0092-8674(03)00308-8.

Creusot F, Acs G, Christman JK. Inhibition of DNA methyltransferase and induction of Friend erythroleukemia cell differentiation by 5-azacytidine and 5-aza-2'-deoxycytidine. *J Biol Chem*. 1982 Feb 25;257(4):2041-8.

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. Histone H3K27ac separates active from

poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A*. 2010 Dec 14;107(50):21931-6. doi: 10.1073/pnas.1016071107.

Crow JF. Just and unjust: E. E. Just (1883-1941). *Genetics*. 2008 Aug;179(4):1735-40. doi: 10.1534/genetics.104.94094.

Dale L, Matthews G, Colman A. Secretion and mesoderm-inducing activity of the TGF-beta-related domain of *Xenopus* Vg1. *EMBO J*. 1993 Dec;12(12):4471-80. doi: 10.1002/j.1460-2075.1993.tb06136.x.

Danilov V, Blum M, Schweickert A, Campione M, Steinbeisser H. Negative autoregulation of the organizer-specific homeobox gene gooseoid. *J Biol Chem*. 1998 Jan 2;273(1):627-35. doi: 10.1074/jbc.273.1.627.

Delaune E, Lemaire P, Kodjabachian L. Neural induction in *Xenopus* requires early FGF signalling in addition to BMP inhibition. *Development*. 2005 Jan;132(2):299-310. doi: 10.1242/dev.01582. Epub 2004 Dec 8.

De Robertis EM, Kuroda H. Dorsal-ventral patterning and neural induction in *Xenopus* embryos. *Annu Rev Cell Dev Biol*. 2004;20:285-308. doi: 10.1146/annurev.cellbio.20.011403.154124.

Ding Y, Colozza G, Zhang K, Moriyama Y, Ploper D, Sosa EA, Benitez MDJ, De Robertis EM. Genome-wide analysis of dorsal and ventral transcriptomes of the *Xenopus laevis* gastrula. *Dev Biol*. 2017 Jun 15;426(2):176-187. doi: 10.1016/j.ydbio.2016.02.032.

Ding Y, Colozza G, Sosa EA, Moriyama Y, Rundle S, Salwinski L, De Robertis EM. Bighead is a Wnt antagonist secreted by the *Xenopus* Spemann organizer that promotes Lrp6 endocytosis. *Proc Natl Acad Sci U S A*. 2018 Sep 25;115(39):E9135-E9144. doi: 10.1073/pnas.1812117115.

Dirksen ML, Morasso MI, Sargent TD, Jamrich M. Differential expression of a Distal-less homeobox gene *XdII-2* in ectodermal cell lineages. *Mech Dev*. 1994 Apr;46(1):63-70. doi: 10.1016/0925-4773(94)90038-8.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013 Jan 1;29(1):15-21. doi: 10.1093/bioinformatics/bts635.

Dovey OM, Foster CT, Cowley SM. Histone deacetylase 1 (HDAC1), but not HDAC2, controls embryonic stem cell differentiation. *Proc Natl Acad Sci U S A*. 2010 May 4;107(18):8242-7. doi: 10.1073/pnas.1000478107. Epub 2010 Apr 19.

Drissi MH, Li X, Sheu TJ, Zuscik MJ, Schwarz EM, Puzas JE, Rosier RN, O'Keefe RJ. Runx2/Cbfa1 stimulation by retinoic acid is potentiated by BMP2 signaling through interaction with Smad1 on the collagen X promoter in chondrocytes. *J Cell Biochem*. 2003 Dec 15;90(6):1287-98. doi: 10.1002/jcb.10677.

Du J, Johnson LM, Jacobsen SE, Patel DJ. DNA methylation pathways and their crosstalk with histone methylation. *Nat Rev Mol Cell Biol*. 2015 Sep;16(9):519-32. doi: 10.1038/nrm4043.

Dupin E, Sommer L. Neural crest progenitors and stem cells: from early development to adulthood. *Dev Biol*. 2012 Jun 1;366(1):83-95. doi: 10.1016/j.ydbio.2012.02.035. Epub 2012 Mar 8.

Ebner R, Chen RH, Lawler S, Zioncheck T, Derynck R. Determination of type I receptor specificity by the type II receptors for TGF-beta or activin. *Science*. 1993 Nov 5;262(5135):900-2. doi: 10.1126/science.8235612.

Efroni S, Duttagupta R, Cheng J, Dehghani H, Hoepfner DJ, Dash C, Bazett-Jones DP, Le Grice S, McKay RD, Buetow KH, Gingeras TR, Misteli T, Meshorer E. Global transcription in pluripotent embryonic stem cells. *Cell Stem Cell*. 2008 May 8;2(5):437-47. doi: 10.1016/j.stem.2008.03.021.

Eimon PM, Harland RM. In *Xenopus* embryos, BMP heterodimers are not required for mesoderm induction, but BMP activity is necessary for dorsal/ventral patterning. *Dev Biol*. 1999 Dec 1;216(1):29-40. doi: 10.1006/dbio.1999.9496.

Eimon PM, Harland RM. Effects of heterodimerization and proteolytic processing on *Derrière* and *Nodal* activity: implications for mesoderm induction in *Xenopus*. *Development*. 2002 Jul;129(13):3089-103. doi: 10.1242/dev.129.13.3089.

Elinson, R P, and B Rowning. "A transient array of parallel microtubules in frog eggs: potential tracks for a cytoplasmic rotation that specifies the dorso-ventral axis." *Developmental biology* vol. 128,1 (1988): 185-97. doi:10.1016/0012-1606(88)90281-3

Falkenberg KJ, Johnstone RW. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat Rev Drug Discov*. 2014 Sep;13(9):673-91. doi: 10.1038/nrd4360. Epub 2014 Aug 18.

Fainsod A, Deissler K, Yelin R, Marom K, Epstein M, Pillemer G, Steinbeisser H, Blum M. The dorsalizing and neural inducing gene *folistatin* is an antagonist of BMP-4. *Mech Dev*. 1997 Apr;63(1):39-50. doi: 10.1016/s0925-4773(97)00673-4.

Fang F, Xu Y, Chew KK, Chen X, Ng HH, Matsudaira P. Coactivators p300 and CBP maintain the identity of mouse embryonic stem cells by mediating long-range chromatin structure. *Stem Cells*. 2014 Jul;32(7):1805-16. doi: 10.1002/stem.1705.

Faure S, Lee MA, Keller T, ten Dijke P, Whitman M. Endogenous patterns of TGFbeta superfamily signaling during early *Xenopus* development. *Development*. 2000 Jul;127(13):2917-31. doi: 10.1242/dev.127.13.2917.

Feledy JA, Beanan MJ, Sandoval JJ, Goodrich JS, Lim JH, Matsuo-Takasaki M, Sato SM, Sargent TD. Inhibitory patterning of the anterior neural plate in *Xenopus* by homeodomain factors *Dlx3* and *Msx1*. *Dev Biol*. 1999 Aug 15;212(2):455-64. doi: 10.1006/dbio.1999.9374.

Feng J, Chang H, Li E, Fan G. Dynamic expression of de novo DNA methyltransferases *Dnmt3a* and *Dnmt3b* in the central nervous system. *J Neurosci Res*. 2005 Mar 15;79(6):734-46. doi: 10.1002/jnr.20404.

Fuks F, Hurd PJ, Deplus R, Kouzarides T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res*. 2003 May 1;31(9):2305-12. doi: 10.1093/nar/gkg332.

Geary L, LaBonne C. FGF mediated MAPK and PI3K/Akt Signals make distinct contributions to pluripotency and the establishment of Neural Crest. *Elife*. 2018 Jan 19;7:e33845. doi: 10.7554/eLife.33845.

Gentsch GE, Owens ND, Martin SR, Piccinelli P, Faial T, Trotter MW, Gilchrist MJ, Smith JC. In vivo T-box transcription factor profiling reveals joint regulation of embryonic neuromesodermal bipotency. *Cell Rep*. 2013 Sep 26;4(6):1185-96. doi: 10.1016/j.celrep.2013.08.012.

- Gentsch GE, Spruce T, Owens ND, Smith JC. Maternal pluripotency factors initiate extensive chromatin remodelling to predefine first response to inductive signals. *Nat Commun.* 2019 Sep 19;10(1):4269. doi: 10.1038/s41467-019-12263-w.
- Glavic A, Silva F, Aybar MJ, Bastidas F, Mayor R. Interplay between Notch signaling and the homeoprotein Xiro1 is required for neural crest induction in *Xenopus* embryos. *Development.* 2004 Jan;131(2):347-59. doi: 10.1242/dev.00945.
- Gore SD, Baylin S, Sugar E, Carraway H, Miller CB, Carducci M, Grever M, Galm O, Dausers T, Karp JE, Rudek MA, Zhao M, Smith BD, Manning J, Jiemjit A, Dover G, Mays A, Zwiebel J, Murgo A, Weng LJ, Herman JG. Combined DNA methyltransferase and histone deacetylase inhibition in the treatment of myeloid neoplasms. *Cancer Res.* 2006 Jun 15;66(12):6361-9. doi: 10.1158/0008-5472.CAN-06-0080.
- Gruenbaum Y, Cedar H, Razin A. Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature.* 1982 Feb 18;295(5850):620-2. doi: 10.1038/295620a0.
- Guenther MG, Jenner RG, Chevalier B, Nakamura T, Croce CM, Canaani E, Young RA. Global and Hox-specific roles for the MLL1 methyltransferase. *Proc Natl Acad Sci U S A.* 2005 Jun 14;102(24):8603-8. doi: 10.1073/pnas.0503072102. Epub 2005 Jun 7.
- Guo X, Wang L, Li J, Ding Z, Xiao J, Yin X, He S, Shi P, Dong L, Li G, Tian C, Wang J, Cong Y, Xu Y. Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature.* 2015 Jan 29;517(7536):640-4. doi: 10.1038/nature13899. Epub 2014 Nov 10.
- Gupta R, Wills A, Ucar D, Baker J. Developmental enhancers are marked independently of zygotic Nodal signals in *Xenopus*. *Dev Biol.* 2014 Nov 1;395(1):38-49. doi: 10.1016/j.ydbio.2014.08.034. Epub 2014 Sep 6.
- Godsave SF, Slack JM. Clonal analysis of mesoderm induction in *Xenopus laevis*. *Dev Biol.* 1989 Aug;134(2):486-90. doi: 10.1016/0012-1606(89)90122-x.
- Goldberg AD, Allis CD, Bernstein E. Epigenetics: a landscape takes shape. *Cell.* 2007 Feb 23;128(4):635-8. doi: 10.1016/j.cell.2007.02.006.
- Graff JM, Thies RS, Song JJ, Celeste AJ, Melton DA. Studies with a *Xenopus* BMP receptor suggest that ventral mesoderm-inducing signals override dorsal signals in vivo. *Cell.* 1994 Oct 7;79(1):169-79. doi: 10.1016/0092-8674(94)90409-x.
- Graff JM, Bansal A, Melton DA. *Xenopus* Mad proteins transduce distinct subsets of signals for the TGF beta superfamily. *Cell.* 1996 May 17;85(4):479-87. doi: 10.1016/s0092-8674(00)81249-0.
- Green JB, New HV, Smith JC. Responses of embryonic *Xenopus* cells to activin and FGF are separated by multiple dose thresholds and correspond to distinct axes of the mesoderm. *Cell.* 1992 Nov 27;71(5):731-9. doi: 10.1016/0092-8674(92)90550-v.
- Groves AK, LaBonne C. Setting appropriate boundaries: fate, patterning and competence at the neural plate border. *Dev Biol.* 2014 May 1;389(1):2-12. doi: 10.1016/j.ydbio.2013.11.027. Epub 2013 Dec 7.
- Grunz H, Tacke L. Neural differentiation of *Xenopus laevis* ectoderm takes place after disaggregation and delayed reaggregation without inducer. *Cell Differ Dev.* 1989 Dec;28(3):211-7. doi: 10.1016/0922-3371(89)90006-3.
- Guo X, Wang XF. Signaling cross-talk between TGF-beta/BMP and other pathways. *Cell Res.* 2009 Jan;19(1):71-88. doi: 10.1038/cr.2008.302.

- Gurdon JB, Harger P, Mitchell A, Lemaire P. Activin signalling and response to a morphogen gradient. *Nature*. 1994 Oct 6;371(6497):487-92. doi: 10.1038/371487a0.
- Hall BK. Germ layers, the neural crest and emergent organization in development and evolution. *Genesis*. 2018 Jun;56(6-7):e23103. doi: 10.1002/dvg.23103.
- Hall BK. The neural crest as a fourth germ layer and vertebrates as quadroblastic not triploblastic. *Evol Dev*. 2000 Jan-Feb;2(1):3-5. doi: 10.1046/j.1525-142x.2000.00032.x.
- Hamazaki T, Oka M, Yamanaka S, Terada N. Aggregation of embryonic stem cells induces Nanog repression and primitive endoderm differentiation. *J Cell Sci*. 2004 Nov 1;117(Pt 23):5681-6. doi: 10.1242/jcs.01489.
- Hansen CS, Marion CD, Steele K, George S, Smith WC. Direct neural induction and selective inhibition of mesoderm and epidermis inducers by Xnr3. *Development*. 1997 Jan;124(2):483-92. doi: 10.1242/dev.124.2.483.
- Hardcastle Z, Chalmers AD, Papalopulu N. FGF-8 stimulates neuronal differentiation through FGFR-4a and interferes with mesoderm induction in *Xenopus* embryos. *Curr Biol*. 2000 Nov 30;10(23):1511-4. doi: 10.1016/s0960-9822(00)00825-3.
- Hardin J, Keller R. The behaviour and function of bottle cells during gastrulation of *Xenopus laevis*. *Development*. 1988 May;103(1):211-30. doi: 10.1242/dev.103.1.211.
- Harland R, Gerhart J. Formation and function of Spemann's organizer. *Annu Rev Cell Dev Biol*. 1997;13:611-67. doi: 10.1146/annurev.cellbio.13.1.611.
- Harvey SA, Tümpel S, Dubrulle J, Schier AF, Smith JC. no tail integrates two modes of mesoderm induction. *Development*. 2010 Apr;137(7):1127-35. doi: 10.1242/dev.046318.
- Hata A, Lagna G, Massagué J, Hemmati-Brivanlou A. Smad6 inhibits BMP/Smad1 signaling by specifically competing with the Smad4 tumor suppressor. *Genes Dev*. 1998 Jan 15;12(2):186-97. doi: 10.1101/gad.12.2.186
- Hata A, Seoane J, Lagna G, Montalvo E, Hemmati-Brivanlou A, Massagué J. OAZ uses distinct DNA- and protein-binding zinc fingers in separate BMP-Smad and Olf signaling pathways. *Cell*. 2000 Jan 21;100(2):229-40. doi: 10.1016/s0092-8674(00)81561-5.
- Hawley SH, Wünnenberg-Stapleton K, Hashimoto C, Laurent MN, Watabe T, Blumberg BW, Cho KW. Disruption of BMP signals in embryonic *Xenopus* ectoderm leads to direct neural induction. *Genes Dev*. 1995 Dec 1;9(23):2923-35.
- Hayashi H, Abdollah S, Qiu Y, Cai J, Xu YY, Grinnell BW, Richardson MA, Topper JN, Gimbrone MA Jr, Wrana JL, Falb D. The MAD-related protein Smad7 associates with the TGFbeta receptor and functions as an antagonist of TGFbeta signaling. *Cell*. 1997 Jun 27;89(7):1165-73. doi: 10.1016/s0092-8674(00)80303-7.
- He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C, Xu GL. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*. 2011 Sep 2;333(6047):1303-7. doi: 10.1126/science.1210944. Epub 2011 Aug 4.

Hemmati-Brivanlou A, Melton DA. A truncated activin receptor inhibits mesoderm induction and formation of axial structures in *Xenopus* embryos. *Nature*. 1992 Oct 15;359(6396):609-14. doi: 10.1038/359609a0.

Hemmati-Brivanlou A, Melton DA. Inhibition of activin receptor signaling promotes neuralization in *Xenopus*. *Cell*. 1994 Apr 22;77(2):273-81. doi: 10.1016/0092-8674(94)90319-0.

Hemmati-Brivanlou A, Kelly OG, Melton DA. Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. *Cell*. 1994 Apr 22;77(2):283-95. doi: 10.1016/0092-8674(94)90320-4.

Hemmati-Brivanlou A, Melton D. Vertebrate embryonic cells will become nerve cells unless told otherwise. *Cell*. 1997 Jan 10;88(1):13-7. doi: 10.1016/s0092-8674(00)81853-x.

Hemmati-Brivanlou A, Thomsen GH. Ventral mesodermal patterning in *Xenopus* embryos: expression patterns and activities of BMP-2 and BMP-4. *Dev Genet*. 1995;17(1):78-89. doi: 10.1002/dvg.1020170109.

Henig C, Elias S, Frank D. A POU protein regulates mesodermal competence to FGF in *Xenopus*. *Mech Dev*. 1998 Feb;71(1-2):131-42. doi: 10.1016/s0925-4773(98)00006-9.

Henningfeld KA, Friedle H, Rastegar S, Knöchel W. Autoregulation of Xvent-2B; direct interaction and functional cooperation of Xvent-2 and Smad1. *J Biol Chem*. 2002 Jan 18;277(3):2097-103. doi: 10.1074/jbc.M108524200. Epub 2001 Nov 9.

Henry GL, Brivanlou IH, Kessler DS, Hemmati-Brivanlou A, Melton DA. TGF-beta signals and a pattern in *Xenopus laevis* endodermal development. *Development*. 1996 Mar;122(3):1007-15. doi: 10.1242/dev.122.3.1007.

Henry GL, Melton DA. Mixer, a homeobox gene required for endoderm development. *Science*. 1998 Jul 3;281(5373):91-6. doi: 10.1126/science.281.5373.91.

Hezroni H, Sailaja BS, Meshorer E. Pluripotency-related, valproic acid (VPA)-induced genome-wide histone H3 lysine 9 (H3K9) acetylation patterns in embryonic stem cells. *J Biol Chem*. 2011 Oct 14;286(41):35977-35988. doi: 10.1074/jbc.M111.266254. Epub 2011 Aug 17.

His W. Untersuchungen über die erste Anlage des Wirbelthierleibes: die erste Entwicklung des Hühnchens im Ei. FCW Vogel; 1868.

Hodges E, Smith AD, Kendall J, Xuan Z, Ravi K, Rooks M, Zhang MQ, Ye K, Bhattacharjee A, Brizuela L, McCombie WR, Wigler M, Hannon GJ, Hicks JB. High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. *Genome Res*. 2009 Sep;19(9):1593-605. doi: 10.1101/gr.095190.109. Epub 2009 Jul 6.

Hollnagel A, Oehlmann V, Heymer J, Rütter U, Nordheim A. Id genes are direct targets of bone morphogenetic protein induction in embryonic stem cells. *J Biol Chem*. 1999 Jul 9;274(28):19838-45. doi: 10.1074/jbc.274.28.19838.

Hong CS, Saint-Jeannet JP. The activity of Pax3 and Zic1 regulates three distinct cell fates at the neural plate border. *Mol Biol Cell*. 2007 Jun;18(6):2192-202. doi: 10.1091/mbc.e06-11-1047.

Hong L, Schroth GP, Matthews HR, Yau P, Bradbury EM. Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *J Biol Chem*. 1993 Jan 5;268(1):305-14.

Honoré SM, Aybar MJ, Mayor R. Sox10 is required for the early development of the prospective neural crest in *Xenopus* embryos. *Dev Biol*. 2003 Aug 1;260(1):79-96. doi: 10.1016/s0012-1606(03)00247-1.

Hoppler S, Moon RT. BMP-2/-4 and Wnt-8 cooperatively pattern the *Xenopus* mesoderm. *Mech Dev*. 1998 Feb;71(1-2):119-29. doi: 10.1016/s0925-4773(98)00004-5.

Hopwood ND, Pluck A, Gurdon JB. A *Xenopus* mRNA related to *Drosophila* twist is expressed in response to induction in the mesoderm and the neural crest. *Cell*. 1989 Dec 1;59(5):893-903. doi: 10.1016/0092-8674(89)90612-0.

Hopwood N. 'Not birth, marriage or death, but gastrulation': the life of a quotation in biology. *Br J Hist Sci*. 2022 Mar;55(1):1-26. doi: 10.1017/S0007087421000790.

Hörmanseder E, Simeone A, Allen GE, Bradshaw CR, Figlmüller M, Gurdon J, Jullien J. H3K4 Methylation-Dependent Memory of Somatic Cell Identity Inhibits Reprogramming and Development of Nuclear Transfer Embryos. *Cell Stem Cell*. 2017 Jul 6;21(1):135-143.e6. doi: 10.1016/j.stem.2017.03.003.

Hsieh J, Nakashima K, Kuwabara T, Mejia E, Gage FH. Histone deacetylase inhibition-mediated neuronal differentiation of multipotent adult neural progenitor cells. *Proc Natl Acad Sci U S A*. 2004 Nov 23;101(47):16659-64. doi: 10.1073/pnas.0407643101. Epub 2004 Nov 10.

Hu N, Strobl-Mazzulla P, Sauka-Spengler T, Bronner ME. DNA methyltransferase3A as a molecular switch mediating the neural tube-to-neural crest fate transition. *Genes Dev*. 2012 Nov 1;26(21):2380-5. doi: 10.1101/gad.198747.112.

Hu N, Strobl-Mazzulla PH, Bronner ME. Epigenetic regulation in neural crest development. *Dev Biol*. 2014a Dec 15;396(2):159-68. doi: 10.1016/j.ydbio.2014.09.034.

Hu N, Strobl-Mazzulla PH, Simoes-Costa M, Sánchez-Vásquez E, Bronner ME. DNA methyltransferase 3B regulates duration of neural crest production via repression of Sox10. *Proc Natl Acad Sci U S A*. 2014b Dec 16;111(50):17911-6. doi: 10.1073/pnas.1318408111.

Hudson C, Clements D, Friday RV, Stott D, Woodland HR. Xsox17alpha and -beta mediate endoderm formation in *Xenopus*. *Cell*. 1997 Oct 31;91(3):397-405. doi: 10.1016/s0092-8674(00)80423-7.

Hyodo-Miura J, Urushiyama S, Nagai S, Nishita M, Ueno N, Shibuya H. Involvement of NLK and Sox11 in neural induction in *Xenopus* development. *Genes Cells*. 2002 May;7(5):487-96. doi: 10.1046/j.1365-2443.2002.00536.x.

Iemura S, Yamamoto TS, Takagi C, Uchiyama H, Natsume T, Shimasaki S, Sugino H, Ueno N. Direct binding of follistatin to a complex of bone-morphogenetic protein and its receptor inhibits ventral and epidermal cell fates in early *Xenopus* embryo. *Proc Natl Acad Sci U S A*. 1998 Aug 4;95(16):9337-42. doi: 10.1073/pnas.95.16.9337.

Imamura T, Takase M, Nishihara A, Oeda E, Hanai J, Kawabata M, Miyazono K. Smad6 inhibits signalling by the TGF-beta superfamily. *Nature*. 1997 Oct 9;389(6651):622-6. doi: 10.1038/39355.  
Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*. 2011 Sep 2;333(6047):1300-3. doi: 10.1126/science.1210597. Epub 2011 Jul 21.

Isaacs HV, Pownall ME, Slack JM. eFGF regulates Xbra expression during *Xenopus* gastrulation. *EMBO J*. 1994 Oct 3;13(19):4469-81. doi: 10.1002/j.1460-2075.1994.tb06769.x.

Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*. 2011 Sep 2;333(6047):1300-3. doi: 10.1126/science.1210597. Epub 2011 Jul 21.

Jamaladdin S, Kelly RD, O'Regan L, Dovey OM, Hodson GE, Millard CJ, Portolano N, Fry AM, Schwabe JW, Cowley SM. Histone deacetylase (HDAC) 1 and 2 are essential for accurate cell division and the pluripotency of embryonic stem cells. *Proc Natl Acad Sci U S A*. 2014 Jul 8;111(27):9840-5. doi: 10.1073/pnas.1321330111.

Jansen C, Paraiso KD, Zhou JJ, Blitz IL, Fish MB, Charney RM, Cho JS, Yasuoka Y, Sudou N, Bright AR, Wlizla M, Veenstra GJC, Taira M, Zorn AM, Mortazavi A, Cho KWY. Uncovering the mesendoderm gene regulatory network through multi-omic data integration. *Cell Rep*. 2022 Feb 15;38(7):110364. doi: 10.1016/j.celrep.2022.110364.

Jonas E, Sargent TD, Dawid IB. Epidermal keratin gene expressed in embryos of *Xenopus laevis*. *Proc Natl Acad Sci U S A*. 1985 Aug;82(16):5413-7. doi: 10.1073/pnas.82.16.5413.

Jones CM, Kuehn MR, Hogan BL, Smith JC, Wright CV. Nodal-related signals induce axial mesoderm and dorsalize mesoderm during gastrulation. *Development*. 1995 Nov;121(11):3651-62. doi: 10.1242/dev.121.11.3651.

Kablar B, Vignali R, Menotti L, Pannese M, Andreazzoli M, Polo C, Giribaldi MG, Boncinelli E, Barsacchi G. *Xotx* genes in the developing brain of *Xenopus laevis*. *Mech Dev*. 1996 Apr;55(2):145-58. doi: 10.1016/0925-4773(96)00497-2.

Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000 Jan 1;28(1):27-30. doi: 10.1093/nar/28.1.27.

Kee Y, Bronner-Fraser M. To proliferate or to die: role of Id3 in cell cycle progression and survival of neural crest progenitors. *Genes Dev*. 2005 Mar 15;19(6):744-55. doi: 10.1101/gad.1257405.

Keller RE. Time-lapse cinemicrographic analysis of superficial cell behavior during and prior to gastrulation in *Xenopus laevis*. *J Morphol*. 1978 Aug;157(2):223-247. doi: 10.1002/jmor.1051570209.

Keller RE. The cellular basis of epiboly: an SEM study of deep-cell rearrangement during gastrulation in *Xenopus laevis*. *J Embryol Exp Morphol*. 1980 Dec;60:201-34.

Keller RE. An experimental analysis of the role of bottle cells and the deep marginal zone in gastrulation of *Xenopus laevis*. *J Exp Zool*. 1981 Apr;216(1):81-101. doi: 10.1002/jez.1402160109.

Keller RE, Danilchik M, Gimlich R, Shih J. The function and mechanism of convergent extension during gastrulation of *Xenopus laevis*. *J Embryol Exp Morphol*. 1985 Nov;89 Suppl:185-209.

Keller R. Early embryonic development of *Xenopus laevis*. *Methods in cell biology*. 1991 Jan 1;36:61-113.

Kelsh RN, Camargo Sosa K, Farjami S, Makeev V, Dawes JHP, Rocco A. Cyclical fate restriction: a new view of neural crest cell fate specification. *Development*. 2021 Nov 15;148(22):dev176057. doi: 10.1242/dev.176057.

Kengaku M, Okamoto H. Basic fibroblast growth factor induces differentiation of neural tube and neural crest lineages of cultured ectoderm cells from *Xenopus gastrula*. *Development*. 1993 Dec;119(4):1067-78. doi: 10.1242/dev.119.4.1067.

Kent B, Magnani E, Walsh MJ, Sadler KC. UHRF1 regulation of Dnmt1 is required for pre-gastrula zebrafish development. *Dev Biol*. 2016 Apr 1;412(1):99-113. doi: 10.1016/j.ydbio.2016.01.036.

- Keshet I, Lieman-Hurwitz J, Cedar H. DNA methylation affects the formation of active chromatin. *Cell*. 1986 Feb 28;44(4):535-43. doi: 10.1016/0092-8674(86)90263-1.
- Kessler DS, Melton DA. Induction of dorsal mesoderm by soluble, mature Vg1 protein. *Development*. 1995 Jul;121(7):2155-64. doi: 10.1242/dev.121.7.2155.
- Kiecker C, Bates T, Bell E. Molecular specification of germ layers in vertebrate embryos. *Cell Mol Life Sci*. 2016 Mar;73(5):923-47. doi: 10.1007/s00018-015-2092-y.
- Kimelman D, Griffin KJ. Vertebrate mesendoderm induction and patterning. *Curr Opin Genet Dev*. 2000 Aug;10(4):350-6. doi: 10.1016/s0959-437x(00)00095-2.
- Kirmizitas A, Gillis WQ, Zhu H, Thomsen GH. Gtpbp2 is required for BMP signaling and mesoderm patterning in *Xenopus* embryos. *Dev Biol*. 2014 Aug 15;392(2):358-67. doi: 10.1016/j.ydbio.2014.05.008.
- Kofron M, Demel T, Xanthos J, Lohr J, Sun B, Sive H, Osada S, Wright C, Wylie C, Heasman J. Mesoderm induction in *Xenopus* is a zygotic event regulated by maternal VegT via TGFbeta growth factors. *Development*. 1999 Dec;126(24):5759-70. doi: 10.1242/dev.126.24.5759.
- Kolde R, Kolde MR. Package 'pheatmap'. R Package. 2018 May 18;1.
- Kretzschmar M, Liu F, Hata A, Doody J, Massagué J. The TGF-beta family mediator Smad1 is phosphorylated directly and activated functionally by the BMP receptor kinase. *Genes Dev*. 1997 Apr 15;11(8):984-95. doi: 10.1101/gad.11.8.984.
- Kumar A, Novoselov V, Celeste AJ, Wolfman NM, ten Dijke P, Kuehn MR. Nodal signaling uses activin and transforming growth factor-beta receptor-regulated Smads. *J Biol Chem*. 2001 Jan 5;276(1):656-61. doi: 10.1074/jbc.M004649200.
- Kurisasi A, Hamazaki TS, Okabayashi K, Iida T, Nishine T, Chonan R, Kido H, Tsunasawa S, Nishimura O, Asashima M, Sugino H. Chromatin-related proteins in pluripotent mouse embryonic stem cells are downregulated after removal of leukemia inhibitory factor. *Biochem Biophys Res Commun*. 2005 Sep 30;335(3):667-75. doi: 10.1016/j.bbrc.2005.07.128.
- Kuroda H, Wessely O, De Robertis EM. Neural induction in *Xenopus*: requirement for ectodermal and endomesodermal signals via Chordin, Noggin, beta-Catenin, and Cerberus. *PLoS Biol*. 2004 May;2(5):E92. doi: 10.1371/journal.pbio.0020092.
- LaBonne C, Burke B, Whitman M. Role of MAP kinase in mesoderm induction and axial patterning during *Xenopus* development. *Development*. 1995 May;121(5):1475-86. doi: 10.1242/dev.121.5.1475.
- LaBonne C, Bronner-Fraser M. Neural crest induction in *Xenopus*: evidence for a two-signal model. *Development*. 1998 Jul;125(13):2403-14. doi: 10.1242/dev.125.13.2403.
- Lachner M, O'Sullivan RJ, Jenuwein T. An epigenetic road map for histone lysine methylation. *J Cell Sci*. 2003 Jun 1;116(Pt 11):2117-24. doi: 10.1242/jcs.00493.
- Ladher R, Mohun TJ, Smith JC, Snape AM. Xom: a *Xenopus* homeobox gene that mediates the early effects of BMP-4. *Development*. 1996 Aug;122(8):2385-94. doi: 10.1242/dev.122.8.2385.
- Lagna G, Hata A, Hemmati-Brivanlou A, Massagué J. Partnership between DPC4 and SMAD proteins in TGF-beta signalling pathways. *Nature*. 1996 Oct 31;383(6603):832-6. doi: 10.1038/383832a0.

- Lamb TM, Knecht AK, Smith WC, Stachel SE, Economides AN, Stahl N, Yancopolous GD, Harland RM. Neural induction by the secreted polypeptide noggin. *Science*. 1993 Oct 29;262(5134):713-8. doi: 10.1126/science.8235591.
- Laslo P, Spooner CJ, Warmflash A, Lancki DW, Lee HJ, Sciammas R, Gantner BN, Dinner AR, Singh H. Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell*. 2006 Aug 25;126(4):755-66. doi: 10.1016/j.cell.2006.06.052.
- Launay C, Fromentoux V, Shi DL, Boucaut JC. A truncated FGF receptor blocks neural induction by endogenous *Xenopus* inducers. *Development*. 1996 Mar;122(3):869-80. doi: 10.1242/dev.122.3.869.
- Laurent MN, Blitz IL, Hashimoto C, Rothbächer U, Cho KW. The *Xenopus* homeobox gene twin mediates Wnt induction of gooseoid in establishment of Spemann's organizer. *Development*. 1997 Dec;124(23):4905-16. doi: 10.1242/dev.124.23.4905.
- Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014 Feb 3;15(2):R29. doi: 10.1186/gb-2014-15-2-r29.
- Lee JH, Hart SR, Skalnik DG. Histone deacetylase activity is required for embryonic stem cell differentiation. *Genesis*. 2004 Jan;38(1):32-8. doi: 10.1002/gene.10250.
- Lee JY, Harland RM. Actomyosin contractility and microtubules drive apical constriction in *Xenopus* bottle cells. *Dev Biol*. 2007 Nov 1;311(1):40-52. doi: 10.1016/j.ydbio.2007.08.010.
- Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, Koseki H, Fuchikami T, Abe K, Murray HL, Zucker JP, Yuan B, Bell GW, Herbolsheimer E, Hannett NM, Sun K, Odom DT, Otte AP, Volkert TL, Bartel DP, Melton DA, Gifford DK, Jaenisch R, Young RA. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell*. 2006 Apr 21;125(2):301-13. doi: 10.1016/j.cell.2006.02.043.
- Lef J, Clement JH, Oswald R, Köster M, Knöchel W. Spatial and temporal transcription patterns of the forkhead related XFD-2/XFD-2' genes in *Xenopus laevis* embryos. *Mech Dev*. 1994 Feb;45(2):117-26. doi: 10.1016/0925-4773(94)90025-6.
- Lemaire P, Darras S, Caillol D, Kodjabachian L. A role for the vegetally expressed *Xenopus* gene Mix.1 in endoderm formation and in the restriction of mesoderm to the marginal zone. *Development*. 1998 Jul;125(13):2371-80. doi: 10.1242/dev.125.13.2371.
- Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol*. 2003 Jul 15;13(14):1192-200. doi: 10.1016/s0960-9822(03)00432-9.
- Levine M, Davidson EH. Gene regulatory networks for development. *Proc Natl Acad Sci U S A*. 2005 Apr 5;102(14):4936-42. doi: 10.1073/pnas.0408031102. Epub 2005 Mar 23.
- Lev Maor G, Yearim A, Ast G. The alternative role of DNA methylation in splicing regulation. *Trends Genet*. 2015 May;31(5):274-80. doi: 10.1016/j.tig.2015.03.002. Epub 2015 Mar 30.
- Li B, Kuriyama S, Moreno M, Mayor R. The posteriorizing gene *Gbx2* is a direct target of Wnt signalling and the earliest factor in neural crest induction. *Development*. 2009 Oct;136(19):3267-78. doi: 10.1242/dev.036954.

- Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011 Aug 4;12:323. doi: 10.1186/1471-2105-12-323.
- Li M, Belmonte JC. Ground rules of the pluripotency gene regulatory network. *Nat Rev Genet*. 2017 Mar;18(3):180-191. doi: 10.1038/nrg.2016.156.
- Liang G, Chan MF, Tomigahara Y, Tsai YC, Gonzales FA, Li E, Laird PW, Jones PA. Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Mol Cell Biol*. 2002 Jan;22(2):480-91. doi: 10.1128/MCB.22.2.480-491.2002.
- Li-Byarlay H, Li Y, Stroud H, Feng S, Newman TC, Kaneda M, Hou KK, Worley KC, Elisk CG, Wickline SA, Jacobsen SE, Ma J, Robinson GE. RNA interference knockdown of DNA methyl-transferase 3 affects gene alternative splicing in the honey bee. *Proc Natl Acad Sci U S A*. 2013 Jul 30;110(31):12750-5. doi: 10.1073/pnas.1310735110. Epub 2013 Jul 12.
- Light W, Vernon AE, Lasorella A, Iavarone A, LaBonne C. *Xenopus* Id3 is required downstream of Myc for the formation of multipotent neural crest progenitor cells. *Development*. 2005 Apr;132(8):1831-41. doi: 10.1242/dev.01734. Epub 2005 Mar 16.
- Linker C, Stern CD. Neural induction requires BMP inhibition only as a late step, and involves signals other than FGF and Wnt antagonists. *Development*. 2004 Nov;131(22):5671-81. doi: 10.1242/dev.01445.
- Little SC, Mullins MC. Bone morphogenetic protein heterodimers assemble heteromeric type I receptor complexes to pattern the dorsoventral axis. *Nat Cell Biol*. 2009 May;11(5):637-43. doi: 10.1038/ncb1870. Epub 2009 Apr 19.
- Liu F, Ventura F, Doody J, Massagué J. Human type II receptor for bone morphogenic proteins (BMPs): extension of the two-kinase receptor model to the BMPs. *Mol Cell Biol*. 1995 Jul;15(7):3479-86. doi: 10.1128/MCB.15.7.3479.
- Liu F, Hata A, Baker JC, Doody J, Cárcamo J, Harland RM, Massagué J. A human Mad protein acting as a BMP-regulated transcriptional activator. *Nature*. 1996 Jun 13;381(6583):620-3. doi: 10.1038/381620a0.
- Liu X, Gao Q, Li P, Zhao Q, Zhang J, Li J, Koseki H, Wong J. UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. *Nat Commun*. 2013;4:1563. doi: 10.1038/ncomms2562.
- Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J, Wong KY, Sung KW, Lee CW, Zhao XD, Chiu KP, Lipovich L, Kuznetsov VA, Robson P, Stanton LW, Wei CL, Ruan Y, Lim B, Ng HH. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet*. 2006 Apr;38(4):431-40. doi: 10.1038/ng1760.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. doi: 10.1186/s13059-014-0550-8.
- Lunyak VV, Rosenfeld MG. Epigenetic regulation of stem cell fate. *Hum Mol Genet*. 2008 Apr 15;17(R1):R28-36. doi: 10.1093/hmg/ddn149.
- Luo T, Matsuo-Takasaki M, Sargent TD. Distinct roles for Distal-less genes *Dlx3* and *Dlx5* in regulating ectodermal development in *Xenopus*. *Mol Reprod Dev*. 2001 Nov;60(3):331-7. doi: 10.1002/mrd.1095.
- Luo T, Lee YH, Saint-Jeannet JP, Sargent TD. Induction of neural crest in *Xenopus* by transcription factor AP2alpha. *Proc Natl Acad Sci U S A*. 2003 Jan 21;100(2):532-7. doi: 10.1073/pnas.0237226100. Epub 2003 Jan 2.

Macías-Silva M, Abdollah S, Hoodless PA, Pirone R, Attisano L, Wrana JL. MADR2 is a substrate of the TGFbeta receptor and its phosphorylation is required for nuclear accumulation and signaling. *Cell*. 1996 Dec 27;87(7):1215-24. doi: 10.1016/s0092-8674(00)81817-6.

Macías-Silva M, Hoodless PA, Tang SJ, Buchwald M, Wrana JL. Specific activation of Smad1 signaling pathways by the BMP7 type I receptor, ALK2. *J Biol Chem*. 1998 Oct 2;273(40):25628-36. doi: 10.1074/jbc.273.40.25628.

Macleod D, Clark VH, Bird A. Absence of genome-wide changes in DNA methylation during development of the zebrafish. *Nat Genet*. 1999 Oct;23(2):139-40. doi: 10.1038/13767.

Maerker M, Getwan M, Dowdle ME, McSheene JC, Gonzalez V, Pelliccia JL, Hamilton DS, Yartseva V, Vejnar C, Tingler M, Minegishi K, Vick P, Giraldez AJ, Hamada H, Burdine RD, Sheets MD, Blum M, Schweickert A. Bicc1 and Dicer regulate left-right patterning through post-transcriptional control of the Nodal inhibitor Dand5. *Nat Commun*. 2021 Sep 16;12(1):5482. doi: 10.1038/s41467-021-25464-z.

Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A*. 1981 Dec;78(12):7634-8. doi: 10.1073/pnas.78.12.7634.

Martins-Taylor K, Schroeder DI, LaSalle JM, Lalande M, Xu RH. Role of DNMT3B in the regulation of early neural and neural crest specifiers. *Epigenetics*. 2012 Jan 1;7(1):71-82. doi: 10.4161/epi.7.1.18750.  
Massagué J. The TGF-beta family of growth and differentiation factors. *Cell*. 1987 May 22;49(4):437-8. doi: 10.1016/0092-8674(87)90443-0.

Massagué J. TGF-beta signal transduction. *Annu Rev Biochem*. 1998;67:753-91. doi: 10.1146/annurev.biochem.67.1.753.

Massagué J, Wotton D. Transcriptional control by the TGF-beta/Smad signaling system. *EMBO J*. 2000 Apr 17;19(8):1745-54. doi: 10.1093/emboj/19.8.1745.

Maunakea AK, Chepelev I, Cui K, Zhao K. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res*. 2013 Nov;23(11):1256-69. doi: 10.1038/cr.2013.110. Epub 2013 Aug 13.

Mayor R, Morgan R, Sargent MG. Induction of the prospective neural crest of *Xenopus*. *Development*. 1995 Mar;121(3):767-77. doi: 10.1242/dev.121.3.767.

McDowell N, Zorn AM, Crease DJ, Gurdon JB. Activin has direct long-range signalling activity and can form a concentration gradient by diffusion. *Curr Biol*. 1997 Sep 1;7(9):671-81. doi: 10.1016/s0960-9822(06)00294-6.

McDowell N, Gurdon JB. Activin as a morphogen in *Xenopus* mesoderm induction. *Semin Cell Dev Biol*. 1999 Jun;10(3):311-7. doi: 10.1006/scdb.1999.0307.

McGrew LL, Hoppler S, Moon RT. Wnt and FGF pathways cooperatively pattern anteroposterior neural ectoderm in *Xenopus*. *Mech Dev*. 1997 Dec;69(1-2):105-14. doi: 10.1016/s0925-4773(97)00160-3.

McKenna A, Findlay GM, Gagnon JA, Horwitz MS, Schier AF, Shendure J. Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science*. 2016 Jul 29;353(6298):aaf7907. doi: 10.1126/science.aaf7907. Epub 2016 May 26.

- McQueen C, Pownall ME. An analysis of MyoD-dependent transcription using CRISPR/Cas9 gene targeting in *Xenopus tropicalis* embryos. *Mech Dev*. 2017 Aug;146:1-9. doi: 10.1016/j.mod.2017.05.002. Epub 2017 May 20.
- Mead PE, Brivanlou IH, Kelley CM, Zon LI. BMP-4-responsive regulation of dorsal-ventral patterning by the homeobox protein Mix.1. *Nature*. 1996 Jul 25;382(6589):357-60. doi: 10.1038/382357a0.
- Mertes C, Scheller IF, Yépez VA, Çelik MH, Liang Y, Kremer LS, Gusic M, Prokisch H, Gagneur J. Detection of aberrant splicing events in RNA-seq data using FRASER. *Nat Commun*. 2021 Jan 22;12(1):529. doi: 10.1038/s41467-020-20573-7. Erratum in: *Nat Commun*. 2022 Jun 16;13(1):3474.
- Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell*. 2006 Jan;10(1):105-16. doi: 10.1016/j.devcel.2005.10.017. Erratum in: *Dev Cell*. Jan 17;22(1):233-4.
- Miller JR, Rowning BA, Larabell CA, Yang-Snyder JA, Bates RL, Moon RT. Establishment of the dorsal-ventral axis in *Xenopus* embryos coincides with the dorsal enrichment of dishevelled that is dependent on cortical rotation. *J Cell Biol*. 1999 Jul 26;146(2):427-37. doi: 10.1083/jcb.146.2.427.
- Minsuk SB, Keller RE. Surface mesoderm in *Xenopus*: a revision of the stage 10 fate map. *Dev Genes Evol*. 1997 Dec;207(6):389-401. doi: 10.1007/s004270050128.
- Mir A, Kofron M, Zorn AM, Bajzer M, Haque M, Heasman J, Wylie CC. Foxl1e activates ectoderm formation and controls cell position in the *Xenopus* blastula. *Development*. 2007 Feb;134(4):779-88. doi: 10.1242/dev.02768. Epub 2007 Jan 17.
- Mizuseki K, Kishi M, Matsui M, Nakanishi S, Sasai Y. *Xenopus* Zic-related-1 and Sox-2, two factors induced by chordin, have distinct activities in the initiation of neural induction. *Development*. 1998 Feb;125(4):579-87. doi: 10.1242/dev.125.4.579.
- Mizuseki K, Kishi M, Shiota K, Nakanishi S, Sasai Y. SoxD: an essential mediator of induction of anterior neural tissues in *Xenopus* embryos. *Neuron*. 1998 Jul;21(1):77-85. doi: 10.1016/s0896-6273(00)80516-4.
- Monk M, Boubelik M, Lehnert S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development*. 1987 Mar;99(3):371-82. doi: 10.1242/dev.99.3.371.
- Monsoro-Burq AH, Fletcher RB, Harland RM. Neural crest induction by paraxial mesoderm in *Xenopus* embryos requires FGF signals. *Development*. 2003 Jul;130(14):3111-24. doi: 10.1242/dev.00531.
- Monsoro-Burq AH, Wang E, Harland R. Msx1 and Pax3 cooperate to mediate FGF8 and WNT signals during *Xenopus* neural crest induction. *Dev Cell*. 2005 Feb;8(2):167-78. doi: 10.1016/j.devcel.2004.12.017.
- Montague TG, Schier AF. Vg1-Nodal heterodimers are the endogenous inducers of mesendoderm. *Elife*. 2017 Nov 15;6:e28183. doi: 10.7554/eLife.28183.
- Moris N, Pina C, Arias AM. Transition states and cell fate decisions in epigenetic landscapes. *Nat Rev Genet*. 2016 Nov;17(11):693-703. doi: 10.1038/nrg.2016.98. Epub 2016 Sep 12.
- Morrison SJ, White PM, Zock C, Anderson DJ. Prospective identification, isolation by flow cytometry, and in vivo self-renewal of multipotent mammalian neural crest stem cells. *Cell*. 1999 Mar 5;96(5):737-49. doi: 10.1016/s0092-8674(00)80583-8.

- Morselli M, Pastor WA, Montanini B, Nee K, Ferrari R, Fu K, Bonora G, Rubbi L, Clark AT, Ottonello S, Jacobsen SE, Pellegrini M. In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. *Elife*. 2015 Apr 7;4:e06205. doi: 10.7554/eLife.06205. Erratum in: *Elife*. 2017 Aug 31;6.
- Mund C, Hackanson B, Stresemann C, Lübbert M, Lyko F. Characterization of DNA demethylation effects induced by 5-Aza-2'-deoxycytidine in patients with myelodysplastic syndrome. *Cancer Res*. 2005 Aug 15;65(16):7086-90. doi: 10.1158/0008-5472.CAN-05-0695.
- Muñoz-Sanjuán I, Brivanlou AH. Neural induction, the default model and embryonic stem cells. *Nat Rev Neurosci*. 2002 Apr;3(4):271-80. doi: 10.1038/nrn786.
- Murgan S, Castro Colabianchi AM, Monti RJ, Boyadjian López LE, Aguirre CE, Stivala EG, Carrasco AE, López SL. FoxA4 favours notochord formation by inhibiting contiguous mesodermal fates and restricts anterior neural development in *Xenopus* embryos. *PLoS One*. 2014 Oct 24;9(10):e110559. doi: 10.1371/journal.pone.0110559.
- Ng RK, Gurdon JB. Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. *Nat Cell Biol*. 2008 Jan;10(1):102-9. doi: 10.1038/ncb1674.
- Nagy PL, Griesenbeck J, Kornberg RD, Cleary ML. A trithorax-group complex purified from *Saccharomyces cerevisiae* is required for methylation of histone H3. *Proc Natl Acad Sci U S A*. 2002 Jan 8;99(1):90-4. doi: 10.1073/pnas.221596698.
- Nakamura Y, de Paiva Alves E, Veenstra GJ, Hoppler S. Tissue- and stage-specific Wnt target gene expression is controlled subsequent to  $\beta$ -catenin recruitment to cis-regulatory modules. *Development*. 2016 Jun 1;143(11):1914-25. doi: 10.1242/dev.131664. Epub 2016 Apr 11.
- Nakata K, Nagai T, Aruga J, Mikoshiba K. *Xenopus* Zic3, a primary regulator both in neural and neural crest development. *Proc Natl Acad Sci U S A*. 1997 Oct 28;94(22):11980-5. doi: 10.1073/pnas.94.22.11980.
- Nakata K, Nagai T, Aruga J, Mikoshiba K. *Xenopus* Zic family and its role in neural and neural crest development. *Mech Dev*. 1998 Jul;75(1-2):43-51. doi: 10.1016/s0925-4773(98)00073-2.
- Nakao A, Afrakhte M, Morén A, Nakayama T, Christian JL, Heuchel R, Itoh S, Kawabata M, Heldin NE, Heldin CH, ten Dijke P. Identification of Smad7, a TGF $\beta$ -inducible antagonist of TGF- $\beta$  signalling. *Nature*. 1997 Oct 9;389(6651):631-5. doi: 10.1038/39369.
- Niehrs C. Regionally specific induction by the Spemann-Mangold organizer. *Nat Rev Genet*. 2004 Jun;5(6):425-34. doi: 10.1038/nrg1347.
- Nieuwkoop PD, Faber J, Gerhart J, Kirschner M. Normal table of *Xenopus laevis* (Daudin): a systematical and chronological survey of the development from the fertilized egg till the end of metamorphosis. Garland Science; 2020 Nov 25.
- Nishimatsu S, Thomsen GH. Ventral mesoderm induction and patterning by bone morphogenetic protein heterodimers in *Xenopus* embryos. *Mech Dev*. 1998 Jun;74(1-2):75-88. doi: 10.1016/s0925-4773(98)00070-7.
- Nishita M, Hashimoto MK, Ogata S, Laurent MN, Ueno N, Shibuya H, Cho KW. Interaction between Wnt and TGF- $\beta$  signalling pathways during formation of Spemann's organizer. *Nature*. 2000 Feb 17;403(6771):781-5. doi: 10.1038/35001602.

Niwa H. How is pluripotency determined and maintained? *Development*. 2007 Feb;134(4):635-46. doi: 10.1242/dev.02787.

Nohno T, Ishikawa T, Saito T, Hosokawa K, Noji S, Wolsing DH, Rosenbaum JS. Identification of a human type II receptor for bone morphogenetic protein-4 that forms differential heteromeric complexes with bone morphogenetic protein type I receptors. *J Biol Chem*. 1995 Sep 22;270(38):22522-6. doi: 10.1074/jbc.270.38.22522.

O'Donnell M, Hong CS, Huang X, Delnicki RJ, Saint-Jeannet JP. Functional analysis of Sox8 during neural crest development in *Xenopus*. *Development*. 2006 Oct;133(19):3817-26. doi: 10.1242/dev.02558. Epub 2006 Aug 30. Erratum in: *Development*. 2006 Oct;133(19):3950.

Okadome T, Yamashita H, Franzén P, Morén A, Heldin CH, Miyazono K. Distinct roles of the intracellular domains of transforming growth factor-beta type I and type II receptors in signal transduction. *J Biol Chem*. 1994 Dec 9;269(49):30753-6.

Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*. 1999 Oct 29;99(3):247-57. doi: 10.1016/s0092-8674(00)81656-6.

Onichtchouk D, Gawantka V, Dosch R, Delius H, Hirschfeld K, Blumenstock C, Niehrs C. The Xvent-2 homeobox gene is part of the BMP-4 signalling pathway controlling [correction of controlling] dorsoventral patterning of *Xenopus* mesoderm. *Development*. 1996 Oct;122(10):3045-53. doi: 10.1242/dev.122.10.3045.

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, Cheng X, Bestor TH. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*. 2007 Aug 9;448(7154):714-7. doi: 10.1038/nature05987.

Osada SI, Wright CV. *Xenopus* nodal-related signaling is essential for mesendodermal patterning during early embryogenesis. *Development*. 1999 Jun;126(14):3229-40. doi: 10.1242/dev.126.14.3229.

Osada SI, Saijoh Y, Frisch A, Yeo CY, Adachi H, Watanabe M, Whitman M, Hamada H, Wright CV. Activin/nodal responsiveness and asymmetric expression of a *Xenopus* nodal-related gene converge on a FAST-regulated module in intron 1. *Development*. 2000 Jun;127(11):2503-14. doi: 10.1242/dev.127.11.2503.

Otani J, Nankumo T, Arita K, Inamoto S, Ariyoshi M, Shirakawa M. Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO Rep*. 2009 Nov;10(11):1235-41. doi: 10.1038/embor.2009.218. Epub 2009 Oct 16.

Owens NDL, Blitz IL, Lane MA, Patrushev I, Overton JD, Gilchrist MJ, Cho KWY, Khokha MK. Measuring Absolute RNA Copy Numbers at High Temporal Resolution Reveals Transcriptome Kinetics in *Development*. *Cell Rep*. 2016 Jan 26;14(3):632-647. doi: 10.1016/j.celrep.2015.12.050. Epub 2016 Jan 7.

Pan G, Tian S, Nie J, Yang C, Ruotti V, Wei H, Jonsdottir GA, Stewart R, Thomson JA. Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell*. 2007 Sep 13;1(3):299-312. doi: 10.1016/j.stem.2007.08.003.

Pander CH. Beiträge zur Entwicklungsgeschichte des Hühnchens im Eye. 1817.

Pannese M, Polo C, Andrezzaoli M, Vignali R, Kablar B, Barsacchi G, Boncinelli E. The *Xenopus* homologue of Otx2 is a maternal homeobox gene that demarcates and specifies anterior body regions. *Development*. 1995 Mar;121(3):707-20. doi: 10.1242/dev.121.3.707.

Paraiso KD, Cho JS, Yong J, Cho KWY. Early *Xenopus* gene regulatory programs, chromatin states, and the role of maternal transcription factors. *Curr Top Dev Biol.* 2020;139:35-60. doi: 10.1016/bs.ctdb.2020.02.009. Epub 2020 Apr 6.

Paranjpe SS, Jacobi UG, van Heeringen SJ, Veenstra GJ. A genome-wide survey of maternal and embryonic transcripts during *Xenopus tropicalis* development. *BMC Genomics.* 2013 Nov 6;14:762. doi: 10.1186/1471-2164-14-762.

Paranjpe SS, Veenstra GJ. Establishing pluripotency in early development. *Biochim Biophys Acta.* 2015 Jun;1849(6):626-36. doi: 10.1016/j.bbagr.2015.03.006. Epub 2015 Apr 7.

Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, McLoughlin EM, Brudno Y, Mahapatra S, Kapranov P, Tahiliani M, Daley GQ, Liu XS, Ecker JR, Milos PM, Agarwal S, Rao A. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature.* 2011 May 19;473(7347):394-7. doi: 10.1038/nature10102. Epub 2011 May 8.

Pechalrieu D, Etievant C, Arimondo PB. DNA methyltransferase inhibitors in cancer: From pharmacology to translational studies. *Biochem Pharmacol.* 2017 Apr 1;129:1-13. doi: 10.1016/j.bcp.2016.12.004.

Pieper M, Ahrens K, Rink E, Peter A, Schlosser G. Differential distribution of competence for panplacodal and neural crest induction to non-neural and neural ectoderm. *Development.* 2012 Mar;139(6):1175-87. doi: 10.1242/dev.074468. Epub 2012 Feb 8.

Penzel R, Oswald R, Chen Y, Tacke L, Grunz H. Characterization and early embryonic expression of a neural specific transcription factor xSOX3 in *Xenopus laevis*. *Int J Dev Biol.* 1997 Oct;41(5):667-77.

Pereira LA, Wong MS, Mei Lim S, Stanley EG, Elefanty AG. The Mix family of homeobox genes--key regulators of mesendoderm formation during vertebrate development. *Dev Biol.* 2012 Jul 15;367(2):163-77. doi: 10.1016/j.ydbio.2012.04.033.

Piccolo S, Sasai Y, Lu B, De Robertis EM. Dorsoventral patterning in *Xenopus*: inhibition of ventral signals by direct binding of chordin to BMP-4. *Cell.* 1996 Aug 23;86(4):589-98. doi: 10.1016/s0092-8674(00)80132-4.

Piccolo S, Agius E, Leyns L, Bhattacharyya S, Grunz H, Bouwmeester T, De Robertis EM. The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. *Nature.* 1999 Feb 25;397(6721):707-10. doi: 10.1038/17820.

Plouhinec JL, Medina-Ruiz S, Borday C, Bernard E, Vert JP, Eisen MB, Harland RM, Monsoro-Burq AH. A molecular atlas of the developing ectoderm defines neural, neural crest, placode, and nonneural progenitor identity in vertebrates. *PLoS Biol.* 2017 Oct 19;15(10):e2004045. doi: 10.1371/journal.pbio.2004045.

Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolzheimer E, Zeitlinger J, Lewitter F, Gifford DK, Young RA. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell.* 2005 Aug 26;122(4):517-27. doi: 10.1016/j.cell.2005.06.026.

Prasad MS, Sauka-Spengler T, LaBonne C. Induction of the neural crest state: control of stem cell attributes by gene regulatory, post-transcriptional and epigenetic interactions. *Dev Biol.* 2012 Jun 1;366(1):10-21. doi: 10.1016/j.ydbio.2012.03.014. Epub 2012 Mar 30.

Qiu X, Hother C, Ralfkiær UM, Søgaard A, Lu Q, Workman CT, Liang G, Jones PA, Grønbaek K. Equitoxic doses of 5-azacytidine and 5-aza-2'deoxyctidine induce diverse immediate and overlapping

heritable changes in the transcriptome. *PLoS One*. 2010 Sep 29;5(9):e12994. doi: 10.1371/journal.pone.0012994.

Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2011 Feb 10;470(7333):279-83. doi: 10.1038/nature09692.

Rao A, LaBonne C. Histone deacetylase activity has an essential role in establishing and maintaining the vertebrate neural crest. *Development*. 2018 Aug 8;145(15):dev163386. doi: 10.1242/dev.163386.

Rashbass J, Taylor MV, Gurdon JB. The DNA-binding protein E12 co-operates with XMyoD in the activation of muscle-specific gene expression in *Xenopus* embryos. *EMBO J*. 1992 Aug;11(8):2981-90. doi: 10.1002/j.1460-2075.1992.tb05368.x.

Rastegar S, Friedle H, Frommer G, Knöchel W. Transcriptional regulation of Xvent homeobox genes. *Mech Dev*. 1999 Mar;81(1-2):139-49. doi: 10.1016/s0925-4773(98)00239-1.

Re'em-Kalma Y, Lamb T, Frank D. Competition between noggin and bone morphogenetic protein 4 activities may regulate dorsalization during *Xenopus* development. *Proc Natl Acad Sci U S A*. 1995 Dec 19;92(26):12141-5. doi: 10.1073/pnas.92.26.12141.

Reich S, Weinstein DC. Repression of Inappropriate Gene Expression in the Vertebrate Embryonic Ectoderm. *Genes (Basel)*. 2019 Nov 6;10(11):895. doi: 10.3390/genes10110895.

Ren W, Fan H, Grimm SA, Guo Y, Kim JJ, Yin J, Li L, Petell CJ, Tan XF, Zhang ZM, Coan JP, Gao L, Cai L, Detrick B, Çetin B, Cui Q, Strahl BD, Gozani O, Wang Y, Miller KM, O'Leary SE, Wade PA, Patel DJ, Wang GG, Song J. Direct readout of heterochromatic H3K9me3 regulates DNMT1-mediated maintenance DNA methylation. *Proc Natl Acad Sci U S A*. 2020 Aug 4;117(31):18439-18447. doi: 10.1073/pnas.2009316117. Epub 2020 Jul 16.

Reissmann E, Jörnvall H, Blokzijl A, Andersson O, Chang C, Minchiotti G, Persico MG, Ibáñez CF, Brivanlou AH. The orphan receptor ALK7 and the Activin receptor ALK4 mediate signaling by Nodal proteins during vertebrate development. *Genes Dev*. 2001 Aug 1;15(15):2010-22. doi: 10.1101/gad.201801.

Reversade B, Kuroda H, Lee H, Mays A, De Robertis EM. Depletion of Bmp2, Bmp4, Bmp7 and Spemann organizer signals induces massive brain formation in *Xenopus* embryos. *Development*. 2005 Aug;132(15):3381-92. doi: 10.1242/dev.01901. Epub 2005 Jun 23.

Robertson A. Conrad Hal Waddington, 8 November 1905--26 September 1975. *Biogr Mem Fellows R Soc*. 1977;23:575-622. doi: 10.1098/rsbm.1977.0022.

Rodaway A, Patient R. Mesendoderm. an ancient germ layer? *Cell*. 2001 Apr 20;105(2):169-72. doi: 10.1016/s0092-8674(01)00307-5.

Roll JD, Rivenbark AG, Jones WD, Coleman WB. DNMT3b overexpression contributes to a hypermethylator phenotype in human breast cancer cell lines. *Mol Cancer*. 2008 Jan 25;7:15. doi: 10.1186/1476-4598-7-15.

Rosa FM. Mix.1, a homeobox mRNA inducible by mesoderm inducers, is expressed mostly in the presumptive endodermal cells of *Xenopus* embryos. *Cell*. 1989 Jun 16;57(6):965-74. doi: 10.1016/0092-8674(89)90335-8.

Rothbart SB, Krajewski K, Nady N, Tempel W, Xue S, Badeaux AI, Barsyte-Lovejoy D, Martinez JY, Bedford MT, Fuchs SM, Arrowsmith CH, Strahl BD. Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat Struct Mol Biol.* 2012 Nov;19(11):1155-60. doi: 10.1038/nsmb.2391. Epub 2012 Sep 30.

Rowning BA, Wells J, Wu M, Gerhart JC, Moon RT, Larabell CA. Microtubule-mediated transport of organelles and localization of beta-catenin to the future dorsal side of *Xenopus* eggs. *Proc Natl Acad Sci U S A.* 1997 Feb 18;94(4):1224-9. doi: 10.1073/pnas.94.4.1224.

Ryan K, Garrett N, Mitchell A, Gurdon JB. Eomesodermin, a key early gene in *Xenopus* mesoderm differentiation. *Cell.* 1996 Dec 13;87(6):989-1000. doi: 10.1016/s0092-8674(00)81794-8.

Saint-Jeannet JP, He X, Varmus HE, Dawid IB. Regulation of dorsal fate in the neuraxis by Wnt-1 and Wnt-3a. *Proc Natl Acad Sci U S A.* 1997 Dec 9;94(25):13713-8. doi: 10.1073/pnas.94.25.13713.

Saka Y, Tada M, Smith JC. A screen for targets of the *Xenopus* T-box gene *Xbra*. *Mech Dev.* 2000 May;93(1-2):27-39. doi: 10.1016/s0925-4773(00)00260-4.

Sano H, Sager R. Tissue specificity and clustering of methylated cytosines in bovine satellite I DNA. *Proc Natl Acad Sci U S A.* 1982 Jun;79(11):3584-8. doi: 10.1073/pnas.79.11.3584.

Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T. Active genes are tri-methylated at K4 of histone H3. *Nature.* 2002 Sep 26;419(6905):407-11. doi: 10.1038/nature01080.

Sasai N, Mizuseki K, Sasai Y. Requirement of FoxD3-class signaling for neural crest determination in *Xenopus*. *Development.* 2001 Jul;128(13):2525-36. doi: 10.1242/dev.128.13.2525.

Sasai Y, Lu B, Steinbeisser H, De Robertis EM. Regulation of neural induction by the *Chd* and *Bmp-4* antagonistic patterning signals in *Xenopus*. *Nature.* 1995 Jul 27;376(6538):333-6. doi: 10.1038/376333a0. Erratum in: *Nature.* 1995 Oct 26;377(6551):757. Erratum in: *Nature.* 1995 Nov 23;378(6555):419.

Sasai Y, Lu B, Piccolo S, De Robertis EM. Endoderm induction by the organizer-secreted factors chordin and noggin in *Xenopus* animal caps. *EMBO J.* 1996 Sep 2;15(17):4547-55.

Sasai Y. Identifying the missing links: genes that connect neural induction and primary neurogenesis in vertebrate embryos. *Neuron.* 1998 Sep;21(3):455-8. doi: 10.1016/s0896-6273(00)80554-1.

Sato SM, Sargent TD. Development of neural inducing capacity in dissociated *Xenopus* embryos. *Dev Biol.* 1989 Jul;134(1):263-6. doi: 10.1016/0012-1606(89)90096-1.

Sato SM, Sargent TD. Localized and inducible expression of *Xenopus*-posterior (*Xpo*), a novel gene active in early frog embryos, encoding a protein with a 'CCHC' finger domain. *Development.* 1991 Jul;112(3):747-53. doi: 10.1242/dev.112.3.747.

Satou-Kobayashi Y, Kim JD, Fukamizu A, Asashima M. Temporal transcriptomic profiling reveals dynamic changes in gene expression of *Xenopus* animal cap upon activin treatment. *Sci Rep.* 2021 Jul 15;11(1):14537. doi: 10.1038/s41598-021-93524-x.

Sauka-Spengler T, Bronner-Fraser M. A gene regulatory network orchestrates neural crest formation. *Nat Rev Mol Cell Biol.* 2008 Jul;9(7):557-68. doi: 10.1038/nrm2428.

Schmidt JE, Suzuki A, Ueno N, Kimelman D. Localized BMP-4 mediates dorsal/ventral patterning in the early *Xenopus* embryo. *Dev Biol.* 1995 May;169(1):37-50. doi: 10.1006/dbio.1995.1124.

Schneider S, Steinbeisser H, Warga RM, Hausen P. Beta-catenin translocation into nuclei demarcates the dorsalizing centers in frog and fish embryos. *Mech Dev.* 1996 Jul;57(2):191-8. doi: 10.1016/0925-4773(96)00546-1.

Schroeder KE, Condic ML, Eisenberg LM, Yost HJ. Spatially regulated translation in embryos: asymmetric expression of maternal Wnt-11 along the dorsal-ventral axis in *Xenopus*. *Dev Biol.* 1999 Oct 15;214(2):288-97. doi: 10.1006/dbio.1999.9426.

Schuler-Metz A, Knöchel S, Kaufmann E, Knöchel W. The homeodomain transcription factor Xvent-2 mediates autocatalytic regulation of BMP-4 expression in *Xenopus* embryos. *J Biol Chem.* 2000 Nov 3;275(44):34365-74. doi: 10.1074/jbc.M003915200.

Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ 3rd. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* 2002 Apr 15;16(8):919-32. doi: 10.1101/gad.973302.

Session AM, Uno Y, Kwon T, Chapman JA, Toyoda A, Takahashi S, Fukui A, Hikosaka A, Suzuki A, Kondo M, van Heeringen SJ, Quigley I, Heinz S, Ogino H, Ochi H, Hellsten U, Lyons JB, Simakov O, Putnam N, Stites J, Kuroki Y, Tanaka T, Michiue T, Watanabe M, Bogdanovic O, Lister R, Georgiou G, Paranjpe SS, van Kruijsbergen I, Shu S, Carlson J, Kinoshita T, Ohta Y, Mawaribuchi S, Jenkins J, Grimwood J, Schmutz J, Mitros T, Mozaffari SV, Suzuki Y, Haramoto Y, Yamamoto TS, Takagi C, Heald R, Miller K, Haudenschield C, Kitzman J, Nakayama T, Izutsu Y, Robert J, Fortriede J, Burns K, Lotay V, Karimi K, Yasuoka Y, Dichmann DS, Flajnik MF, Houston DW, Shendure J, DuPasquier L, Vize PD, Zorn AM, Ito M, Marcotte EM, Wallingford JB, Ito Y, Asashima M, Ueno N, Matsuda Y, Veenstra GJ, Fujiyama A, Harland RM, Taira M, Rokhsar DS. Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature.* 2016 Oct 20;538(7625):336-343. doi: 10.1038/nature19840.

Shamblott MJ, Axelman J, Wang S, Bugg EM, Littlefield JW, Donovan PJ, Blumenthal PD, Huggins GR, Gearhart JD. Derivation of pluripotent stem cells from cultured human primordial germ cells. *Proc Natl Acad Sci U S A.* 1998 Nov 10;95(23):13726-31. doi: 10.1073/pnas.95.23.13726.

Sharma S, De Carvalho DD, Jeong S, Jones PA, Liang G. Nucleosomes containing methylated DNA stabilize DNA methyltransferases 3A/3B and ensure faithful epigenetic inheritance. *PLoS Genet.* 2011 Feb 3;7(2):e1001286. doi: 10.1371/journal.pgen.1001286.

Shayevitch R, Askayo D, Keydar I, Ast G. The importance of DNA methylation of exons on alternative splicing. *RNA.* 2018 Oct;24(10):1351-1362. doi: 10.1261/rna.064865.117. Epub 2018 Jul 12.

Shi Y, Massagué J. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell.* 2003 Jun 13;113(6):685-700. doi: 10.1016/s0092-8674(03)00432-x.

Shih J, Keller R. Cell motility driving mediolateral intercalation in explants of *Xenopus laevis*. *Development.* 1992a Dec;116(4):901-14. doi: 10.1242/dev.116.4.901.

Shih J, Keller R. Patterns of cell motility in the organizer and dorsal mesoderm of *Xenopus laevis*. *Development.* 1992b Dec;116(4):915-30. doi: 10.1242/dev.116.4.915.

Shih J, Keller R. Gastrulation in *Xenopus laevis*: Involution—a current view. In *Seminars in DEVELOPMENTAL BIOLOGY* 1994 Apr 1 (Vol. 5, No. 2, pp. 85-90). Academic Press.

Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, Oberdoerffer P, Sandberg R, Oberdoerffer S. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature.* 2011 Nov 3;479(7371):74-9. doi: 10.1038/nature10442.

- Slack JM, Darlington BG, Heath JK, Godsave SF. Mesoderm induction in early *Xenopus* embryos by heparin-binding growth factors. *Nature*. 1987 Mar 12-18;326(6109):197-200. doi: 10.1038/326197a0
- Smith JC, Price BM, Van Nimmen K, Huylebroeck D. Identification of a potent *Xenopus* mesoderm-inducing factor as a homologue of activin A. *Nature*. 1990 Jun 21;345(6277):729-31. doi: 10.1038/345729a0.
- Smith WC, Harland RM. Expression cloning of noggin, a new dorsalizing factor localized to the Spemann organizer in *Xenopus* embryos. *Cell*. 1992 Sep 4;70(5):829-40. doi: 10.1016/0092-8674(92)90316-5.
- Smith JC, Price BM, Green JB, Weigel D, Herrmann BG. Expression of a *Xenopus* homolog of Brachyury (T) is an immediate-early response to mesoderm induction. *Cell*. 1991 Oct 4;67(1):79-87. doi: 10.1016/0092-8674(91)90573-h.
- Smukler SR, Runciman SB, Xu S, van der Kooy D. Embryonic stem cells assume a primitive neural stem cell fate in the absence of extrinsic influences. *J Cell Biol*. 2006 Jan 2;172(1):79-90. doi: 10.1083/jcb.200508085.
- Snape A, Wylie CC, Smith JC, Heasman J. Changes in states of commitment of single animal pole blastomeres of *Xenopus laevis*. *Dev Biol*. 1987 Feb;119(2):503-10. doi: 10.1016/0012-1606(87)90053-4.
- Solnica-Krezel L, Sepich DS. Gastrulation: making and shaping germ layers. *Annu Rev Cell Dev Biol*. 2012;28:687-717. doi: 10.1146/annurev-cellbio-092910-154043.
- Sokol SY, Melton DA. Interaction of Wnt and activin in dorsal mesoderm induction in *Xenopus*. *Dev Biol*. 1992 Dec;154(2):348-55. doi: 10.1016/0012-1606(92)90073-p.
- Sokol SY. Analysis of Dishevelled signalling pathways during *Xenopus* development. *Curr Biol*. 1996 Nov 1;6(11):1456-67. doi: 10.1016/s0960-9822(96)00750-6.
- Soriano AO, Yang H, Faderl S, Estrov Z, Giles F, Ravandi F, Cortes J, Wierda WG, Ouzounian S, Quezada A, Pierce S, Estey EH, Issa JP, Kantarjian HM, Garcia-Manero G. Safety and clinical activity of the combination of 5-azacytidine, valproic acid, and all-trans retinoic acid in acute myeloid leukemia and myelodysplastic syndrome. *Blood*. 2007 Oct 1;110(7):2302-8. doi: 10.1182/blood-2007-03-078576. Epub 2007 Jun 27.
- Spemann H, Mangold H. über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Archiv für mikroskopische Anatomie und Entwicklungsmechanik*. 1924 Sep;100(3):599-638.
- Spokony RF, Aoki Y, Saint-Germain N, Magner-Fink E, Saint-Jeannet JP. The transcription factor Sox9 is required for cranial neural crest development in *Xenopus*. *Development*. 2002 Jan;129(2):421-32. doi: 10.1242/dev.129.2.421.
- Stennard F, Zorn AM, Ryan K, Garrett N, Gurdon JB. Differential expression of VegT and Antipodean protein isoforms in *Xenopus*. *Mech Dev*. 1999 Aug;86(1-2):87-98. doi: 10.1016/s0925-4773(99)00119-7.
- Stern CD. Neural induction: old problem, new findings, yet more questions. *Development*. 2005 May;132(9):2007-21. doi: 10.1242/dev.01794.
- Strahl BD, Ohba R, Cook RG, Allis CD. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in *Tetrahymena*. *Proc Natl Acad Sci U S A*. 1999 Dec 21;96(26):14967-72. doi: 10.1073/pnas.96.26.14967.

- Strahl BD, Grant PA, Briggs SD, Sun ZW, Bone JR, Caldwell JA, Mollah S, Cook RG, Shabanowitz J, Hunt DF, Allis CD. Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol Cell Biol.* 2002 Mar;22(5):1298-306. doi: 10.1128/MCB.22.5.1298-1306.2002.
- Streit A, Lee KJ, Woo I, Roberts C, Jessell TM, Stern CD. Chordin regulates primitive streak development and the stability of induced neural cells, but is not sufficient for neural induction in the chick embryo. *Development.* 1998 Feb;125(3):507-19. doi: 10.1242/dev.125.3.507.
- Streit A, Stern CD. Mesoderm patterning and somite formation during node regression: differential effects of chordin and noggin. *Mech Dev.* 1999 Jul;85(1-2):85-96. doi: 10.1016/s0925-4773(99)00085-4.
- Streit A, Stern CD. Establishment and maintenance of the border of the neural plate in the chick: involvement of FGF and BMP activity. *Mech Dev.* 1999 Apr;82(1-2):51-66. doi: 10.1016/s0925-4773(99)00013-1.
- Streit A, Berliner AJ, Papanayotou C, Sirulnik A, Stern CD. Initiation of neural induction by FGF signalling before gastrulation. *Nature.* 2000 Jul 6;406(6791):74-8. doi: 10.1038/35017617.
- Stresemann C, Lyko F. Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine. *Int J Cancer.* 2008 Jul 1;123(1):8-13. doi: 10.1002/ijc.23607.
- Sumanas S, Strege P, Heasman J, Ekker SC. The putative wnt receptor *Xenopus* frizzled-7 functions upstream of beta-catenin in vertebrate dorsoventral mesoderm patterning. *Development.* 2000 May;127(9):1981-90. doi: 10.1242/dev.127.9.1981.
- Sun BI, Bush SM, Collins-Racie LA, LaVallie ER, DiBlasio-Smith EA, Wolfman NM, McCoy JM, Sive HL. *derrière*: a TGF-beta family member required for posterior development in *Xenopus*. *Development.* 1999 Apr;126(7):1467-82. doi: 10.1242/dev.126.7.1467.
- Suri C, Harembaki T, Weinstein DC. Inhibition of mesodermal fate by *Xenopus* HNF3beta/FoxA2. *Dev Biol.* 2004 Jan 1;265(1):90-104. doi: 10.1016/j.ydbio.2003.09.017.
- Suzuki A, Ueno N, Hemmati-Brivanlou A. *Xenopus* *msx1* mediates epidermal induction and neural inhibition by BMP4. *Development.* 1997a Aug;124(16):3037-44. doi: 10.1242/dev.124.16.3037.
- Suzuki A, Kaneko E, Maeda J, Ueno N. Mesoderm induction by BMP-4 and -7 heterodimers. *Biochem Biophys Res Commun.* 1997b Mar 6;232(1):153-6. doi: 10.1006/bbrc.1997.6219.
- Suzuki A, Chang C, Yingling JM, Wang XF, Hemmati-Brivanlou A. *Smad5* induces ventral fates in *Xenopus* embryo. *Dev Biol.* 1997c Apr 15;184(2):402-5. doi: 10.1006/dbio.1997.8548.
- Suzuki A, Kaneko E, Ueno N, Hemmati-Brivanlou A. Regulation of epidermal induction by BMP2 and BMP7 signaling. *Dev Biol.* 1997d Sep 1;189(1):112-22. doi: 10.1006/dbio.1997.8652.
- Suzuki A, Raya A, Kawakami Y, Morita M, Matsui T, Nakashima K, Gage FH, Rodríguez-Esteban C, Izpisua Belmonte JC. Maintenance of embryonic stem cell pluripotency by Nanog-mediated reversal of mesoderm specification. *Nat Clin Pract Cardiovasc Med.* 2006a Mar;3 Suppl 1:S114-22. doi: 10.1038/ncpcardio0442.
- Suzuki A, Raya Á, Kawakami Y, Morita M, Matsui T, Nakashima K, Gage FH, Rodríguez-Esteban C, Izpisua Belmonte JC. Nanog binds to *Smad1* and blocks bone morphogenetic protein-induced

- differentiation of embryonic stem cells. *Proc Natl Acad Sci U S A*. 2006b Jul 5;103(27):10294-10299. doi: 10.1073/pnas.0506945103.
- Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, Fukuda M, Takeda N, Niida H, Kato H, Shinkai Y. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev*. 2002 Jul 15;16(14):1779-91. doi: 10.1101/gad.989402.
- Tada M, Casey ES, Fairclough L, Smith JC. Bix1, a direct target of *Xenopus* T-box genes, causes formation of ventral mesoderm and endoderm. *Development*. 1998 Oct;125(20):3997-4006. doi: 10.1242/dev.125.20.3997.
- Tada M, Smith JC. Xwnt11 is a target of *Xenopus* Brachyury: regulation of gastrulation movements via Dishevelled, but not through the canonical Wnt pathway. *Development*. 2000 May;127(10):2227-38. doi: 10.1242/dev.127.10.2227.
- Tada S, Era T, Furusawa C, Sakurai H, Nishikawa S, Kinoshita M, Nakao K, Chiba T, Nishikawa S. Characterization of mesendoderm: a diverging point of the definitive endoderm and mesoderm in embryonic stem cell differentiation culture. *Development*. 2005 Oct;132(19):4363-74. doi: 10.1242/dev.02005.
- Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 2009 May 15;324(5929):930-5. doi: 10.1126/science.1170116. Epub 2009 Apr 16.
- Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006 Aug 25;126(4):663-76. doi: 10.1016/j.cell.2006.07.024.
- Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007 Nov 30;131(5):861-72. doi: 10.1016/j.cell.2007.11.019.
- Tang Y, Horikoshi M, Li W. ggfortify: unified interface to visualize statistical results of popular R packages. *R J*. 2016 Dec 1;8(2):474.
- Tao J, Kulyiyev E, Wang X, Li X, Wilanowski T, Jane SM, Mead PE, Cunningham JM. BMP4-dependent expression of *Xenopus* Grainyhead-like 1 is essential for epidermal differentiation. *Development*. 2005 Mar;132(5):1021-34. doi: 10.1242/dev.01641.
- Tavares AT, Andrade S, Silva AC, Belo JA. Cerberus is a feedback inhibitor of Nodal asymmetric signaling in the chick embryo. *Development*. 2007 Jun;134(11):2051-60. doi: 10.1242/dev.000901.
- Thomsen G, Woolf T, Whitman M, Sokol S, Vaughan J, Vale W, Melton DA. Activins are expressed early in *Xenopus* embryogenesis and can induce axial mesoderm and anterior structures. *Cell*. 1990 Nov 2;63(3):485-93. doi: 10.1016/0092-8674(90)90445-k.
- Thomsen GH, Melton DA. Processed Vg1 protein is an axial mesoderm inducer in *Xenopus*. *Cell*. 1993 Aug 13;74(3):433-41. doi: 10.1016/0092-8674(93)80045-g
- Thomsen GH. *Xenopus* mothers against decapentaplegic is an embryonic ventralizing agent that acts downstream of the BMP-2/4 receptor. *Development*. 1996 Aug;122(8):2359-66. doi: 10.1242/dev.122.8.2359.

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Embryonic stem cell lines derived from human blastocysts. *Science*. 1998 Nov 6;282(5391):1145-7. doi: 10.1126/science.282.5391.1145.

Tolhuis B, de Wit E, Muijters I, Teunissen H, Talhout W, van Steensel B, van Lohuizen M. Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in *Drosophila melanogaster*. *Nat Genet*. 2006 Jun;38(6):694-9. doi: 10.1038/ng1792. Epub 2006 Apr 20.

Trentin A, Glavieux-Pardanaud C, Le Douarin NM, Dupin E. Self-renewal capacity is a widespread property of various types of neural crest precursor cells. *Proc Natl Acad Sci U S A*. 2004 Mar 30;101(13):4495-500. doi: 10.1073/pnas.0400629101. Epub 2004 Mar 15.

Tribulo C, Aybar MJ, Nguyen VH, Mullins MC, Mayor R. Regulation of *Msx* genes by a *Bmp* gradient is essential for neural crest specification. *Development*. 2003 Dec;130(26):6441-52. doi: 10.1242/dev.00878. Epub 2003 Nov 19.

Trindade M, Tada M, Smith JC. DNA-binding specificity and embryological function of *Xom* (*Xvent-2*). *Dev Biol*. 1999 Dec 15;216(2):442-56. doi: 10.1006/dbio.1999.9507.

Tronick E, Hunter RG. Waddington, Dynamic Systems, and Epigenetics. *Front Behav Neurosci*. 2016 Jun 10;10:107. doi: 10.3389/fnbeh.2016.00107.

Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S, Sakaue M, Matsuoka C, Shimotohno K, Ishikawa F, Li E, Ueda HR, Nakayama J, Okano M. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases *Dnmt1*, *Dnmt3a* and *Dnmt3b*. *Genes Cells*. 2006 Jul;11(7):805-14. doi: 10.1111/j.1365-2443.2006.00984.x.

Umbhauer M, Marshall CJ, Mason CS, Old RW, Smith JC. Mesoderm induction in *Xenopus* caused by activation of MAP kinase. *Nature*. 1995 Jul 6;376(6535):58-62. doi: 10.1038/376058a0.

Vallier L, Reynolds D, Pedersen RA. Nodal inhibits differentiation of human embryonic stem cells along the neuroectodermal default pathway. *Dev Biol*. 2004 Nov 15;275(2):403-21. doi: 10.1016/j.ydbio.2004.08.031.

Veenstra GJ, Wolffe AP. Constitutive genomic methylation during embryonic development of *Xenopus*. *Biochim Biophys Acta*. 2001 Oct 31;1521(1-3):39-44. doi: 10.1016/s0167-4781(01)00280-9.

Verschueren K, Remacle JE, Collart C, Kraft H, Baker BS, Tylzanowski P, Nelles L, Wuytens G, Su MT, Bodmer R, Smith JC, Huylebroeck D. SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5'-CACCT sequences in candidate target genes. *J Biol Chem*. 1999 Jul 16;274(29):20489-98. doi: 10.1074/jbc.274.29.20489.

Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat İ, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P; SciPy 1.0 Contributors. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020 Mar;17(3):261-272. doi: 10.1038/s41592-019-0686-2. Epub 2020 Feb 3. Erratum in: *Nat Methods*. 2020 Feb 24;:

Vonica A, Brivanlou AH. The left-right axis is regulated by the interplay of *Coco*, *Xnr1* and *derrière* in *Xenopus* embryos. *Dev Biol*. 2007 Mar 1;303(1):281-94. doi: 10.1016/j.ydbio.2006.09.039.

Vincent, J P, and J C Gerhart. "Subcortical rotation in *Xenopus* eggs: an early step in embryonic axis specification." *Developmental biology* vol. 123,2 (1987): 526-39. doi:10.1016/0012-1606(87)90411-8

Waddington CH. Organisers and genes. Organisers and genes.. 1940.

Waddington CH. The epigenotype. 1942. *Int J Epidemiol.* 2012 Feb;41(1):10-3. doi: 10.1093/ije/dyr184.

Waddington CH. A discussion of some aspects of theoretical biology. *The Strategy of the Genes*, ed. CH Waddington (London: London George Allen and Unwin). 1957.

Waddington CH. Towards a theoretical biology. *Nature.* 1968 May 11;218(5141):525-7. doi: 10.1038/218525a0.

Wallingford JB, Rowling BA, Vogeli KM, Rothbacher U, Fraser SE, Harland RM. Dishevelled controls cell polarity during *Xenopus* gastrulation. *Nature.* 2000 May 4;405(6782):81-5. doi: 10.1038/35011077.

Wallingford JB, Harland RM. *Xenopus* Dishevelled signaling regulates both neural and mesodermal convergent extension: parallel forces elongating the body axis. *Development.* 2001 Jul;128(13):2581-92. doi: 10.1242/dev.128.13.2581.

Wang J, Zhang K, Xu L, Wang E. Quantifying the Waddington landscape and biological paths for development and differentiation. *Proc Natl Acad Sci U S A.* 2011 May 17;108(20):8257-62. doi: 10.1073/pnas.1017017108.

Wang Q, Wei X, Zhu T, Zhang M, Shen R, Xing L, O'Keefe RJ, Chen D. Bone morphogenetic protein 2 activates Smad6 gene transcription through bone-specific transcription factor Runx2. *J Biol Chem.* 2007 Apr 6;282(14):10742-8. doi: 10.1074/jbc.M610997200. Epub 2007 Jan 10.

Wardle FC, Smith JC. Transcriptional regulation of mesendoderm formation in *Xenopus*. *Semin Cell Dev Biol.* 2006 Feb;17(1):99-109. doi: 10.1016/j.semcd.2005.11.008. Epub 2005 Dec 13.

Warga RM, Nüsslein-Volhard C. Origin and development of the zebrafish endoderm. *Development.* 1999 Feb;126(4):827-38. doi: 10.1242/dev.126.4.827.

Watabe T, Kim S, Candia A, Rothbacher U, Hashimoto C, Inoue K, Cho KW. Molecular mechanisms of Spemann's organizer formation: conserved growth factor synergy between *Xenopus* and mouse. *Genes Dev.* 1995 Dec 15;9(24):3038-50. doi: 10.1101/gad.9.24.3038.

Watanabe D, Suetake I, Tada T, Tajima S. Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. *Mech Dev.* 2002 Oct;118(1-2):187-90. doi: 10.1016/s0925-4773(02)00242-3.

Watanabe D, Suetake I, Tajima S, Hanaoka K. Expression of Dnmt3b in mouse hematopoietic progenitor cells and spermatogonia at specific stages. *Gene Expr Patterns.* 2004 Nov;5(1):43-9. doi: 10.1016/j.modgep.2004.06.008.

Watanabe D, Uchiyama K, Hanaoka K. Transition of mouse de novo methyltransferases expression from Dnmt3b to Dnmt3a during neural progenitor cell development. *Neuroscience.* 2006 Oct 27;142(3):727-37. doi: 10.1016/j.neuroscience.2006.07.053.

Watanabe M, Whitman M. FAST-1 is a key maternal effector of mesoderm inducers in the early *Xenopus* embryo. *Development.* 1999 Dec;126(24):5621-34. doi: 10.1242/dev.126.24.5621.

Wawersik, S., Evola, C., & Whitman, M. (2005). Conditional BMP inhibition in *Xenopus* reveals stage-specific roles for BMPs in neural and neural crest induction. *Developmental biology*, 277(2), 425-442.

- Weber AR, Krawczyk C, Robertson AB, Kuśnierczyk A, Vågbo CB, Schuermann D, Klungland A, Schär P. Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. *Nat Commun*. 2016 Mar 2;7:10806. doi: 10.1038/ncomms10806.
- Weber H, Symes CE, Walmsley ME, Rodaway AR, Patient RK. A role for GATA5 in *Xenopus* endoderm specification. *Development*. 2000 Oct;127(20):4345-60. doi: 10.1242/dev.127.20.4345.
- Weinberg DN, Papillon-Cavanagh S, Chen H, Yue Y, Chen X, Rajagopalan KN, Horth C, McGuire JT, Xu X, Nikbakht H, Lemiesz AE, Marchione DM, Marunde MR, Meiners MJ, Cheek MA, Keogh MC, Bareke E, Djedid A, Harutyunyan AS, Jabado N, Garcia BA, Li H, Allis CD, Majewski J, Lu C. The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. *Nature*. 2019 Sep;573(7773):281-286. doi: 10.1038/s41586-019-1534-3. Epub 2019 Sep 4.
- Wessely O, Agius E, Oelgeschläger M, Pera EM, De Robertis EM. Neural induction in the absence of mesoderm: beta-catenin-dependent expression of secreted BMP antagonists at the blastula stage in *Xenopus*. *Dev Biol*. 2001 Jun 1;234(1):161-73. doi: 10.1006/dbio.2001.0258.
- Wills AE, Choi VM, Bennett MJ, Khokha MK, Harland RM. BMP antagonists and FGF signaling contribute to different domains of the neural plate in *Xenopus*. *Dev Biol*. 2010 Jan 15;337(2):335-50. doi: 10.1016/j.ydbio.2009.11.008. Epub 2009 Nov 10.
- Wilson PA, Hemmati-Brivanlou A. Induction of epidermis and inhibition of neural fate by Bmp-4. *Nature*. 1995 Jul 27;376(6538):331-3. doi: 10.1038/376331a0.
- Wilson PA, Lagna G, Suzuki A, Hemmati-Brivanlou A. Concentration-dependent patterning of the *Xenopus* ectoderm by BMP4 and its signal transducer Smad1. *Development*. 1997 Aug;124(16):3177-84. doi: 10.1242/dev.124.16.3177.
- Winning RS, Shea LJ, Marcus SJ, Sargent TD. Developmental regulation of transcription factor AP-2 during *Xenopus laevis* embryogenesis. *Nucleic Acids Res*. 1991 Jul 11;19(13):3709-14. doi: 10.1093/nar/19.13.3709.
- Wu J, Yang J, Klein PS. Neural crest induction by the canonical Wnt pathway can be dissociated from anterior-posterior neural patterning in *Xenopus*. *Dev Biol*. 2005 Mar 1;279(1):220-32. doi: 10.1016/j.ydbio.2004.12.016.
- Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet*. 2017 Sep;18(9):517-534. doi: 10.1038/nrg.2017.33. Epub 2017 May 30.
- Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH, Allis CD. WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell*. 2005 Jun 17;121(6):859-72. doi: 10.1016/j.cell.2005.03.036.
- Xie W, Song C, Young NL, Sperling AS, Xu F, Sridharan R, Conway AE, Garcia BA, Plath K, Clark AT, Grunstein M. Histone h3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells. *Mol Cell*. 2009 Feb 27;33(4):417-27. doi: 10.1016/j.molcel.2009.02.004.
- Xu RH, Kim J, Taira M, Lin JJ, Zhang CH, Sredni D, Evans T, Kung HF. Differential regulation of neurogenesis by the two *Xenopus* GATA-1 genes. *Mol Cell Biol*. 1997 Jan;17(1):436-43. doi: 10.1128/MCB.17.1.436.
- Yang AS, Doshi KD, Choi SW, Mason JB, Mannari RK, Gharybian V, Luna R, Rashid A, Shen L, Estecio MR, Kantarjian HM, Garcia-Manero G, Issa JP. DNA methylation changes after 5-aza-2'-deoxycytidine

therapy in patients with leukemia. *Cancer Res.* 2006 May 15;66(10):5495-503. doi: 10.1158/0008-5472.CAN-05-2385.

Yang X, Lay F, Han H, Jones PA. Targeting DNA methylation for epigenetic therapy. *Trends Pharmacol Sci.* 2010 Nov;31(11):536-46. doi: 10.1016/j.tips.2010.08.001. Epub 2010 Sep 16.

Yearim A, Gelfman S, Shayevitch R, Melcer S, Glaich O, Mallm JP, Nissim-Rafinia M, Cohen AH, Rippe K, Meshorer E, Ast G. HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. *Cell Rep.* 2015 Feb 24;10(7):1122-34. doi: 10.1016/j.celrep.2015.01.038.

Yeo CY, Chen X, Whitman M. The role of FAST-1 and Smads in transcriptional regulation by activin during early *Xenopus* embryogenesis. *J Biol Chem.* 1999 Sep 10;274(37):26584-90. doi: 10.1074/jbc.274.37.26584.

Ying QL, Nichols J, Chambers I, Smith A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell.* 2003 Oct 31;115(3):281-92. doi: 10.1016/s0092-8674(03)00847-x.

Young RA. Control of the embryonic stem cell state. *Cell.* 2011 Mar 18;144(6):940-54. doi: 10.1016/j.cell.2011.01.032.

Yu Z, Xiao Q, Zhao L, Ren J, Bai X, Sun M, Wu H, Liu X, Song Z, Yan Y, Mi X, Wang E, Jin F, Wei M. DNA methyltransferase 1/3a overexpression in sporadic breast cancer is associated with reduced expression of estrogen receptor-alpha/breast cancer susceptibility gene 1 and poor prognosis. *Mol Carcinog.* 2015 Sep;54(9):707-19. doi: 10.1002/mc.22133. Epub 2014 Jan 25.

Zalc A, Sinha R, Gulati GS, Wesche DJ, Daszczuk P, Swigut T, Weissman IL, Wysocka J. Reactivation of the pluripotency program precedes formation of the cranial neural crest. *Science.* 2021 Feb 5;371(6529):eabb4776. doi: 10.1126/science.abb4776.

Zhang C, Basta T, Fawcett SR, Klymkowsky MW. SOX7 is an immediate-early target of VegT and regulates Nodal-related gene expression in *Xenopus*. *Dev Biol.* 2005 Feb 15;278(2):526-41. doi: 10.1016

Zhang H, Reynaud S, Kloc M, Etkin LD, Spohr G. Id gene activity during *Xenopus* embryogenesis. *Mech Dev.* 1995 Apr;50(2-3):119-30. doi: 10.1016/0925-4773(94)00329-l.

Zhang J, Houston DW, King ML, Payne C, Wylie C, Heasman J. The role of maternal VegT in establishing the primary germ layers in *Xenopus* embryos. *Cell.* 1998 Aug 21;94(4):515-24. doi: 10.1016/s0092-8674(00)81592-5.

Zhang YW, Yasui N, Ito K, Huang G, Fujii M, Hanai J, Nogami H, Ochi T, Miyazono K, Ito Y. A RUNX2/PEBP2alpha A/CBFA1 mutation displaying impaired transactivation and Smad interaction in cleidocranial dysplasia. *Proc Natl Acad Sci U S A.* 2000 Sep 12;97(19):10549-54. doi: 10.1073/pnas.180309597.

Zimmerman LB, De Jesús-Escobar JM, Harland RM. The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4. *Cell.* 1996 Aug 23;86(4):599-606. doi: 10.1016/s0092-8674(00)80133-6.