Quantitative transcriptional input/output relationships in the *Ciona* notochord


by


Matthew J. Harder


B.S., Hope College, 2015


AN ABSTRACT OF A DISSERTATION


submitted in partial fulfillment of the requirements for the degree


DOCTOR OF PHILOSOPHY


Division of Biology
College of Arts and Sciences


KANSAS STATE UNIVERSITY
Manhattan, Kansas


2020

# Abstract

Embryonic development proceeds under the control of an intricate network of signaling molecules and transcription factors that control the spatial and temporal dynamics of gene expression. These gene regulatory networks depend on enhancers and related cis-regulatory DNA sequences that integrate information from multiple upstream regulators to control the transcription of individual downstream genes. This integration involves precise relationships between input signals and output responses, and while the direct activators and repressors of many enhancers are known, the quantitative details of these input/output relationships are poorly understood. The invertebrate tunicate chordate, *Ciona robusta*, is an excellent model system for examining transcriptional input/output relationships thanks to a small sequenced genome, rapid embryogenesis, simple and stereotyped cell lineages, and straightforward transgenesis by electroporation. In this dissertation, I leverage these advantages to address multiple questions about tissue-specific transcriptional regulation using the *Ciona* notochord as a model. RNA-seq on notochord-enriched cell populations and *in situ* hybridization validation of those data is used to establish a comprehensive *Ciona* notochord transcriptome and identify new expression domains in subregions of the notochord. I dissected the enhancer regions of a gene specific to the posterior notochord and identified upstream regulators of this novel expression pattern. This cis-regulatory analysis identified multiple regions capable of driving the same expression pattern, consistent with growing evidence that seemingly redundant distributed enhancers are common. I developed a reporter-based assay based on the graded pharmacological inhibition of an upstream signal to quantify input/output relationships for two distributed enhancers of the key notochord regulator, *Brachyury*. Despite driving similar notochord-specific expression patterns, I found that these two enhancers have fundamentally different dose-response relationships to MAPK

signaling activity.  This indicates that they are not truly redundant genetic elements and supports

a model in which distributed enhancers play an important role in shaping transcriptional

input/output curves.

Quantitative transcriptional input/output relationships in the *Ciona* notochord

by

Matthew J. Harder

B.S., Hope College, 2015

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology

College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Approved by:

Major Professor
Michael Veeman

# Copyright

# Abstract

Embryonic development proceeds under the control of an intricate network of signaling molecules and transcription factors that control the spatial and temporal dynamics of gene expression. These gene regulatory networks depend on enhancers and related cis-regulatory DNA sequences that integrate information from multiple upstream regulators to control the transcription of individual downstream genes. This integration involves precise relationships between input signals and output responses, and while the direct activators and repressors of many enhancers are known, the quantitative details of these input/output relationships are poorly understood. The invertebrate tunicate chordate, *Ciona robusta*, is an excellent model system for examining transcriptional input/output relationships thanks to a small sequenced genome, rapid embryogenesis, simple and stereotyped cell lineages, and straightforward transgenesis by electroporation. In this dissertation, I leverage these advantages to address multiple questions about tissue-specific transcriptional regulation using the *Ciona* notochord as a model. RNA-seq on notochord-enriched cell populations and *in situ* hybridization validation of those data is used to establish a comprehensive *Ciona* notochord transcriptome and identify new expression domains in subregions of the notochord. I dissected the enhancer regions of a gene specific to the posterior notochord and identified upstream regulators of this novel expression pattern. This cis-regulatory analysis identified multiple regions capable of driving the same expression pattern, consistent with growing evidence that seemingly redundant distributed enhancers are common. I developed a reporter-based assay based on the graded pharmacological inhibition of an upstream signal to quantify input/output relationships for two distributed enhancers of the key notochord regulator, *Brachyury*. Despite driving similar notochord-specific expression patterns, I found that these two enhancers have fundamentally different dose-response relationships to MAPK

signaling activity.  This indicates that they are not truly redundant genetic elements and supports

a model in which distributed enhancers play an important role in shaping transcriptional

input/output curves.

# Table of Contents

# List of Figures

# List of Tables

# Acknowledgements

# Dedication

This work is dedicated to all the teachers in my life, and their lessons that built twin passions for science and helping others.

# Chapter 1 - Introduction

## Transcriptional Regulation in Development

### Gene Regulatory Networks and Boolean Models

Embryonic development relies on an intricate clockwork of signaling molecules, second messengers, and transcription factors (TFs) that coordinate their action to promote cell division, determine cell fates, direct migratory behaviors, and give shape to the organism. Understanding how different gene products regulate the expression of other genes and transmit signals throughout the embryo is one of the central questions of developmental biology. In order to conceptualize the thousands of genes and regulatory steps involved in development, biologists map these relationships into gene regulatory networks (GRNs) that define all of the known regulatory steps and connections in both time and space in the developing embryo or organ (Levine and Davidson, 2005).

In constructing a GRN, each gene forms a 'node,' or a point in the network that can interact with other points, while regulatory interactions form an 'edge' between nodes (Li and Davidson, 2009, Wilczynski and Furlong, 2009). Depending on the type of GRN model, these regulatory interactions may represent direct activation or repression of target genes by transcription factors, or they may represent statistical patterns of correlated expression. GRN models can also incorporate inductive signaling events between different cell types that regulate specific transcription factors.

The genes in a GRN fall into three main categories: signaling factors, transcription factors, and downstream effectors. Signaling factors allow cells to communicate to one another, and are necessary for inductive events to occur. Transcription factors are key convergence points in the networks, and are the direct activators and repressors of transcription in GRNs. Each

transcription factor may have several hundred to well over a thousand potential target genes (Kubo et al., 2010). Downstream effectors are terminal nodes of the network, responsible for executing cellular functions and behaviors, but not directly involved in the regulation of other genes in the network. Developmental GRNs can incorporate linear cascades and relays and also various sorts of positive and negative feedback and feedforward loops (Fig 1.1). Depending on how they are configured, these may act as switches, amplifiers, oscillators, noise filters, persistence detectors, or have other complex effects on gene expression (Davidson et al., 2002).

Given the complexity and scale of these networks, it has been helpful for developmental biologists to consider these interactions in terms of Boolean logic, in which particular combinations of factors turn expression to a fundamentally ON state, or switch expression OFF. This logical simplification has assisted developmental biology for decades, allowing researchers to build detailed models of GRNs in a variety of model systems (Koide et al., 2005, Imai et al., 2006, Oliveri et al., 2008, Ririe et al., 2008, Wilczynski and Furlong, 2009). In these Boolean terms, the particular combination of transcription factors necessary to induce the expression of each target gene represents input into a logic function, which could function on AND logic (factor X and factor Y must both be present for induction), OR logic (factor X or factor Y must be present), NOT logic (factor X must not be present), or any combination of a variety of different transcription factors and possible logical functions (Li and Davidson, 2009, Wilczynski and Furlong, 2009) (Fig 1.1).

## Enhancers

The integration of transcription factor inputs on the expression of a particular target gene occurs at enhancers. Enhancers are noncoding *cis*-regulatory elements that contain clusters of

sequence-specific transcription factor binding sites (TFBSs), recruiting the respective transcription factors that mediate the induction or repression of their target genes. Different combinations of TFBSs can mediate different expression patterns, but the underlying *cis*-regulatory logic is often elusive. TFBSs are short sequences, usually 6-10 bp in length, with each transcription factor having a particular sequence or set of sequences that it can specifically bind (Khan et al., 2018). The overall length of enhancers is highly variable and can range from as little as a few dozen base pairs (Blackwood and Kadonaga, 1998) to several kb (Benton et al., 2019).

Despite their functional importance, enhancers do not have a fixed rule to describe their location relative to the transcription start site (TSS) of the gene that they regulate. Although they are commonly found between the TSS and the next gene upstream, examples of enhancers in locations other than the immediate upstream region abound. They can be megabases away from the TSS in either direction, in introns, or on the far side of other coding genes (Schoenfelder and Fraser, 2019). Some data even suggest that interchromosomal interactions between DNase I hypersensitive sites, a signature of enhancers, could affect transcription (Spilianakis et al., 2005). While the exact distance between the enhancer and the promoter has been shown to sometimes have a moderate effect on transcriptional output (Scholes et al., 2019), enhancers typically mediate normal cell type-specific expression when their distance or orientation to the TSS has been experimentally changed. This has proven useful in detailed studies of enhancer function (See Identification of Enhancers section of this chapter).

Enhancers do not have a one to one relationship with the genes that they regulate. Enhancers vastly outnumber coding sequences (Zhu et al., 2013, Madgwick et al., 2019), with estimates reaching as high as ten to one (Shen et al., 2012). Several well-studied cases have

established that discrete expression domains of a given gene are often controlled by individual enhancers, such as the individual stripes of *Drosophila* pair-rule genes like *even-skipped* (Macdonald et al., 1986, Frasch et al., 1987, Goto et al., 1989, Harding et al., 1989, Fukioka et al., 1999). Many genes also have multiple distinct enhancers that regulate common or highly similar expression patterns (Barolo, 2012, See Distributed Enhancers section of this chapter). Enhancers can even act promiscuously, regulating the expression of multiple target genes (Fukaya et al., 2016).

### Molecular mechanisms of enhancers

Enhancer activity relies on the sequence-specific binding of transcription factors to their TFBSs to mediate their function. Because the tight packing of chromatin hides much of the sequence information, the chromatin containing an enhancer may need to be 'opened' by pioneer transcription factors that recruit protein complexes to modify histones and allow a second wave of transcription factors to bind to the DNA (Mayran and Drouin, 2018). The second wave of transcription factors can mediate a variety of functions to further control gene expression as activators or repressors of transcription. Transcriptional activators have been shown to assist in the formation of the preinitiation complex at the basal promoter via binding to transcriptional machinery and the Mediator complex (Jeronimo and Robert, 2017), and transcriptional activators can also release 'poised' RNA-polymerase from the promoter to trigger elongation (Gaertner and Zeitlinger, 2014). Transcriptional repressors have been shown to function through inhibitory interactions with the transcriptional machinery to block it from transcribing the DNA, or through remodeling of chromatin to a closed state (Gaston and Jayaraman, 2003).

Transcription factors generally possess two domains, one for DNA binding, and one which mediates their transcriptional activity. The DNA binding domains are commonly used to identify and classify transcription factors, as they tend to have much higher conservation than activation domains (Staby et al., 2017). Decades of research into binding domains and the sequences that they recognize has resulted in extensive databases of position weight matrices representing the binding preferences of many transcription factors (Matys et al., 2006, Khan et al., 2018). Transcription factors generally have somewhat degenerate binding and are capable of recognizing a variety of sequences with different affinities (Khan et al., 2018). While many studies have been interested in identifying the very highest affinity interactions, sub-optimal TFBSs have been shown to be critical to enhancer function, striking a balance between robust expression in the desired tissue while avoiding ectopic expression (Farley et al., 2015). Orientation and spacing of TFBSs can also have an impact on the function of the enhancer as a whole, which is likely a result of interactions between the TFs themselves when bound to the DNA (Muhlethaler-Mottet et al., 2004, Kulkarni and Arnosti, 2005, Swanson et al., 2010, Farley et al., 2015, Farley et al., 2016).

The presence or absence of a particular TFBS in an enhancer provides a simple explanation for aspects of the Boolean logic commonly used to describe GRNs. Where edges connect transcription factors to target genes, there should be a necessary TFBS in an enhancer for that gene to mediate the direct effect. That said, most TFBSs are short, commonly occurring sequences that are widespread across genomes and not unique to functional enhancers. It is generally unclear why some combinations of TFBSs have enhancer activity and others do not. This might reflect aspects of enhancer 'grammar' involving TFBS spacing, order, affinity and

orientation that are more complex than their simple presence or absence (Farley et al. 2015, Farley et al., 2016).

Until recently, the mechanism by which transcription factor activation domains function to promote transcription had been a mystery. While the DNA-binding domains of transcription factors have clear sequence conservation with related transcription factors, activation domains typically consist of intrinsically disordered regions (IDRs) that lack a consistent molecular structure or sequence conservation (Staby et al., 2017). Recent studies using live imaging suggest that enhancers help to establish phase-separated liquid-liquid droplets nucleated by these intrinsically disordered activation domains, where the clustering of TFBSs helps to promote locally high concentrations of TFs. This model suggests that these phase-separated TF activation domains interact nonspecifically with the Mediator transcriptional coactivation complex and potentially other coactivators to establish a semi-stable structure that can allow transcription to initiate (Boija et al., 2018, Zamudio et al., 2019).

**Identification of enhancers**

Given the great importance of enhancers as mediators of spatiotemporally specific patterns of gene expression, researchers have put extensive effort into identifying enhancers in a variety of organisms. The gold-standard approach is to use reporter assays in which a putative enhancer sequence is placed upstream of an exogenous coding sequence such as LacZ or GFP and reintroduced into the organism either transiently or through the development of a stable transgenic line. Through repeated dissection and mutation of different regions of putative enhancer sequences, precise maps of enhancers can be established and critical binding sites can be identified (Katikala et al., 2013). Reporter assays are powerful but they are also slow,

expensive and low-throughput, making large-scale unbiased searches of noncoding DNA for enhancers impractical. While individual reporter assays are commonly used as a validation step, a variety of approaches have been developed to screen genomes for possible enhancer sequences.

Unlike protein coding sequences, where open reading frames and protein sequence conservation provide key sequence-based signatures to search for genes, enhancers lack a 'smoking gun' signature in the genome. The lack of a fixed distance or orientation to the TSS of the target gene prevents either criterion from being used in isolation to identify enhancers (Schoenfelder and Fraser, 2019). Searching for clusters of TFBSs can help narrow down lists of putative enhancer regions, and numerous algorithmic approaches have been developed to query putative enhancers (Hashim et al., 2019), but the great abundance of known TFs and the promiscuity of TFs in binding to multiple TFBSs produces many false positive matches of TFBSs in putative enhancer sequences (Khamis et al., 2018). Raising the stringency of the search to require best-quality matches can assist in reducing the number of irrelevant hits, as can eliminating matching TFBSs for TFs not present in the tissue of interest. However, increasing the stringency of the search comes with the drawback that biasing the search to only the best possible matches limits the possibility that potentially critical less than optimal TFBSs may be in use (Farley et al., 2015).

Enhancers can sometimes be identified based on sequence conservation with related species (Johnson et al., 2004, Dickel et al., 2018). However, there are several complications. Reporter assays using partially randomized synthetic enhancers indicate that the sequences in between actual TFBSs are not critical for tissue-specific expression and may not be conserved (Farley et al., 2015). Many enhancers with minimal sequence conservation, or even completely different repertoires of TFBSs, are capable of producing identical expression patterns and

morphologies (Hare et al., 2008, Wunderlich et al., 2015, Madgwick et al., 2019). Even the assumption of conserved mechanisms for related enhancers in different species is not a given, as a common ancestral enhancer can evolve to be regulated by entirely different mechanisms in related species in a process known as developmental systems drift (True and Haag, 2001).

Sequencing-based approaches have also been employed to identify enhancers. Assays for Transposase-Accessible Chromatin using sequencing (ATAC-seq, Buenrostro et al., 2013) leverage the need for active enhancers to be in an open chromatin configuration. Chromatin Immunoprecipitation followed by sequencing (ChIP-seq, Johnson et al., 2007) can be used to identify TFs bound to DNA or chromatin modification such as H3K27 acetylation (Creyghton et al., 2010) that should be enriched in enhancer regions. Massively parallel reporter assays (MPRAs) introduce thousands of barcoded reporters at once, with RNA-seq then being used to determine which sequences have enhancer activity (Melnikov et al., 2012). MPRAs are an exciting technological development, but they are currently restricted to testing very short regulatory sequences and have only been implemented using bulk tissue lysates. All of these genome-scale methods require validation and more detailed individual reporter assays in order to confirm spatiotemporal specificity and characterize specific enhancers in greater mechanistic detail.

### Super-enhancers

Some studies suggest that there is a distinct subset of enhancers dubbed 'super-enhancers' that typically involve large, clustered enhancers containing particularly high numbers of occupied TFBSs. As their name would imply, these enhancers stand out from 'normal' enhancers in several ways that make them particularly notable and powerful regulators of gene

expression. Super-enhancers are generally found directing the expression of potent regulators of cell fate, such as Oct4, Sox2, and Nanog (Hnisz et al., 2013, Whyte et al., 2013).

Super-enhancers were first defined in mouse embryonic stem cells according to a systematic classification system by Whyte et al. (2013). Regions with overlapping ChIP-seq peaks for Nanos, Oct4, and Sox2 were defined as eligible enhancers, and were clustered together if enhancers were within 12.5 kb of one another. All enhancers and enhancer clusters were rank-ordered by Med1 occupancy, as measured by ChIP-seq, to determine the transcriptional output and plotted on 0-1 normalized axes where the x-axis was the enhancer rank and the y-axis was Med1 occupancy. Where the slope of the line was greater than 1, the corresponding enhancers were defined as super-enhancers. Subsequent studies have expanded the definition and used different master regulators in different cell types as the enhancer markers (reviewed by Pott and Lieb, 2015). Clustered enhancers are a common feature of many of the sequences defined as super-enhancers, but clusters of discrete enhancer modules are neither required for nor exclusive to super-enhancers. Similarly clustered enhancers have been found that do not possess the Med1 enrichment necessary to qualify as super-enhancers, while a not insignificant percentage of super-enhancers are singletons (Whyte et al., 2013, Pott and Lieb, 2015).

The localization of super-enhancers near key regulatory genes, as well as the enrichment of TFBSs for key or master regulatory TFs within super-enhancers appears to help maintain critical cell identities (Pott and Lieb, 2015, Peng and Zhang, 2018). Extensive cooperativity both between TFBSs within individual enhancers of the super-enhancer and between the enhancers of the super-enhancer (Hnisz et al., 2013, Whyte et al., 2013, Lovén et al., 2013) may also make transcriptional responses more Boolean. This is discussed in greater detail in the GRF section of this chapter.

The definition of what is and is not a super-enhancer is, however, quite arbitrary. Defining super-enhancers purely by clustering neglects the cases in which researchers have found clustered enhancers, but deemed them to not be super-enhancers, including the original identification of super-enhancers (Whyte et al., 2013). A clustering-based definition is also difficult to apply across metazoan genomes, as organisms with more compact genomes will necessarily have to cluster their enhancers more, such as in *Ciona robusta*, where gene density is estimated to be about 13 times higher than in humans (Dehal et al., 2002) and the super-enhancer concept has not been applied. Further, the enrichment cutoff has no particular biological basis, and only serves to identify enhancers with the very highest levels of marker enrichment values (Pott and Lieb, 2015). Although a useful category for driving research in enhancer function and identification, it should be remembered that the distinctions allowing an enhancer to be classified as a super-enhancer are made by placing a semi-arbitrary threshold on a fundamentally continuous set of distributions, and that this definition is therefore not an absolute.

### Distributed enhancers

The term 'shadow enhancer' originally comes from a specific set of requirements in which a newly-discovered enhancer drives expression of a given gene in the same tissues as a known enhancer, but that this newly-discovered enhancer was on the other side, or 'in the shadow' of another gene (Barolo, 2012). The term has since been coopted, with some objections, to more broadly describe all cases in which multiple enhancers are capable of promoting transcription of a given gene in the same expression pattern (Barolo, 2012). Several similar terms have also been proposed to describe similar cases of separable enhancers driving comparable expression patterns of a common gene, including 'redundant' and 'distributed' enhancers.

Although the term 'shadow' enhancer enjoys considerable popularity, it also carries connotations implying the existence of a 'main' enhancer, and it is not clear whether more proximal enhancers are always more powerful. Similarly, the term 'redundant' makes the implicit assumption that there are no functional differences at all between the two enhancers, which may not be the case. To avoid such implications, I will generally attempt to avoid the terms 'shadow' or 'redundant' in this dissertation, instead prioritizing the term 'distributed enhancers' to refer to multiple enhancers that promote the same or a highly similar expression pattern of a given gene.

The widespread existence of distributed enhancers in a variety of genomes (Cannavò et al., 2016, Osterwalder et al., 2018, Madgwick et al., 2019) immediately poses an important question. If these enhancers were truly redundant, conventional evolutionary logic suggests that they would be under less evolutionary constraint, and therefore lost. It is thus likely that there are separable functions for distributed enhancers, placing them under selection that preserves them over time. These separable functions have been hotly debated (Barolo, 2012), and may not have a single answer. The proposed hypotheses fall into three general categories: 1) Distributed enhancers are necessary for evolvability; 2) Distributed enhancers are used as a buffer against genetic variation, adverse environmental conditions and stochastic transcriptional noise; and 3) Distributed enhancers are used to generate precise boundaries of gene expression. It should be noted that these hypotheses are not mutually exclusive, nor should it be expected that all cases of distributed enhancers adhere to the same reasons (i.e., distributed enhancers for gene X might be used as a buffer, while distributed enhancers for gene Y might be used for precision of expression boundaries).

The evolvability hypothesis for distributed enhancers rests on the logic that if two or more enhancers drive truly redundant gene expression patterns, only one such enhancer would be

required to keep that expression pattern, leaving the remaining enhancers free to mutate and take on new functions, leading to the evolution of new or altered regulatory pathways (Hong et al., 2008, Wunderlich et al., 2015, Letelier et al., 2018). This hypothesis is challenged, however, by findings that distributed enhancers can actually have greater evolutionary constraint than enhancers that act alone in regulating their associated gene (Cannavò et al., 2016). The evolvability hypothesis also rests on the assumption that these enhancers are truly redundant, which is a concept with dubious support (Barolo, 2012).

Contrary to the ideal and stable environments in which laboratory model organisms are generally kept, organisms in nature are subjected to a wide range of variability in their environmental conditions, to say nothing of the variability in their own genomes, which may or may not interfere with proper gene expression in embryonic development. The buffer hypothesis for distributed enhancers posits that in order for embryos to develop properly in the face of transcriptional/molecular noise, environmental perturbation or mutation, enhancers that might be functionally redundant in ideal conditions would prove nonredundant in adverse conditions. When an adverse condition might threaten the embryo, distributed enhancers could protect the expression of critical developmental genes through additive effects or through responding to slightly different input factors (Dunipace et al., 2019). Considerable evidence exists to support this hypothesis, with examples in which researchers found that deletion of distributed enhancers only produced phenotypes in temperature-stressed *Drosophila* embryos (Frankel et al., 2010), or when the enhancer deletions were placed in the context of sensitized genetic backgrounds (Antosova et al., 2016, Osterwalder et al., 2018). Aspects of the buffer hypothesis will also be addressed further in Chapter 4.

Limiting ectopic expression is a critical concern in embryonic development, especially concerning master regulators and other key transcription factors where ectopic expression could have a catastrophic effect. Distributed enhancers have been proposed to help define sharp boundaries of expression, either directly through repression in ectopic tissues (Perry et al., 2011, El-Sherif and Levine, 2016), or through the additive effects of individually weak but highly tissue specific enhancers (Farley et al., 2015). Enhancer function in the precision of expression patterns is discussed further in Chapter 3.

# Gene Regulatory Functions

Boolean logic is quite useful in modeling the inductive events in GRNs, but is ultimately a simplification of the far more graded signals, responses, and interactions that occur in real biological systems. The *Drosophila* embryo initially patterns its anterior-posterior axis with opposing gradients of the transcription factors *bicoid* (Driever and Nüsslein-Volhard, 1988) and *nanos* (Wang and Lehmann, 1991) in the syncytial blastoderm, eventually leading to extraordinarily detailed and tightly regulated boundaries of gene expression to define precise segments within the developing organism (Nüsslein-Volhard and Weischaus, 1980, Schupbach and Weischaus, 1986). Different gap genes acting downstream of Bicoid and Nanos are activated and/or repressed by specific concentrations of these gradients and not by their simple presence vs absence. Similarly, intercellular signaling molecules, including Wnt (Reviewed by Yang, 2012), FGF (Reviewed by Balasubramanian and Zhang, 2016), and Hedgehog (Reviewed by Varjosalo and Taipale, 2008) can have variable effects on gene expression and fate choices in their target tissues based on the local concentration of the signaling factor, itself at least partially a function of the distance between the source cell and target cell, providing an elegant regulation mechanism for the embryo.

That different molecular concentrations of transcription factors or signaling molecules can produce different transcriptional responses suggests that simple Boolean logic that describes gene expression in fundamental ON or OFF terms cannot fully describe the integration of inputs on an enhancer. A more complete understanding requires the incorporation of quantitative information about how varying concentrations of upstream factors influence the expression of the regulated gene. Quantitative dose-response relationships between the activity or

concentration of upstream regulators and the expression of a gene of interest are known as Gene Regulatory Functions (GRFs) (Rosenfeld et al., 2005).

GRFs are inherently sigmoidal relationships, and can be modeled using Hill equations, in which the concentration of a transcription factor or signaling molecule (the dose) is on the x-axis and the transcriptional output (the response) is on the y-axis (Bost and Veitia, 2013, Veitia, 2018) (Figure 1.2). For any given enhancer, as the concentration of transcriptional activators increases, so too will the transcriptional output, but it will do so nonlinearly. At very low concentrations, output will generally be minimal, and at a high enough concentration, the enhancer would become 'saturated,' already inducing a maximum output of transcription and unable to increase it any further, regardless of increases in TF concentration. In between these plateaus is one (or more) range of concentrations in which the transcriptional output increases dramatically.

The two key properties of this transition zone are the slope and the $EC_{50}$. The slope is measured as a Hill Coefficient, and is a generally accepted implicit measure of the cooperativity of TF binding to the enhancer, with higher Hill Coefficients indicating greater cooperativity. If the Hill Coefficient is high enough, this transition zone in the input/output relationship will be particularly steep, and will closely approximate a Boolean ON/OFF response. Lower Hill Coefficients indicate much shallower transition zones (Veitia, 2018) (Fig 1.2A). The $EC_{50}$ is the point on the x-axis at which the transcriptional response is halfway between its maximum and minimum plateaus (Fig 1.2B), and is thus a direct measure of the concentration of input factor needed to elicit a half-maximal response.

Both the Hill Coefficient and the $EC_{50}$ are critical parameters for understanding the regulation of gene expression. The Hill Coefficient represents the extent to which an enhancer

controls bistable, switch-like responses versus graded sensor-like responses (Veitia, 2018). The $EC_{50}$ represents the actual threshold concentration of active upstream factor needed to 'flip the switch' and is critical to how morphogen gradients and other patterns of upstream factors lead to exact spatial boundaries of gene expression (Veitia, 2018). The quantitative details of enhancer sensitivity and cooperativity are also likely to explain interesting genetic phenomena like the haploinsufficient dominance seen for many human transcription factors. If gene regulatory networks depend on threshold levels of active transcription factors to activate specific enhancers, a ~50% decrease from losing one allele may be enough to drop its effective level below the $EC_{50}$ (Bost and Veitia, 2013).

Despite their importance for understanding developmental regulatory networks, very few GRFs have been quantified in developing embryos. This is because it is generally difficult to systematically titrate the concentration of upstream transcription factors. Some work has begun to quantify input/output relationships in *Drosophila* enhancers, taking advantage of the early *Drosophila* embryo's *bicoid* gradient, in which the anterior-posterior (AP) position can be used to infer the local TF concentration. These studies have explored a variety of deeply quantitative transcriptional questions, including the establishment of sharp transcriptional boundaries (Perry et al., 2011, El-Sherif and Levine, 2016), additivity functions of multiple enhancers (Bothma et al., 2015, Scholes et al., 2019), and cooperativity of TF-enhancer interactions (Park et al., 2019). These studies have provided new insight into the molecular mechanisms that underlie GRFs, but it is unclear to what extent these insights are applicable to other organisms using fundamentally different mechanisms of embryonic cell fate specification.

## *Ciona robusta* as a Model Organism

### *Ciona* History and Embryology

Tunicates are the closest invertebrate outgrouping of the vertebrates (Delsuc et al., 2006), and were used in some of the earliest studies of developmental biology in the late 1800s and early 1900s (Satoh, 2014). Although they fell out of favor for much of the 1900s, they experienced a resurgence in the 1990s due to emerging molecular tools and several key advantages for studying developmental biology. This resurgence began with research groups in Japan studying tunicates from the genus *Halocynthia*, but has since expanded to include a variety of different tunicate species. Among the most popular is *Ciona robusta* (formerly *Ciona intestinalis* Type A) (Pennati et al., 2015) (Fig 1.3A-B).

Adult *Ciona* are sessile filter-feeding hermaphrodites with cosmopolitan distribution throughout the world's oceans. Although the adults bear little resemblance to their vertebrate cousins, their embryos and larvae are stereotypically chordate, with a hollow dorsal neural tube, muscular tail, and a notochord (Fig 1.3B). Fertilization occurs externally, with sperm and eggs spawned into the environment shortly after sunrise, and hatching occurring approximately 24 hours later (Hotta et al., 2007). *Ciona* embryos progress rapidly through invariant, simple, and stereotyped development. The first cleavage divides the left and right halves of the embryo, with anterior-posterior and animal-vegetal axes defined by the second and third divisions, respectively. By the time of gastrulation at the 110-cell stage, almost all cells have become fate-restricted to a single tissue type (Hotta et al., 2007).

Cell fate decisions in *Ciona* are invariant and stereotyped, and are made through a combination of maternal factors and cell signaling. β-catenin is deposited as a maternal determinant into the egg, and brought to the vegetal pole after fertilization via cortical

rearrangements (Imai et al., 2000, Nishida, 2005). GATAa is also maternally deposited, but its activity is restricted to the animal pole by β-catenin (Rothbächer et al., 2007). Inductive events in *Ciona* development are common, but in contrast to the long-range gradients present in vertebrate development, these occur via direct cell-cell contacts with specific cellular responses being dependent on relative surface areas of contact regions (Tassy et al., 2006).

In addition to the external fertilization and rapid invariant development, *Ciona* embryos possess several additional advantages for use in research. At ~140 µm in diameter (Gregory and Veeman, 2013), the embryos are large enough for manipulations such as microsurgery or injection, but small enough that the entire embryo can be imaged at high resolution in a single field of view (Veeman and Reeves, 2015). Large-scale experiments are relatively simple, as thousands of synchronized embryos can be obtained quickly through dissection of adults for sperm and eggs, chemically dechorionated (Mita-Miyazawa et al., 1985) and transgenesis can be performed by simple electroporation (Corbo et al., 1997). *Ciona* adults are somewhat self-fertile, but do have a genetic-self incompatibility system that reduces the efficiency of successfully selfing adults (Harada et al., 2008). While this can reduce some of the labor involved in maintaining stable transgenic lines, it is much more common for researchers to use wild-caught *Ciona*.

## The Notochord

As the eponymous organ of the chordate phylum, the notochord is a useful organ to study for both functional and evolutionary reasons. Although vertebrates make extensive use of the notochord for embryonic patterning, most famously in the induction of the floor plate on the ventral side of the neural tube (Yamada et al., 1991, Placzek et al., 1993), a signaling role for the

notochord in *Ciona* has not been established (Passamaneck and Di Gregorio, 2005). During development, the extension of the tail is driven primarily through the convergent extension of the notochord (Nakatani et al., 1999), and the notochord acts as the hydrostatic skeleton of the larval tail, allowing the larva to swim before finding a substrate suitable for attachment and metamorphosis into the adult form. The notochord is lost during metamorphosis, most likely through apoptotic mechanisms (Chambon et al., 2002). The notochord of *Ciona*, with only forty cells, is comparatively simpler than the notochord of vertebrates, with several hundred or more cells even in smaller vertebrates such as zebrafish (Passamaneck and Di Gregorio, 2005). These traits make the *Ciona* notochord an attractive model for studying the mechanisms of notochord formation and gene regulation in great detail.

In *Ciona* and other ascidians, the notochord is derived from two separate developmental lineages, a 32-cell 'primary' notochord derived from the A7.3 and A7.7 blastomeres, and an 8-cell 'secondary' notochord derived from the B8.6 blastomeres. These two lineages form a contiguous single-file rod, with the primary notochord cells taking the anterior 32 positions, and the secondary notochord cells taking the posterior 8 (Nishida, 1987), and there are some differences in how these fates are induced. Primary notochord fate is induced by Fibroblast Growth Factor (FGF) 9/16/20 from the endoderm, starting at the 32-cell stage (Nakatani and Nishida, 1994, Yasuo and Hudson, 2007). This signal polarizes the A6.2 and A6.4 blastomeres, producing an asymmetric division in which the more anterior A7.4 and A7.8 daughters adopt an A-line neural fate, and the more posterior A7.3 and A7.7 daughters become primary notochord (Yasuo and Hudson, 2007). The FGF signal is absolutely required for notochord fate, as removal or interference of this pathway results in both sets of daughter cells adopting neural fates (Minokawa et al., 2001, Yasuo and Hudson, 2007). The FGF signaling leads to activation of Ets

19

family transcription factors, which together with the transcription factor Zic-r.b (formerly ZicL) are necessary for inducing the expression of *Brachyury* at the 64-cell stage (Imai et al., 2002, Imai et al., 2006). *Brachyury* is responsible for the induction of a large number of notochord genes (Takahashi et al., 1999, Hotta et al., 1999, Kubo et al., 2010, Katikala et al., 2013, Reeves et al., 2020), and has long been presumed to be a master regulator of notochord fate (see Chapter 2, however, for a discussion of this topic). Secondary notochord induction is slightly different from primary notochord induction. Nodal signaling from b6.5 induces expression of Delta2 in the A6.3 lineage, whose daughter cell A7.6 also expresses Delta2 and signals to B7.3, which divides asymmetrically to produce the mesenchyme precursor B8.5 and the secondary notochord precursor B8.6 (Hudson and Yasuo, 2006). Blastomere names and positions are shown in Fig. 1.4

At specification, there are 10 notochord precursor cells in total (8 primary, 2 secondary). These cells each undergo two rounds of cell division to produce a 40-cell notochord primordium, before this primordium undergoes a dramatic intercalation event, in which the 40 cells rearrange from a flat, heart-shaped sheet into a single-file rod centrally located in the extending tail, partially concurrent with the embryo's neurulation (Munro et al., 2006). Following intercalation, the notochord continues to undergo dramatic shape changes, as the disc-shaped cells further constrict their diameters and elongate along their AP axis, narrowing and extending the tail (Veeman and Smith, 2013). Closer to hatching, notochord cells secrete a lumen into the space between each cell, then rearrange as these spaces merge into a single compartment spanning the length of the tail to produce the hydrostatic skeleton (Dong et al., 2009).

The dramatic shape changes that occur during notochord development are highly dependent upon several tightly controlled processes. The final tapered shape of the intercalated

notochord, in which the anterior and posterior ends are both much narrower than the middle, is initially set up during the two rounds of cell division that convert the initial ten notochord precursor cells into the 40-cell primordium, which are asymmetric with respect to the size of the daughter cells (Winkley et al., 2019). Intercalation of the notochord depends on both an intact notochordal sheath and polarity cues within the notochord that bias the intercalation in the mediolateral direction (Veeman et al., 2008). The neighbor exchanges during mediolateral intercalation of the notochord help to distribute the asymmetrically sized precursors (Carlson et al., 2015), providing an initial level of tapering that is further expanded due to the narrowing and elongating of the disc-shaped cells, and temporal differences in when cells finish their intercalation and undergo their shape change (Veeman and Smith, 2013).

The active and dynamic processes that shape the notochord rely on the proper expression of a vast array of genes. Coupled with the interest in the *Ciona* notochord as a model for vertebrate organ formation and morphogenesis, this has led to numerous studies that have explored the expression patterns and regulatory mechanisms of a wide variety of notochord genes (Hotta et al., 1999, Takahashi et al., 1999, Hotta et al., 2000, Hotta et al., 2008, Kugler et al., 2008, Passamaneck et al., 2009, Kubo et al., 2010, José-Edwards et al., 2011, José-Edwards et al., 2013, Katikala et al., 2013, Reeves et al., 2014, José-Edwards et al., 2015, Kugler et al., 2019, Reeves et al., 2020). Many long-held assumptions of the *Ciona* notochord regulatory network have been challenged in recent years, as recent studies have shown some genes in the notochord to have decidedly non-uniform expression (see Chapter 3, Reeves et al., 2014), and questioned the assumption that *Brachyury* is a true master regulator of the notochord (see Chapter 2, Reeves et al., 2020). Together with the multiple layers of regulation involved in

establishing notochord shape, the case is building that the *Ciona* notochord is much more finely patterned than was initially suspected.

## *Ciona* Genome and Evolution

The Tunicates diverged from vertebrates approximately 520 million years ago, prior to a genome duplication event in the vertebrate lineage (Reviewed by Lemaire, 2011). With approximately 15,000 genes, there are considerably fewer instances of redundant proteins in *Ciona* than in vertebrates (Dehal et al., 2002), greatly simplifying molecular research, especially work involving loss-of-function experiments. *Ciona* genetics are complicated, however, by an extreme degree of polymorphism, as much as 1.5% (Kim et al., 2007).

Tunicates diverged from vertebrates prior to several genome duplication events in the vertebrate lineage, and *Ciona* therefore lacks the large groups of paralogous genes present in the vertebrates (Passamaneck and Di Gregorio, 2005). The overall size of the *Ciona* genome is also considerably smaller, roughly 1/20$^{th}$ the size of the human genome (Dehal et al., 2002). This compactness of the *Ciona* genome has also proven beneficial in the study of noncoding regulatory DNA in *Ciona*. With more genes and coding DNA per Mb, there is comparatively less noncoding DNA, simplifying the search for *bona fide* enhancer regions. Additionally, enhancers in *Ciona* are also generally quite close to the TSS, with the majority of known enhancers located less than 2 kb upstream of the TSS or in introns (Irvine, 2013, Madgwick et al., 2019). Together, these traits greatly simplify the discovery and functional analysis of *Ciona* enhancers.


## FGF Signaling in *Ciona*

FGF signaling follows a classic receptor tyrosine kinase pathway model (Eswarakumar et al., 2005). Briefly, the FGF ligand is secreted from a signaling cell and travels to the receiving

cell via diffusion, which may sometimes involve mechanisms, such as Heparan sulfate proteoglycan binding, which can modulate the rate and therefore distance of diffusion (Müller and Schier, 2011). Upon reaching the receiving cell, the ligand forms a complex with the FGF receptor that dimerizes the receptor, causing the intracellular domains of the receptors to phosphorylate and activate one another (Lemmon and Schlessinger, 2010). The activated FGF receptors proceed to act on their intracellular second messengers through several possible pathways, but canonically this is the MAPK pathway, in which the FGFR complex phosphorylates Ras, which phosphorylates Raf, which phosphorylates MEK, which phosphorylates ERK (Morrison, 2012). Phosphorylated ERK is the terminal kinase in this cascade, which in turn activates various transcription factors by phosphorylation (Morrison, 2012). In both *Ciona* and other species these are commonly members of the Erythroblast Transformation Specific (Ets) family (Sharrocks, 2001, Khoueiry et al., 2010).

In *Ciona*, FGF signaling is used in the induction or maintenance of numerous tissues. Notable induction events early in the embryo include induction of notochord fate, mesenchyme fate, and a-line neural fate. These tissues are induced simultaneously, and by the same FGF signal, FGF9/16/20. The differentiator between these three tissues comes from determinant factors present only in the specific receiving cells; FoxAa and Zic-r.b for notochord, Macho-1 for the muscle/mesenchyme precursors, and GATA for a-line neural (Reviewed by Lemaire et al., 2008). The FGF signal comes from the vegetal blastomeres of the embryo, where the expression of FGF was induced by maternal β-catenin (Bertrand et al., 2003).

In the notochord, the initial step of induction of *Brachyury* at the 64-cell stage by FGF9/16/20 is not, by itself, enough to fully induce stable notochord fate. Before *Brachyury* is fully capable of maintaining its own expression, further signaling of FGF9/16/20 and

FGF8/18/19 must continue in order to maintain *Brachyury* transcription and fully activate the notochord GRN (Yasuo and Hudson, 2007) and produce a functional notochord.

## Dissertation Overview

In this dissertation, I take advantage of the simple, stereotyped morphology, rapid transgenesis, and compact genome of the simple ascidian chordate, *Ciona robusta*, to address quantitative questions about the regulation of gene expression. In Chapter 3, the relatively simple anatomy of the *Ciona* notochord enabled detailed quantification of regionalized gene expression patterns with single-cell precision. In Chapter 4, the large numbers of embryos that could be obtained allowed for systematic titration of upstream signaling factors in order to model GRFs mediated by a pair of quantitatively distinct distributed enhancers.

The generation and validation of the *Ciona* notochord transcriptome (Reeves et al., 2017, see Chapter 2) has enabled identification of a number of genes that had nonuniform AP patterning of their expression within the notochord. The most striking of these was a pattern in which gene expression was highly enriched in the secondary notochord, but not detectable in the primary notochord. We found evidence that this expression pattern was dependent on regulation by both Wnt and FGF signaling, and extensively dissected the upstream DNA to identify a minimal enhancer mediating this expression. To our surprise, we found multiple, non overlapping regions capable of producing notochord expression with specific enrichment in the secondary notochord (See Chapter 3).

With mounting evidence of distributed enhancers in the *Ciona* genome (See Chapter 3, Farley et al., 2016, Madgwick et al., 2019), we selected a pair of distributed enhancers of *Brachyury*, and titrated the input of FGF via disruption of the MEK signal transduction pathway to identify quantitative differences between them, establishing that these are fundamentally not

redundant genetic elements. Fine-scale quantitative assessment reveals that the expression loss

due to the disruption of the MEK signal transduction pathway is due to both total loss of

expression in some cells and graded loss in others. Endogenous *Brachyury* expression under

similar dose-response conditions revealed different sensitivities of the notochord founder cells to

FGF inhibition while also validating our reporter assay results.

# References

Antosova, B., Smolikova, J., Klimova, L., Lachova, J., Bendova, M., Kozmikova, I., Macon, O., Kozmik, Z., 2016. The Gene Regulatory Network of Lens Induction Is Wired through Meis-Dependent Shadow Enhancers of Pax6. PLoS Genetics 12, e1006441.

Balasubramanian, R., Zhang, X., 2016. Mechanisms of FGF gradient formation during embryogenesis. Sem. in Cell and Dev. Biol. 53, 94-100.

Barolo, S., 2012. Shadow enhancers: Frequently asked questions about distributed cis-regulatory information and enhancer redundancy. BioEssays 34, 135-141.

Benton, M.L., Talipineni, S.C., Kostka, D., Capra, J.A., 2019. Genome-wide enhancer annotations differ significantly in genomic distribution, evolution, and function. BMC Genomics 20, 511

Bertrand, V., Hudson, C., Caillol, D., Popovici, C., Lemaire, P., 2003. Neural Tissue in Ascidian Embryos Is Induced by FGF9/16/20, Acting via a Combination of Maternal GATA and Ets Transcription Factors. Cell 115, 615-627.

Blackwood, E.M., Kadonaga, J.T., 1998. Going the Distance: A Current View of Enhancer Action. Science 281, 60-63.

Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Mantegna, J.C., Hannett, N.M., Abraham, B.J., Afeyan, L.K., Guo, Y.E., Rimel, J.K., Fant, C.B., Schuijers, J., Lee, T.I., Taatjes, D.J., Young, R.A., 2018. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell 175, 1842-1855.

Bost, B., Veitia, R.A., 2013. Dominance and interloci interactions in transcriptional activation cascades: Models explaining compensatory mutations and inheritance patterns. Bioessays 36, 84-92.

Bothma, J.P., Garcia, H.G., Ng, S., Perry, M.W., Gregor, T., Levine, M., 2015. Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo. eLife 4, e07956.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., Greenleaf, W.J., 2013. Transposition of native chromatin for multimodal regulatory analysis and personal epigenomics. Nat. Meth. 10, 1213-1218.

Cannavò, E., Khoueiry, P., Garfield, D.A., Geeleher, P., Zichner, T., Gustafson, E.H., Ciglar, L., Kornberg, J.O., Furlong, E.E.M., 2016. Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks. Curr. Biol. 26, 38-51.

Carlson, M., Reeves, W., Veeman, M., 2015. Stochasticity and stereotypy in the Ciona notochord. Dev. Biol. 397, 248-256.

Chambon, J.-P., Soule, J., Pomies, P., Fort, P., Sahuquet, A., Alexandre, D., Mangeat, P.-H., Baghdiguian, S., 2002. Tail regression in Ciona intestinalis (Prochordate) involves a Caspase-dependent apoptosis event associated with ERK activation. Development 129, 3105-3114.

Corbo, J., Levine, M., Zeller, R., 1997. Characterization of a notochord-specific enhancer from the Brachyury promoter region of the ascidian, Ciona intestinalis. Development 124, 589-602.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., Boyer, L.A., Young, R.A., Jaenisch, R., 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. PNAS 107, 21931-21936.

Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.-H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., Otis, O., Brown, C.T., Livi, C.B., Lee, P.Y., Revila, R., Rust, A.G., jun Pan, Z., Schilstra, M.J., Clarke, P.J.C., Arnone, M.I., Rowen, L., Cameron, R.A., McClay, D.R., Hood, L., Bolouri, H., 2002. A Genomic Regulatory Network for Development. Science 295, 1669-1678.

Dehal, P., Satou, Y., Campbell, R.K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D.M., Harafuji, N., Hastings, K.E.M., Ho, I., Hotta, K., Huang, W., Kawashima, T., Lemaire, P., Martinez, D., Meinertzhagen, I.A., Necula, S., Nonaka, M., Putnam, N., Rash, S., Saiga, H., Satake, M., Terry, A., Yamada, L., Wang, H.-G., Awazu, S., Azuma, K., Boore, J., Brando, M., Chin-bow, S., DeSantis, R., Doyle, S., Francine, P., Keys, D.N., Haag, S., Hayashi, H., Hino, K., Imai, K.S., Inaba, K., Kano, S., Kobayashi, K., Kobayashi, M., Lee, B.-I., Makabe, K.W., Manohar, C., Matisse, G., Medina, M., Mochizuki, Y., Mount, S., Morishita, T., Maura, S., Nakayama, A., Nishizaka, S., Nomoto, H., Ohta, F., Oisin, K., Rigoutsos, I., Sano, M., Sasaki, A., Sasakura, Y., Shoguchi, E., Shin-i, T., Spagnuolo, A., Stainier, D., Suzuki, M.M., Tassy, O., Takatori, N., Tokuoka, M., Yagi, K., Yoshizaki, F., Wanda, S., Zhang, C., Hyatt, P.D., Larimer, F., Detter, C., Doggett, N., Glavina, T., Hawkins, T., Richardson, P., Lucas, S., Kohara, Y., Levine, M., Satoh, N., Rokhsar, D.S., 2002. The Draft Genome of Ciona intestinalis: Insights into Chordate and Vertebrate Origins. Science 298, 2157-2167.

Delsuc, F., Brinkmann, H., Chourrout, D., Philippe, H., 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature 439, 965-968.

Dickel, D.E., Ypsilanti, A.R., Pla, R., Zhu, Y., Barozzi, I., Mannion, B.J., Khin, Y.S., Fukuda-Yuzawa, Y., Placzek-Frick, I., Pickle, C.S., Lee, E.A., Harrington, A.N., Pham, Q.T., Garvin, T.H., Kate, M., Osterwalder, M., Akiyama, J.A., Afzal, V., Rubinstein, J.L.R.,

Pennacchio, L.A., Visel, A., 2018. Ultraconserved Enhancers Are Required for Normal Development. Cell 172, 491-499.

Dong, B., Horie, T., Denver, E., Kusakabe, T., Tsuda, M., Smith, W.C., Jiang, D., 2009. Tube formation by complex cellular processes in Ciona intestinalis notochord. Dev. Biol. 330, 237-249.

Driever, W., Nüsslein-Volhard, C., 1988. The bicoid Protein Determines Position in the Drosophila Embryo in a Concentration-Dependent Manner. Cell 54, 95-104.

Dunipace, L., Ákos, Z., Stathopoulos, A., 2019. Coacting enhancers can have complementary functions within gene regulatory networks and promote canalization. PLoS Genetics 15, e1008525.

El-Sherif, E., Levine, M., 2016. Shadow Enhancers Mediate Dynamic Shifts of Gap Gene Expression in the Drosophila Embryo. Current Biol. 26, 1164-1169.

Eswarakumar, V.P., Lax, I., Schlessinger, J., 2005. Cellular signaling by fibroblast growth factor receptors. Cytokines and Growth Factor Reviews. 16(SI), 139-149

Farley, E.K., Olson, K.M., Zhang, W., Brandt, A.J., Rokhsar, D.S., Levine, M.S., 2015. Suboptimization of developmental enhancers. Science 350, 325-328.

Farley, E.K., Olson, K.M., Zhang, W., Rokhsar, S.D., Levine, M.S., 2016. Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. PNAS 113, 6508-6513.

Frankel, N., Davis, G.K., Vargas, D., Wang, S., Payre, F., Stern, D.L., 2010. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. Nature 466, 490-493.

Frash, M., Hoey, T., Rushlow, C., Doyle, H., Levine, M., 1987. Characterization and localization of the even-skipped protein of Drosophila. The EMBO Journal 6, 749-759.

Fujioka, M., Emi-Sarker, Y., Yusibova, G.L., Goto, T., Jaynes, J.B., 1999. Analysis of an even-skipped rescue transgene reveals both composite and discrete neuronal and early blastoderm enhancers, and multi-stripe positioning by gap gene repressor gradients. Development 126, 2527-2538.

Fukaya, T., Liam, B., Levine, M., 2016. Enhancer Control of Transcriptional Bursting. Cell 166, 358-368.

Gaertner, B., Zeitlinger, J., 2014. RNA polymerase II pausing during development. Development 141, 1179-1183.

Gaston, K., Jayaraman, P.-S., 2003. Transcriptional repression in eukaryotes: repressors and repression mechanisms. Cell Mol. Life Sci. 60, 721-741.

Goto, T., Macdonald, P., Maniatis, T., 1989. Early and Late Periodic Patterns of even skipped Expression Are Controlled by Distinct Regulatory Elements That Respond to Different Spatial Cues. Cell 57, 413-422.

Gregory, C., Veeman, M., 2013. 3D-printed microwell arrays for Ciona microinjection and timelapse imaging. PLoS ONE 8, e82307.

Harada, Y., Takahagi, Y., Sunagawa, M., Saito, T., Yamada, L., Taniguchi, H., Shoguchi, E., Sawada, H., 2008. Mechanism of Self-Sterility in a Hermaphroditic Chordate. Science 320, 548-550.

Harding, K., Hoey, T., Warrior, R., Levine, M., 1989. Autoregulatory and gap gene response elements of the even-skipped promoter of Drosophila. The EMBO Journal 8, 1205-1212.

Hare, E.E., Peterson, B.K., Iyer, V.N., Meier, R., Eisen, M.B., 2008. Sepsid even-skipped Enhancers Are Funcitonally Conserved in Drosophila Despite Lack of Sequence Conservation. PLoS Genetics 4, e1000106.

Hashim, F.A., Mabrouk, M.S., Al-Atabany, W., 2019. Review of Different Sequence Motif Finding Algorithms. Avicenna J. Med. Biotech. 11, 130-148.

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

Hong, J.-W., Hendrix, D.A., Levine, M.S., 2008. Shadow Enhancers as a Source of Evolutionary Novelty. Science 321, 1314.

Hotta, K., Takahashi, H., Erives, A., Levine, M., Satoh, N., 1999. Temporal expression patterns of 39 Brachyury-downstream genes associated with notochord formation in the Ciona intestinalis embryo. Dev. Growth Diff. 41, 657-664.

Hotta, K., Takahashi, H., Asakura, T., Satoh, B., Takatori, N., Satou, Y., Satoh, N., 2000. Characterization of Brachyury-Downstream Notochord Genes in the Ciona intestinalis Embryo. Dev. Biol. 224, 69-80.

Hotta, K., Mitsuhara, K., Takahashi, H., Inaba, K., Oka, K., Gojobori, T., Ikeo, K., 2007. A Web-Based Interactive Developmental Table for the Ascidian Ciona intestinalis, Including 3D Real-Image Embryo Reconstructions: I. From Fertilized Egg to Hatching Larva. Dev. Dyn. 236, 1790-1805.

Hotta, K., Takahashi, H., Satoh, N., Gojobori, T., 2008. Brachyury-downstream gene sets in a chordate, Ciona intestinalis: integrating notochord specification, morphogenesis and chordate evolution. Evol. and Dev. 10, 37-51.

Hudson, C., Yasuo, H., 2006. A signalling relay involving Nodal and Delta ligands acts during secondary notochord induction in Ciona embryos. Development 133, 2855-2864.

Imai, K., Takada, N., Satoh, N., Satou, Y., 2000. β-catenin mediates the specification of endoderm cells in ascidian embryos. Development 127, 3009-3020.

Imai, K.S., Satou, Y., Satoh, N., 2002. Multiple functions of a Zic-like gene in the differentiation of notochord, central nervous system and muscle in Ciona savignyi embryos. Development 129, 2723-2732.

Imai, K.S., Levine, M., Satoh, N., Satou, Y., 2006. Regulatory Blueprint for a Chordate Embryo. Science 312, 1183-1187.

Irvine, S.Q., 2013. Study of Cis-regulatory Elements in the Ascidian Ciona intestinalis. Curr. Gen. 14, 56-67.

Jeronimo, C., Robert, F., 2017. The Mediator Complex: At the Nexus of RNA Polymerase II Transcription. Trends in Cell Biol. 27, 765-783.

Johnson, D.S., Davidson, B., Brown, C.D., Smith, W.C., Sidow, A., 2004. Noncoding regulatory sequences of Ciona exhibit strong correspondence between evolutionary constraint and functional importance. Gen. Res. 14, 2448-2456.

Johnson, D.S., Mortazavi, A., Myers, R.M., Wold, B., 2007. Genome-Wide Mapping of in Vivo Protein-DNA Interactions. Science 316, 1497-1502.

José-Edwards, D.S., Kerner, P., Kugler, J.E., Deng, W., Jaing, D., Di Gregorio, A., 2011. The identification of transcription factors expressed in the notochord of Ciona intestinalis adds new potential players to the Brachyury gene regulatory network. Dev. Dyn. 240, 1793-1805.

José-Edwards, D.S., Oda-Ishii, I., Nibu, Y., Di Gregorio, A., 2013. Tbx2/3 is an essential mediator within the Brachyury gene network during Ciona notochord development. Development 140, 2422-2433.

José-Edwards, D.S., Oda-Ishii, I., Kugler, J.E., Passamaneck, Y.J., Katikala, L., Nibu, Y., Di Gregorio, A., 2015. Brachyury, Foxa2 and the cis-Regulatory Origins of the Notochord. PLoS Gen. 11, 1-16.

Katikala, L., Aihara, H., Passamaneck, Y.J., Gazdoiu, S., José-Edwards, D.S., Kugler, J.E., Oda-Ishii, I., Imai, J.H., Nibu, Y., Di Gregorio, A., 2013. Functional Brachyury Binding Sites

Establish a Temporal Read-out of Gene Expression in the Ciona Notochord. PLoS Biol. 11, e1001697.

Khamis, A.M., Motwalli, O., Oliva, R., Jankovic, B.R., Medvedeva, Y.A., Ashoor, H., Essack, M., Gao, X., Bajic, V.B., 2018. A novel method for improved accuracy of transcription factor binding site prediction. Nucl. Acids Res. 46, e72.

Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J.A., van der Lee, R., Bessy, A., Chèneby, J., Kulkarni, S.R., Tan, G., Baranasic, D., Arenillas, D.J., Sandelin, A., Vandepeole, K., Lenhard, B., Ballester, B., Wasserman, W.W., Parcy, F., Mathelier, A., 2018. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nuc. Acids Res. 46(D1), D250-D266.6

Khoueiry, P., Rothbächer, U., Ohtsuka, Y., Daian, F., Frangulian, E., Roure, A., Dubchak, I., Lemaire, P., 2010. A cis-Regulatory Signature in Ascidians and Flies, Independent of Transcription Factor Binding Sites. Curr. Biol. 20, 792-802.

Kim, J.H., Waterman, M.S., Li, L.M., 2007. Diploid genome reconstruction of Ciona intestinalis and comparative analysis with Ciona savignyi. Gen. Res. 17, 1101-1110.

Koide, T., Hayata, T., Cho, K.W.Y., 2005. Xenopus as a model system to study transcriptional regulatory networks. PNAS 102, 4943-4948.

Kourakis, M.J., Smith, W.C., 2015. An organismal perspective on C. intestinalis development, origins and diversification. eLife 4, e06024.

Kubo, A., Suzuki, N., Yuan, X., Nakatani, K., Satoh, N., Imai, K.S., Satou, Y., 2010. Genomic cis-regulatory networks in the early Ciona intestinalis embryo. Development 137, 1613-1623.

Kugler, J.E., Passamaneck, Y.J., Feldman, T.G., Beh, J., Regnier, T.W., Di Gregorio, A., 2008. Evolutionary Conservation of Vertebrate Notochord Genes in the Ascidian Ciona intestinalis. Genesis 46, 697-710.

Kugler, J.E., Wu, Y., Katikala, L., Passamaneck, Y.J., Addy, J., Caballero, N., Oda-Ishii, I., Maguire, J.E., Li, R., Di Gregorio, A., 2019. Positioning a multifunctional basic helix-loop-helix transcription factor within the Ciona notochord gene regulatory network. Dev. Biol. 448, 119-135.

Kulkarni, M.M., Arnosti, D.N., 2005. Cis-Regulatory Logic of Short-Range Transcriptional Repression in Drosophila melanogaster. Mol. And Cell. Biol. 25, 3411-3420.

Lemaire, P., 2011. Evolutionary crossroads in developmental biology: the tunicates. Development 138, 2143-2152.

Lemaire, P., Smith, W.C., Nishida, H., 2008. Ascidians and the Plasticity of the Chordate Developmental Program. Curr. Biol. 18, 620-631.

Lemmon, M.A., Schlessinger, J., 2010. Cell Signaling by Receptor Tyrosine Kinases. Cell 141, 1117-1134.

Letelier, J., de la Calle-Mustienes, E., Pieretti, J., Naranjo, S., Maeso, I., Nakamura, T., Pascual-Anaya, J., Shubin, N.H., Schneider, I., Martinez-Morales, J.R., Gómez-Skarmeta, J.L., 2018. A conserved Shh cis-regulatory module highlights a common developmental origin of unpaired and paired fins. Nature Gen. 50, 504-509.

Levine, M., Davidson, E.H., 2005. Gene regulatory networks for development. PNAS 102, 4936-4942.

Li, E., Davidson, E.H., 2009. Building Developmental Gene Regulatory Networks. Brith Defects Res C Embryo Today. 87, 123-130.

Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., Young, R.A., 2013. Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers. Cell 153, 320-334.

Macdonald, P.M., Ingham, P., Struhl, G., 1986. Isolation, Structure, and Expression of even-skipped: A Second Pair-Rule Gene of Drosophila Containing a Homeo Box. Cell 47, 721-734.

Madgwick, A., Magri, M.S., Dantec, C., Gailly, D., Fiuza, U.-M., Guignard, L., Hettinger, S., Gomez-Skarmeta, J.L., Lemaire, P., 2019. Evolution of embryonic cis-regulatory landscapes between divergent Phallusia and Ciona ascidians. Dev. Biol. 448, 71-87.

Matys, V., Kel-Margoulis, O.V., Fricke, E., Liebich, I., Land, S., Barre-Dirre, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K., Voss, N., Stegmaier, P., Lewicki-Potapov, B., Saxel, H., Kel, A.E., Wingender, E., 2006. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nuc. Acids Res. 34(D1), D108-D110.

Mayran, A., Drouin, J., 2018. Pioneer transcription factors shape the epigenetic landscape. J. Biol. Chem. 293, 13795-13804.

Melnikov, A., Murugan, A., Zhang, X., Tesileanu, T., Wang, L., Rogov, P., Feizi, S., Gnirke, A., Callan Jr., C.G., Kinney, J.B., Kellis, M., Lander, E.S., Mikkelsen, T.S., 2012. Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. Nat. Biotech. 30, 271-277.

Minokawa, T., Yagi, K., Makabe, K.W., Nishida, H., 2001. Binary specification of nerve cord and notochord cell fates in ascidian embryos. Development 128, 2007-2017.

Mita-Miyazawa, I., Ikegami, S., Satoh, N., 1985. Histospecific acetylcholinesterase development in the presumptive muscle cells isolated from 16-cell-stage ascidian embryos with respect to the number of DNA replications. J. Embryol. Exp. Morph. 87, 1-12.

Morrison, D.K., 2012. MAP Kinase Pathways. Cold Spring Harb. Perspect. Biol. 4, a011254.

Muhlethaler-Mottet, A., Krawczyk, M., Masternak, K., Spilianakis, C., Kretsovali, A., Papamatheakis, J., Reith, W., 2004. The S Box of Major Histocompatibility Complex Class II Promoters Is a Key Determinant for Recruitment of the Transcriptional Co-activator CIITA*. J. Biol. Chem. 279, 40529-40535.

Müller, P., Schier, A.F., 2011. Extracellular Movement of Signaling Molecules. Dev. Cell 21, 145-158.

Munro, E., Robin, F., Lemaire, P., 2006. Cellular morphogenesis in ascidians: how to shape a simple tadpole. Curr. Op. in Gen. and Dev. 16, 399-405.

Nakatani, Y., Nishida, H., 1994. Induction of notochord during ascidian embryogenesis. Dev. Biol. 166, 289-299.

Nakatani, Y., Moody, R., Smith WC. 1999. Mutations affecting tail and notochord development in the ascidian Ciona savignyi. Development 126, 3293-3301.

Nishida, H., 1987. Cell Lineage Analysis in Ascidian Embryos by Intracellular Injection of a Tracer Enzyme. III. Up to the Tissue Restricted Stage. Dev. Biol. 121, 526-541.

Nishida, H., 2005. Specification of Embryonic Axis and Mosaic Development in Ascidians. Dev. Dyn. 233, 1177-1193.

Nüsslein-Volhard, C., Weischaus, E., 1980. Mutations affecting segment number and polarity in Drosophila. Nature 287, 795-801.

Oliveri, P., Qiang, T., Davidson, E.H., 2008. Global regulatory logic for specification of an embryonic cell lineage. PNAS 105, 5955-5962.

Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B.J., Afzal, S.Y., Lee, E.A., Zhu, Y., Placzek-Frick, I., Pickle, C.S., Kato, M., Garvin, T.H., Pham, Q.T., Harrington, A.N., Akiyama, J.A., Afzal, V., Lopez-Rios, J., Dickey, D.E., Visel, A., Pennacchio, L.A., 2018. Enhancer redundancy provides phenotypic robustness in mammalian development. Nature 554, 239-243.

Park, J., Estrada, J., Johnson, G., Vincent, B., Ricci-Tam, C., Bragdon, M., Shulgina, Y., Cha, A., Wunderlich, Z., Gunawardena, J., DePace, A., 2019. Dissecting the sharp response of a canonical developmental enhancer reveals multiple sources of cooperativity. eLife 8, e41266.

Passamaneck, Y.J., Di Gregorio, A., 2005. Ciona intestinalis: Chordate development made simple. Dev. Dyn. 233, 1-19.

Passamaneck, Y.J., Katikala, L., Perrone, L., Dunn, M.P., Oda-Ishii, I., Di Gregorio, A., 2009. Direct activation of a notochord cis-regulatory module by Brachyury and FoxA in the ascidian Ciona intestinalis. Development 136, 3679-3689.

Peng, Y., Zhang, Y., 2018. Enhancer and super-enhancer: Positive regulators in gene transcription. Animal Models and Exp. Med. 1, 169-179.

Pennati, R., Ficetola, G.F., Brunetti, R., Caicci, F., Gasparini, F., Griggio, F., Sato, A., Stach, T., Kabul-Strehlow, S., Gissi, C., Manni L. 2015. Morphological Differences between Larvae of the Ciona intestinalis Species Complex: Hints for a Valid Taxonomic Definition of Distinct Species. PLoS One 10, e0122879.

Perry, M., Boettiger, A., Levine, M., 2011. Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. PNAS 108, 13570-13575.

Placzek, M., Jessell, T.M., Dodd, J., 1993. Induction of floor plate differentiation by contact-dependent homeogenetic signals. Development 117, 205-218.

Pott, S., Lieb, J.D., 2015. What are super-enhancers? Nat. Gen. 47, 8-12.

Reeves, W., Thayer, R., Veeman, M., 2014. Anterior-Posterior Regionalized Gene Expression in the Ciona Notochord. Dev. Dyn. 243, 612-620.

Reeves, W.M., Wu, Y., Harder, M.J., Veeman, M.T., 2017. Functional and evolutionary insights from the Ciona notochord transcriptome. Development 144, 3375-3387.

Reeves, W.M., Shimai, K., Winkley, K.M., Veeman, M.T., 2020. Brachyury controls Ciona notochord fate as part of a feedforward network and not as a unitary master regulator. bioRxiv, https://doi.org/10.1101/2020.05.29.12402.

Ririe, T.O., Fernandes, J.S., Sternberg, P.W., 2008. The Caenorhabditis elegant vulva: A post-embryonic gene regulatory network controlling organogenesis. PNAS 105, 20095-20099.

Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S., Elowitz, M.B., 2005. Gene Regulation at the Single-Cell Level. Science 307, 1962-1965

Rothbächer, U., Bertrand, V., Lamy, C., Lemaire, P., 2007. A combinatorial code of maternal GATA, Ets, and β-catenin-TCF transcription factors specifies and patterns the early ascidian ectoderm. Development 134, 4023-4032.

Satoh, N., 2014. 'A Brief Introduction to Ascidians' in Developmental genomics of ascidians, 15-24.

Schoenfelder, S., Fraser, P., 2019. Long-range enhancer-promoter contacts in gene expression control. Nat. Rev. Gen. 20, 437-455

Scholes, C., Biette, K., Harden, T., DePace, A., 2019. Signal Integration by Shadow Enhancers and Enhancer Duplications Varies across the Drosophila Embryo. Cell Reports 26, 2407-2418.

Schupbach, T., Weischaus, E., 1986. Germline autonomy of maternal-effect mutations altering the embryonic body pattern of Drosophila. Dev. Biol. 113, 443-448.

Sharrocks, A.D., 2001. The Ets-domain transcription factor family. Nat. Rev. 2, 827-837.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., Ren, B., 2012. A map of the cis-regulatory sequences in the mouse genome. Nature 488, 116-120.

Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R., Flavell, R.A., 2005. Interchromosomal associations between alternatively expressed loci. Nature 435, 637-645.

Staby, L., O'Shea, C., Willemoës, M., Theiler, F., Kragelund, B.B., Skriver, K., 2017. Eukaryotic transcription factors: paradigms of protein intrinsic disorder. Biochem. Journal 474, 2509-2532.

Swanson, C.I., Evans, N.C., Barolo, S., 2010. Structural Rules and Complex Regulatory Circuitry Constrain Expression of a Notch- and EGFR-Regulated Eye Enhancer. Dev. Cell 18, 359-376.

Tassy, O., Daian, F., Hudson, C., Bertrand, V., Lemaire, P., 2006. A Quantitative Approach to the Study of Cell Shapes and Interactions during Early Chordate Embryogenesis. Curr. Biol. 16, 345-358.

Takahashi, H., Hotta, K., Erives, A., Di Gregorio, A., Zeller, R.W., Levine, M., Satoh, N., 1999. Brachyury downstream notochord differentiation in the ascidian embryo. Genes and Dev. 13, 1519-1523.

True, J.R., Haag, E.S., 2001. Developmental system drift and flexibility in evolutionary trajectories. Evol. and Dev. 3, 109-119.

Varjosalo, M., Taipale, J., 2008. Hedgehog: functions and mechanisms. Genes and Development 22, 2454-2472.

Veeman, M., Reeves, W., 2015. Quantitative and in toto imaging in ascidians: Working toward an image-centric systems biology of chordate morphogenesis. Genesis 53, 143-159.

Veeman, M.T., Smith, W.C., 2013. Whole-organ cell shape analysis reveals the developmental basis of ascidian notochord taper. Dev. Biol. 373, 281-289.

Veeman, M.T., Nakatani, Y., Hendrickson, C., Ericsson, V., Lin, C., Smith, W.C., 2008. Chongmague reveals an essential role for laminitis-mediated boundary formation in chordate convergence and extension movements. Development 135, 33-41.

Veitia, R.A., 2018. Dosage effects in morphogenetic gradients of transcription factors: insights from a simple mathematical model. Journal of Gen. 97, 365-370.

Wang, C., Lehmann, R., 1991. Nanos Is the Localized Posterior Determinant in Drosophila. Cell 66, 637-647.

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

Wilczynski, B., Furlong, E.E.M., 2009. Challenges for modeling global gene regulatory networks during development: Insights from Drosophila. Dev. Biol. 340, 161-169.

Winkley, K., Ward, S., Reeves, W., Veeman, M., 2019. Iterative and Complex Asymmetric Divisions Control Cell Volume Differences in Ciona Notochord Tapering. Curr. Biol. 29, 3466-3477.

Wunderlich, Z., Bragdon, M.D.J., Vincent, B.J., White, J.A., Estrada, J., DePace, A.H., 2015. Krüppel Expression Levels Are Maintained through Compensatory Evolution of Shadow Enhancers. Cell Rep. 12, 1740-1747.

Yamada, T., Placzek, M., Tanaka, H., Dodd, J., Jessell, T.M., 1991. Control of cell pattern in the developing nervous system: polarizing activity of the floor plate and notochord. Cell 64, 635-647.

Yang, Y., 2012. Wnt signaling in development and disease. Cell & Bioscience 2, 14.

Yasuo, H., Hudson, C., 2007. FGF8/17/18 functions together with FGF9/16/20 during formation of the notochord in Ciona embryos. Dev. Biol. 302, 92-103.

Zamudio, A.V., Dall'Agnese, A., Heyningen, J.E., Mantegna, J.C., Afeyan, L.K., Hannett, N.M., Coffey, E.L., Li, C.H., Oksuz, O., Sabari, B.R., Boija, A., Klein, I.A., Hawken, S.W., Spille, J.-H., Decker, T.-M., Cisse,, I.I., Abraham, B.J., Lee, T.I., Taatjes, D.J., Schuijers, J., Young, R.A., 2019. Mediator Condensates Localize Signaling Factors to Key Cell Identity Genes. Mol. Cell 76, 1-14.

Zhu, J., Adlai, M., Zou, J.Y., Verstappen, G., Coyne, M., Zhang, X., Durham, T., Miri, M., Deshpande, V., De Jager, P.L., Bennett, D.A., Houmard, J.A., Muoio, D.M., Onder, T.T., Camahort, R., Cowan, C.A., Meissen, A., Epstein, C.B., Shoresh, N., Bernstein, B.E., 2013. Genome-wide Chromatin State Transitions Associated with Developmental and Environmental Cues. Cell 152, 642-654.
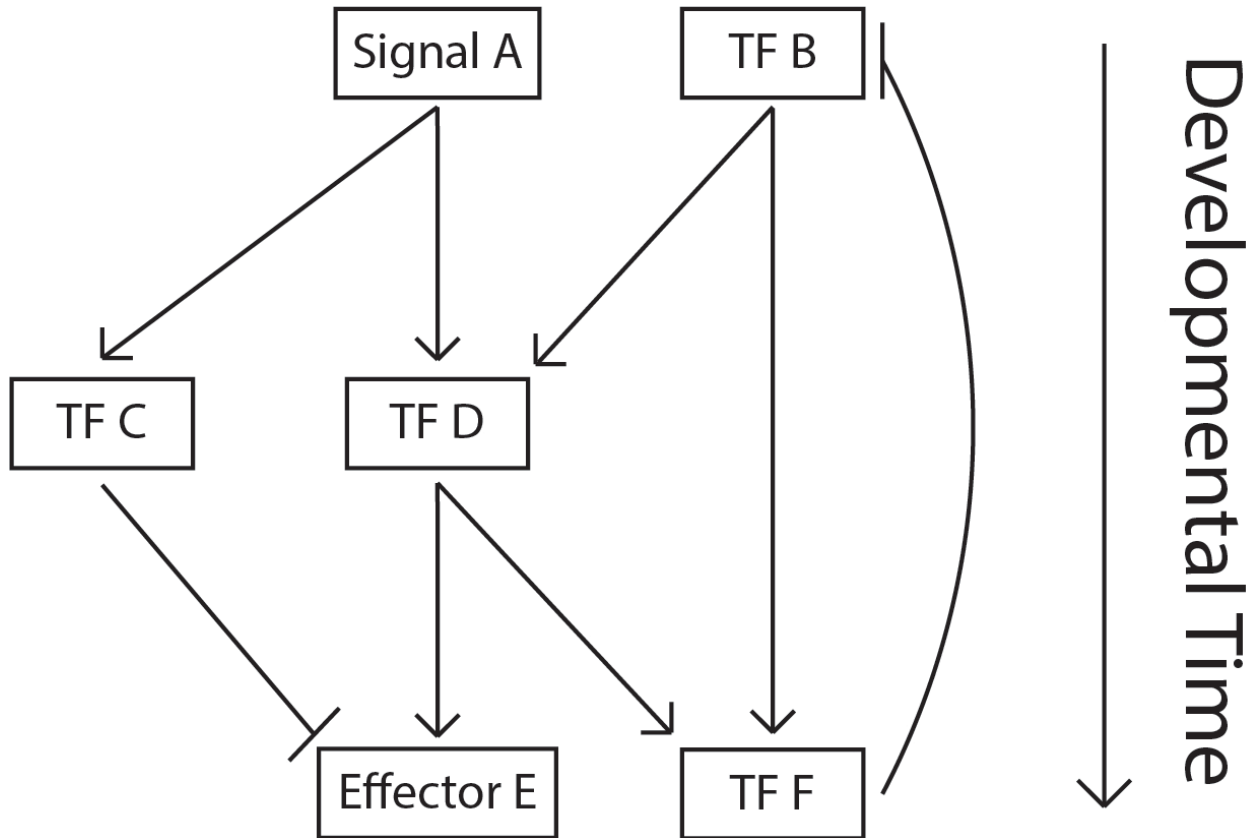
**Figure 1.1 Simple model of a GRN.**

A model GRN featuring six genes with several common regulatory mechanisms. AND logic is used to induce the expression of TF genes D and F, NOT logic is used to limit the expression of Effector gene E, and TFs B, D, and F participate in feedback and feedforward loops.
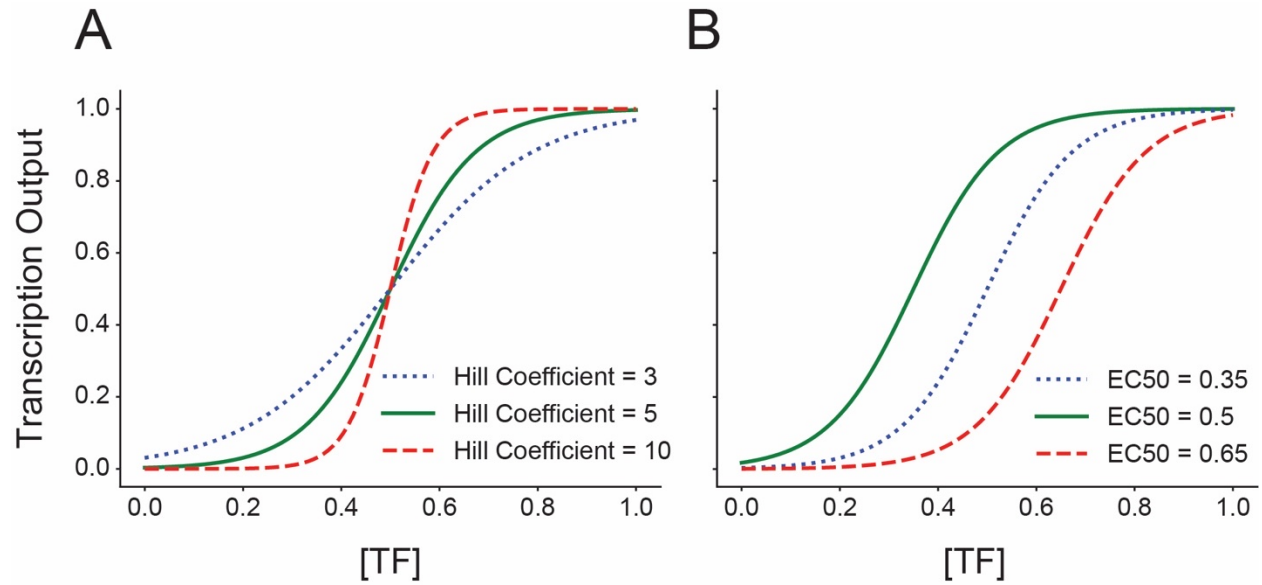
**Figure 1.2 Changes in the key Hill Equation parameters.**

(**A**) Increasing the Hill Coefficient increases the steepness of the response without changing the overall sensitivity. (**B**) Increasing the EC50 decreases the sensitivity without changing the cooperativity.
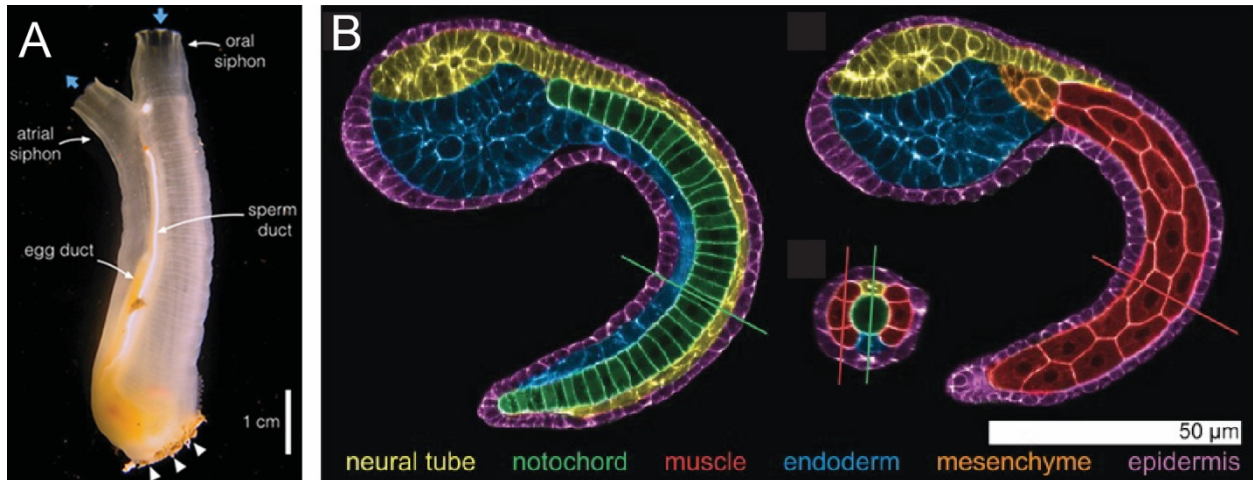
**Figure 1.3 *Ciona* morphology.**

(**A**) Adult *Ciona* are small filter feeders. White arrowheads indicate attachment point to the substrate. Image credit: Kourakis and Smith, 2015. (**B**) *Ciona* embryos have a stereotypically chordate body plan, with a dorsal nerve cord (yellow), central notochord (green) and tail muscle (red) flanking the notochord. Image credit: Veeman and Reeves, 2015.
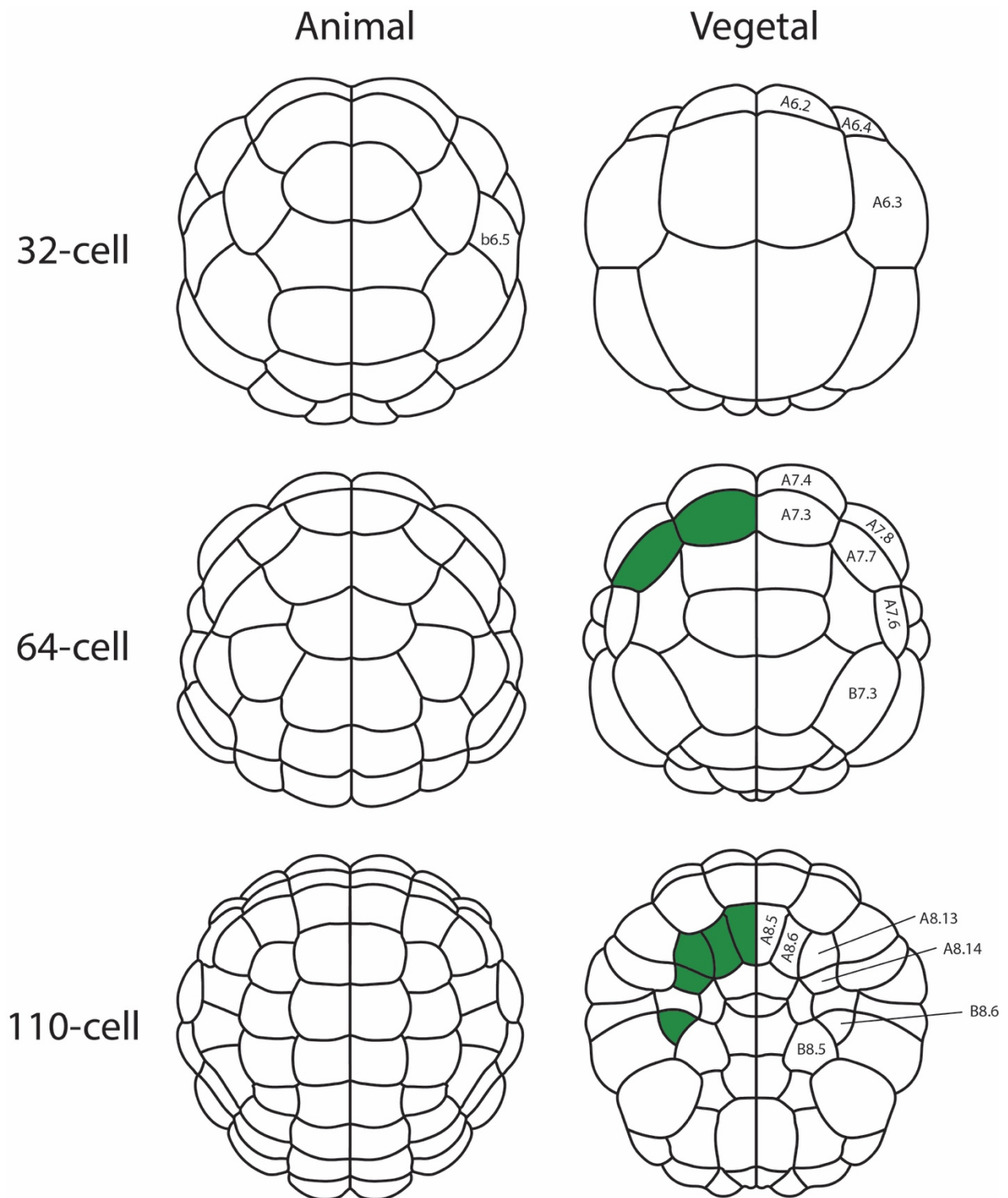
**Figure 1.4 Relevant blastomere identities**

*Ciona* embryos at 32-, 64-, and 110-cell stages, animal view at left, vegetal view at right. Embryos are left/right bilaterally symmetrical. For simplicity, only the left side blastomeres are labeled by name. Green color on right side indicates fate-restricted notochord precursor cells.

# Chapter 2 - The *Ciona* notochord transcriptome

Although there have been many publications over the last twenty years using the *Ciona* notochord as a model to understand tissue-specific gene expression (Hotta et al., 1999, Takahashi et al., 1999, Hotta et al., 2000, Hotta et al., 2008, Kugler et al., 2008, Passamaneck et al., 2009, Kubo et al., 2010, José-Edwards et al., 2011, José-Edwards et al., 2013, Katikala et al., 2013, Reeves et al., 2014, José-Edwards et al., 2015, Kugler et al., 2019), these studies lacked the genome-wide, 'omics'-scale perspective feasible using modern transcriptional profiling by RNAseq. In 2017, Reeves et al. defined a *Ciona* notochord transcriptome through RNAseq of dissociated *Ciona* embryos electroporated with a notochord-specific *Brachyury*>GFP reporter and separated into notochord-enriched and notochord-depleted populations by Fluorescence-Activated Cell Sorting (FACS). Extensive *in situ* hybridization confirmed the overall validity of the flow-sorted RNAseq data, and was also useful in identifying genes with intriguing nonuniform expression patterns within the notochord. Comparison of this dataset with further RNAseq of embryos ectopically mis-expressing *Brachyury* revealed surprisingly little overlap, suggesting that *Brachyury* may not be a true master regulator of notochord fate in *Ciona* as had long been assumed.

I am an author on this paper but not one of the lead authors, so I am giving a brief overview of it here that focuses on the aspects that I was most involved in and that are most relevant to the rest of my dissertation work. My direct involvement was largely in the summer of 2014 when I was an undergraduate REU student working in the lab. My project that summer was to conduct some of the *in situ* validation and identify new regionalized patterns of gene expression, but I did not resume this project at the official start of my graduate career. It did, however, lead directly into the project detailed in the next chapter of this dissertation.

## Summary of Reeves et al., 2017

*Ciona robusta* embryos were electroporated with *Brachyury*>GFP, and then dissociated at late neurula (Hotta stage 16) and early tailbud (Hotta stage 19.5) (Hotta et al., 2007). An additional timepoint was collected at late tailbud (Hotta stage 23), but was ultimately excluded from analysis because of imperfect dissociation. Dissociated cells were sorted by FACS to obtain notochord-enriched and notochord-depleted populations of cells at each timepoint, and RNAseq was performed on each sorted cell population to quantify the transcriptome of the different populations at the different times and test for differential gene expression. 1364 genes were identified in total as being enriched in the notochord cell population, with a wide range of enrichment levels. 460 of these genes were enriched at both stages.

Validation by *in situ* hybridization was performed on 111 genes identified as being notochord-enriched, staining for expression at the three timepoints collected for RNA-seq, representing the total range of notochord expression values, as determined by RNA-seq. 98 of these genes were detectable, and 88 were expressed in notochord in particular. Although the most common expression pattern for notochord genes was to have uniform expression throughout the notochord, several genes were found to have regionalized expression patterns within the notochord, including C8.891 (tip-depleted), GPA1 (tip-enriched), MLKL (stochastic), and *fibulin-like* (secondary-enriched) (Fig. 2.1). The gene *fibulin-like*, a putative ECM component, was particularly enriched in the secondary notochord at the very posterior end of the intercalated notochord. The mechanisms controlling the posterior-enriched expression of *fibulin-like* became one of my major research topics and is detailed in Chapter 3.

*fibulin-like* was not the only ECM gene with significant notochord enrichment. At both timepoints, Gene Ontology (GO) codes associated with various ECM functions were found to be

deeply enriched in the notochord gene sets. Several transcription factors and genes involved in signaling pathways were also found to be upregulated in the notochord, providing a more detailed picture of potential players in the *Ciona* notochord GRN.

*Brachyury* had long been considered to be a notochord master regulator in *Ciona*, both necessary for notochord fate (Yamada et al., 2003, Chiba et al., 2009) and sufficient to transform other cell types to notochord when misexpressed (Takahashi et al., 1999). Most of the previously known notochord genes, however, were first identified in a Brachyury misexpression screen and there had been limited independent testing of this hypothesis. With a notochord transcriptome now available, there was an opportunity to address this on a much larger scale. *Brachyury* was ectopically expressed outside of the notochord under the control of a *FoxAa* enhancer, similar to (Takahashi et al., 1999), but using RNAseq as a sensitive and genome-wide transcriptional readout. Surprisingly, there was only a small overlap between the notochord enriched gene sets and the genes induced by ectopic Bra expression. Many notochord genes are not induced by ectopic Bra expression and many genes induced by ectopic Bra expression are not normally expressed in the notochord. The Veeman lab has subsequently confirmed that Bra is not a unitary master regulator but instead acts in parallel to FoxAa as part of a more complex feedforward network (Reeves et al., 2020).

# Connection to this Dissertation

Although not the sole focus of Reeves et al. (2017), nonuniform gene expression patterns within the notochord identified by *in situ* hybridization were particularly interesting. Considerable evidence is emerging that the *Ciona* notochord is a finely-patterned organ with differences in both cell behavior and gene expression between different notochord cells (Veeman and Smith, 2013, Reeves et al., 2014, Carlson et al., 2015, Winkley et al., 2019). Therefore, genes identified as having regionalized gene expression patterns are of particular interest, both for the purpose of finding candidate genes mediating these different cell behaviors, and for revealing the gene regulatory networks that can mediate finely patterned regionalized expression within this organ that was previously thought to be spatially uniform.

The secondary notochord enriched expression pattern of *fibulin-like* was discovered shortly after my REU internship concluded, but I was able to take over the project not long after the start of my graduate program. The primary and secondary notochord lineages are separated several cell divisions before notochord specification (Hotta et al., 2007), yet form a contiguous organ shortly before notochord intercalation. Despite the elements of stochasticity in the exact final positioning of the notochord cells during the notochord intercalation, secondary notochord cells remain at the posterior-most position (Carlson et al., 2015). Despite these clear differences in the two cell lineages, as well as differences in how these lineages are induced to take on notochord fate (Hudson and Yasuo, 2006), the primary and secondary notochord lineages had not been found to have sharp differences in gene expression until *fibulin-like*, prompting its selection as a candidate gene for dissecting the mechanisms driving differences in the transcriptional regulation between primary and secondary notochord fate.

# References

Carlson, M., Reeves, W., Veeman, M., 2015. Stochasticity and stereotypy in the Ciona notochord. Dev. Biol. 397, 248-256.

Chiba, S., Jiang, D., Satoh, N., Smith, W.C., 2009. Brachyury null mutant-induced defects in juvenile ascidian endodermal organs. Development 136, 35-39.

Hotta, K., Takahashi, H., Erives, A., Levine, M., Satoh, N., 1999. Temporal expression patterns of 39 Brachyury-downstream genes associated with notochord formation in the Ciona intestinalis embryo. Dev. Growth Diff. 41, 657-664.

Hotta, K., Takahashi, H., Asakura, T., Satoh, B., Takatori, N., Satou, Y., Satoh, N., 2000. Characterization of Brachyury-Downstream Notochord Genes in the Ciona intestinalis embryo. Dev. Biol. 224, 69-80.

Hotta, K., Mitsuhara, K., Takahashi, H., Inaba, K., Oka, K., Gojobori, T., Ikeo, K., 2007. A Web-Based Interactive Developmental Table for the Ascidian Ciona intestinalis, Including 3D Real-Image Embryo Reconstructions: I. From Fertilized Egg to Hatching Larva. Dev. Dyn. 236, 1790-1805.

Hotta, K., Takahashi, H., Satoh, N., Gojobori, T., 2008. Brachyury-downstream gene sets in a chordate, Ciona intestinalis: integrating notochord specification, morphogenesis and chordate evolution. Evol. and Dev. 10, 37-51.

Hudson, C., Yasuo, H., 2006. A signalling relay involving Nodal and Delta ligands acts during secondary notochord induction in Ciona embryos. Development 133, 2855-2864.

José-Edwards, D.S., Kerner, P., Kugler, J.E., Deng, W., Jaing, D., Di Gregorio, A., 2011. The identification of transcription factors expressed in the notochord of Ciona intestinalis adds new potential players to the Brachyury gene regulatory network. Dev. Dyn. 240, 1793-1805.

José-Edwards, D.S., Oda-Ishii, I., Nibu, Y., Di Gregorio, A., 2013. Tbx2/3 is an essential mediator within the Brachyury gene network during Ciona notochord development. Development 140, 2422-2433.

José-Edwards, D.S., Oda-Ishii, I., Kugler, J.E., Passamaneck, Y.J., Katikala, L., Nibu, Y., Di Gregorio, A., 2015. Brachyury, Foxa2 and the cis-Regulatory Origins of the Notochord. PLoS Gen. 11, e1005730.

Katikala, L., Aihara, H., Passamaneck, Y.J., Gazdoiu, S., José-Edwards, D.S., Kugler, J.E., Oda-Ishii, I., Imai, J.H., Nibu, Y., Di Gregorio, A., 2013. Functional Brachyury Binding Sites Establish a Temporal Read-Out of Gene Expression in the Ciona Notochord. PLoS Biol. 11, e1001697.

Kubo, A., Suzuki, N., Yuan, X., Nakatani, K., Satoh, N., Imai, K.S., Satou, Y., 2010. Genomic cis-regulatory networks in the early Ciona intestinalis embryo. Development 137, 1613-1623.

Kugler, J.E., Passamaneck, Y.J., Feldman, T.G., Beh, J., Regnier, T.W., Di Gregorio, A., 2008. Evolutionary Conservation of Vertebrate Notochord Genes in the Ascidian Ciona intestinalis. Genesis 46, 697-710.

Kugler, J.E., Wu, Y., Katikala, L., Passamaneck, Y.J., Addy, J., Caballero, N., Oda-Ishii, I., Maguire, J.E., Li, R., Di Gregorio, A., 2019. Positioning a multifunctional basic helix-loop-helix transcription factor within the Ciona notochord gene regulatory network. Dev.Biol. 448, 119-135.

Passamaneck, Y.J., Katikala, L., Perrone, L., Dunn, M.P., Oda-Ishii, I., Di Gregorio, A., 2009. Direct activation of a notochord cis-regulatory module by Brachyury and FoxA in the ascidian Ciona intestinalis. Development 136, 3679-3689.

Reeves, W., Thayer, R., Veeman, M., 2014. Anterior-Posterior Regionalized Gene Expression in the Ciona Notochord. Dev. Dyn. 243, 612-620.

Reeves, W., Wu, Y., Harder, M., Veeman, M., 2017. Functional and evolutionary insights from the Ciona notochord transcriptome. Development 144, 3375-3387.

Reeves, W.M., Shimai, K., Winkley, K.M., Veeman, M.T., 2020. Brachyury controls Ciona notochord fate as part of a feedforward network and not as a unitary master regulator. bioRxiv, https://doi.org/10.1101/2020.05.29.12402.

Takahashi, H., Hotta, K., Erives, A., Di Gregorio, A., Zeller, R.W., Levine, M., Satoh, N., 1999. Brachyury downstream notochord differentiation in the ascidian embryo. Genes and Dev. 13, 1519-1523.

Veeman, M.T., Smith, W.C., 2013. Whole-organ cell shape analysis reveals the developmental basis of ascidian notochord taper. Dev. Biol. 373, 281-289.

Winkley, K., Ward, S., Reeves, W., Veeman, M., 2019. Iterative and Complex Asymmetric Division sControl Cell Volume Differences in Ciona Notochord Tapering. Curr. Biol. 29, 3466-3477.

Yamada, L., Shoguchi, E., Wanda, S., Kobayashi, K., Mochizuki, Y., Satou, Y., Satoh, N., 2003. Morpholino-based gene knockdown screen of novel genes with developmental function in Ciona intestinalis. Development 130, 6485-6495.
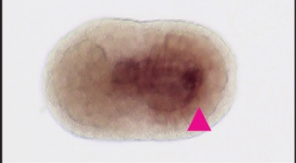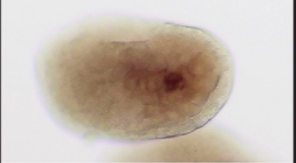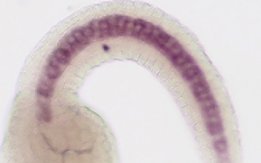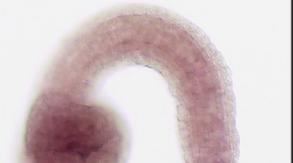
| C8.749 | GPA1 C2.872 | MLKL C4.411 | Fibulin-like C11.331 |
|---|---|---|---|
| 345 (33.0x)* | 209 (47.5x)* | 30 (18.4x)* | 45 (29.1x)* |
| 1878 (24.1x)* | 253 (20.4x)* | 36 (15.0x)* | 76 (14.2x)* |
| 1174 (30.4x)* | 198 (10.3x)* | 145 (13.7x)* | 68 (47.1x)* |

**Figure 2.1 Regionalized gene expression patterns of notochord-enriched genes.**
Gene expression patterns determined by *in situ* hybridization at stages 16 (top row), 19/20
(middle row), and 23 (bottom row) (Hotta et al., 2007). Green arrows indicate tip depletion,
magenta arrows indicate tip enrichment, and magenta brackets indicate secondary notochord
enrichment. Numbers indicate FKPM expression values and (notochord fold-enrichment) at the
indicated stage and gene. Asterisk indicates statistically-significant enrichment. Figure
annotations by W. Reeves, *in situ* hybridization for C8.749, GPA1, and MLKL by M. Harder,
and *in situ* hybridization for Fibulin-like by W. Reeves.

# Chapter 3 - Multiple inputs into a posterior-specific regulatory network in the *Ciona* notochord

Matthew Harder[1], Wendy Reeves[1], Chase Byers, Mercedes Santiago, Michael Veeman*

[1]Equal contribution

*Corresponding author (veeman@ksu.edu)

# Abstract

The gene regulatory networks underlying *Ciona* notochord fate specification and differentiation have been extensively investigated, but the regulatory basis for regionalized expression within the notochord is not understood. Here we identify three notochord-expressed genes, *C11.331*, *C12.115* and *C8.891*, with strongly enriched expression in the secondary notochord cells at the posterior tip of the tail. *C11.331* and *C12.115* share a distinctive expression pattern that is highly enriched in the secondary notochord lineage but also graded within that lineage with the strongest expression at the posterior tip. Both genes show similar responses to pharmacological perturbations of Wnt and FGF signaling, consistent with an important role for Wnt and FGF ligands expressed at the tail tip. Reporter analysis indicates that the *C11.331 cis*-regulatory regions are extensively distributed, with multiple non-overlapping regions conferring posterior notochord-enriched expression. Fine-scale analysis of a minimal *cis*-regulatory module (CRM) identifies discrete positive and negative elements including a strong silencer. Truncation of the silencer region leads to increased expression in the primary notochord, indicating that *C11.331* expression is influenced by putative regulators of primary *versus* secondary notochord fate. The minimal CRM contains predicted ETS, GATA, LMX and Myb sites, all of which lead to reduced expression in secondary notochord when mutated. These results show that the posterior-enriched notochord expression of *C11.331* depends on multiple inputs, including Wnt and FGF signals from the tip of the tail, multiple notochord-specific regulators, and yet-to-be identified regulators of regional identity within the notochord.

# Introduction

The *Ciona* notochord consists of only 40 cells that intercalate to form a tapered, single-file column. Cell fate specification in the *Ciona* notochord has been extensively studied. In ascidians, the primary notochord cells are derived from blastomere pairs A7.3 and A7.7 and give rise to the anterior 32 cells in the intercalated notochord, whereas the secondary notochord cells are derived from B8.6 and give rise to the posterior 8 cells (Nishida, 1987). In *Ciona*, A7.3 and A7.7 become fate-restricted to notochord at the 64-cell stage downstream of both beta-catenin mediated mesendodermal specification and FGF signaling (Hudson et al., 2016, Hudson et al., 2013, Yasuo and Hudson, 2007). B8.6 becomes fate-restricted at the 112-cell stage *via* different mechanisms that involve a Nodal/Delta2 signaling relay initiated by the lateral b6.5 cells (Hudson and Yasuo, 2006, Hudson and Yasuo, 2005, Imai et al., 2006, Imai et al., 2002).

Both primary and secondary notochord cells express the key notochord-specific transcription factor Brachyury (Bra) (Corbo et al., 1997, Yasuo and Satoh, 1993), and the majority of notochord-enriched genes are expressed uniformly throughout the notochord (Hotta et al., 2000, Hotta et al., 1999, José-Edwards et al., 2011, Kugler et al., 2008, Reeves et al., 2017). We have previously shown, however, that notochord cell sizes, shapes and behaviors vary depending on both lineage and anterior-posterior position (Carlson et al., 2015, Veeman and Smith, 2013). We have also shown that a subset of notochord-enriched genes are differentially expressed within subregions of the notochord (Reeves et al., 2014). The regulatory basis for regionalized notochord expression has received no previous attention.

Differential gene expression between primary and secondary notochord is of particular interest because secondary notochord cells are quite distinct from primary notochord cells in their morphology and behavior. Secondary notochord cells are more stereotyped in their patterns

of intercalation than primary cells and give rise to a particularly distinct taper towards the tail's distal tip (Carlson et al., 2015, Veeman and Smith, 2013). Unlike primary cells, they still intercalate effectively even in mutants for the planar cell polarity pathway gene *Prickle* (Jiang et al., 2005). Elements of differential gene expression have previously been noted between primary and secondary notochord cells, including a gene upregulated at both the anterior and posterior tips of the notochord, one expressed in a posterior to anterior gradient, and another expressed throughout the notochord but with visible enrichment in the secondary lineage (Reeves et al., 2014). Despite these indications of underlying patterning, markers that were strongly differentially expressed between primary and secondary notochord cells were previously elusive.

We recently used RNAseq on flow-sorted notochord cells to identify a large set of 1364 genes predicted to be enriched in the notochord. Extensive *in situ* validation of 151 genes showed that ~90% were indeed notochord enriched (Reeves et al., 2017 and unpublished data). In the course of this we identified several new genes showing regionalized notochord expression, including three that are strongly enriched at the notochord's posterior tip. Here we characterize the expression patterns of these posterior notochord tip-specific genes with fine detail and investigate the molecular mechanisms controlling their regionalized expression.

# Results

## Secondary-enriched notochord expression

Although the majority of notochord enriched genes we identified by RNAseq are expressed uniformly across the notochord (Reeves et al., 2017), a subset display regionalized expression within the notochord. Fig. 3.1 shows *in situ* expression patterns for three genes with strongly differential expression between primary and secondary notochord. *C11.331* has homology to the fibulin/hemicentin family of secreted extracellular matrix proteins (Timpl et al., 2003), but is not a clear ortholog of any vertebrate fibulin. *C12.115* and *C8.891* are both predicted SLC family solute carriers. We refer to all three here by their KH2012 gene model names (Brozovic et al., 2018, Satou et al., 2008) for unambiguity. During early notochord intercalation at mid neurula (Hotta stage 16) (Hotta et al., 2007), *C11.331* is barely detectable (Fig. 3.1A), but is distinctly expressed at the notochord's posterior tip late in intercalation at early tailbud stage 20 (Fig. 3.1B) and during tail elongation at late tailbud stage 23 (Fig. 3.1C). *C12.115* has posterior-enriched notochord expression at all three stages (Fig. 3.1D-F) which resolves over time to be increasingly specific to the notochord's posterior tip. *C8.891* is qualitatively different than the other two genes and shows a more graded expression pattern at the first two stages (Fig. 3.1G,H) before developing a sharper transition between primary and secondary at stage 23 (Fig. 3.1I).

To better quantify their expression patterns, we also imaged fluorescent *in situ* hybridizations against *C11.331* and *C12.115* by confocal microscopy. This allowed the boundary between primary and secondary notochord to be clearly identified. Both genes showed expression in secondary but not primary notochord (Fig. 3.1J,K). For both *C11.331* and *C12.115*, this secondary-specific expression was stronger at the notochord's posterior tip than at the

boundary between primary and secondary notochord. Fig. 3.1L,M show quantitative analyses of these expression patterns across nine embryos each, confirming that both genes have graded expression in the secondary notochord that is strongest at the posterior tip.

## Perturbation of Wnt and FGF signaling

These expression patterns suggested the potential involvement of a secreted signaling molecule expressed at the posterior tip of the tail. Candidate molecules for such a signal include Wnt5, FGF8/17/18 and FGF9/16/20, all of which are expressed in tail tip ectoderm or posterior tail muscles (Hudson et al., 2007, Hudson and Yasuo, n.d, Ikuta et al., 2010, Imai et al., 2004). We tested their involvement pharmacologically using the FGF pathway inhibitor U0126 and the canonical Wnt pathway activator BIO. Both of these drugs have been extensively used in *Ciona* and we used doses previously established to be specific and effective (Dumollard et al., 2013, Hudson et al., 2013, Racioppi et al., 2014, Sakabe et al., 2006). We treated the embryos with drugs starting at mid gastrula (Hotta stage 12). This is after notochord has been specified and has completed all cell divisions (Hotta et al., 2007), but before the earliest detectable expression of *C11.331*, *C12.115* or *C8.891* by *in situ* hybridization (data not shown). A Wnt pathway inhibitor would be conceptually appealing but no such reagents have been validated in *Ciona* embryos.

Both drugs caused mild defects in tail morphogenesis, as has been described previously for U0126 (Ikuta et al., 2010). As expected, the BIO and U0126 treatments had no effect on the expression of the key notochord transcriptional regulator *Brachyury* (Fig. 3.2A-D), but the drugs caused strong effects on the expression of the three posterior-specific notochord markers. *C11.331* and *C12.115* responded to both drugs in similar ways. BIO treatment led to the expansion of *C11.331* and *C12.115* expression throughout primary notochord (compare Fig.

3.2E and I to Fig. 3.2F and J). Interestingly, this ectopic expression was often strongest at the notochord's anterior tip, giving rise to an overall expression pattern reminiscent of *BCamL* (*KH.C2.209*) (Reeves et al., 2014) in terms of being enriched at both the anterior and posterior tips of the notochord. FGF pathway inhibition with U0126, however, led to a major decrease in the expression of both markers (Fig. 3.2G,K). This was particularly evident for *C12.115* expression, which was virtually eliminated by U0126 treatment despite being much more robustly expressed than *C11.331* in untreated embryos. *C8.891* expression expanded throughout the notochord in response to Wnt pathway activation by BIO, but this ectopic expression was uniform and did not show the 'both tips' pattern usually seen with the other two genes (Fig. 3.2N). *C8.891* also differed from *C11.331* and *C12.115* in that FGF pathway inhibition by U0126 only caused a modest decrease in expression (Fig. 3.2O).

The BIO and U0126 treatments indicated that both Wnt and FGF signals may act as positive regulators of *C11.331*, *C12.115* and *C8.891* expression. As a preliminary test of the potential relationships between Wnt and FGF signaling in this context, we also treated embryos with both drugs simultaneously. Both *C11.331* and *C12.115* showed an intermediate phenotype in which expression expanded throughout the notochord but was much fainter than in response to BIO treatment alone (Fig. 3.2H,L). In addition, no enrichment in the anterior tip was seen and posterior enrichment was also lost in most embryos. This lack of a clear epistatic relationship suggests that Wnt and FGF signals act in parallel upon these two genes. *C8.891* again behaved differently; treatment with combined BIO and U0126 led to expanded expression throughout the notochord, but this ectopic expression was strong and indistinguishable from BIO treatment alone (Fig. 3.2P). FGF signaling appears to play a less important role in the regulation of *C8.891* compared to the other two posterior enriched genes.

## Quantitative *cis*-regulatory analysis

To narrow in on the molecular mechanisms responsible for posterior-specific notochord expression, we developed a dual reporter strategy. Electroporated transgenes in *Ciona* are expressed mosaically, and it is common for individual embryos electroporated with reporter constructs for uniformly expressed notochord genes such as *Brachyury* to exhibit expression in secondary notochord but not primary and *vice-versa* (Carlson et al., 2015, Corbo et al., 1997). Different transgenes electroporated at the same time, however, typically show common patterns of mosaicism (Zeller et al., 2006). To control for differential reporter expression due to transgene mosaicism, as opposed to *bona fide* differences in regionalized expression, we co-electroporated HA-tagged Histone H2B under the control of the *Bra* enhancer/promoter together with Venus fluorescent protein downstream of candidate regulatory regions and a basal promoter (Fig. 3.3A).

We used this dual reporter approach to investigate the *cis*-regulation of *C11.331*. All embryos were imaged by confocal microscopy using uniform imaging parameters. We developed a quantitative analysis approach based on measuring Venus and HA intensity along the AP axis of computationally straightened and flattened notochords. For each embryo analyzed, we measured the mean background-corrected Venus reporter signal for *Bra*>H2B: HA expressing cells in primary notochord (cells 1–32) and secondary notochord (cells 33–40), and normalized the results to the *Bra*>H2B: HA intensity in each region (Fig. 3.3B). We also scored a larger number of embryos with a qualitative system, ranging from 0 (no detectable expression) to 4 (oversaturated) (Tables A.1 and A.2), which largely matched the quantitative analysis. An ectopic expression score was generated based on qualitative expression levels in several non-notochord tissues (Tables A.3 and A.4). Full details of the scoring systems are given in the Methods section.

*C11.331* is separated from the 3′ end of the closest upstream gene by ~1.5 kb, and we found that this intergenic region (-1565 to +13) gave rise to strong expression in secondary notochord, much weaker expression in primary notochord, and relatively little ectopic expression (Fig. 3.3B). We dissected this region with a series of truncations and found evidence for extensive functional redundancy. On a coarse scale, both the distal −1565 to −758 and the proximal −787 to +13 region gave rise to secondary-enriched notochord expression, with the distal region showing somewhat less ectopic expression but the proximal region showing stronger notochord expression and greater statistical evidence for enrichment in the secondary notochord. We further dissected the −787 to +13 region and again found aspects of redundancy, with both the −787 to −405 and the minimally overlapping −488 to +13 region showing secondary-enriched notochord expression.

Finer-scale analysis of the −488 to +13 region allowed us to identify distinct regulatory regions. A construct spanning the −488 to −165 region was still strongly expressed in secondary notochord. Unlike the previous constructs, however, it was strongly expressed in primary notochord as well, indicating the presence of a silencer element for primary notochord expression proximal to −165. A further truncation to give the −488 to −245 construct eliminated expression altogether, indicating the presence of essential enhancer elements in the −245 to −165 interval. Truncating from the other side, the −322 to +13 construct was highly expressed and strongly enriched in secondary notochord, whereas the −245 to +13 construct was enriched in secondary notochord, but much more weakly expressed overall, indicating that there are additional enhancer elements between −322 and −245. Our subsequent efforts therefore focused on the −322 to +13 interval.

Fig. 3.4A shows normalized *C11.331*(-322 to +13)>Venus expression as a function of AP position for 19 imaged notochords. The normalized mean intensity (green line) reveals a sharp transition from low to high expression precisely at the boundary between primary and secondary notochord. This is in contrast to the *Bra*>H2B: HA control that is expressed uniformly along the notochord's AP axis (Fig. 3.4B).

For these finer-scale dissections, we quantified reporter expression separately for the anterior (cells 33–36) and posterior (cells 37–40) secondary notochord. For the −322 to +13 construct, the sample mean of this quantitative metric of reporter expression was 19.7 fold higher in anterior secondary notochord as compared to primary notochord, and 26.6 fold higher in posterior secondary notochord as compared to primary notochord. These differences were highly significant as measured by the Wilcoxon sign-rank test. Two truncated constructs, −322 to −92 and −322 to −127, retained significant enrichment in secondary notochord but were less highly expressed overall, indicating that there are enhancer elements promoter-proximal to −92. A further truncation to give the −322 to −165 construct was significantly derepressed in primary notochord, confirming the presence of important silencer elements promoter-proximal to −165. Combined with the previous constructs (Fig. 3.3B), this indicates that the primary notochord silencer has essential elements between −165 and −127.

### Identification and testing of candidate TFBSs

We searched the −165 to −127 interval and flanking regions for putative transcription factor binding sites (TFBSs) using both TRANSFAC position weight matrices (PWMs) (Matys et al., 2006) *via* the LASAGNA web tool (Lee and Huang, 2013) and also using SELEX-seq-derived (Slattery et al., 2011) consensus sequences for known *Ciona* notochord TFs downloaded

from the ANISEED community database (Brozovic et al., 2018). We found several predicted

sites of interest, including an ETS site just outside of the silencer in a region required for strong

expression (Fig. 3.5A). This was of particular interest given that ETS family transcription factors

are key mediators of ascidian FGF signaling (Gainous et al., 2015, Miya and Nishida, 2003). We

did not identify any predicted LEF/TCF sites (Brannon et al., 1997), suggesting that any Wnt

inputs into this particular *cis*-regulatory module are likely to be indirect. Other sites of interest

included a GATA site, an LMX site and a Myb site that were all in very close proximity within

the putative silencer region. An LMX ortholog (KH.C9.485) is known to be expressed in the

*Ciona* notochord (José-Edwards et al., 2011), and functionally important Myb sites have been

identified in several *Ciona* notochord CRMs (José-Edwards et al., 2015). GATA TFs have not

been implicated in notochord-specific gene expression, but immunostaining suggests there may

be *GATA.a* expression in the posterior notochord (Oda-Ishii et al., 2016).

We individually mutated the core motif for each of these TFBSs in the −322 to +13

construct. The GATA and Myb core sequences are almost overlapping, and the LMX site is also

nearby, so we took care to mutate these core sequences without affecting the adjacent core sites

(Fig. 3.5A). We cannot exclude, however, that these mutations may have promiscuous effects *via*

important but uncharacterized flanking sequences. Our hypothesis was that one or more of these

mutations would have a phenotype similar to the −322 to −165 sequence in which expression in

primary notochord is derepressed. We instead found that all 4 mutations led to decreased

expression overall, particularly in secondary notochord cells (Fig. 3.5B). This was particularly

evident for the ETS mutation, which is not unexpected given its location outside of the silencer

region, the U0126 phenotype and the expression of FGF8/17/18 and FGF9/16/20 in the tip of the

tail. It remains unclear, however, whether repression in the primary notochord involves cryptic

combinatorial effects of LMX, Myb and/or GATA, or whether there are essential binding sites

for unknown factors that have yet to be identified in the −164 to −127 region.

## Discussion

Here we characterized the expression patterns and dependencies on Wnt and FGF signaling for three newly identified genes with strongly enriched expression in the posterior notochord. These fell into two classes: *C11.331* and *C12.115* were consistently differentially expressed between primary and secondary notochord, whereas *C8.891* showed a graded expression pattern at several stages extending past the primary/secondary boundary; *C11.331* and *C12.115* were strongly and consistently affected by manipulations of both Wnt and FGF signaling, whereas *C8.891* was strongly affected by ectopic Wnt activation but only modestly affected by FGF inhibition. This indicates that there may be at least two functionally distinct mechanisms of posterior-specific notochord gene regulation involving multiple signaling molecules expressed in the posterior tip of the tail.

Both *C11.331* and *C12.115* showed a surprisingly complex expression pattern when the canonical Wnt pathway was ectopically activated by BIO treatment. In both cases, the ectopic expression induced by BIO treatment is highly specific to notochord and is not seen in other tissues, confirming that there must be distinct tissue-specific inputs into the expression of these genes. Both expression patterns expanded throughout the notochord upon BIO treatment, but with distinct enrichment in most embryos at both the anterior and posterior tips. This is reminiscent of the wildtype expression pattern for *BCamL* which is similarly enriched in the anterior and posterior notochord. FGFs are expressed in the trunk as well as the posterior tail tip (Hudson and Yasuo, n.d, Imai et al., 2002), suggesting that they might play a key role in generating these bipolar expression patterns. This is supported by the weak uniform notochord expression of *C11.331* and *C12.115* seen in embryos treated with both BIO and U0126, which shows that induction of tip specific expression by ectopic Wnt pathway activation requires

functional FGF signaling. We predict that the combination of both Wnt and FGF signaling at the tail tip is essential for the robust expression of *C11.331* and *C12.115* in posterior notochord, but that *BCamL* likely responds to FGF alone. If this hypothesis is correct, *BCamL* expression should be reduced or eliminated by U0126 treatment but not affected by BIO.

We have also extensively probed the *cis*-regulatory architecture underlying the posterior-specific expression of *C11.331* and find evidence for extensively distributed enhancer function. There are multiple non-overlapping or minimally overlapping regions in the *C11.331* upstream intergenic region that confer posterior notochord-enriched expression in reporter assays. Relatively few 'shadow/redundant/distributed' enhancers have been identified to date in *Ciona* (Farley et al., 2016), but they are extremely common in other model systems (Barolo, 2012, Cannavò et al., 2016, Perry et al., 2011) and easily missed in studies focused on rapidly identifying minimal CRM regions. A recent study using ATAC-seq to identify candidate enhancer regions genome-wide suggests that they may be more common in *Ciona* as well (Madgwick et al., 2018).

The functional significance of shadow/redundant/distributed enhancers remains unclear, but multiple non-exclusive hypotheses involving the strength, robustness, tissue-specificity and evolvability of gene expression have been proposed (Barolo, 2012). While not extensively investigated here, our data are consistent with distributed enhancer function being important for both the strength and the tissue-specificity of expression. The smaller constructs that recapitulated the secondary enriched expression seen with the full length $-1565$ to $+13$ region tended to have greater ectopic expression and/or weaker expression overall (Fig. 3.3B).

While the *C11.331* upstream regions showed extensively redundant *cis*-regulatory activity, further dissections of the minimal $-488$ to $+13$ regulatory module identified discrete

regions with predominant roles in positive and negative regulation. These include a region between −165 and −127 essential for repressing expression in primary notochord, and a region between −245 and −165 inferred to contain essential elements required for both notochord-specific and ectopic expression. The −165 to −127 silencer region is of particular interest, because it implies that there is a key role for negatively-acting factors that prevent *C11.331* expression in primary notochord. We accordingly searched the −165 to −127 region for TFBSs of interest, but mutations in several predicted sites led to decreased expression in secondary notochord as opposed to increased expression in primary notochord. It remains unclear whether there are yet to be identified sites with clear repressive function, or whether some more complex regulatory scheme might be at work.

It was not clear from the initial *in situ* hybridization experiments whether *C11.331* and *C12.115* are truly specific to secondary notochord, or whether they are induced in the posterior notochord by FGFs and Wnts expressed in the posterior tail tip and coincidentally have an anterior limit of detectable expression that is closely aligned with the primary/secondary boundary. Both the chromogenic and fluorescent *in situs* made it clear that there is a graded quality to *C11.331* and *C12.115* expression patterns, which are consistently strongest in the posteriormost notochord cells and drop off rapidly towards the primary/secondary boundary. It was unclear, however, whether there was a sharp transition in expression levels at the primary/secondary boundary given the limited sensitivity and modest dynamic range of traditional (non-single molecule) *in situ* hybridization. Quantitative analysis of *C11.331* reporter constructs revealed, however, that although there is graded expression within the secondary notochord, there is a sharp and abrupt transition between very low expression levels in the primary notochord and much higher levels in secondary notochord (Fig. 3.4A). This suggests

that there are inputs into *C11.331* expression based on primary vs secondary notochord fate and that it is not solely based on proximity to Wnt and FGF signals in the tip of the tail.

Fig. 3.6 shows a provisional model for *C11.331* expression in the secondary notochord. FGF and Wnt signals from a signaling center at the tip of the tail help to induce regionalized *C11.331* expression, which is also dependent on notochord-specific transcriptional regulators (possibly including LMX1 and potential Myb family members) and an as-yet unidentified transcriptional repressor that is predicted to be differentially expressed in primary but not secondary notochord. The identification of this putative primary notochord repressor is of considerable interest, but it will also be important to determine if similar mechanisms are at work in other *C11.331* CRMs as well as in the regulatory regions for other regionally expressed notochord genes. The 1 kb upstream regions of *C12.115* and *C8.891* both contain predicted Ets, LMX, GATA and Myb sites that define candidate regions of interest.

## Methods

### *Ciona* husbandry and embryology

*Ciona robusta* (formerly known as *Ciona intestinalis* type A) (Pennati et al., 2015) were collected in San Diego and shipped to KSU by Marine Research and Educational Products Inc. (M-REP, San Diego, CA). Adult *Ciona* were maintained in a recirculating aquarium. Standard fertilization, dechorionation and electroporation protocols were used (Veeman et al., 2011). Staging is based upon the series of Hotta (Hotta et al., 2007).

Drug treatments: BIO (GSK inhibitor IX, CAS 667463–62-9, Sigma-Aldrich) was dissolved in DMSO at 5 mM and used at a final concentration of 5 micromolar. MEK1/MEK2 inhibitor U0126 (CAS109511-58, Sigma-Aldrich) was dissolved in DMSO at 4 mM and used at a final concentration of 4 micromolar. DMSO alone (at 1:1000 dilution) was used as a control. Drugs were added to artificial seawater (ASW) at Hotta stage 12 and embryos were grown in drug-treated ASW until fixation at Hotta stage 23.

### *in situ* hybridization

Probe synthesis, embryo collection, *in situ* hybridization and imaging were performed essentially as described previously (Reeves, 2014), except that embryos for fluorescent *in situs* were electroporated with 60 micrograms *Bra*>hCD4: mCherry (Gline et al., 2015) to visualize notochord cell boundaries post-hybridization. mCherry antibody (Biovision, #5993) was included at 1:500 during the anti-DIG-AP incubation. After SIGMA-FAST FastRed staining (Sigma-Aldrich; F4648) was stopped by washing 5 times in PBS-Tween20, embryos were incubated overnight at 4 °C with 1:1000 anti-rabbit AlexaFluor 488 (Invitrogen, A-11034). Embryos were washed 5 times in PBS-Tween20, then mounted, cleared, imaged and

fluorescence was quantified in cells 30–40 as described previously (Reeves et al., 2014). All fluorescence intensities were background subtracted and then normalized to a 0–1 scale. Expression data was centered on the primary/secondary boundary and scaled to a common length.

The *Brachyury in situ* probe was previously described (Reeves et al., 2014). Regions of *C11.331*, *C12.115* and *C8.891* were amplified from cDNA and cloned into pBSII-SK(-), using the following primers: *C11.331* (993 bp, Forward: ACGGGACTCACACAACTTCC, Reverse: GTCTCCAATCGCTTGCTGTT); *C12.115* (865 bp, Forward: ACGCCATTAACACCGGTTC, Reverse: CAAATGTTTAGAAAACTGATTTTGAC), *C8.891* (900 bp, Forward: GTGCTGATGCCAAGAATGC, Reverse: GTTTCACACAGCTGGTAGGC).

## Reporter cloning and mutagenesis

The genomic region spanning from −1565 to + 13 of KH2012: KH.C11.331.v1.A.SL2-1 was PCR amplified from genomic DNA (Forward primer: TAAAATGGCGCGCCCAGGTGCCACAAATAAACC; Reverse primer: CCTCCGTCTAGACCTATTTGTCCTTCTGAAATAACAG and cloned into the *Asc*I and *Xba*I sites of pX2 +bpFOG>UNC76: Venus (Stolfi et al., 2015). All subsequent reporters were subcloned from this original plasmid (see Table A.5 for primers and genomic coordinates). TFBS single mutants were generated through the Q5 Site-Directed Mutagenesis Kit (E0554S, NEB), following manufacturer directions. Mutagenesis primers are listed in Table A.5. The *Bra*-H2B: HA reporter was generated by standard Gateway cloning of pENTR-H2B into pSP72BSSPE-*Swa*I::RFA-HA (Roure et al., 2007) A 2.2 kb *Brachyury* enhancer was cloned into the plasmid's *Xho*I/*Hin*dIII sites.

**Reporter expression, staining, and imaging**

Fertilized dechorionated eggs were co-electroporated with 30 μg *Bra*>H2B: HA plasmid

and 60 μg *C11.331* reporter plasmid (region of interest-bpFOG>Venus). At least three separate

replicates were performed for each reporter. Embryos were fixed at Hotta stage 21–22, stained

and prepared for imaging as previously described (Carlson et al., 2015). GFP polyclonal

antibody (Invitrogen, A-11122) was used at 1:000 and HA-Tag mouse monoclonal antibody

(Cell Signaling Technology, 2367S) at 1:750. Secondary antibodies (Invitrogen; anti-mouse

AlexaFluor 488, A-11029 and anti-rabbit AlexaFluor 555, A-21420) were used at 1:1000, and

AlexaFluor 633 Phalloidin (Invitrogen, A22284) was used at 1:150. Embryo identities were

blinded prior to mounting and clearing embryos. Mounted embryos were imaged on a Zeiss 700

laser scanning confocal microscope, using a 40 × 1.3NA objective. Z-stacks of each embryo were

collected at a slice interval of 0.5 μm and a pixel size of 0.24 μm, using consistent settings for

laser power, PMT gain and scan speed. Embryos were selected for imaging based on expression

of the *Bra*>H2B: HA control in both primary and secondary notochord, but without respect for

the expression of the *C11.331* reporter plasmid.

**Quantitative reporter analysis**

Confocal stacks of embryos electroporated with the indicated reporter plasmids were

analyzed interactively using FIJI/ImageJ (Schindelin et al., 2012). The tail is rarely flat enough

or sufficiently parallel to the coverslip to visualize the entire notochord in one 2D image. To

allow a simple 2D analysis of reporter intensity, the notochord was first computationally

flattened in the Z axis to bring all the *Bra*>H2B: HA labeled nuclei into a single plane. This

involved reslicing the image to view the image volume as XZ or YZ slices, manually tracing a

polyline that followed the notochord in Z, and then reslicing again along the polyline to reconstruct a flattened plane through the full length of the notochord. Embryos that could not be cleanly resliced in this manner to generate a clear image of the embryo were rejected from quantitative analysis. A 40 pixel wide polyline was then traced along the notochord's curvature in X and Y to capture the full width of each labeled nucleus from cell 1 to cell 40. Intensity data from the Venus and HA reporter channels was collected along that line using the 'plot profile' function. All of these steps used standard FIJI/ImageJ tools.

We then derived a normalized expression metric implemented as a MATLAB function. Inputs to the function include the AP intensity profiles of Venus and HA, manually estimated values for Venus and HA background levels, and the inferred positions of the primary/secondary notochord boundary and the cell 36/37 boundary (anterior *versus* posterior secondary notochord). The intensity profiles were segmented into the three regions of interest. For each region, a normalized expression score was generated by dividing background-subtracted reporter>Venus intensity by control background-subtracted *Bra*>HA intensity, but only for the positions with greater than threshold levels of *Bra*>HA expression.

Wilcoxon signed-rank tests were used to compare expression levels between notochord regions for the quantitative analysis. ANOVA followed by Tukey's Honest Significant Difference post-hoc test was used to compare differences in expression between different reporter constructs for each notochord subregion to better control for multiple comparisons. All statistical tests used standard MATLAB functions. We used the Python Pandas, Numpy, Matplotlib and Seaborn packages for plotting and data visualization.

## Qualitative reporter analysis

Confocal stacks of embryos from the reporter dissections were scored for expression in the primary notochord, anterior four cells of the secondary notochord, and posterior four cells of the secondary notochord on a scale from zero (no visible expression) to four (oversaturated expression). Scores were assigned based on the cell in the notochord region with the strongest Venus expression that also expressed *Bra*>H2B: HA. Scores were also given for ectopic expression in CNS, epidermis, endoderm, muscle, and mesenchyme, but without the requirement for *Bra*>H2B: HA expression in the same cell. Little evidence was seen for distinct tissue-specific silencers outside the notochord, so a combined ectopic expression score was derived by averaging the scores across the 5 tissues examined. Scoring was performed separately and without knowledge of the specific secondary reporter plasmid by both MH and CB prior to reconciling any score differences and recording consensus scores for each embryo. For simpler presentation, we averaged the anterior and posterior secondary notochord scores into a single combined value.

## TFBS analysis

The predicted ETS, GATA and Myb sites were identified using TRANSFAC PWMs (Matys et al., 2006) *via* the LASAGNA web tool (http://biogrid-lasagna.engr.uconn.edu/lasagna_search/) (Lee and Huang, 2013). The LMX site was identified by manually searching the minimal region of interest for matches to the core sequence identified by SELEX against *Ciona* Lmx1A (Brozovic et al., 2018). We also searched for matches to SELEX-derived motifs for Brachyury (shared by Tbx2/3) and FoxAa but did not locate any within this interval.

**Resources**

Information on key resources are provided in Table 3.1

# Acknowledgements

# References

Barolo, S., 2012. Shadow enhancers: frequently asked questions about distributed cis-regulatory information and enhancer redundancy. Bioessays 34, 135-141.

Brannon, M., Gomperts, M., Sumoy, L., Moon, R.T., Kimelman, D., 1997. A beta-catenin/XTcf-3 complex binds to the siamois promoter to regulate dorsal axis specification in Xenopus. Genes Dev. 11, 2359-2370.

Brozovic, M., Dantec, C., Dardaillon, J., Dauga, D., Faure, E., Gineste M., Louis, A., Naville, M., Nitta, K.R., Piette, J., Reeves, W., Scornavacca, C., Simion, P., Vincentelli, R., Bellec, M., Aicha, S., Ben, Fagotto, M., Guéroult-Bellone, M., Haeussler, M., Jacox, E., Lowe, E.K., Mendez, M., Roberge, A., Stolfi, A., Yokomori, R., Brown, C.T., Cambillau, C., Christiaen, L., Delsuc, F., Douzery, E., Dumollard, R., Kusakabe, T., Nakai, K., Nishida, H., Satou, Y., Swalla, B., Veeman, M., Volff, J.-N., Lemaire, P., 2018. ANISEED 2017: extending the integrated ascidian database to the exploration and evolutionary comparison of genome-scale datasets. Nucleic Acids Res. 46, D718-D725.

Cannavò, E., Khoueiry, P., Garfield, D.A., Geeleher, P., Zichner, T., Gustafson, E.H., Ciglar, L., Korbel, J.O., Furlong, E.E.M., 2016. Shadow enhancers are pervasive features of developmental regulatory networks. Curr. Biol. 26, 38-51.

Carlson, M., Reeves, W., Veeman, M., 2015. Stochasticity and stereotypy in the Ciona notochord. Dev. Biol. 397, 248-256.

Corbo, J.C., Levine, M., Zeller, R.W., 1997. Characterization of a notochord-specific enhancer from the Brachyury promoter reigon of the ascidian, Ciona intestinalis. Development 124, 589-602.

Dumollard, R., Hebras, C., Besnardeau, L., McDougall, A., 2013. Beta-catenin patterns the cell cycle during maternal-to-zygotic transition in urochordate embryos. Dev. Biol. 384, 331-342.

Farley, E.K., Olson, K.M., Zhang, W., Rokhsar, D.S., Levine, M.S., 2016. Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. PNAS 113, 6508-6513.

Gainous, T.B., Wagner, E., Levine, M., 2015. Diverse ETS transcription factors mediate FGF signaling in the Ciona anterior neural plate. Dev. Biol. 399, 218-225.

Gline, S., Kaplan, N., Bernadskaya, Y., Abdu, Y., Christiaen, L., 2015. Surrounding tissues canalize motile cardiopharyngeal progenitors towards collective polarity and directed migration. Development 142, 544-554.

Hotta, K., Takahashi, H., Erives, A., Levine, M., Satoh, N., 1999. Temporal expression patterns of 39 Brachyury-downstream genes associated with notochord formation in the Ciona intestinalis embryo. Dev. Growth Differ. 41, 657-664.

Hotta, K., Takahashi, H., Asakura, T., Saitoh, B., Takatori, N., Satou, Y., Satoh, N., 2000. Characterization of Brachyury-downstream notochord genes in the Ciona intestinalis embryo. Dev. Biol. 224, 69-80.

Hotta, K., Mitsuhara, K., Takahashi, H., Inaba, K., Oka, K., Gojobori, T., Ikeo, K., 2007. SPECIAL issue research article A web-based interactive developmental table for the ascidian Ciona intestinalis, including 3D real-image embryo reconstructions: I. From fertilized egg to hatching larva. Dev. Biol. 236, 1790-1805.

Hudson, C., Yasuo, H., 2005. Patterning across the ascidian neural plate by lateral Nodal signalling sources. Development 132, 1199-1210.

Hudson, C., Yasuo, H., 2006. A signalling relay involving Nodal and Delta ligands acts during secondary notochord induction in Ciona embryos. Development 133, 2855-2864.

Hudson, C., Yasuo, H., n.d. C2.125 in situ hybridization [WWW Document]. ANISEED. URL (https://www.aniseed.cnrs.fr/aniseed/gene/show_expression?Unique_id=Cirobu.g00004295&offset=60) (Accessed 31 May 2018).

Hudson, C., Lotito, S., Yasuo, H., 2007. Sequential and combinatorial inputs from Nodal, Delta2/Notch and FGF/MEK/ERK signalling pathways establish a grid-like organization of distinct cell identities in the ascidian neural plate. Development 134, 3527-3537.

Hudson, C., Kawai, N., Negishi, T., Yasuo, H., 2013. β-Catenin-driven binary fate specification segregates germ layers in ascidian embryos. Curr. Biol. 23, 491-495.

Hudson, C., Sirour, C., Yasuo, H., 2016. Co-expression of Foxa.a, Foxd and Fgf9/16/20 defines a transient mesendoderm regulatory state in ascidian embryos. Elife, 5.

Ikuta, T., Satoh, N., Saiga, H., 2010. Limited functions of Hox genes in the larval development of the ascidian Ciona intestinalis. Development 137, 1505-1513.

Imai, K.S., Satou, Y., Satoh, N., 2002. Multiple functions of a Zic-like gene in the differentiation of notochord, central nervous system and muscle in Ciona savignyi embryos. Development 129, 2723-2732.

Imai, K.S., Hino, K., Yagi, K., Satoh, N., Satou, Y., 2004. Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. Development 131, 4047-4058.

Imai, K.S., Levine, M., Satoh, N., Satou, Y., 2006. Regulatory blueprint for a chordate embryo. Science 312, 1183-1187.

Jiang, D., Munro, E.M., Smith, W.C., 2005. Ascidian prickle regulates both mediolateral and anterior-posterior cell polarity of notochord cells. Curr. Biol. 15, 79-85.

José-Edwards, D.S., Kerner, P., Kugler, J.E., Deng, W., Jiang, D., Di Gregorio, A., 2011. The identification of transcription factors expressed in the notochord of Ciona intestinalis adds new potential players to the brachyury gene regulatory network. Dev. Dyn. 240, 1793-1805.

José-Edwards, D.S., Oda-Ishii, I., Kugler, J.E., Passamaneck, Y.J., Katikala, L., Nibu, Y., Di Gregorio, A., 2015. Brachyury, Foxa2 and the cis-regulatory origins of the notochord. PLoS Genet. 11, 1-16.

Kugler, J.E., Passamaneck, Y.J., Feldman, T.J., Beh, J., Regnier, T.W., Di Gregorio, A., 2008. Evolutionary conservation of vertebrate notochord genes in the ascidian Ciona intestinalis. Genesis 46, 697-710.

Lee, C., Huang, C.-H., 2013. LASAGNA-Search: an integrated web tool for transcription factor binding site search and visualization. Biotechniques, 54.

Madgwick, A., Magri, M.S., Dantec, C., Gailly, D., Fiuza, U.-M., Guignard, L., Hettinger, S., Gomez-skarmeta, J.L., Lemaire, P., 2018. Evolution of the embryonic cis-regulatory landscapes between divergent Phallusia and Ciona ascidians. bioRxiv.

Matyx, V., Kel-Margoulis, O.V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K., Voss, N., Stegmaier, P., Lewicki-Potapov, B., Saxel, H., Kel, A.E., Wingender, E., 2006. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res. 34, D108-D110.

Miya, T., Nishida, H., 2003. An Ets transcription factor, HrEts, is target of FGF signaling and involved in induction of notochord, mesenchyme, and brain in ascidian embryos. Dev. Biol. 261, 25-38.

Nishida, H., 1987. Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. III. Up to the tissue restricted stage. Dev. Biol. 121, 526-541.

Oda-Ishii, I., Kubo, A., Kari, W., Suzuki, N., Rothbächer, U., Satou, Y., 2016. A maternal system initiating the zygotic developmental program through combinatorial repression in the Ascidian Embryo. PLoS Genet. 12, e1006045.

Pennati, R., Ficetola, G.F., Brunetti, R., Caicci, F., Gasparini, F., Griggio, F., Sato, A., Stach, T., Kaul-Strehlow, S., Gissi, C., Manni, L., 2015. Morphological differences between larvae of the ciona intestinalis species complex: hints for a valid taxonomic definition of distinct species. PLoS One 10, e1022879.

Perry, M.W., Boettiger, A.N., Levine, M., 2011. Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. PNAS 108, 13570-13575.

Racioppi, C., Kamal, A.K., Razy-Krajka, F., Gambardella, G., Zanetti, L., di Bernardo, D., Sanges, R., Christiaen, L.A., Ristoratore, F., 2014. Fibroblast growth factor signalling controls nervous system patterning and pigment cell formation in Ciona intestinalis. Nat. Commun. 5, 4830.

Reeves, W., Thayer, R., Veeman, M., 2014. Anterior-posterior regionalized gene expression in the Ciona notochord. Dev. Dyn. 243, 612-620.

Reeves, W. M., Wu, Y., Harder, M.J., Veeman, M.T., 2017. Funcitonal and evolutionary insights from the Ciona notochord transcriptome. Development 144, 3375-3387.

Roure, A., Rothbächer, U., Robin, F., Kalmar, E., Ferone, G., Lamy, C., Missero, C., Mueller, F., Lemaire, P., 2007. A multicassette Gateway vector set for high throughput and comparative analyses in ciona and vertebrate embryos. PLoS One 2, e916.

Sakabe, E., Tanaka, N., Shimozono, N., Gojobori, T., Fujiwara, S., 2006. Effects of U0126 and fibroblast growth factor on gene expression profile in Ciona intestinalis embryos as revealed by microarray analysis. Dev. Growth Differ. 48, 391-400.

Satou, Y., Mineta, K., Ogasawara, M., Sasakura, Y., Shoguchi, E., Ueno, K., Yamada, L., Matsumoto, J., Wasserscheid, J., Dewar, K., Wiley, G.B., Macmil, S.L., Roe, B.A., Zeller, R.W., Hastings, K.E.M., Lemaire, P., Lindquist, E., Endo, T., Hotta, K., Inaba, K., 2008. Improved genome assembly and evidence-based global gene model set for the chordate Ciona intestinalis: new insight into intron and operon populations. Genom Biol. 9, R152.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676-682.

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., Zhou, T., Rohs, R., Honig, B., Bussemaker, H.H., Mann, R.S., 2011. Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. Cell 147, 1270-1282.

Stolfi, A., Ryan, K., Meinertzhagen, I.A., Christiaen, L., 2015. Migratory neuronal progenitors arise from the neural plate borders in tunicates. Nature 527, 371-374.

Timpl, R., Sasaki, T., Kostka, G., Chu, M.L., 2003. Fibulins: a versatile family of extracellular matrix proteins. Nat. Rev. Mol. Cell Biol. 4, 479-489.

Veeman, M.T., Smith, W.C., 2013. Whole-organ cell shape analysis reveals the developmental basis of ascidian notochord taper. Dev. Biol. 373, 281-289.

Veeman, M.T., Chiba, S., Smith, W.C., 2011. Ciona genetics. Vertebr. Embryogenesis, 401-422.

Yasuo, H., Hudson, C., 2007. FGF8/17/18 functions together with FGF9/16/20 during formation of the notochord in Ciona embryos. Dev. Biol. 302, 92-103.

Yasuo, H., Satoh, N., 1993. Function of vertebrate T gene. Nature 364, 582-583.

Zeller, R.W., Virata, M.J., Cone, A.C., 2006. Predictable mosaic transgene expression in ascidian embryos produced with a simple electroporation device. Dev. Dyn. 235, 1921-1932.

# Tables and Figures – Chapter 3

**Table A.1 Key Resources for Chapter 3**

| Reagent or resource | Source | Identifier |
|---|---|---|
| **Antibodies** | | |
| Rabbit polyclonal anti-mCherry | Biovision | Cat #5993 |
| Goat anti-rabbit polyclonal AlexaFluor488 | Invitrogen | Cat #A11034 |
| Goat anti-mouse polyclonal AlexaFluor488 | Invitrogen | Cat #A11029 |
| Goat anti-rabbit polyclonal AlexaFluor555 | Invitrogen | Cat #A22420 |
| Rabbit polyclonal anti-GFP | Invitrogen | Cat #A11122 |
| Mouse monoclonal anti-HA | Cell Signaling Technology | Cat #2367 S |
| **Bacterial and Virus Strains** | | |
| n/a | | |
| **Biological Samples** | | |
| n/a | | |
| **Chemicals, Peptides, and Recombinant Proteins** | | |
| GSK inhibitor IX (BIO) | Sigma-Aldrich | Cat #361550; CAS667463-62-9 |
| U0126 | Sigma-Aldrich | Cat #662005; CAS109511-58-2 |
| SIGMA-FAST FastRed | Sigma-Aldrich | Cat #F4648 |
| AlexaFluor633 Phalloidin | Invitrogen | Cat #A22284 |
| **Critical Commercial Assays** | | |
| n/a | | |
| **Deposited Data** | | |
| *Ciona* SELEX data | ANISEED | https://www.aniseed.cnrs.fr/ |

| Reagent or resource | Source | Identifier |
|---|---|---|
| **Experimental Models: Cell Lines** | | |
| n/a | | |
| **Experimental Models: Organisms/Strains** | | |
| Wild *Ciona robusta* | M-REP, San Diego, CA | N/A |
| **Oligonucleotides** | | |
| See Table A.1 | | |
| **Recombinant DNA** | | |
| *Bra* > hCD4:mCherry | Gline et al. (2015) | N/A |
| All *C11.331* > UNC76:Venus reporter plasmids | This paper | N/A |
| *Bra* > H2B:HA | This paper | N/A |
| **Software and Algorithms** | | |
| MATLAB | The Mathworks, Inc | Release 2012b |
| FIJI/ImageJ | Shindelin et al. (2012) | https://fiji.sc/ |
| LASAGNA-GRID | Lee and Huang (2013) | http://biogrid-lasagna.engr.uconn.edu/lasagna_search |
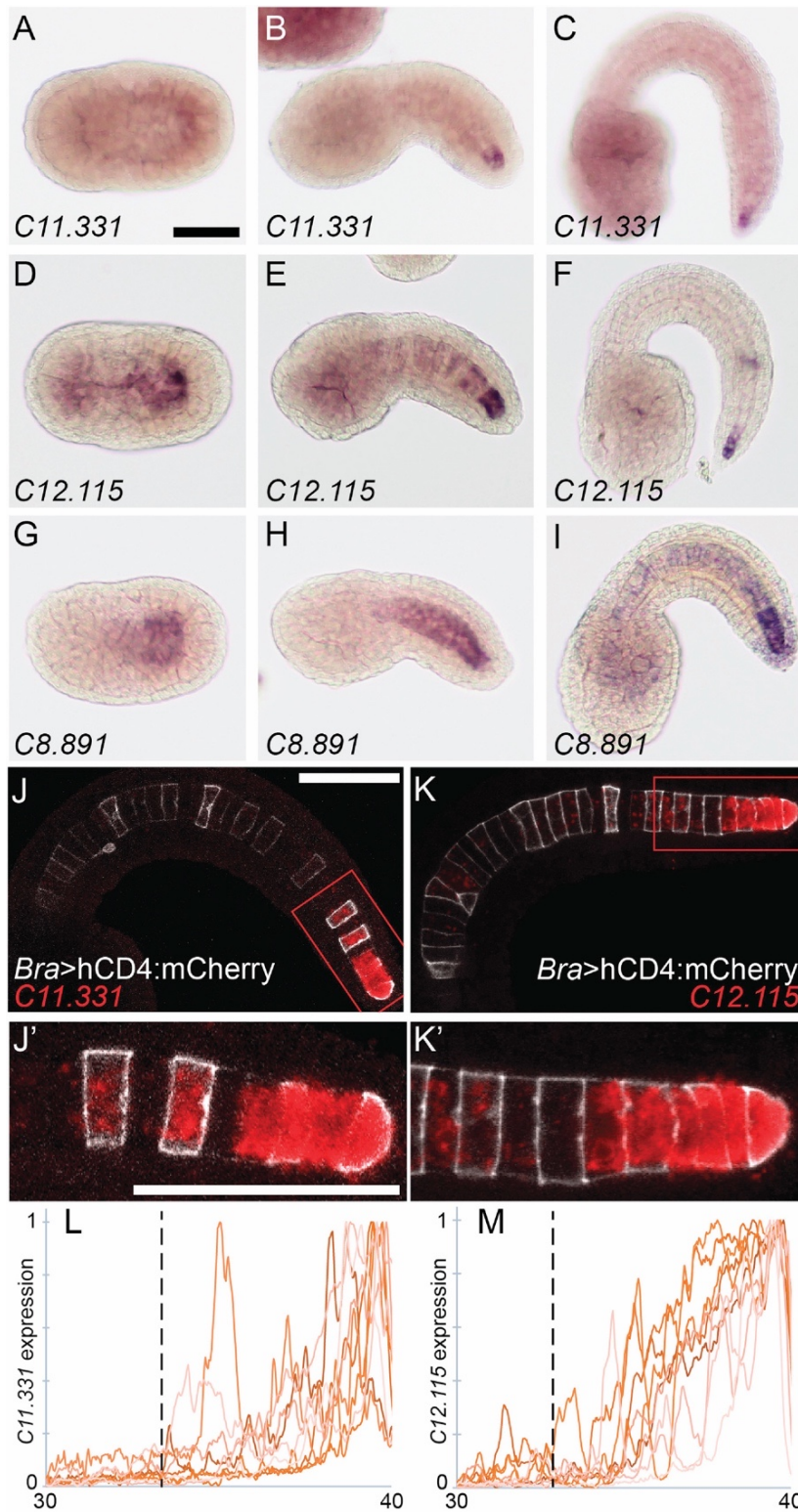
**Figure 3.1 Secondary-enriched notochord genes.**

*In situ* hybridization of probes against (**A-C**) *KH.C11.331*, (**D-F**) *KH.C12.115* and (**G-I**) *KH.C8.891*. Expression was tested during early intercalation at Hotta stage 16 (**A, D, G**), late in

intercalation at Hotta stage 20 (**B, E, H**) and during notochord elongation at Hotta stage 23 (**C, F, I**). (**J-K**) Fluorescent *in situ* hybridization against *C11.331* (**J**) and *C12.115* (**K**) in red at Hotta stage 22. Embryos were electroporated with *Bra*>hCD4: mCherry (white). (**J'-K'**) Closeup of secondary notochord region indicated by red box in (**J, K**). (**L-M**) Background subtracted and normalized expression of *C11.331* (**L**) and *C12.115* (**M**) in notochord cells 30–40. Nine embryos for each gene are shown. The location of the primary/secondary border is indicated by the vertical dotted line. Scale bars = 50 μm.
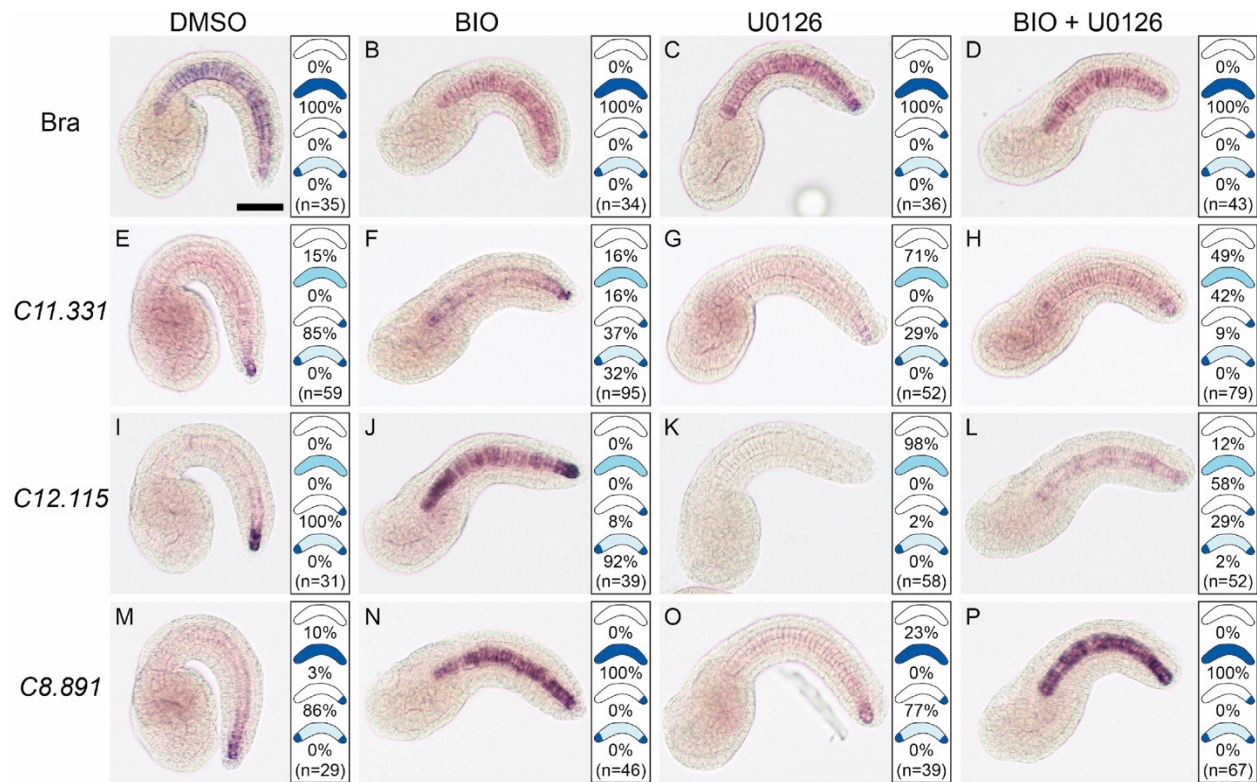
**Figure 3.2 Wnt and FGF signaling regulate posterior-enriched notochord expression.**
*In situ* hybridization against *Brachyury* (**A-D**), *C11.331* (**E-H**), *C12.115* (**I-L**) and *C8.891* (**M-P**)
on Hotta stage 23 embryos treated starting at mid gastrula (Hotta stage 12) with 1:1000 DMSO
(**A, E, I, M**), 5 µM Wnt activator BIO (**B, F, J, N**), 4 µM FGF pathway inhibitor U0126 (**C, G,
K, O**) or BIO + U0126 (**D, H, L, P**). Representative images of the most common expression
pattern for each gene/treatment are shown. Panels to the right of each image show the
distribution of expression patterns observed. Scale bar = 50 µm.
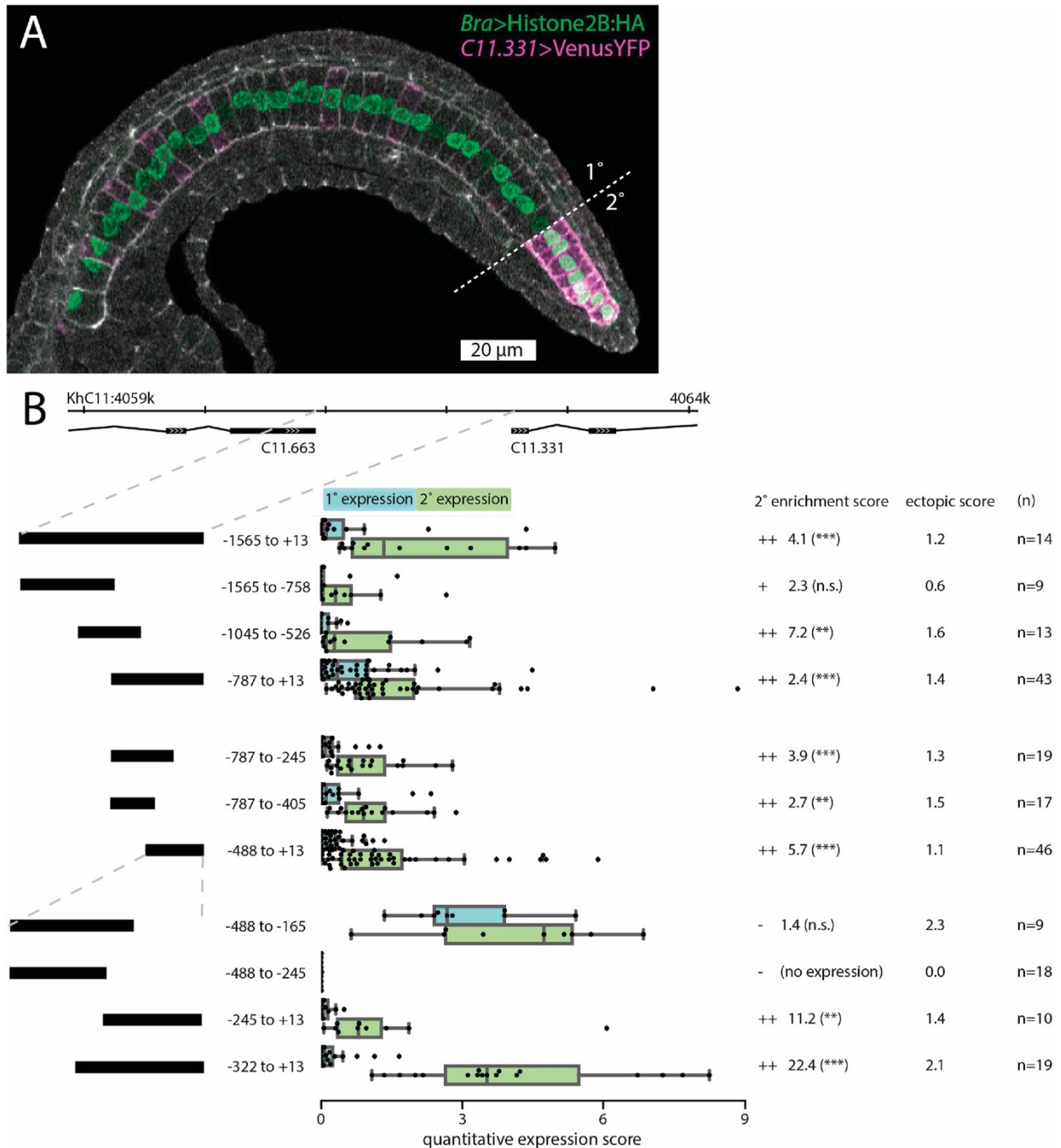
**Figure 3.3 *Cis*-regulatory analysis of the *C11.331* upstream intergenic region.**
(**A**) Representative embryo expressing *C11.331* (-787 to +13)>Venus and *Bra*>Histone2B: HA
fixed during tail elongation at Hotta stage 22. Cell cortices are labeled with Phalloidin-Alexa633
(white). Primary/secondary notochord boundary indicated with dashed line. Anterior is to the left
and dorsal is to the top of the image. (**B**) Summary statistics from reporter assays for the full-
length *C11.331* upstream region and a series of truncation mutants. Construct coordinates refer

to the predicted transcriptional start site for *C11.331*. *C11.331* quantitative expression score distributions are shown for primary (blue) and secondary (green) notochord. The superimposed box plots for each distribution show the median flanked by the interquartile range, with whiskers extending to the full range excluding outliers> 1.5X IQR. Expression differences between notochord regions for each reporter are shown as enrichment scores, calculated as the mean secondary expression divided by the mean primary expression. ++ indicates that the construct was enriched in secondary notochord with statistical significance< 0.05. + indicates that the construct showed signs of enrichment in secondary notochord but it was not statistically significant. – indicates that the construct was not enriched in secondary notochord. Secondary enrichment was tested by one-tailed Wilcoxon signed-rank tests (*, $p < 0.05$; **, $p < 0.005$; ***, $p < 0.0005$). The ectopic expression scores represent the average expression in 5 non-notochord tissues using a 0–4 qualitative expression scale described in the methods. For a tissue-specific breakdown of the scores, see Tables A.3,A.4.
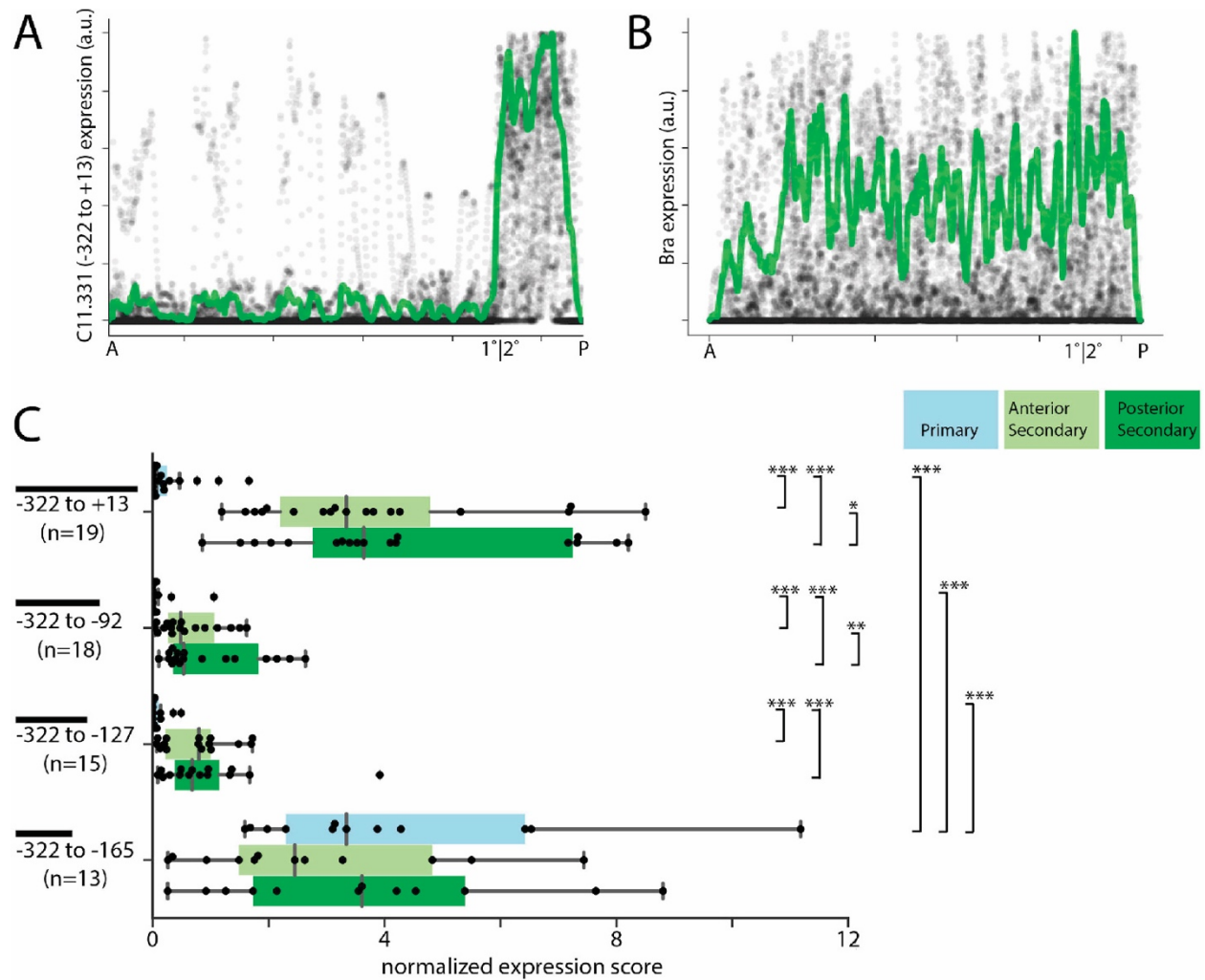
**Figure 3.4 Quantitative analysis of the *C11.331*(-322 to +13) CRM.**
(**A-B**) Quantitative traces of *C11.331* (-322 to +13) (**A**) and *Bra* (**B**) expression from 19 embryos as a function of anterior-posterior position. Expression data from each embryo was scaled in intensity to {0 1}, scaled to a common length, aligned by the primary/secondary notochord boundary and plotted as transparent gray to help visualize overlapping data points. Normalized mean intensity is overlaid in green. (**C**) *C11.331* quantitative expression score distributions are shown for primary (blue) anterior secondary (light green) and posterior secondary (dark green) notochord. The superimposed box plots for each distribution show the median flanked by the interquartile range, with whiskers extending to the full range excluding outliers> 1.5X IQR. Expression differences between notochord regions for a given reporter were tested by one-tailed Wilcoxon signed-rank tests. Expression differences for each notochord region between the

different reporter constructs were tested by ANOVA followed by Tukey's Honest Significant Difference post-hoc test (*, $p < 0.05$; **, $p < 0.005$; ***, $p < 0.0005$).
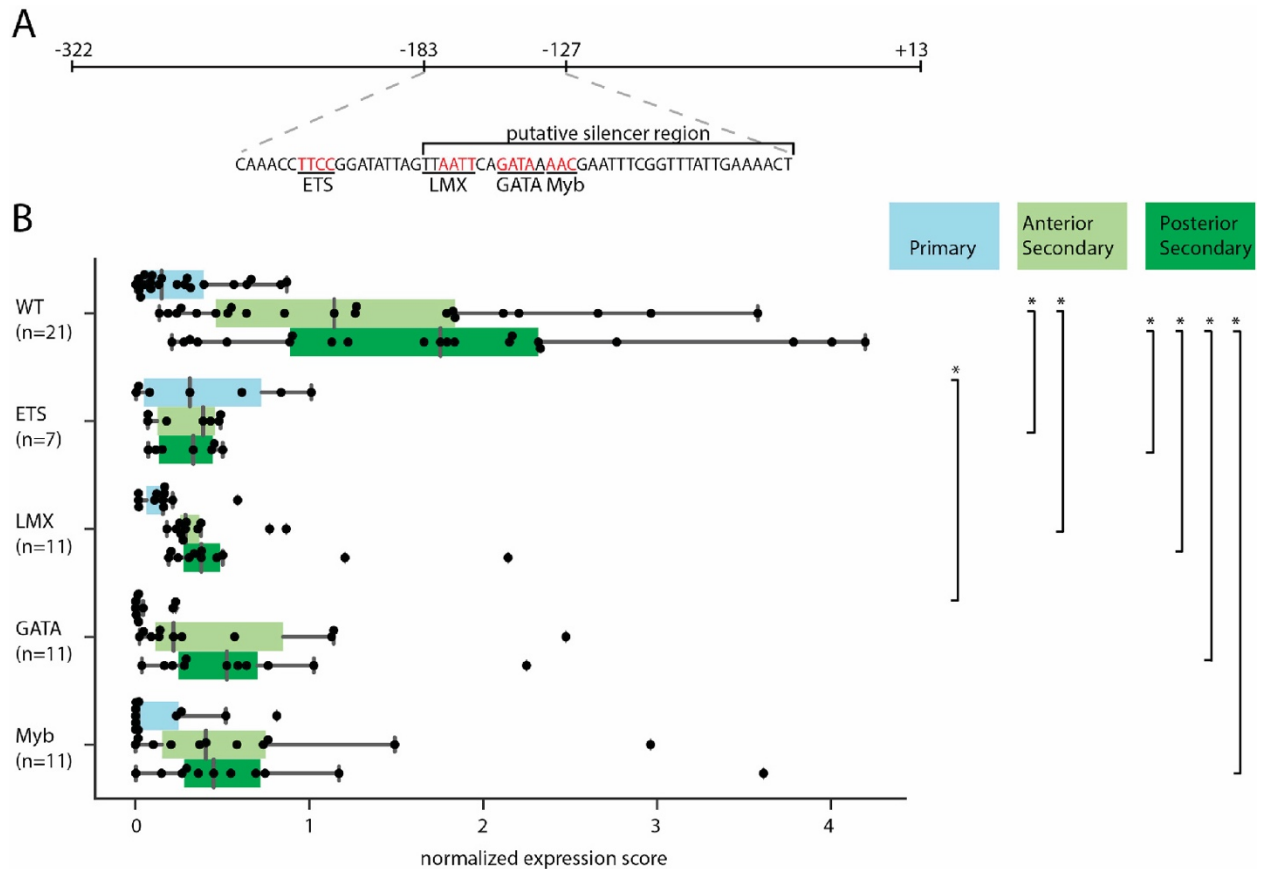
**Figure 3.5 Mapping TFBSs in the *C11.331* (-322 to +13) CRM.**

(**A**) Schematic of *C11.331* (-322 to +13) with the sequence of the −183 to −127 region shown and the −165 to −127 putative silencer region indicated. Core binding sites for the transcription factors ETS, LMX, GATA, and Myb are underlined, with the specific bases mutated in derivative constructs shown in red. (**B**) Quantitative analysis of regionalized notochord expression for the *C11.331*(-322 to +13) construct and derivatives in which each of these predicted TFBSs have been mutated. Quantitative expression scores for individual embryos are shown for primary notochord (blue) anterior secondary (light green) and posterior secondary (dark green). The superimposed box plots for each distribution show the median flanked by the interquartile range, with whiskers extending to the full range excluding outliers> 1.5X IQR. Between-reporter differences in expression for each notochord region were tested by ANOVA followed by Tukey's Honest Significant Difference post-hoc test (*, $p < 0.05$).
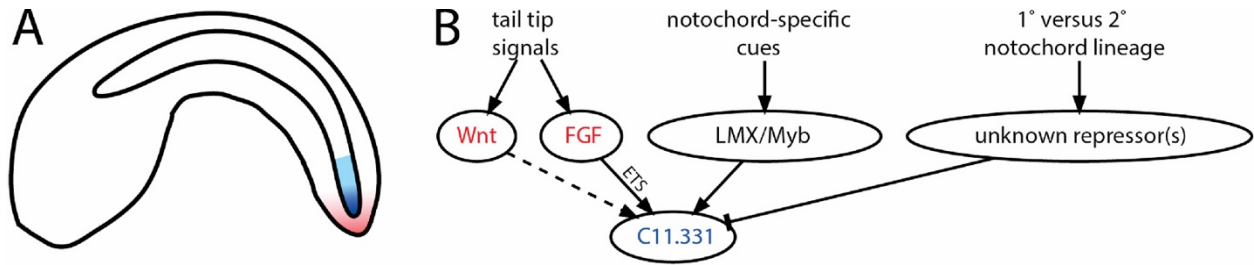
**Figure 3.6 Regulatory network model for *C11.331* expression.**

(**A**) Schematic of a *Ciona* embryo showing Wnt and FGF expression in the posterior tail tip (red) and *C11.331* expression in blue. *C11.331* expression is both highly enriched in secondary notochord compared to primary notochord and also graded within the secondary notochord. (**B**) Simple network model of inferred inputs into *C11.331* expression.

# Chapter 4 - *Ciona Brachyury* proximal and distal enhancers have different FGF dose-response relationships

Matthew J. Harder, Julie Hix, Wendy M. Reeves, and Michael T. Veeman*

*Corresponding author (veeman@ksu.edu)

# Abstract

Many genes are regulated by two or more enhancers that drive similar expression patterns. Evolutionary theory suggests that these seemingly redundant enhancers must have functionally important differences. In the simple ascidian chordate *Ciona*, the transcription factor Brachyury is induced exclusively in the presumptive notochord downstream of lineage specific regulators and FGF-responsive Ets family transcription factors. Here we exploit the ability to finely titrate FGF signaling activity via the MAPK pathway using the MEK inhibitor U0126 to quantify the dependence of transcription driven by different *Brachyury* reporter constructs on this direct upstream regulator. We find that the more powerful promoter-adjacent proximal enhancer and a weaker distal enhancer have fundamentally different dose-response relationships to MAPK inhibition. The Distal enhancer is more sensitive to MAPK inhibition but shows a less cooperative response, whereas the Proximal enhancer is less sensitive and more cooperative. A longer construct containing both enhancers has a complex dose-response curve that supports the idea that the proximal and distal enhancers are moderately super-additive. We show that the overall expression loss from intermediate doses of U0126 is not only a function of the fraction of cells expressing these reporters, but also involves graded decreases in expression at the single-cell level. Expression of the endogenous gene shows a comparable dose-response relationship to the full length reporter, and we find that different notochord founder cells are differentially sensitive to MAPK inhibition. Together, these results indicate that although the two *Brachyury* enhancers have qualitatively similar expression patterns, they respond to FGF in quantitatively different ways and act together to drive high levels of *Brachyury* expression with a characteristic input/output relationship. This indicates that they are fundamentally not redundant genetic elements.

# Introduction

Embryonic development depends on the precise spatial and temporal regulation of gene expression. Enhancers and other *cis*-regulatory elements embody the logic of the regulatory genome via their specific sets of binding motifs for different sequence-specific transcriptional activators and inhibitors (Zamudio et al., 2019). It is now clear that 'shadow,' 'distributed,' or 'redundant' enhancers are ubiquitous features of many genomes (Cannavò et al., 2016, Osterwalder et al., 2018, Madgwick et al., 2019). These terms refer to cases in which a gene has multiple non-overlapping regulatory elements that control seemingly identical expression patterns. The prevalence of these seemingly redundant elements suggests that they must have important separable functions or else they would not be evolutionarily conserved. Here we refer to these as 'distributed' enhancers to avoid implications about the relative importance of 'main' versus 'shadow' enhancers as well as implications as to whether they are truly redundant.

Distributed enhancers are thought to act together to drive high levels of gene expression that can buffer gene regulatory networks against problems resulting from stochastic transcriptional noise, mutation, or environmental perturbation. In support of this, there are several examples in *Drosophila* and vertebrates in which deletions of single distributed enhancers only show a phenotype when grown under heat stress or in sensitized genetic backgrounds. (Frankel et al., 2010, Antosova et al., 2016, Osterwalder et al., 2018). Predicted distributed enhancers in *Drosophila*, however, show increased sequence conservation compared to solitary enhancers (Cannavò et al., 2016), suggesting that they may have separable functions beyond jointly driving higher levels of expression of their regulated gene.

Distributed enhancers in some cases control similar but not completely overlapping expression patterns, with the differences in expression being functionally important. A related

idea is that cooperative and/or inhibitory interactions between multiple enhancers can create sharper boundaries of expression than single enhancers (Perry et al., 2011, El-Sherif and Levine, 2016). Some distributed enhancers have been shown to drive comparable expression patterns using fundamentally different *cis*-regulatory logic involving distinct upstream transcription factors (Wunderlich et al., 2015), which may provide another aspect of developmental robustness.

In addition to these questions about the overall roles of distributed enhancers, there are also major questions about how regulatory information from multiple enhancers becomes integrated into the expression of the regulated gene. Distributed enhancer pairs have been found to function in sub-additive, additive and super-additive regimes (Bothma et al., 2015), although sub-additive and additive relationships have been most common in the limited number of cases where this has been addressed. Scholes et al. (2019) recently found that the additivity of different combinations of *Krüppel* enhancers is not uniform as a function of the different concentrations of upstream activators present at different AP positions in the early *Drosophila* embryo. These particular enhancers are thought to respond to different combinations of upstream transcription factors, but it is also possible that distributed enhancers might have quantitatively different responses to the same upstream regulators.

A mechanistic understanding of how distributed enhancers work together to control gene expression depends on being able to quantify transcriptional input/output relationships in ways that are difficult in most model organisms. Quantitative analyses of distributed enhancer function have generally used the early *Drosophila* embryo as a model, where the AP patterning system provides natural gradients of key transcription factors (Perry et al., 2011, Bothma et al., 2015, El-

Sherif and Levine, 2016, Park et al., 2019, Scholes et al., 2019). In most cases, however, the quantitative details of distributed enhancer function are completely unknown.

The embryo of the ascidian tunicate *Ciona robusta* (formerly *Ciona intestinalis* Type A) provides an excellent alternate model system in which to study transcriptional input/output relationships. Thousands of transgenic embryos can be quickly obtained through simple electroporation methods (Corbo et al., 1997), allowing rapid dissection and analysis of *cis*-regulatory elements in reporter assays. Development proceeds via stereotyped and well-characterized lineages from fertilized egg to a swimming chordate tadpole larva with a muscular tail, dorsal neural tube, and notochord in less than 24 hours (Hotta et al., 2007). Its sequenced genome is only 1/20th the size of the human genome, with around 15 thousand genes (Dehal et al., 2002).

The T-box transcription factor Brachyury (Bra) is a major regulator of notochord fate in *Ciona* (Takahashi et al., 1999, Chiba et al., 2009, Reeves et al., 2017). Unlike in vertebrates, it is only expressed in the notochord and does not have broader roles in the posterior mesoderm (Yasuo and Satoh, 1993, Corbo et al., 1997, Yasuo and Satoh, 1998). *Bra* expression is induced in the presumptive notochord starting at the 64-cell stage through the intersection of lineage-specific transcription factors, including Zic-r.b (Yagi et al., 2004), and the activity of FGF-regulated Ets family transcription factors (Imai et al., 2006). Two *Bra* enhancers have been identified, including one proximal to the transcription start site (Corbo et al., 1997) and a more distal enhancer several hundred bp upstream (Farley et al., 2016). The *Bra* distal enhancer was the first reported shadow enhancer in *Ciona*, but a recent reporter assay survey of open chromatin regions suggests that they are quite common (Madgwick et al., 2019). Reporter constructs for both of these *Bra* enhancers have very similar expression patterns including

specific expression in both the primary and secondary notochord and ectopic expression in the mesenchyme (Corbo et al., 1997, Farley et al., 2016).

FGF is expressed on the vegetal side of the embryo under the control of maternal beta-catenin signaling, and plays a key role in establishing a large number of distinct vegetal and marginal cell fates including notochord (reviewed by Lemaire et al., 2008). The MEK inhibitor U0126 has been widely used in ascidians to interfere with these inductive interactions (Darras and Nishida, 2001, Kim and Nishida, 2001, Hudson et al., 2003, Sakabe et al., 2006, Hudson et al., 2007, Picco et al., 2007, Yasuo and Hudson, 2007, Pasini et al., 2012, Hudson et al., 2016) No off-target effects have been described and no other MAPK-dependent ligands have been identified at these stages (Imai et al., 2004, Imai et al., 2006). Simultaneous knockdown of FGF9/16/20 and FGF8/17/18 eliminates *Bra* expression similarly to U0126 treatment, confirming that U0126 effects on notochord fate are specific to the inhibition of the FGF signaling pathway (Yasuo and Hudson, 2007).

Previous studies using U0126 in ascidian embryos have always used high doses with the goal of completely blocking MAPK pathway activity. Here we systematically titrate MAPK pathway activity using finely graded doses to test the hypothesis that the *Bra* Proximal and Distal enhancers have quantifiably different input/output relationships. We find that the *Bra* Proximal and Distal enhancers each have a distinct, characteristic response to graded FGF pathway inhibition, and that they appear to act in a weakly super-additive fashion. Using *in situ* hybridization, we show that expression of endogenous *Bra* has largely similar responses to those seen in the reporter assays, but also that the different precursor cells of the notochord at gastrulation have subtle but detectable differences in their response to FGF inhibition.

# Results

## Qualitative responses to U0126

Recent ATACseq data (Madgwick et al., 2019) identifies several open chromatin regions upstream of the *Bra* transcription start site at timepoints that overlap the onset of *Bra* expression (Fig 4.1A). We cloned the three open chromatin regions closest to the transcriptional start site into a standard *Ciona* reporter vector containing a minimal basal promoter from *Ciona Friend of Gata* (bpFOG) and a Venus YFP reporter gene. We found that the peak overlapping the start site and the small peak ~500 bp upstream each drove reasonably strong reporter expression in the notochord and weaker expression in the mesenchyme. These correspond to previously established enhancer regions (Corbo et al., 1997, and Farley et al., 2016, respectively), and we refer to them here as the Bra 'Proximal' and 'Distal' enhancers. We also generated a Full Length reporter that included both enhancers and the genome sequence spanning the gap between them (Fig 4.1B). The third, even more distal ATACseq peak located around the KH2012.S1404: 4800-5100 bp positions did not drive reporter expression in any of our assays (not shown), but we cannot exclude it having some form of yet to be determined regulatory activity.

Bra expression requires both the transcription factor Zic-r.b as well as active FGF signaling mediated by Ets family transcription factors (Yagi et al., 2004, Imai et al., 2006, Fig 4.1C). These interactions are likely direct because mutation of predicted Zic and Ets binding sites in both the Proximal (Reeves and Shimai et al., 2020) and Distal (Farley et al., 2016) enhancers abrogates Bra expression. Essential Zic and Ets sites have also been identified in the *Bra* enhancer in the stolidobranch ascidian *Halocynthia* (Matsumoto et al., 2007). To identify quantitative differences between the *Bra* Proximal and Distal enhancers, we systematically titrated the FGF signaling pathway by the addition of varying doses of the MEK inhibitor U0126

at the 16-cell stage to embryos electroporated with the Proximal, Distal, or Full Length reporters.

Embryos were fixed at the early-tailbud stage (Hotta stage 19) (Hotta et al., 2007), then stained,

cleared and imaged *in toto* by confocal microscopy (Fig 4.1D). Embryos were stained by

antibody for the Venus reporter to provide a bright signal of reporter expression that was not

confounded by fluorescent protein maturation times and had minimal photobleaching. Embryos

were also stained with phalloidin to visualize embryonic morphology.

Electroporated transgene expression in *Ciona* is mosaic and subject to variable

transfection efficiency between different electroporations. We controlled for this in several ways.

We only included replicates where the capacitance reported by the electroporator in time

constant mode was between 900 and 1300 µFd. We imaged a random sample of ~10 embryos for

each construct/dose/replicate that was selected before looking at the reporter channel. We also

performed at least three independent biological replicates based on separate fertilizations and

electroporations for each construct/dose combination, and included a DMSO control treatment

for all electroporations.

Very low doses of U0126 have minimal effects on embryo morphology, but defects

become more common and pronounced as the dose increases, with severe notochord

malformation becoming frequent at the 0.34 µM dose, and notochord cells becoming almost

impossible to identify morphologically by the 0.63 µM dose. (Fig 4.1D). In embryos

electroporated with either the Proximal or Full Length reporters, expression persisted to some

extent even as the notochord itself became otherwise unrecognizable. Conversely, the Distal

reporter appeared far more sensitive to U0126 treatment, with expression becoming dramatically

reduced even at very low doses that did not produce frequent notochord malformation

phenotypes. At the highest doses of 1-4 µM, expression of all three reporters was essentially

eliminated. While we did not directly quantify any aspect of MAPK pathway activity across this range of doses, our assumption is that MAPK pathway activity is monotonically inhibited by U0126. A quantitative understanding of the U0126 dose/response curve with respect to ERK phosphorylation or other readouts of pathway activity is not needed to identify quantitative differences between the three *Bra* enhancer constructs.

### *Bra* enhancers have distinct dose-response relationships to U0126

To quantify reporter expression within each embryo, we summed total background-subtracted expression over the whole of each imaged embryo and normalized and scaled the data based on vehicle (DMSO) controls to account for variable electroporation efficiency (Fig 4.2A, details in Methods). We used bootstrap estimates for all statistical comparisons with no underlying assumptions about normality or homoskedasticity. We first compared the quantitative expression of the three constructs in the DMSO controls. This confirmed our qualitative assessment that the Proximal enhancer is stronger than the Distal enhancer, and also showed that the Full Length reporter expression is slightly higher than the sum of the Proximal and Distal expression values (Fig 4.2B). This difference was statistically significant and suggests that the Proximal and Distal enhancers are weakly synergistic with one another to drive higher levels of expression.

While enhancers are commonly conceptualized as Boolean logic gates that switch expression from OFF to ON based on the binary presence/absence of upstream factors, on a quantitative level this is thought to reflect fundamentally sigmoid transcriptional responses (Veitia, 2018). The functioning of gene regulatory networks is thus likely to be critically dependent on the quantitative details of these sigmoid transitions. As expected, we found that the

Proximal, Distal and Full Length constructs all show roughly sigmoid responses to graded

MAPK pathway inhibition (Fig 4.2C-G). All three dose-response curves were, however, quite

different from one another.

The Proximal and Distal enhancers both exhibited a simple monophasic relationship to

MAPK pathway inhibition (Fig 4.2C), though with apparent differences in sensitivity and

cooperativity. To quantify these differences, we fitted Hill functions (4-parameter logistic

curves) to the dose-response data. As the expression data is skewed and heteroskedastic, we

bootstrapped the nonlinear regression residuals to generate median parameter estimates and

confidence intervals (Supplemental Fig B.1). Two key parameters of these models are the $EC_{50}$

and the Hill Coefficient. The $EC_{50}$ is the drug concentration giving a half-maximal response, and

is a direct measure of sensitivity. The Hill Coefficient is an exponential term describing the

steepness of the transition between upper and lower plateaus, and is an implicit measure of

cooperativity in which higher absolute values indicate greater cooperativity.

As predicted from our qualitative assessment of the reporter imaging (Fig 4.1), the $EC_{50}$

values of the Proximal and Distal reporters were very different from one another (Fig 4.2D). The

Distal reporter's $EC_{50}$ of 0.0686 µM was less than $1/7^{th}$ of the $EC_{50}$ of the Proximal reporter, and

these differences had strong statistical support. The Hill Coefficients were also very different

(Fig 4.2D), with a median parameter estimate of -5.00 for Proximal vs. -1.43 for Distal. There

was again strong statistical support for these parameters being different. Together, these

differences in expression strength, sensitivity and cooperativity all indicate that the Proximal and

Distal enhancers are fundamentally different from one another in their quantitative responses to

MAPK pathway activity despite their dependence on the same Zic and Ets input factors.

The Full Length construct showed a more complex U0126 dose-response curve that suggested a potentially biphasic relationship. To investigate this, we fitted both a monophasic Hill function as well as a biphasic double Hill function to the data. We also fitted a simple linear relationship (Fig 4.2E). We used the Akaike Information Criterion (AIC) to estimate the relative likelihood of the three different models. The biphasic model was preferred, but the relative likelihood differences were modest (Fig 4.2F). Parameter estimates for the Hill Coefficients of the two phases of the biphasic fit were quite high but had very broad confidence intervals (Fig 4.2G, Supplemental Fig B.1C). A more finely graded series of U0126 doses would be needed to gain confidence as to whether these putative individual transitions are more or less cooperative than the single enhancer constructs. The $EC_{50}$ values of the two phases were quite close, however, to the $EC_{50}$ values of the Proximal and Distal constructs, supporting the idea that the two enhancers act in a quasi-additive fashion. A complete table of estimates of all bootstrap curve fit parameters can be found in Supplemental Table 1.

### Binary and graded responses at the single-cell level

The dose-response curves in Figure 4.2 are based on summed expression across the entire embryo, which is straightforward to automate and robust to the loss of embryonic morphology at higher doses of U0126 in which notochord cells cannot be reliably identified. Embryo-level measurements, however, may potentially obscure important differences in transcriptional responses at the level of individual cells. One possibility is that the transcriptional responses to MAPK inhibition might be more switch-like and quasi-Boolean at the single-cell level, and that the graded decreases in whole-embryo reporter expression seen at intermediate doses of U0126 might be largely a function of the fraction of cells in 'ON' versus 'OFF ' states. Alternatively,

there could be a uniform but graded loss of expression across all the notochord cells (Fig 4.3A). These 'switch' and 'fade' mechanisms are not mutually exclusive and might represent two extremes on a spectrum of possibilities that ultimately reflect the cooperativity of transcriptional responses at the single cell level.

To address this, we measured reporter expression in over 16,000 individual notochord cells from our confocal dose-response dataset at the DMSO, 0.1 μM, and 0.34 μM doses. These doses were selected to be as close as possible to $EC_{50}$ for the different reporter constructs while still being able to reliably identify all notochord cells for quantitation. Mean whole-embryo expression of the Distal reporter at 0.1 μM dose represents 51% of its expression in the DMSO controls. The Full Length construct is expressed at 66% of control expression at 0.1 μM U0126 and 61% at 0.34 μM. These doses were also at approximately the start and end of the putative intermediate plateau for this construct. The $EC_{50}$ for the Proximal construct is beyond the dose at which notochord cells become unrecognizable, so the mean expression at 0.34 μM is only modestly decreased to 80% of control expression.

For each embryo, we manually identified the midpoint of each notochord cell nucleus in the Z dimension and then positioned a circular region of interest (ROI) of fixed area on top of it. We used FIJI to measure various metrics of signal intensity from each nucleus midplane ROI, including the mean, median and modal grey values. Notochord cells were only measured if all notochord cells from the same embryo could be identified.

Given the mosaicism of electroporated transgene expression in *Ciona*, even DMSO control embryos typically have a considerable fraction of non-expressing cells and extensive variation in the brightness of expressing cells. We found that there was sufficient variation in both background intensity and the intensity of expressing cells that the mean gray value of each

circular ROI did not clearly distinguish between ON and OFF states (Fig 4.3B). The distribution of modal gray values of these ROIs, however, was distinctly bimodal and manual inspection of cells in the two categories confirmed that this metric cleanly separated high background non-expressing cells from weak but *bona fide* expressing cells.

The observed distributions of reporter intensity values are shown in Fig 4.3C (all cells) and Fig 4.3D (only ON cells). The distributions in Fig 4.3D have also been scaled to the proportion of cells that are ON relative to the DMSO dose, as shown in (Fig 4.3F). Differences between U0126 doses are apparent, but these distributions are relatively noisy and complex. To better understand the phenomena driving expression loss at the single-cell level, we first quantified the overall response of the three reporter constructs in terms of the mean expression of all measured cells to confirm that this cell-based metric was comparable to the automated whole-embryo reporter quantitation (Fig 4.3E). As expected, the Proximal reporter showed the least change over the doses measured, with the 0.34 µM dose having a small but significant loss of expression compared to DMSO or 0.1 µM, but no significant difference between DMSO and 0.1 µM. The Distal reporter had highly significant losses of expression at each increase in U0126 dose. The Full Length reporter had a large, highly significant loss in mean cell expression between the DMSO and 0.1 µM doses, but no significant difference between 0.1 and 0.34 µM doses. These findings were largely consistent with the whole-embryo data (Fig 4.2A), indicating that the embryo-level data are not biased by ectopic expression outside the notochord.

We next tested the possibility that expression loss at intermediate doses might involve a complete loss of expression in at least some cells by quantifying the fraction of cells that were inferred to be ON vs OFF based on the bimodal distribution of modal gray values. Analyses of the ON fraction were quite sensitive to experiment to experiment variation in electroporation

efficiency, so we calculated bootstrap confidence intervals based on changes in the ON fraction between matched DMSO controls and the 0.1 µM and 0.34 µM doses (Fig 4.3F). We found that both these doses caused a statistically significant (p<0.001) decrease in the proportion of cells detectably expressing all three reporter constructs. This effect was weaker in magnitude for the Proximal and Full Length reporters, which showed only a 10-20% decrease in the ON fraction in U0126-treated cells compared to DMSO controls. This effect was stronger in magnitude for the Distal reporter, which showed a ~40% decrease in the ON fraction at 0.1 µM U0126, and a ~55% decrease at 0.34 µM. This demonstrates that at least part of the graded decrease in bulk expression seen at intermediate U0126 doses is due to an apparently complete loss of expression in a subset of notochord cells.

We then tested the possibility that expression loss at intermediate doses might also involve a graded loss of expression in individual ON cells by quantifying both the mean and the 90th percentile of mean cell expression values of just the cells inferred to be ON (Fig 4.3G-H). We again used bootstrap estimates of these parameter values and their associated confidence intervals. The mean cell expression values of just the ON cells showed very similar trends at increasing U0126 doses to the changes in expression overall (Fig 4.3G). This indicates that the graded decreases in bulk expression seen at intermediate U0126 doses are at least in part a function of graded decreases in expression levels at the single-cell level and not merely a function of changes in the fraction of cells expressing the reporter at all. The 90th percentile values also showed similar trends, indicating that intermediate doses of U0126 cause graded decreases in the expression of even the brightest expressing cells (Fig 4.3H). It was also notable that the mean and 90th percentile cell expression of the Proximal reporter ON cells was close to or even higher than the mean or 90th percentile cell expression of the Full Length ON cells at all

measured doses. Together, it is clear from these analyses that the loss of reporter expression at increasing doses of U0126 involves both an increase in the fraction of cells that fail to detectably express the reporter and also a graded decrease in the intensity of expressing cells. This indicates that Bra reporter expression is not intrinsically Boolean over the time-scale examined. Graded responses to intermediate levels of MAPK pathway inhibition are seen at the level of single notochord cells.

### Endogenous *Bra* expression follows similar responses to U0126

While our reporter experiments revealed characteristic dose-response behaviors for each enhancer, reporter assays may be confounded by the lack of genomic context or normal chromatin structure. To test whether endogenous *Bra* expression has a similar U0126 dose-response behavior to the Full Length reporter construct, we treated unelectroporated embryos with a range of U0126 doses, fixing the embryos at the mid-gastrula stage (Hotta stage 12), and staining for *Bra* mRNA by *in situ* hybridization (Fig 4.4A-H). Fixing the embryos at mid-gastrula allows for each notochord precursor blastomere to be accurately identified, thus allowing us to score each cell for *Bra* expression on a semi-quantitative scale, ranging from 0 (no expression) through 3 (robust expression).

When the average scores from each embryo were rescaled to the same DMSO average of 1 as the Full Length quantitative reporter data and plotted together, the two dose-response curves were quite similar (Fig 4.4I). The in situ curve was not as distinctly biphasic as the Full Length reporter curve, but the 95% confidence intervals were largely overlapping and the in situ data showed the same trend of a modest decrease over lower doses followed by a precipitous decline

between 0.5 µm and 1 µm. This suggests that transient reporter assays are indeed a reasonable proxy for the effects of U0126 treatment on endogenous *Bra* expression.

Grouping the *in situ* scores into medial primary, lateral primary, and secondary notochord cell groups revealed small but clear differences in their responses to U0126, particularly at moderate doses between 0.25 µM and 1.0 µM (Fig 4.4J-L). The medial A8.5 and A8.6 cell pairs have significantly higher expression at 0.5 µM U0126 than the more lateral primary notochord precursors A8.13 and A8.14 and the secondary notochord precursor B8.6. (Fig 4.4K-L). This indicates that the medial primary notochord cells are less sensitive to MAPK pathway inhibition than the more lateral notochord cells. A heatmap visualization of the *Bra* in situ dose-response data also demonstrates this effect and confirms that the loss of expression at intermediate doses of U0126 involves both graded decreases in expression in some cells as well as an apparently complete loss in others (Fig 4.4M-R). This lateral-to-medial loss of expression is particularly evident when these heatmaps are compared to heatmaps in which the positions of the blastomeres have been randomly shuffled within each embryo (Fig 4.4M'-R'). These founder cell-specific differences would not have been evident in our reporter assays, which were performed at a later stage where the progeny of the 4 primary notochord founder blastomeres are not unambiguously distinguishable.

## Discussion

Some distributed enhancers are thought to respond to different combinations of direct upstream regulators (Wunderlich et al., 2015). Distributed enhancers that are dependent on different upstream activators could potentially increase the robustness of development to many types of genetic, stochastic and environmental perturbation, and could also help to shape more complex transcriptional input/output relationships. An alternate but not mutually exclusive possibility is that distributed enhancers may shape transcriptional input/output relationships by having quantitatively different responses to the same directly upstream transcription factors. Our results here strongly support this hypothesis, as we find that the *Bra* Proximal and Distal enhancers have fundamentally different quantitative responses in a transient reporter assay to the graded inhibition of a MAPK-dependent signal acting directly upstream. The Proximal enhancer is less sensitive to MAPK inhibition but shows a sharper, more cooperative response, whereas the Distal enhancer is more sensitive but shows a more graded response.

This study is subject to certain caveats. It is based on a transient reporter assay and not deletions of individual enhancers in the context of normal chromatin. It uses a protein reporter and not the MS2 RNA tagging system, which has not yet been implemented in *Ciona*. Pharmacological inhibition of MAPK-dependent FGF signaling will interfere not just with the direct induction of *Bra* by Ets family TFs but also any relevant FGF-dependent feedback or feedforward loops. These concerns apply mostly to questions of the additivity of the *Bra* Proximal and Distal enhancers. Although they appear to be slightly super-additive in our experiments, this was not our major focus and a more elaborate set of controls akin to Scholes et al. (2019) would be needed to fully explore this question. We have not, for example, excluded the possibility that the region between the Proximal and Distal enhancer elements might contain

relevant transcription factor binding sites. Our major conclusion, however, that the *Bra* Proximal and Distal enhancers have different dose-response relationships to MAPK pathway inhibition is largely robust to these concerns. It is clear that they act very differently from one another in this assay, regardless of whether the assay perfectly recapitulates all of the properties of the endogenous genetic elements.

Given that the *Bra* Proximal and Distal enhancers have quantifiably different responses to MAPK pathway inhibition, this raises intriguing questions about the functions of these different elements. One possibility is that they might be differentially involved in the initiation versus the maintenance of *Bra* expression. Different FGF ligands have been shown to have separable roles in *Bra* induction and maintenance (Yasuo and Hudson, 2007), so this is certainly plausible. This could potentially be tested by dose-response reporter assays in which U0126 was applied and embryos were fixed across different stages, or by CRISPR or morpholino disruption of individual FGF ligands. Another possibility is that the two enhancers might have subtly different roles in different notochord lineages, though both enhancers are clearly able to drive expression in both primary and secondary notochord.

Another question is why, on a mechanistic level, these two different enhancers have such distinct dose-response relationships to MAPK pathway inhibition. Small differences in the number, affinity, order and spacing of transcription factor binding motifs have been shown to have major effects on the strength and tissue-specificity of expression in *Ciona* (Bertrand et al., 2003, Passamaneck et al., 2009, Khoueiry et al., 2010, Roure et al., 2014, Farley et al., 2015), including various mutations of the *Bra* Distal enhancer (Farley et al., 2016). Differences in the number, affinity, order and spacing of Ets sites and/or binding motifs for other transcription factors acting cooperatively with Ets family TFs presumably account for the quantitative

differences between the two *Bra* enhancers. This could be tested experimentally but the combinatorial space of relevant mutations and treatments is quite large and it would benefit from a more scalable reporter assay than the *in toto* confocal imaging used here.

In addition to quantifying differences in the sensitivity and cooperativity of the *Bra* Proximal and Distal reporter constructs in response to MAPK pathway inhibition, we also explored the U0126 dose-response curve of endogenous *Bra* expression. This supported our conclusion from the reporter assays that the two enhancers acting together have a complex dose-response curve that is not well represented by a simple monophasic Hill function. Unexpectedly, this identified a differential sensitivity to U0126 between different notochord founder cells, with the medial primary notochord cells being particularly resistant to MAPK pathway inhibition. This could potentially involve differences in cell-cell contact surface areas between notochord founder cells and their FGF-expressing endodermal neighbors. Several cell fate decisions in the early *Ciona* embryo are thought to involve quantitative differences in cell contact surface between adjacent cells (Tassy et al., 2006), so it may be that transcriptional input/output relationships are quite finely tuned even in the compact and stereotyped *Ciona* embryo where inductive events are thought to largely involve direct cell contacts and not long-range gradients.

Given the ubiquity of distributed enhancers across different animal species that often have quite different mechanisms for cell fate specification, it is important to understand how distributed enhancers contribute to *cis*-regulatory logic in multiple contexts. It is possible, for example, that fundamental aspects of enhancer cooperativity may be different between cell fate decisions involving direct cell contacts versus the unusual long-range gradients of maternal TFs seen in insects with syncytial early development or other long-range gradients of signaling molecules commonly seen in vertebrate development. The simple but stereotypically chordate

*Ciona* embryo provides a new model for quantitative studies of *cis*-regulatory input/output

relationships.

## Methods

### *Ciona* husbandry and embryology

Adult *Ciona robusta* (formerly *C. intestinalis* type A) were collected from San Diego harbor and shipped to KSU by Marine Research and Educational Products, Inc. (M-REP, San Diego, CA), before being kept in a recirculating tank filled with artificial seawater (ASW) until use. Fertilized embryos were obtained by sacrificing 3 adult animals for their eggs and sperm, which were then mixed for *in vitro* fertilization, and immediately dechorionated using standard procedures (Veeman et al., 2011). Dechorionated embryos were grown in ASW treated with 0.1% Bovine Serum Albumin (ASW+BSA) to minimize clumping. Embryos were incubated at 19.5 degrees Celsius and staged according to (Hotta, 2007).

### Enhancer identification and cloning

Publicly available ATAC-seq data from *Ciona robusta* embryos (Madgwick et al., 2019) viewed through the genome browser on the ANISEED database (Dardaillon et al., 2020) were used to identify the boundaries of the enhancers used in this study. Enhancer regions were PCR-amplified from a 2.2 kb enhancer/promoter region of plasmid *Bra*>Rfa-Venus (Newman-Smith et al., 2015). Primer sequences (underline indicates genomic sequence): Distal Forward (ACGTCTCGAGTCATTGAGGTTTTGTCGCCC), Distal Reverse (ACGTAAGCTTCTCCCCTTTTTAGTTTGATTGATG), Proximal Forward (ACGTCTCGAGTCACAATACAAACAAAATATTTTGAC), Proximal Reverse (ACGTAAGCTTTATAGGTTTGTAACTCGCACTGAG), Third Forward (ACGTCTCGAGTGCTAGACCGCCATCGC), and Third Reverse (ACGTAAGCTTCCTAATGACGTCACGAAACG). The Full Length reporter was generated by

PCR amplification using the Distal Forward and Proximal Reverse primers. All PCR-amplified enhancer regions were cloned into the XhoI/HindIII sites of pX2+bpFOG>UNC76:Venus (Stolfi et al., 2015).

### Reporter assays

Fertilized, dechorionated embryos were electroporated with 45 µg *Bra*(Region of interest)-bpFOG>Venus, then washed into ASW+BSA. For drug treatments, plates of 3 mL ASW+BSA were supplemented with 1:1000 dilution of U0126 stock solutions (Sigma Cat. #662005-1MG) dissolved in DMSO to produce the indicated concentrations of U0126. Embryos were added to drug-treated seawater in 100 µl volumes at the 16-cell stage and fixed at the early tailbud I stage (Hotta Stage 19) in 2% paraformaldehyde/ASW overnight at room temperature. It was not feasible to split electroporated embryos between more than 8 drug treatments. The first round of experiments treated embryos with the doses between 0.1 µM and 4 µM. When it became clear that lower doses were needed, a second round of experiments with U0126 doses between 0.029 and 0.1 µM were run for each reporter. Because it was impractical to image more than 16 reporter/drug combinations in one experiment, the 3 different reporters were tested in overlapping sets of 2. At least three independent electroporations were performed for each reporter/drug dose combination.

Fixed embryos were stained using a rabbit polyclonal anti-GFP primary antibody (Fisher Cat. #A-11122) and goat anti-rabbit-555 secondary antibody (Fisher Cat. #A-21429). Phalloidin-488 (Fisher Cat. #A-12379) was used to stain cortical actin in the embryos. Stained embryos were mounted to poly-L-lysine-coated coverslips and cleared in Murray's Clear.

Embryos were imaged on a Zeiss 880 Laser Scanning Confocal Microscope using a 40X

1.3NA oil immersion objective under constant scan speed and laser power settings. Gain settings

for the 555 reporter channel were also held constant across all embryos. The imaging settings

were carefully optimized to have only minimal saturation of the brightest expressing cells while

still being able to detect very faint expression. Pixel sizes were set to 0.32 µm/pixel, and z slices

were made at an interval of 1.5 µm. All images were collected in 12-bit mode.


## Whole-embryo reporter quantification

Files containing the confocal stacks of embryos were passed to an in-house Python

function that subtracted a fixed background level, applied a light median filtering to the reporter

channel, and sum-projected the values of the phalloidin and reporter channels in Z. These

flattened images sometimes contained false signal from specks of dust or other embryos

intruding into the field of view. To ensure that we only quantified reporter signal from within

single embryos, the script thresholded the phalloidin channel and used binary morphology to

generate a binary mask approximating each embryo. All masks were individually checked, then

edited by hand in FIJI/ImageJ (Schindelin et al., 2012) if they did not accurately capture the

boundary of the embryo. A summed enhancer value (SEV) was calculated for each embryo by

summing all pixel values within each final mask region.

To account for variation of electroporation efficiency, the SEV scores of the DMSO

control embryos for each electroporation were averaged, and the individual SEV score of each

embryo in that same electroporation was divided by this average to obtain an experiment

normalized score. This experiment normalized score was then multiplied by the average SEV of

all DMSO control embryos for each reporter across all experiments to obtain a normalized scaled

SEV score (NSSEV), and expressed as a fraction of the Full Length DMSO mean NSSEV. These normalized expression values were used in all plotting and curve-fitting calculations. Bootstrapping of DMSO expression scores was performed by resampling normalized expression values of each reporter 1000 times, ensuring that scores were sampled equally from each electroporation. Sum of Proximal and Distal reporters was determined by adding the Proximal and Distal average normalized expression values once for each bootstrap replicate to obtain a set of 1000 sums. Differences between the Proximal/Distal and Sum/Full Length pairs were tested by independent sample t-test. 95% bootstrap confidence intervals were calculated as the 2.5th and 97.5th percentiles for all of the bootstrapped parameter estimates.

Curves describing linear, monophasic, or biphasic relationships were fit according to the following formulas, where y is normalized expression, and x is the log10 of the U0126 concentration (DMSO was coded as two log-steps lower than the lowest U0126 dose to avoid a log of zero error):

Linear: $y = mx + b$

(m is the slope and b is the y-intercept)

Monophasic: $y = \frac{A-D}{1+10^{(C-x)*B}} + D$

(A and D are the top and bottom plateaus, respectively, B is the Hill Coefficient, and C is the $EC_{50}$.)

Biphasic: $y = \frac{A-D}{1+10^{(C-x)*B}} + \frac{D-G}{1+10^{((F-x)*E)}} + G$

(A, D, and G are the top, middle, and bottom plateaus, respectively; B and C are the Hill Coefficient and $EC_{50}$ of the first phase, respectively; and E and F are the Hill Coefficient and $EC_{50}$ of the second phase, respectively.)

Nonlinear regression fitting monophasic, biphasic, and linear curves was performed using the optimize.curve_fit function of the Scipy Python library (Oliphant, 2007, Millman and Aivazis, 2011). We used minimally restrictive parameter bounds and crudely estimated parameter seeds to obtain a matrix of residuals from initial curve fits. These residuals were bootstrapped 1000 times to obtain bootstrap replicates, which were each used to fit a set of parameters by nonlinear regression. Differences in parameter values for Proximal and Distal reporter monophasic curves were determined by Wilcoxon Rank-Sum test. For Full Length reporter bootstrapping, bootstrap replicates that failed to be fit to a curve were dropped. Median values from each parameter distribution were used to plot the best fit curves for the Proximal and Distal reporters in Fig 4.2C.

**Individual cell measurement reporter quantification**

Individual notochord cell measurements were made by opening confocal stacks in FIJI/ImageJ, identifying the Z plane representing the approximate midpoint in Z of each notochord cell nucleus, and using the ROI Manager tool to manually place circular Regions of Interest (ROIs) with 10-pixel diameters over the nucleus of each notochord cell. ROI measurements included reporter channel mean, standard deviation, mode, maximum, minimum, median, and total expression values for each cell, and were aggregated across all measured embryos for each given dose/reporter combination. Mean expression values were normalized and scaled in the same manner as whole-embryo expression. We added a pseudocount of 1 prior to all log transformations of various expression metrics to avoid log of zero errors. Cells with a ROI mode pixel value of 10 or less were classed as being 'off' based on the bimodal distribution of the ROI modes and manual inspection of a subset of cells.

Bootstrapping of the ON/OFF ratios was performed by resampling cells from each U0126 treatment for each reporter 1000 times. Bootstrapping of matching DMSO-treated cells was performed by resampling matching numbers of DMSO cells from each experiment in the corresponding U0126 dose, generating a unique matched control set for each U0126/reporter combination for each bootstrap replicate. The ON ratio was calculated as the proportion of measured cells that were on in the U0126 bootstrap replicate divided by the same proportion in the matching DMSO replicate. Differences between these ratios and a ratio of one were tested by t-tests comparing each reporter/dose combination's bootstrap distribution to 1.

Bootstrapping of normalized mean expression values was performed by resampling cells at each dose 1000 times, ensuring that the number of cells resampled from each electroporation equaled the number of cells measured in the same electroporation, bootstrapping separately for the ALL cells and ON cells only groups. Statistical tests for differences between doses for the mean and $90^{th}$ percentile values were performed by matching bootstrap replicates for each pairwise U0126 dose comparison within each reporter, counting the number of cases in which expression at the higher dose exceeded expression at the lower dose, and dividing by 1000 to obtain a p-value.

### *in situ* hybridization

Fertilized, dechorionated embryos were allowed to develop in ASW+BSA at 19.5 degrees Celsius until the 16-cell stage, at which point they were drug treated in the same manner as for the reporter assays. Embryos were fixed in MEM-PFA at Mid-Gastrula (Hotta stage 12) on ice 10 minutes before overnight storage at 4 degrees. Probe synthesis and *in situ* hybridization was conducted as in (Reeves et al., 2014). Mounted embryos were visualized on an Olympus

BX61WI compound microscope using a 10X 0.3NA objective and imaged with a Canon EOS Rebel T3i digital camera under constant illumination conditions. Images were opened in FIJI/ImageJ and each notochord precursor cell was scored on a whole-number scale from 0 (no expression) to 3 (robust expression). Line plots and bar plots show mean scores with 95% confidence intervals. Differences in mean expression scores were calculated by one-way ANOVA followed by Tukey's HSD, looking within each dose for differences between cell pair scores.

Expression heatmaps were generated by ordering embryos from highest average expression to lowest average expression across all ten precursor cells. Embryos that had uniform scores across all ten precursor cells were not included in the heatmaps. Randomized heat maps were generated by shuffling the position of each precursor cell without replacement in each embryo.

## Data visualization and analysis

All statistical tests, plotting, curve fitting and visualization of quantitative expression data were performed in Python, using standard Pandas, Numpy, Matplotlib, Scipy, and Seaborn packages.

# Acknowledgements

This chapter is currently in preparation for submission for publication.

# References

Antosova, B., Smolikova, J., Klimova, L., Lachova, J., Bendova, M., Kozmikova, I., Macon, O., Kozmik, Z., 2016. The Gene Regulatory Network of Lens Induction Is Wired through Meis-Dependent Shadow Enhancers of Pax6. PLoS Genetics 12, e1006441.

Bertrand, V., Hudson, C., Caillol, D., Popovici, C., Lemaire, P., 2003. Neural Tissue in Ascidian Embryos Is Induced by FGF9/16/20, Acting via a Combination of Maternal GATA and Ets Transcription Factors. Cell 115, 615-627.

Bothma, J.P., Garcia, H.G., Ng, S., Perry, M.W., Gregor, T., Levine, M., 2015. Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo. eLife 4, e07956.

Cannavò, E., Khoueiry, P., Garfield, D.A., Geeleher, P., Zichner, T., Gustafson, E.H., Ciglar, L., Kornberg, J.O., Furlong, E.E.M., 2016. Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks. Curr. Biol. 26, 38-51.

Chiba, S., Jiang, D., Satoh, N., Smith, W., 2009. Brachyury null mutant-induced defects in juvenile ascidian endodermal organs. Development 136, 35-39.

Corbo, J., Levine, M., Zeller, R., 1997. Characterization of a notochord-specific enhancer from the Brachyury promoter region of the ascidian, Ciona intestinalis. Development 124, 589-602.

Dardaillon, J., Dauga, D., Simion, P., Faure, E., Onuma, T.A., DeBiasse, M.B., Louis, A., Nitta, K.R., Naville, M., Besnardeau, L., Reeves, W., Wang, K., Fagotto, M., Guéroult-Bellone, M., Fujiwara, S., Dumollard, R., Veeman, M., Volff, J.-N., Crollius, H.R., Douzery, E., Ryan, J.F., Davidson, B., Nishida, H., Dantec, C., Lemaire, P., 2020. ANISEED 2019: 4D exploration of genetic data for an extended range of tunicates. Nucl. Acids Research 48, D668-D675.

Darras, S., Nishida, H., 2001. The BMP signaling pathway is required together with the FGF pathway for notochord induction in the ascidian embryo. Development 128, 2629-2638.

Dehal, P., Satou, Y., Campbell, R.K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D.M., Harafuji, N., Hastings, K.E.M., Ho, I., Hotta, K., Huang, W., Kawashima, T., Lemaire, P., Martinez, D., Meinertzhagen, I.A., Necula, S., Nonaka, M., Putnam, N., Rash, S., Saiga, H., Satake, M., Terry, A., Yamada, L., Wang, H.-G., Awazu, S., Azuma, K., Boore, J., Brando, M., Chin-bow, S., DeSantis, R., Doyle, S., Francine, P., Keys, D.N., Haag, S., Hayashi, H., Hino, K., Imai, K.S., Inaba, K., Kano, S., Kobayashi, K., Kobayashi, M., Lee, B.-I., Makabe, K.W., Manohar, C., Matisse, G., Medina, M., Mochizuki, Y., Mount, S., Morishita, T., Maura, S., Nakayama, A., Nishizaka, S., Nomoto, H., Ohta, F., Oisin, K., Rigoutsos, I., Sano, M., Sasaki, A., Sasakura, Y., Shoguchi, E., Shin-i, T., Spagnuolo, A., Stainier, D., Suzuki, M.M., Tassy, O., Takatori, N., Tokuoka, M., Yagi, K., Yoshizaki, F., Wanda, S., Zhang,

C., Hyatt, P.D., Larimer, F., Detter, C., Doggett, N., Glavina, T., Hawkins, T., Richardson, P., Lucas, S., Kohara, Y., Levine, M., Satoh, N., Rokhsar, D.S., 2002. The Draft Genome of Ciona intestinalis: Insights into Chordate and Vertebrate Origins. Science 298, 2157-2167.

El-Sherif, E., Levine, M., 2016. Shadow Enhancers Mediate Dynamic Shifts of Gap Gene Expression in the Drosophila Embryo. Current Biol. 26, 1164-1169.

Farley, E.K., Olson, K.M., Zhang, W., Brandt, A.J., Rokhsar, D.S., Levine, M.S., 2015. Suboptimization of developmental enhancers. Science 350, 325-328.

Farley, E.K., Olson, K.M., Zhang, W., Rokhsar, S.D., Levine, M.S., 2016. Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. PNAS 113, 6508-6513.

Frankel, N., Davis, G.K., Vargas, D., Wang, S., Payre, F., Stern, D.L., 2010. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. Nature 466, 490-493.

Hotta, K., Mitsuhara, K., Takahashi, H., Inaba, K., Oka, K., Gojobori, T., Ikeo, K., 2007. A Web-Based Interactive Developmental Table for the Ascidian Ciona intestinalis, Including 3D Real-Image Embryo Reconstructions: I. From Fertilized Egg to Hatching Larva. Dev. Dyn. 236, 1790-1805.

Hudson, C., Darras, S., Caillol, D., Yasuo, H., Lemaire, P., 2003. A conserved role for the MEK signalling pathway in neural tissue specification and posteriorisation in the invertebrate chordate, the ascidian Ciona intestinalis. Development 130, 147-159.

Hudson, C., Lotito, S., Yasuo, H., 2007. Sequential and combinatorial inputs from Nodal, Delta2/Notch and FGF/MEK/ERK signalling pathways establish a grid-like organization of distinct cell identities in the ascidian neural plate. Development 134, 3527-3537.

Hudson, C., Sirour, C., Yasuo, H., 2016. Co-expression of Foxa.a, Foxd, and Fgf9/16/20 defines a transient mesendoderm regulatory stat in ascidian embryos. eLife 5, e14692.

Imai, K.S., Hino, K., Yagi, K., Satoh, N., Satou, Y., 2004. Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. Development 131, 4047-4058.

Imai, K.S., Levine, M., Satoh, N., Satou, Y., 2006. Regulatory Blueprint for a Chordate Embryo. Science 312, 1183-1187.

Khoueiry, P., Rothbächer, U., Ohtsuka, Y., Daian, F., Frangulian, E., Roure, A., Dubchak, I., Lemaire, P., 2010. A cis-Regulatory Signature in Ascidians and Flies, Independent of Transcription Factor Binding Sites. Curr. Biol. 20, 792-802.

Kim, G.J., Nishida, H., 2001. Role of the FGF and MEK signaling pathway in the ascidian embryo. Develop. Growth Differ. 43, 521-533.

Lemaire, P., Smith, W.C., Nishida, H., 2008. Ascidians and the Plasticity of the Chordate Developmental Program. Curr. Biol. 18, 620-631.

Madgwick, A., Magri, M.S., Dantec, C., Gailly, D., Fiuza, U.-M., Guignard, L., Hettinger, S., Gomez-Skarmeta, J.L., Lemaire, P., 2019. Evolution of embryonic cis-regulatory landscapes between divergent Phallusia and Ciona ascidians. Dev. Biol. 448, 71-87.

Matsumoto, J., Kumano, G., Nishida, H., 2007. Direct activation by Ets and Zic is required for initial expression of the Brachyury gene in the ascidian notochord. Dev. Biol. 306, 870-882.

Millman, K.J., Aivazis, M., 2011. Python for scientists and engineers. Comput. Sci. & Eng. 13, 9-12.

Newman-Smith, E., Kourakis, M., Reeves, W., Veeman, M., Smith, W.C., 2015. Reciprocal and dynamic polarization of planar cell polarity core components and myosin. Elife 4, e05361.

Oliphant, T.E., 2007. Python for Scientific Computing. Comput. Sci. & Eng. 9, 10-20.

Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B.J., Afzal, S.Y., Lee, E.A., Zhu, Y., Placzek-Frick, I., Pickle, C.S., Kato, M., Garvin, T.H., Pham, Q.T., Harrington, A.N., Akiyama, J.A., Afzal, V., Lopez-Rios, J., Dickey, D.E., Visel, A., Pennacchio, L.A., 2018. Enhancer redundancy provides phenotypic robustness in mammalian development. Nature 554, 239-243.

Park, J., Estrada, J., Johnson, G., Vincent, B., Ricci-Tam, C., Bragdon, M., Shulgina, Y., Cha, A., Wunderlich, Z., Gunawardena, J., DePace, A., 2019. Dissecting the sharp response of a canonical developmental enhancer reveals multiple sources of cooperativity. eLife 8, e41266.

Pasini, A., Manenti, R., Rothbächer, U., Lemaire, P., 2012. Antagonizing Retinoic Acid Dan FGF/MAPK Pathways Control Posterior Body Patterning in the invertebrate Chordate Ciona intestinalis. PLoS ONE 7, e46193.

Passamaneck, Y.J., Katikala, L., Perrone, L., Dunn, M.P., Oda-Ishii, I., Di Gregorio, A., 2009. Direct activation of a notochord cis-regulatory module by Brachyury and FoxA in the ascidian Ciona intestinalis. Development 136, 3679-3689.

Perry, M., Boettiger, A., Levine, M., 2011. Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. PNAS 108, 13570-13575.

Picco, V., Hudson, C., Yasuo, H., 2007. Ephrin-Eph signalling drives the asymmetric division of notochord/neural precursors in Ciona embryos. Development 134, 1491-1497.

Reeves, W., Thayer, R., Veeman, M., 2014. Anterior-Posterior Regionalized Gene Expression in the Ciona Notochord. Dev. Dyn. 243, 612-620.

Reeves, W.M., Wu, Y., Harder, M.J., Veeman, M.T., 2017. Functional and evolutionary insights from the Ciona notochord transcriptome. Development 144, 3375-3387.

Reeves, W.M., Shimai, K., Winkley, K.M., Veeman, M.T., 2020. Brachyury controls Ciona notochord fate as part of a feedforward network and not as a unitary master regulator. bioRxiv, https://doi.org/10.1101/2020.05.29.12402.

Roure, A., Lemaire, P., Darras, S., 2014. An Otx/Nodal Regulatory Signature for Posterior Neural Development in Ascidians. PLoS Genetics 10, e1004548.

Sakabe, E., Tanaka, N., Shimozono, N., Gojobori, T., Fujiwara, S., 2006. Effects of U0126 and fibroblast growth factor on gene expression profile in Ciona intestinalis embryos as revealed by microarray analysis. Dev. Growth and Differ. 48, 391-400.

Schindelin, J., Armando-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueben, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 767-682.

Scholes, C., Biette, K., Harden, T., DePace, A., 2019. Signal Integration by Shadow Enhancers and Enhancer Duplications Varies across the Drosophila Embryo. Cell Reports 26, 2407-2418.

Stolfi, A., Ryan, K., Meinertzhagen, I.A., Christiaen, L., 2015. Migratory neuronal progenitors arise from the neural plate borders in tunicates. Nature 527, 371-374.

Takahashi, H., Hotta, K., Erives, A., Di Gregorio, A., Zeller, R.W., Levine, M., Satoh, N., 1999. Brachyury downstream notochord differentiation in the ascidian embryo. Genes and Dev. 13, 1519-1523.

Tassy, O., Daian, F., Hudson, C., Bertrand, V., Lemaire, P., 2006. A Quantitative Approach to the Study of Cell Shapes and Interactions during Early Chordate Embryogenesis. Curr. Biol. 16, 345-358.

Veeman, M.T., Chiba, S., Smith, W.C., 2011. Ciona Genetics. Methods in Mol. Biol. 770, 401-422.

Veitia, R.A., 2018. Dosage effects in morphogenetic gradients of transcription factors: insights from a simple mathematical model. Journal of Gen. 97, 365-370.

Wunderlich, Z., Bragdon, M.D.J., Vincent, B.J., White, J.A., Estrada, J., DePace, A.H., 2015. Krüppel Expression Levels Are Maintained through Compensatory Evolution of Shadow Enhancers. Cell Rep. 12, 1740-1747.

Yagi, K., Satou, Y., Satoh, N., 2004. A zinc finger transcription factor, ZicL, is a direct activator of Brachyury in the notochord specification of Ciona intestinalis. Development 131, 1279-1288.

Yasuo, H., Hudson, C., 2007. FGF8/17/18 functions together with FGF9/16/20 during formation of the notochord in Ciona embryos. Dev. Biol. 302, 92-103.

Yasuo, H., Satoh, N., 1993. Function of vertebrate T gene. Nature 364, 582-583

Yasuo, H., Satoh, N., 1998. Conservation of the developmental role of Brachyury in notochord formation in a Urochordate, the ascidian Halocynthia roretzi. Dev. Biol. 200, 158-170

Zamudio, A.V., Dall'Agnese, A., Heyningen, J.E., Mantegna, J.C., Afeyan, L.K., Hannett, N.M., Coffey, E.L., Li, C.H., Oksuz, O., Sabari, B.R., Boija, A., Klein, I.A., Hawken, S.W., Spille, J.-H., Decker, T.-M., Cisse,, I.I., Abraham, B.J., Lee, T.I., Taatjes, D.J., Schuijers, J., Young, R.A., 2019. Mediator Condensates Localize Signaling Factors to Key Cell Identity Genes. Mol. Cell 76, 1-14.

# Figures – Chapter 4

A

5000    5500    6000  KH2012.S1404

ATAC-seq (112-cell)

ATAC-seq (late gastrula)

ATAC-seq (mid neurula)

Brachyury transcript models

B

Proximal    bpFOG>Venus

Distal    bpFOG>Venus

Distal    Proximal    bpFOG>Venus

C

FGF
↓
MEK ⊢ (U0126)
↓
Zic    Ets
↓    ↓
Bra

D

|  | DMSO | 0.029 uM | 0.054 uM | 0.10 uM | 0.18 uM |
| --- | --- | --- | --- | --- | --- |
| Proximal | | | | | |
| Distal | | | | | |
| Full Length | | | | | |

|  | 0.34 uM | 0.63 uM | 1.2 uM | 2.2uM | 4.0 uM |
| --- | --- | --- | --- | --- | --- |
| Proximal | | | | | |
| Distal | | | | | |
| Full Length | | | | | |

121
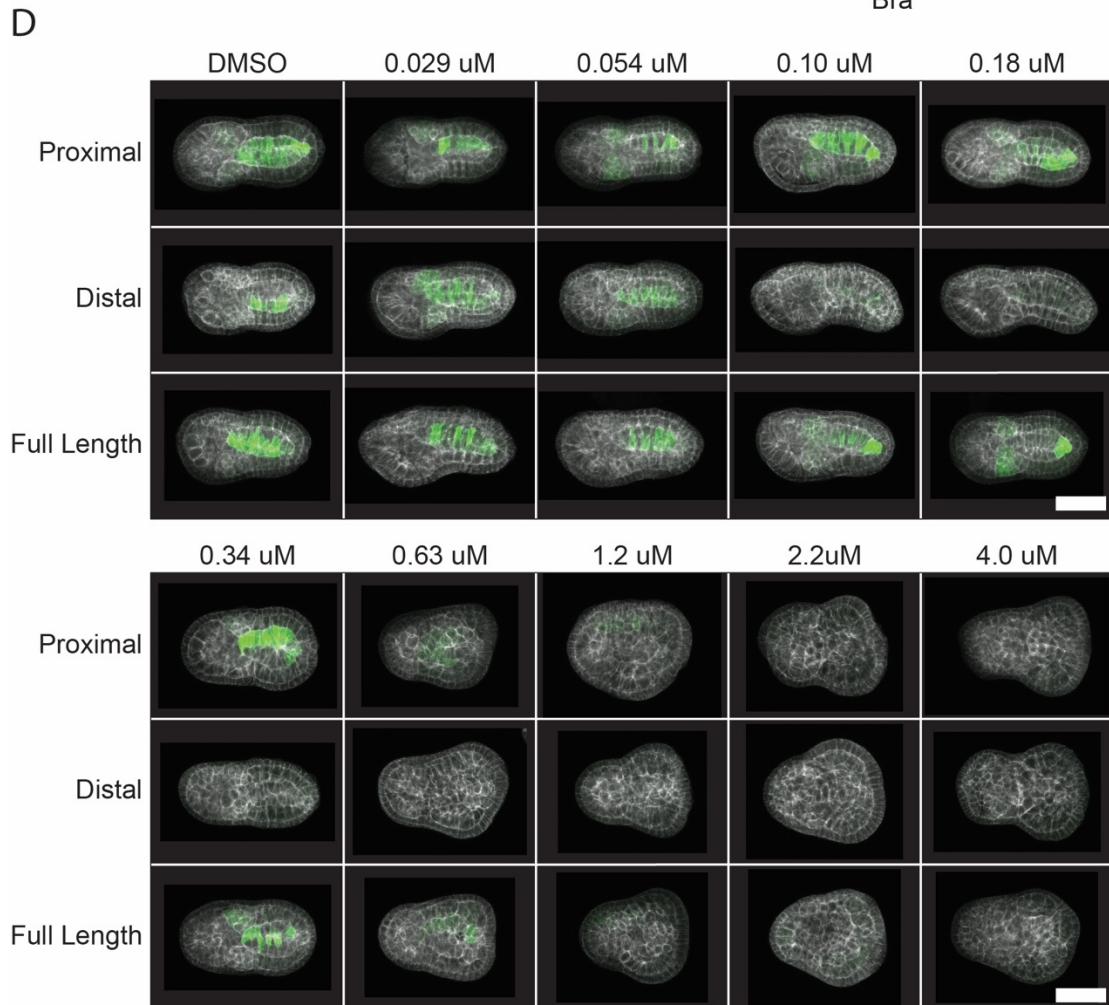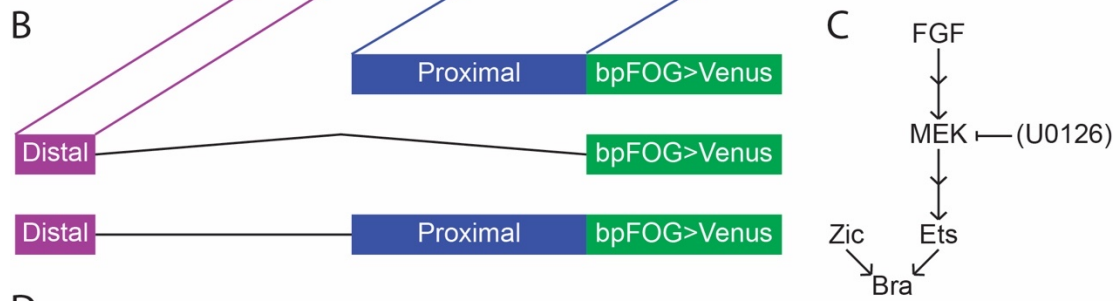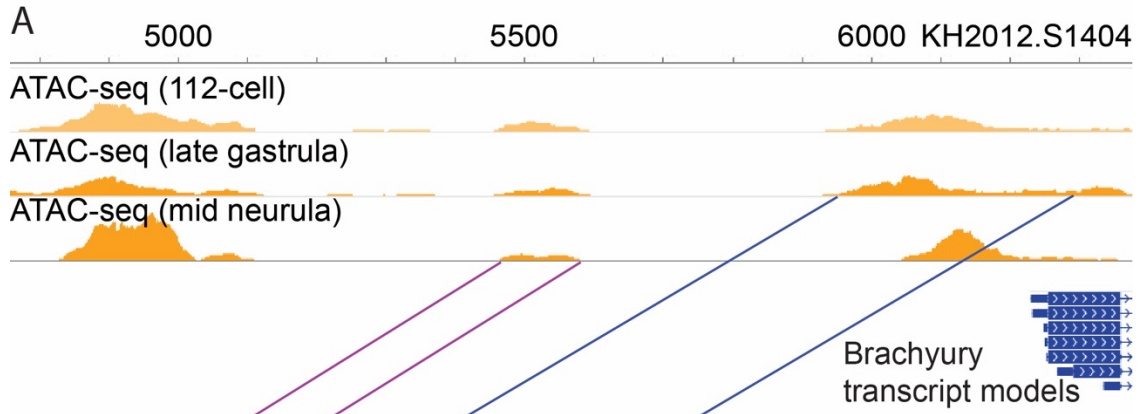
**Figure 4.1 U0126 inhibition of *Bra* reporter constructs.**

**(A)** ATACseq tracks for the genomic region immediately upstream of *Bra* in wildtype *Ciona* embryos at the 112-cell, late gastrula, and mid neurula stages. **(B)** Schematic of reporter constructs used. A 340-bp Proximal enhancer, 116-bp Distal enhancer, and a construct spanning both enhancers with the wild-type 371-bp spacer region between them, each fused directly to the FOG basal promoter in front of a Venus YFP coding sequence. The lines connecting the Proximal and Distal enhancers to the ATACseq tracks in **(A)** indicate the genomic location of each enhancer. **(C)** Simplified schematic of the upstream regulators of *Bra* expression. U0126 inhibits FGF signaling at the step of MEK phosphorylating ERK. Double arrows indicate omitted pathway steps. **(D)** Representative images of the three reporter constructs at the indicated U0126 doses. Each image represents a shallow sum of slices through the notochord or the equivalent depth of the embryo at doses where notochord morphology has been lost. Green: reporter. White: phalloidin. Scale bar: 50 microns.
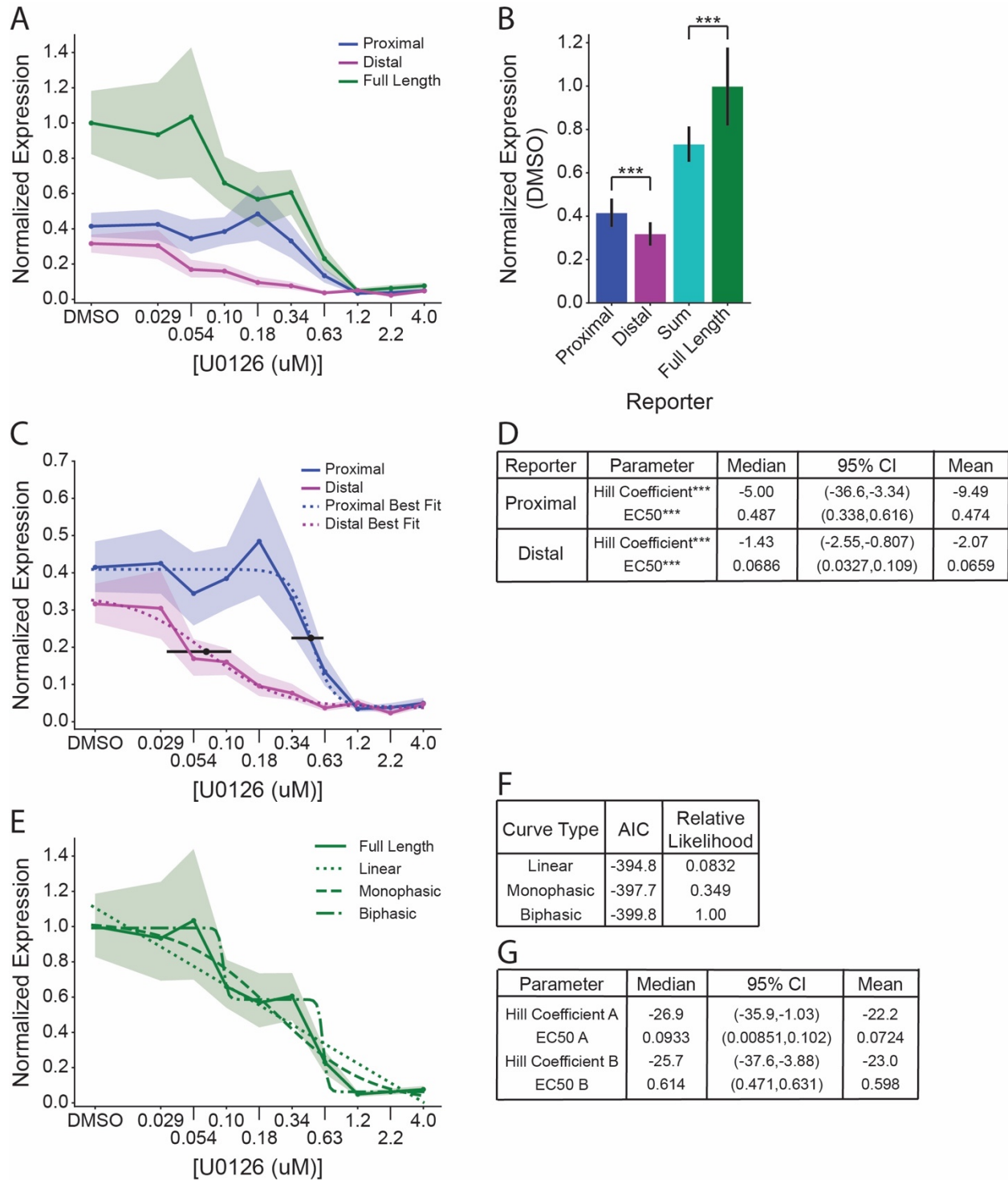
**Figure 4.2 Different *Bra* enhancers have distinct U0126 dose-response curves.**

**(A)** Normalized expression values for each reporter plotted against U0126 dose. Shading indicates 95% confidence intervals of the mean. **(B)** Bootstrapped mean normalized expression of each reporter at the DMSO control dose, as well as the sum of the Proximal and Distal

reporter. Error bars indicate 95% bootstrap confidence intervals (independent samples t-test; ***: p<0.0001). **(C)** Normalized expression values of the Proximal and Distal reporters, with best fit curves plotted using the median parameter values from bootstrapped curve fitting. Black point and error bars indicate the median $EC_{50}$ and its 95% bootstrap confidence interval for each construct. **(D)** Summary of Hill Coefficients and $EC_{50}$ parameter estimates for the curves in **(C)**. Differences in parameters for the Proximal and Distal reporters were compared by Wilcoxon Rank-Sum test (***: p<0.0001). **(E)** Full Length reporter normalized expression values with linear, monophasic, and biphasic models fit to the data by nonlinear regression. **(F)** AIC and relative likelihood values for each model shown in panel **(E)**. **(G)** Summary of Hill Coefficients and $EC_{50}$ parameter estimates of the A and B phases of the biphasic curve fit for the Full Length Reporter. $EC_{50}$ values in **(D)** and **(G)** were calculated in log space, but are shown here after conversion back to a linear scale.
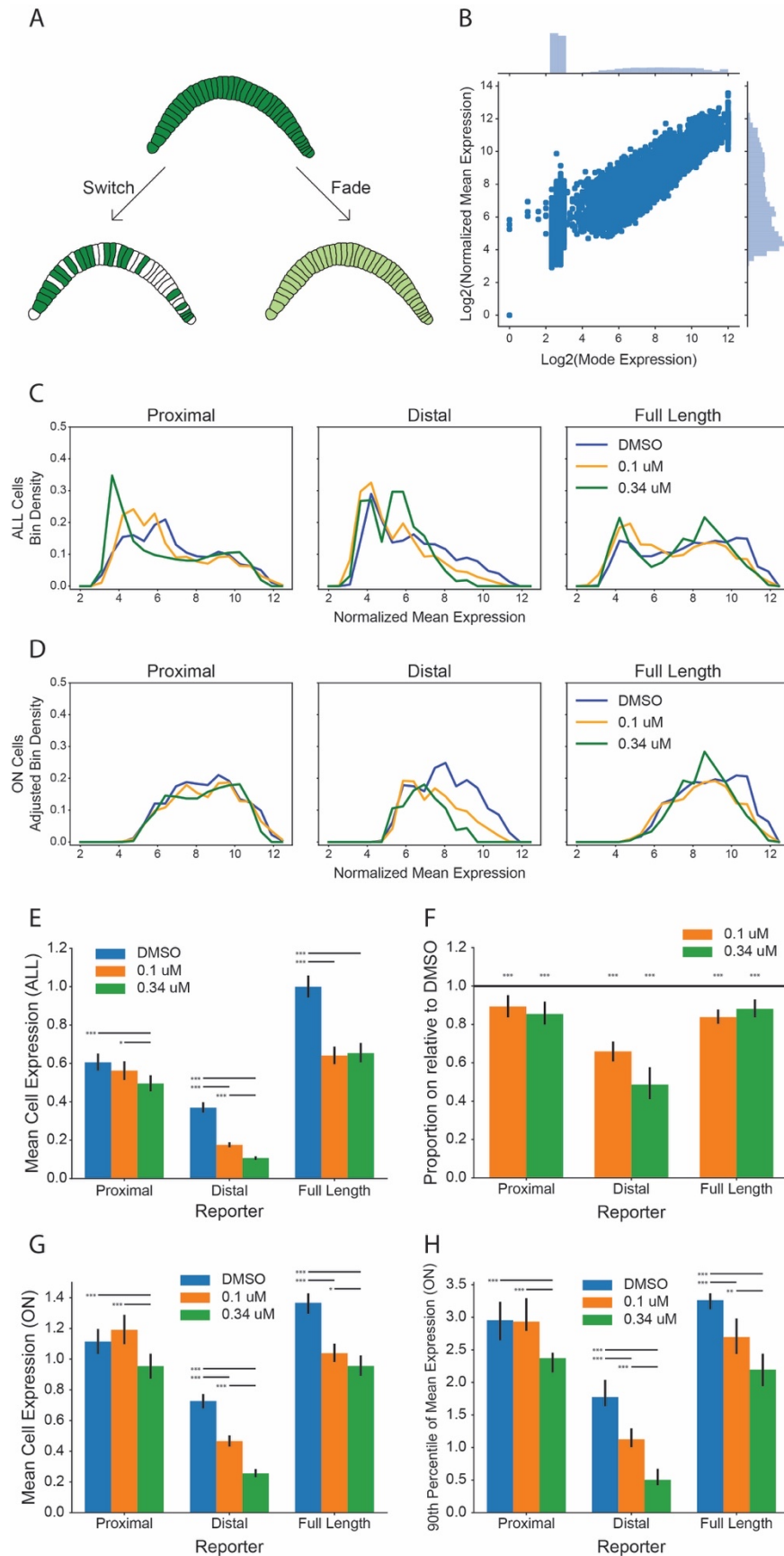
**Figure 4.3 Quantifying responses to MAPK inhibition at the single-cell level.**

**(A)** Cartoon model of switch and fade mechanisms of single-cell expression loss at EC50. For simplicity, this cartoon ignores the mosaic expression of electroporated transgenes. **(B)** Scatterplot of mode pixel values versus normalized mean pixel values for disk-shaped regions of interest spanning each notochord cell nucleus. **(C)** Normalized mean cell expression distributions for all cells, regardless of whether they were classed as ON or OFF. **(D)** Normalized mean cell expression distributions for ON cells, scaled according to the proportion of cells that are ON. **(E)** Bar plot of mean normalized expression of ALL cells for each reporter at DMSO, 0.1 μM U0126, and 0.34 μM U0126. Error bars indicate 95% bootstrap confidence intervals; horizontal bars indicate significantly different pairwise comparisons of the bootstrap distributions, *p<.05; **p<.005; ***p<0.001. **(F)** Bootstrap estimates of the decrease in the fraction of cells classed as ON at a given dose compared to matched DMSO controls. Error bars are 95% confidence intervals. The horizontal reference line is at a ratio of 1, which would indicate no change in the proportion of expressing cells. **(G)** Bootstrap estimates of changes in the mean of the distribution of single-cell normalized mean expression values. **(H)** Bootstrap estimates of changes in the $90^{th}$ percentile of the distribution of single-cell normalized mean expression values. Error bars in G and H are 95% bootstrap confidence intervals of the median bootstrap value; horizontal bars indicate significantly different pairwise comparisons of the bootstrap distributions, *p<0.05; **p<0.005; ***p<0.001.
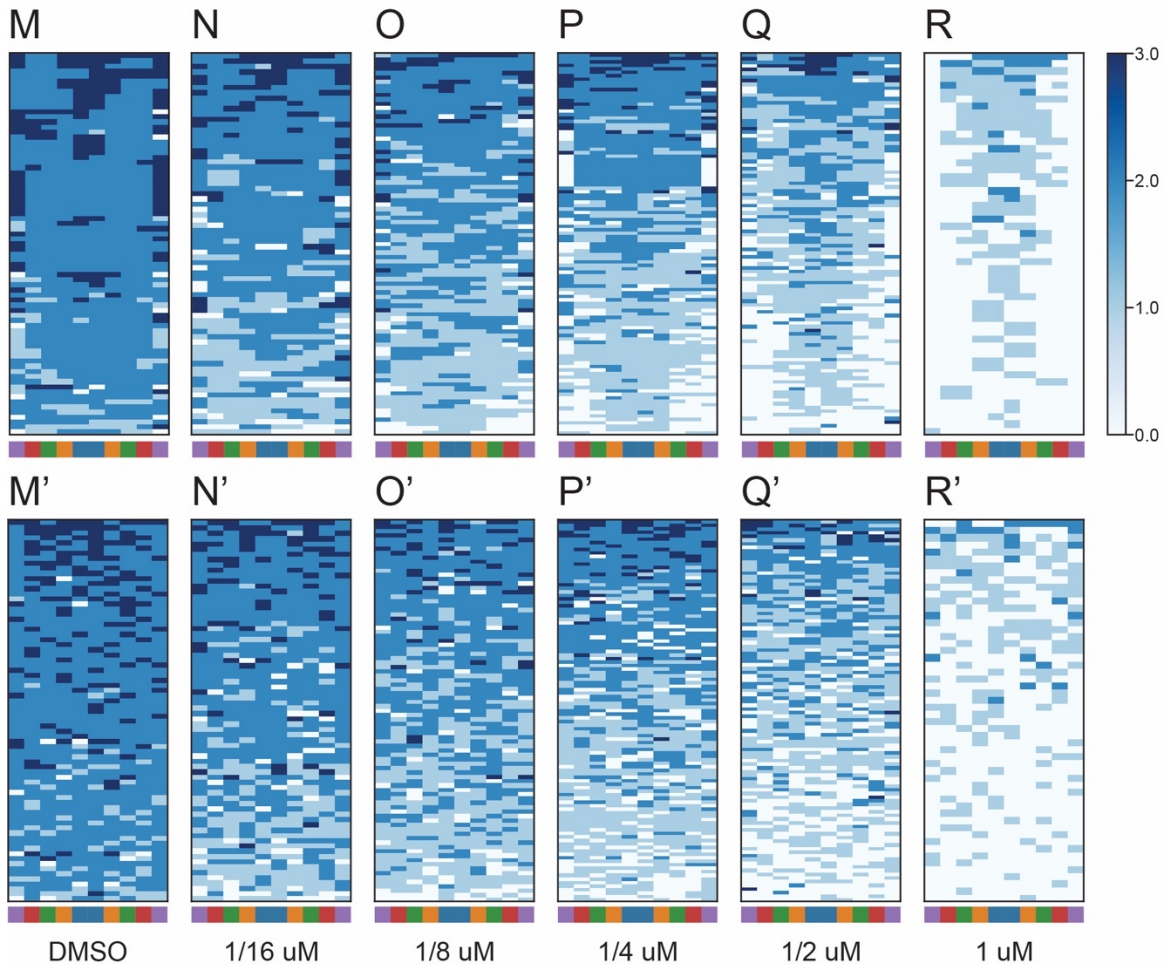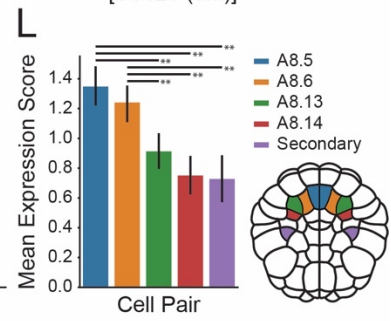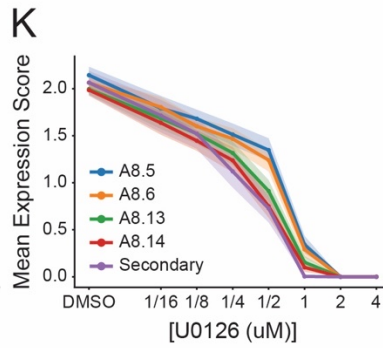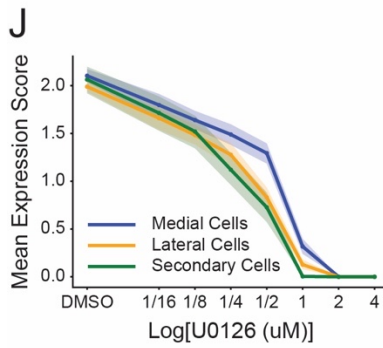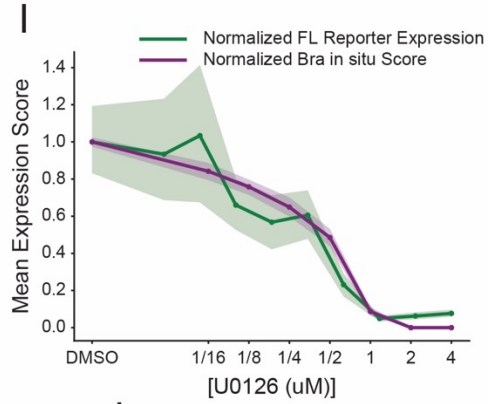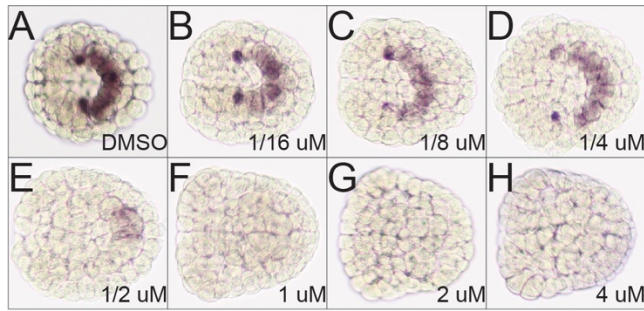
**Figure 4.4 Endogenous *Bra* expression in response to graded MAPK inhibition.**
**(A-H)** Representative images of mid-gastrula stage embryos at the indicated doses stained by *in situ* hybridization for endogenous *Bra* expression. **(I)** U0126 dose-response curves for semiquantitative scores of endogenous *Bra* expression (purple) overlaid on the Full Length reporter whole-embryo expression data (green). Both dose-response curves are scaled to 1 at the DMSO control dose. Shading indicates 95% confidence intervals. **(J-L)** Dose response curves for the semiquantitatively scored endogenous *Bra* expression separated by founder cell identity. The scoring system involved subjectively classing individual cells on an integer scale between 0 (no expression) and 3 (very strong expression). Shading and error bars indicate 95% confidence intervals. **(J)** Cell scores for blastomeres A8.5 and A8.6 were plotted as medial, while scores for blastomeres A8.13 and A8.14 were plotted as lateral. **(K)** Plotting of individual blastomeres. **(L)** Clustered bar plots of the 0.5 µM dose. Horizontal bars indicate pairwise comparisons found to be significantly different by ANOVA followed by Tukey's HSD (**, $p<0.005$). Cartoon indicates location of cell pairs in the 110-cell embryo, anterior at top. **(M-R)** Heatmaps of embryos with nonuniform expression scores across the ten blastomeres, ordered from highest average score at top to lowest average score at bottom for **(M)** DMSO (n=75), **(N)** 0.0625 µM (n=72), **(O)** 0.125 µM (n=87), **(P)** 0.25 µM (n=109), **(Q)** 0.5 µM (n=104), and **(R)** 1.0 µM (n=54). Left to right ordering matches cartoon in **(L)**. Each row in these heat maps represents a different embryo and each column a different notochord founder cell. Embryos with uniform expression or lack thereof across all 10 cells were excluded. **(M'-R')** Similar heatmaps to **(M-R)** except that left-right cell order was randomized for each embryo.

# Chapter 5 - Conclusions and Future Directions

In Chapter 3, we identified mechanisms driving regionalized expression within the notochord, finding important roles for Wnt and FGF signaling as well as for regulation by distributed enhancers. In Chapter 4, we demonstrated that two distributed enhancers of *Brachyury*, despite very similar expression patterns overall, are distinct from one another in their quantitative dependence on upstream MAPK signaling. Along with the insights from these studies, several new questions are raised.

# Regionalized Notochord Gene Expression

The most obvious open question remaining from the regionalized gene expression project is to identify additional specific TFBSs mediating the expression patterns driven by the *fibulin-like* enhancer. Dissection of the minimal reporter indicated the action of a transcriptional repressor that kept expression in the primary notochord off. However, mutation of individual putative TFBSs for Ets, LMX, GATA, and Myb in or near this silencer region resulted in large but general losses of expression, rather than a removal of the silencing effect. Identifying the elusive transcriptional repressor mediating this effect could prove fruitful as a line of future research, and it would be interesting to identify other possible target genes of this repressor.

The finding that *fibulin-like* expression appears to be regulated by distributed enhancers was one of the major sources of inspiration for the study of *Brachyury* distributed enhancers in Chapter 4, but questions raised by the *Brachyury* distributed enhancer project can be applied to the *fibulin-like* enhancers. In the long term, many distributed enhancer sets will need to be analyzed to develop a comprehensive understanding of how they contribute to developmental transcriptional regulatory networks. The *Brachyury* distributed enhancers are quantitatively distinct from one another in several ways, and finer mapping of the *fibulin-like* distributed enhancers may reveal similar quantitative differences between them as well. The *Brachyury* enhancers both respond to common input factors, but that may not necessarily be the case with the *fibulin-like* enhancers, as some distributed enhancers have been found to use different upstream transcription factors to mediate the same expression pattern (Wunderlich et al., 2015).

# Distributed Enhancers

Distributed enhancers have been proposed to play roles in generating robustness of expression to environmental perturbations or genetic polymorphism (Frankel et al., 2010, Antosova et al., 2016, Osterwalder et al., 2018, Dunipace et al., 2019), the establishment or maintenance of precise spatiotemporal expression domains (Perry et al., 2011, Farley et al., 2015, El-Sherif and Levine, 2016), or as raw material for evolution (Hong et al., 2008, Wunderlich et al., 2015, Letelier et al., 2018). These are not mutually exclusive hypotheses, either within a particular set of distributed enhancers for a single gene or across distributed enhancers as a whole. The U0126 treatments I used to titrate the effective levels of FGF signaling can be thought of as comparable to hypomorphic alleles of FGF or signal transduction molecules, or simply molecular 'noise' that produced an effective level of FGF signaling below the ideal state. The finding that the two enhancers have quite different sensitivities and cooperativities, as well as the persistence of proper notochord formation at U0126 doses that eliminate expression from the Distal reporter, provides some evidence in support of a robustness argument. Although we did not rigorously test roles for precision of expression, all three reporters were comparably notochord specific. This suggests that spatial precision may not be a major factor in explaining this enhancer pair, though we cannot rule out a subtle role. We did not ask questions related to the evolvability hypothesis, but note that the theoretical support for this hypothesis is somewhat tenuous (Barolo, 2012, Cannavò et al., 2016). Further studies will be necessary to more rigorously test these hypotheses in this enhancer pair, including individual mutations of the endogenous enhancers to directly determine their necessity and quantification of phospho-ERK levels under similar U0126 dose/response treatments to obtain a more direct measure of the inputs to the system.

We have some data supporting the concept that the Proximal and Distal enhancers are weakly super additive under ideal conditions, and their additivity appears to contribute to the complex shape of the Full Length dose/response curve. However, we were unable to satisfactorily model the apparently biphasic dose/response relationships of the Full Length enhancer, preventing clear identification of the individual $EC_{50}$ and Hill Coefficient parameters that may reveal higher-order interactions between the Proximal and Distal enhancers in the Full Length context. In order to provide more complete answers about the combined function of the Proximal and Distal enhancers, additional testing of the Full Length reporter with a finer scale of doses will be necessary to fit a dose/response curve with higher confidence. A rigorous understanding of the additivity of these two enhancers will also require determining how their relative positions, orientations, and spacing affect their transcriptional outputs, as has been tested with the CD1 and CD2 enhancers of *Krüppel* in *Drosophila* (Scholes et al., 2019). Cryptic sequences in the spacer region may also play a role in Full Length reporter expression, and could be tested by replacing that sequence with random variants.

The *cis*-regulatory features driving the differences between the Proximal and Distal enhancers also remain a mystery. Both the Proximal (Reeves et al., 2020) and Distal (Farley et al., 2016) enhancers have functionally important TFBSs for both Zic and Ets. Differences in the 'grammar' of TFBSs have been shown to have a large influence on expression levels and tissue specificity (Farley et al., 2015; Farley et al., 2016), and are likely to be generating both the different levels of wild type expression and the different responses to U0126 inhibition. Further experiments using reporters with systematically mutated numbers, affinities, orders, spacing, or orientation of TFBSs would be useful to identify underlying rules for how *cis*-regulatory sequences embody the logic of the regulatory genome. Tissue specificity is ultimately a question

of enhancer sensitivity and cooperativity in response to upstream regulators, so the quantitative dose-response framework may be a powerful way to address fundamental questions in *cis*-regulatory biology.

Most of the studies of *cis*-regulatory modules in *Ciona* have used transient reporter assays (This dissertation, Corbo et al., 1997, Passamaneck et al., 2009, José-Edwards et al., 2013, Katikala et al., 2013, Farley et al., 2015, José-Edwards et al., 2015, Farley et al., 2016, Kugler et al., 2019), and although they have been extremely informative, reporter assays alone cannot test the functional necessity of enhancer elements for normal development. Mutation of the endogenous Proximal or Distal enhancers would be difficult, but CRISPR has been successfully applied to *Ciona* (Sasaki et al., 2014, Stolfi et al., 2014, Gandhi et al., 2017, Reeves et al., 2020), and pairs of guide RNAs could potentially be used to selectively delete individual enhancers. If they did not have overt phenotypes, these mutants could be challenged with U0126 dose/response treatments, heat shock or other perturbations to test the necessity of the two enhancers under diverse conditions. It is entirely possible that the Distal enhancer, although far more sensitive to FGF inhibition, might be more robust to other types of perturbation that may be experienced by the *Ciona* embryo, and that the Proximal enhancer could be more sensitive to those perturbations. In that model, the embryo would be able to develop under a wider variety of sub-ideal conditions, with the weaknesses of each enhancer compensated by the strengths of the other to keep the notochord GRN intact.

There are also questions of the evolution of enhancer additivity in this system. The weak super-additivity of the Proximal and Distal enhancers could have evolved to push *Brachyury* expression as high as possible in the notochord, with the side effect that these enhancers have distinct sensitivities to FGF signaling and thus a complex overall dose-response relationship.

Alternatively, it is possible that this biphasic dose-response relationship is of fundamental importance to how the system operates and was the key aspect under selection. Different FGF ligands play different roles in the induction and maintenance of *Brachyury* expression (Yasuo and Hudson, 2007), and the different FGF sensitivities of the Proximal and Distal enhancers and complex shape of the overall response could be connected to these separable roles.

# Conclusion

This dissertation addressed two quite different topics in transcriptional regulation. I identified important inputs into how different anterior/posterior regions in the *Ciona* notochord express different genes. I also identified how two distributed enhancers that drive comparable notochord-specific expression patterns actually have fundamentally different quantitative relationships with upstream MAPK signaling. These projects were united in exploiting the simple embryo and straightforward transgenesis of *Ciona* to work towards an increasingly quantitative understanding of enhancer function. New developments in automated imaging and single-cell genomics will make it possible to address increasingly exciting questions about *cis*-regulatory function in this tractable model chordate.

# References

Antosova, B., Smolikova, J., Klimova, L., Lachova, J., Bendova, M., Kozmikova, I., Macon, O., Kozmik, Z., 2016. The Gene Regulatory Network of Lens Induction Is Wired through Meis-Dependent Shadow Enhancers of Pax6. PLoS Genetics 12, e1006441.

Barolo, S., 2012. Shadow enhancers: Frequently asked questions about distributed cis-regulatory information and enhancer redundancy. BioEssays 34, 135-141.

Cannavò, E., Khoueiry, P., Garfield, D.A., Geeleher, P., Zichner, T., Gustafson, E.H., Ciglar, L., Kornberg, J.O., Furlong, E.E.M., 2016. Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks. Curr. Biol. 26, 38-51.

Corbo, J., Levine, M., Zeller, R., 1997. Characterization of a notochord-specific enhancer from the Brachyury promoter region of the ascidian, Ciona intestinalis. Development 124, 589-602.

Dunipace, L., Ákos, Z., Stathopoulos, A., 2019. Coacting enhancers can have complementary functions within gene regulatory networks and promote canalization. PLoS Genetics 15, e1008525.

El-Sherif, E., Levine, M., 2016. Shadow Enhancers Mediate Dynamic Shifts of Gap Gene Expression in the Drosophila Embryo. Current Biol. 26, 1164-1169.

Farley, E.K., Olson, K.M., Zhang, W., Brandt, A.J., Rokhsar, D.S., Levine, M.S., 2015. Suboptimization of developmental enhancers. Science 350, 325-328.

Farley, E.K., Olson, K.M., Zhang, W., Rokhsar, S.D., Levine, M.S., 2016. Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. PNAS 113, 6508-6513.

Frankel, N., Davis, G.K., Vargas, D., Wang, S., Payre, F., Stern, D.L., 2010. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. Nature 466, 490-493.

Gandhi, S., Haeussler, M., Razy-Krajka, F., Christian, L., Stolfi, A., 2017. Evaluation and rational design of guide RNAs for efficient CRISPR/Cas9-mediated mutagenesis in Ciona. Dev. Biol. 425, 8-20.

Hong, J.-W., Hendrix, D.A., Levine, M.S., 2008. Shadow Enhancers as a Source of Evolutionary Novelty. Science 321, 1314.

José-Edwards, D.S., Oda-Ishii, I., Nibu, Y., Di Gregorio, A., 2013. Tbx2/3 is an essential mediator within the Brachyury gene network during Ciona notochord development. Development 140, 2422-2433.

José-Edwards, D.S., Oda-Ishii, I., Kugler, J.E., Passamaneck, Y.J., Katikala, L., Nibu, Y., Di Gregorio, A., 2015. Brachyury, Foxa2 and the cis-Regulatory Origins of the Notochord. PLoS Gen. 11, 1-16.

Katikala, L., Aihara, H., Passamaneck, Y.J., Gazdoiu, S., José-Edwards, D.S., Kugler, J.E., Oda-Ishii, I., Imai, J.H., Nibu, Y., Di Gregorio, A., 2013. Functional Brachyury Binding Sites Establish a Temporal Read-out of Gene Expression in the Ciona Notochord. PLoS Biol. 11, e1001697.

Kugler, J.E., Wu, Y., Katikala, L., Passamaneck, Y.J., Addy, J., Caballero, N., Oda-Ishii, I., Maguire, J.E., Li, R., Di Gregorio, A., 2019. Positioning a multifunctional basic helix-loop-helix transcription factor within the Ciona notochord gene regulatory network. Dev. Biol. 448, 119-135.

Letelier, J., de la Calle-Mustienes, E., Pieretti, J., Naranjo, S., Maeso, I., Nakamura, T., Pascual-Anaya, J., Shubin, N.H., Schneider, I., Martinez-Morales, J.R., Gómez-Skarmeta, J.L., 2018. A conserved Shh cis-regulatory module highlights a common developmental origin of unpaired and paired fins. Nature Gen. 50, 504-509.

Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B.J., Afzal, S.Y., Lee, E.A., Zhu, Y., Placzek-Frick, I., Pickle, C.S., Kato, M., Garvin, T.H., Pham, Q.T., Harrington, A.N., Akiyama, J.A., Afzal, V., Lopez-Rios, J., Dickey, D.E., Visel, A., Pennacchio, L.A., 2018. Enhancer redundancy provides phenotypic robustness in mammalian development. Nature 554, 239-243.

Passamaneck, Y.J., Katikala, L., Perrone, L., Dunn, M.P., Oda-Ishii, I., Di Gregorio, A., 2009. Direct activation of a notochord cis-regulatory module by Brachyury and FoxA in the ascidian Ciona intestinalis. Development 136, 3679-3689.

Perry, M., Boettiger, A., Levine, M., 2011. Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. PNAS 108, 13570-13575.

Reeves, W.M., Shimai, K., Winkley, K.M., Veeman, M.T., 2020. Brachyury controls Ciona notochord fate as part of a feedforward network and not as a unitary master regulator. bioRxiv, https://doi.org/10.1101/2020.05.29.12402.

Sasaki, H., Yoshida, K., Hozumi, A., Sasakura, Y., 2014. CRISPR/Cas9-mediated gene knockout in the ascidian Ciona intestinalis. Develop. Growth Differ. 56, 499-510.

Scholes, C., Biette, K., Harden, T., DePace, A., 2019. Signal Integration by Shadow Enhancers and Enhancer Duplications Varies across the Drosophila Embryo. Cell Reports 26, 2407-2418.

Stolfi, A., Gandhi, S., Salek, F., Christian, L., 2014. Tissue-specific genome editing in Ciona embryos by CRISPR/Cas9. Development 141, 4115-4120.

Wunderlich, Z., Bragdon, M.D.J., Vincent, B.J., White, J.A., Estrada, J., DePace, A.H., 2015. Krüppel Expression Levels Are Maintained through Compensatory Evolution of Shadow Enhancers. Cell Rep. 12, 1740-1747.

Yasuo, H., Hudson, C., 2007. FGF8/17/18 functions together with FGF9/16/20 during formation of the notochord in Ciona embryos. Dev. Biol. 302, 92-103.

# Appendix A - Chapter 3 Supplement

**Table A.1 Individual Qualitative Scores**

| Reporter | Experiment | Embryo | Primary Score | Secondary Anterior Score | Secondary Posterior Score | Secondary (Average) |
|---|---|---|---|---|---|---|
| Minus 1565 to +13 | 1 | 1 | 1.5 | 3 | 3 | 3 |
| Minus 1565 to +13 | 1 | 3 | 3 | 3 | 4 | 3.5 |
| Minus 1565 to +13 | 1 | 4 | 1.5 | 2.5 | 2 | 2.25 |
| Minus 1565 to +13 | 1 | 5 | 2 | 1 | 3 | 2 |
| Minus 1565 to +13 | 1 | 6 | 3 | 3 | 3 | 3 |
| Minus 1565 to +13 | 1 | 7 | 0 | 2.5 | 2.5 | 2.5 |
| Minus 1565 to +13 | 1 | 9 | 4 | 1 | 2 | 1.5 |
| Minus 1565 to +13 | 1 | 11 | 2.5 | 2 | 3 | 2.5 |
| Minus 1565 to +13 | 1 | 13 | 1.5 | 4 | 3 | 3.5 |
| Minus 1565 to +13 | 1 | 14 | 0 | 1.5 | 3 | 2.25 |
| Minus 1565 to +13 | 1 | 15 | 2.5 | 2.5 | 2.5 | 2.5 |
| Minus 1565 to +13 | 1 | 16 | 2.5 | 2 | 1.5 | 1.75 |
| Minus 1565 to +13 | 1 | 18 | 4 | 0 | 0 | 0 |
| Minus 1565 to +13 | 2 | 1 | 3 | 2 | 2 | 2 |
| Minus 1565 to +13 | 2 | 2 | 1.5 | 1 | 2 | 1.5 |
| Minus 1565 to +13 | 2 | 4 | 1 | 4 | 4 | 4 |
| Minus 1565 to +13 | 2 | 5 | 2.5 | 1 | 2.5 | 1.75 |
| Minus 1565 to +13 | 2 | 6 | 2.5 | 2.5 | 4 | 3.25 |
| Minus 1565 to +13 | 2 | 7 | 1 | 4 | 4 | 4 |
| Minus 1565 to +13 | 2 | 8 | 1 | 4 | 4 | 4 |
| Minus 1565 to +13 | 2 | 9 | 1 | 2 | 4 | 3 |
| Minus 1565 to +13 | 2 | 11 | 1.5 | 1 | 2 | 1.5 |
| Minus 1565 to +13 | 2 | 12 | 2.5 | 2.5 | 3 | 2.75 |
| Minus 1565 to +13 | 2 | 13 | 2 | 4 | 4 | 4 |
| Minus 1565 to +13 | 2 | 14 | 4 | 2.5 | 1.5 | 2 |
| Minus 1565 to +13 | 2 | 15 | 1.5 | 2 | 4 | 3 |
| Minus 1565 to +13 | 2 | 16 | 3 | 1 | 1.5 | 1.25 |
| Minus 1565 to +13 | 2 | 17 | 2.5 | 2 | 1 | 1.5 |
| Minus 1565 to +13 | 2 | 18 | 1 | 4 | 4 | 4 |
| Minus 1565 to +13 | 2 | 19 | 3 | 3 | 3 | 3 |
| Minus 1565 to +13 | 2 | 20 | 1 | 4 | 1.5 | 2.75 |
| Minus 1565 to +13 | 2 | 21 | 1 | 3 | 2 | 2.5 |
| Minus 1565 to +13 | 3 | 1 | 4 | 2.5 | 2.5 | 2.5 |
| Minus 1565 to +13 | 3 | 2 | 1 | 4 | 4 | 4 |
| Minus 1565 to +13 | 3 | 3 | 4 | 1.5 | 1 | 1.25 |
| Minus 1565 to +13 | 3 | 4 | 2 | 4 | 3 | 3.5 |

| Reporter | Experiment | Embryo | Primary Score | Secondary Anterior Score | Secondary Posterior Score | Secondary (Average) |
|---|---|---|---|---|---|---|
| Minus 1565 to +13 | 3 | 5 | 1.5 | 1 | 1 | 1 |
| Minus 1565 to +13 | 3 | 6 | 1 | 2.5 | 4 | 3.25 |
| Minus 1565 to +13 | 3 | 7 | 1 | 2 | 3 | 2.5 |
| Minus 1565 to +13 | 3 | 8 | 4 | 1 | 1 | 1 |
| Minus 1565 to +13 | 3 | 9 | 2.5 | 2.5 | 1.5 | 2 |
| Minus 1565 to +13 | 3 | 10 | 1.5 | 2.5 | 1 | 1.75 |
| Minus 1565 to +13 | 3 | 11 | 1 | 2.5 | 2.5 | 2.5 |
| Minus 1565 to +13 | 3 | 12 | 2.5 | 1 | 2.5 | 1.75 |
| Minus 1565 to +13 | 3 | 13 | 1.5 | 4 | 4 | 4 |
| Minus 1565 to +13 | 3 | 14 | 1 | 2 | 4 | 3 |
| Minus 1565 to +13 | 3 | 15 | 1 | 4 | 4 | 4 |
| Minus 1565 to +13 | 3 | 16 | 3 | 1 | 1.5 | 1.25 |
| Minus 1565 to +13 | 3 | 17 | 1 | 4 | 3 | 3.5 |
| Minus 1565 to +13 | 3 | 18 | 1 | 3 | 4 | 3.5 |
| Minus 1565 to -758 | 1 | 2 | 2.5 | 1 | 1 | 1 |
| Minus 1565 to -758 | 1 | 4 | 0 | 1 | 1 | 1 |
| Minus 1565 to -758 | 1 | 5 | 1 | 2 | 2.5 | 2.25 |
| Minus 1565 to -758 | 1 | 7 | 1 | 1.5 | 1.5 | 1.5 |
| Minus 1565 to -758 | 1 | 9 | 0 | 1 | 1.5 | 1.25 |
| Minus 1565 to -758 | 1 | 10 | 0 | 2 | 2 | 2 |
| Minus 1565 to -758 | 1 | 11 | 0 | 1.5 | 2 | 1.75 |
| Minus 1565 to -758 | 1 | 12 | 1 | 1.5 | 1.5 | 1.5 |
| Minus 1565 to -758 | 1 | 13 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 1 | 14 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 1 | 15 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 1 | 16 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 1 | 17 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 2 | 1 | 0 | 1 | 1 | 1 |
| Minus 1565 to -758 | 2 | 2 | 1.5 | 2.5 | 3 | 2.75 |
| Minus 1565 to -758 | 2 | 3 | 0 | 4 | 4 | 4 |
| Minus 1565 to -758 | 2 | 4 | 3 | 1 | 1 | 1 |
| Minus 1565 to -758 | 2 | 5 | 1 | 0 | 0 | 0 |
| Minus 1565 to -758 | 2 | 6 | 1 | 2 | 3 | 2.5 |
| Minus 1565 to -758 | 2 | 8 | 1 | 1 | 1 | 1 |
| Minus 1565 to -758 | 2 | 9 | 1 | 2 | 2 | 2 |
| Minus 1565 to -758 | 2 | 10 | 0 | 1 | 2 | 1.5 |
| Minus 1565 to -758 | 2 | 12 | 0 | 1 | 1 | 1 |
| Minus 1565 to -758 | 2 | 14 | 0 | 1 | 1 | 1 |
| Minus 1565 to -758 | 2 | 15 | 0 | 2 | 2 | 2 |
| Minus 1565 to -758 | 2 | 16 | 3 | 2 | 2 | 2 |

| Reporter | Experiment | Embryo | Primary Score | Secondary Anterior Score | Secondary Posterior Score | Secondary (Average) |
|---|---|---|---|---|---|---|
| Minus 1565 to -758 | 2 | 18 | 1 | 4 | 4 | 4 |
| Minus 1565 to -758 | 2 | 19 | 1 | 2 | 4 | 3 |
| Minus 1565 to -758 | 2 | 20 | 4 | 2.5 | 2.5 | 2.5 |
| Minus 1565 to -758 | 2 | 21 | 0 | 1.5 | 1.5 | 1.5 |
| Minus 1565 to -758 | 2 | 22 | 1 | 1.5 | 2 | 1.75 |
| Minus 1565 to -758 | 2 | 23 | 1 | 1 | 1 | 1 |
| Minus 1565 to -758 | 3 | 2 | 0 | 1 | 1 | 1 |
| Minus 1565 to -758 | 3 | 4 | 0 | 2 | 2 | 2 |
| Minus 1565 to -758 | 3 | 5 | 1 | 1 | 1 | 1 |
| Minus 1565 to -758 | 3 | 6 | 0 | 2.5 | 2.5 | 2.5 |
| Minus 1565 to -758 | 3 | 7 | 1 | 1 | 1 | 1 |
| Minus 1565 to -758 | 3 | 8 | 1 | 1 | 1 | 1 |
| Minus 1565 to -758 | 3 | 9 | 0 | 1 | 1 | 1 |
| Minus 1565 to -758 | 3 | 11 | 0 | 1 | 1 | 1 |
| Minus 1565 to -758 | 3 | 12 | 2 | 1 | 1 | 1 |
| Minus 1565 to -758 | 3 | 14 | 0 | 1 | 1 | 1 |
| Minus 1565 to -758 | 3 | 15 | 1 | 0 | 0 | 0 |
| Minus 1565 to -758 | 3 | 16 | 1 | 1 | 1 | 1 |
| Minus 1565 to -758 | 3 | 17 | 1 | 0 | 0 | 0 |
| Minus 1565 to -758 | 3 | 18 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 3 | 19 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 3 | 20 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 3 | 21 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 3 | 22 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 3 | 23 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 3 | 24 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 3 | 25 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 3 | 26 | 0 | 0 | 0 | 0 |
| Minus 1045 to -526 | 1 | 1 | 1 | 3 | 3 | 3 |
| Minus 1045 to -526 | 1 | 2 | 1 | 0 | 1 | 0.5 |
| Minus 1045 to -526 | 1 | 3 | 0 | 1 | 2.5 | 1.75 |
| Minus 1045 to -526 | 1 | 4 | 0 | 1 | 1 | 1 |
| Minus 1045 to -526 | 1 | 5 | 0 | 2 | 2 | 2 |
| Minus 1045 to -526 | 1 | 6 | 0 | 2 | 1.5 | 1.75 |
| Minus 1045 to -526 | 1 | 10 | 1 | 2.5 | 2.5 | 2.5 |
| Minus 1045 to -526 | 1 | 11 | 1 | 1.5 | 1.5 | 1.5 |
| Minus 1045 to -526 | 1 | 12 | 1 | 2 | 2 | 2 |
| Minus 1045 to -526 | 1 | 13 | 0 | 2 | 2 | 2 |
| Minus 1045 to -526 | 1 | 15 | 0 | 1 | 1 | 1 |
| Minus 1045 to -526 | 2 | 1 | 1 | 2 | 2.5 | 2.25 |

| Reporter | Experiment | Embryo | Primary Score | Secondary Anterior Score | Secondary Posterior Score | Secondary (Average) |
|---|---|---|---|---|---|---|
| Minus 1045 to -526 | 2 | 2 | 0 | 1.5 | 1.5 | 1.5 |
| Minus 1045 to -526 | 2 | 3 | 0 | 3 | 3 | 3 |
| Minus 1045 to -526 | 3 | 1 | 1.5 | 1.5 | 1.5 | 1.5 |
| Minus 1045 to -526 | 3 | 2 | 1 | 2 | 2 | 2 |
| Minus 1045 to -526 | 3 | 3 | 0 | 1 | 1 | 1 |
| Minus 1045 to -526 | 3 | 4 | 0 | 1 | 1.5 | 1.25 |
| Minus 1045 to -526 | 3 | 6 | 0 | 1 | 1 | 1 |
| Minus 1045 to -526 | 3 | 7 | 1.5 | 1 | 1 | 1 |
| Minus 1045 to -526 | 3 | 8 | 2 | 1 | 1 | 1 |
| Minus 1045 to -526 | 3 | 9 | 1 | 1.5 | 1.5 | 1.5 |
| Minus 1045 to -526 | 3 | 10 | 0 | 1 | 1 | 1 |
| Minus 1045 to -526 | 3 | 11 | 1 | 1.5 | 1.5 | 1.5 |
| Minus 1045 to -526 | 3 | 12 | 1.5 | 1 | 1 | 1 |
| Minus 1045 to -526 | 3 | 14 | 1.5 | 0 | 0 | 0 |
| Minus 1045 to -526 | 3 | 15 | 0 | 1 | 1 | 1 |
| Minus 1045 to -526 | 3 | 16 | 0 | 0 | 0 | 0 |
| Minus 1045 to -526 | 3 | 17 | 0 | 0 | 0 | 0 |
| Minus 1045 to -526 | 3 | 18 | 0 | 0 | 0 | 0 |
| Minus 1045 to -526 | 3 | 19 | 0 | 0 | 0 | 0 |
| Minus 787 to +13 | 1 | 1 | 1.5 | 1 | 3 | 2 |
| Minus 787 to +13 | 1 | 3 | 1 | 1.5 | 0 | 0.75 |
| Minus 787 to +13 | 1 | 4 | 1 | 2 | 3 | 2.5 |
| Minus 787 to +13 | 1 | 5 | 0 | 3 | 3 | 3 |
| Minus 787 to +13 | 1 | 6 | 1 | 3 | 3 | 3 |
| Minus 787 to +13 | 1 | 7 | 1.5 | 3 | 2.5 | 2.75 |
| Minus 787 to +13 | 1 | 9 | 1 | 3 | 1 | 2 |
| Minus 787 to +13 | 1 | 10 | 1 | 2.5 | 1 | 1.75 |
| Minus 787 to +13 | 1 | 11 | 2.5 | 2 | 2.5 | 2.25 |
| Minus 787 to +13 | 1 | 12 | 3 | 2 | 4 | 3 |
| Minus 787 to +13 | 1 | 13 | 1 | 3 | 3 | 3 |
| Minus 787 to +13 | 1 | 14 | 3 | 3 | 3 | 3 |
| Minus 787 to +13 | 1 | 16 | 1 | 2 | 3 | 2.5 |
| Minus 787 to +13 | 1 | 17 | 1 | 2 | 0 | 1 |
| Minus 787 to +13 | 1 | 18 | 0 | 0 | 0 | 0 |
| Minus 787 to +13 | 2 | 1 | 1 | 4 | 2.5 | 3.25 |
| Minus 787 to +13 | 2 | 2 | 1.5 | 1 | 1.5 | 1.25 |
| Minus 787 to +13 | 2 | 3 | 1.5 | 2 | 2.5 | 2.25 |
| Minus 787 to +13 | 2 | 4 | 4 | 4 | 4 | 4 |
| Minus 787 to +13 | 2 | 5 | 3 | 1 | 4 | 2.5 |
| Minus 787 to +13 | 2 | 12 | 1 | 2 | 2 | 2 |

| Reporter | Experiment | Embryo | Primary Score | Secondary Anterior Score | Secondary Posterior Score | Secondary (Average) |
|---|---|---|---|---|---|---|
| Minus 787 to +13 | 2 | 13 | 4 | 2.5 | 2.5 | 2.5 |
| Minus 787 to +13 | 2 | 14 | 0 | 4 | 4 | 4 |
| Minus 787 to +13 | 2 | 15 | 1 | 2.5 | 1.5 | 2 |
| Minus 787 to +13 | 2 | 17 | 3 | 1 | 2 | 1.5 |
| Minus 787 to +13 | 2 | 19 | 0 | 4 | 4 | 4 |
| Minus 787 to +13 | 3 | 1 | 3 | 3 | 3 | 3 |
| Minus 787 to +13 | 3 | 5 | 3 | 1 | 2 | 1.5 |
| Minus 787 to +13 | 3 | 6 | 1 | 1 | 4 | 2.5 |
| Minus 787 to +13 | 3 | 7 | 3 | 2 | 1.5 | 1.75 |
| Minus 787 to +13 | 3 | 8 | 1 | 2 | 2 | 2 |
| Minus 787 to +13 | 3 | 10 | 2 | 2 | 3 | 2.5 |
| Minus 787 to +13 | 3 | 11 | *1.5* | 4 | 1.5 | 2.75 |
| Minus 787 to +13 | 3 | 13 | 1 | 1.5 | 2 | 1.75 |
| Minus 787 to +13 | 3 | 14 | 2 | 4 | 4 | 4 |
| Minus 787 to +13 | 3 | 15 | 2.5 | 4 | 4 | 4 |
| Minus 787 to +13 | 3 | 17 | 1.5 | 4 | 4 | 4 |
| Minus 787 to +13 | 3 | 18 | 2.5 | 1 | 1 | 1 |
| Minus 787 to +13 | 4 | 1 | 4 | 2 | 2 | 2 |
| Minus 787 to +13 | 4 | 2 | 1 | 3 | 3 | 3 |
| Minus 787 to +13 | 4 | 3 | 1 | 2 | 2 | 2 |
| Minus 787 to +13 | 4 | 4 | 4 | 3 | 3 | 3 |
| Minus 787 to +13 | 4 | 5 | 1 | 3 | 3 | 3 |
| Minus 787 to +13 | 4 | 6 | 2.5 | 3 | 3 | 3 |
| Minus 787 to +13 | 4 | 7 | 2 | 2.5 | 2.5 | 2.5 |
| Minus 787 to +13 | 4 | 8 | 1 | 2.5 | 2.5 | 2.5 |
| Minus 787 to +13 | 4 | 10 | 4 | 3 | 4 | 3.5 |
| Minus 787 to +13 | 4 | 11 | 4 | 4 | 4 | 4 |
| Minus 787 to +13 | 4 | 12 | 2 | 4 | 4 | 4 |
| Minus 787 to +13 | 5 | 1 | 1 | 4 | 4 | 4 |
| Minus 787 to +13 | 5 | 2 | 1 | 1.5 | 2.5 | 2 |
| Minus 787 to +13 | 5 | 3 | 1.5 | 2 | 2 | 2 |
| Minus 787 to +13 | 5 | 6 | 4 | 1 | 2 | 1.5 |
| Minus 787 to +13 | 5 | 7 | 2 | 1.5 | 1 | 1.25 |
| Minus 787 to +13 | 5 | 8 | 4 | 4 | 4 | 4 |
| Minus 787 to +13 | 5 | 9 | 1.5 | 2.5 | 2.5 | 2.5 |
| Minus 787 to +13 | 5 | 10 | 2 | 2 | 2 | 2 |
| Minus 787 to +13 | 5 | 12 | 2.5 | 2.5 | 2.5 | 2.5 |
| Minus 787 to +13 | 5 | 13 | 3 | 3 | 1.5 | 2.25 |
| Minus 787 to +13 | 5 | 15 | 1 | 2 | 3 | 2.5 |
| Minus 787 to +13 | 5 | 16 | 1 | 1 | 1.5 | 1.25 |

| Reporter | Experiment | Embryo | Primary Score | Secondary Anterior Score | Secondary Posterior Score | Secondary (Average) |
|---|---|---|---|---|---|---|
| Minus 787 to +13 | 6 | 2 | 2.5 | 3 | 3 | 3 |
| Minus 787 to +13 | 6 | 3 | 4 | 3 | 3 | 3 |
| Minus 787 to +13 | 6 | 4 | 4 | 3 | 3 | 3 |
| Minus 787 to +13 | 6 | 5 | 4 | 2 | 2 | 2 |
| Minus 787 to +13 | 6 | 6 | 1 | 2.5 | 2.5 | 2.5 |
| Minus 787 to +13 | 6 | 7 | 3 | 0 | 3 | 1.5 |
| Minus 787 to +13 | 6 | 8 | 1.5 | 1 | 1 | 1 |
| Minus 787 to +13 | 6 | 9 | 0 | 2 | 3 | 2.5 |
| Minus 787 to +13 | 6 | 10 | 2 | 4 | 4 | 4 |
| Minus 787 to +13 | 6 | 11 | 1.5 | 4 | 4 | 4 |
| Minus 787 to +13 | 6 | 12 | 2.5 | 1 | 1 | 1 |
| Minus 787 to +13 | 6 | 13 | 2.5 | 3 | 3 | 3 |
| Minus 787 to -245 | 4 | 1 | 1 | 2 | 1.5 | 1.75 |
| Minus 787 to -245 | 4 | 2 | 1 | 2 | 2 | 2 |
| Minus 787 to -245 | 4 | 3 | 2.5 | 2.5 | 2.5 | 2.5 |
| Minus 787 to -245 | 4 | 6 | 2 | 3 | 3 | 3 |
| Minus 787 to -245 | 4 | 9 | 1 | 2.5 | 2 | 2.25 |
| Minus 787 to -245 | 4 | 10 | 1.5 | 2.5 | 2.5 | 2.5 |
| Minus 787 to -245 | 4 | 11 | 3 | 0 | 0 | 0 |
| Minus 787 to -245 | 4 | 13 | 1 | 4 | 4 | 4 |
| Minus 787 to -245 | 5 | 1 | 0 | 1 | 1 | 1 |
| Minus 787 to -245 | 5 | 2 | 0 | 1 | 1 | 1 |
| Minus 787 to -245 | 5 | 3 | 3 | 3 | 2.5 | 2.75 |
| Minus 787 to -245 | 5 | 4 | 0 | 2 | 2 | 2 |
| Minus 787 to -245 | 5 | 7 | 1 | 1.5 | 2.5 | 2 |
| Minus 787 to -245 | 5 | 8 | 1 | 3 | 3 | 3 |
| Minus 787 to -245 | 5 | 9 | 2 | 2 | 3 | 2.5 |
| Minus 787 to -245 | 5 | 12 | 1 | 1.5 | 1.5 | 1.5 |
| Minus 787 to -245 | 5 | 13 | 1 | 1 | 1 | 1 |
| Minus 787 to -245 | 5 | 15 | 1 | 2 | 2 | 2 |
| Minus 787 to -245 | 5 | 16 | 1 | 1.5 | 1 | 1.25 |
| Minus 787 to -245 | 6 | 2 | 1.5 | 2.5 | 2.5 | 2.5 |
| Minus 787 to -245 | 6 | 3 | 2 | 4 | 4 | 4 |
| Minus 787 to -245 | 6 | 8 | 1 | 3 | 3 | 3 |
| Minus 787 to -245 | 6 | 9 | 2 | 4 | 1.5 | 2.75 |
| Minus 787 to -245 | 6 | 11 | 1.5 | 3 | 4 | 3.5 |
| Minus 787 to -245 | 6 | 13 | 0 | 1.5 | 1.5 | 1.5 |
| Minus 787 to -405 | 4 | 2 | 1 | 2 | 1.5 | 1.75 |
| Minus 787 to -405 | 4 | 5 | 1 | 2 | 2 | 2 |
| Minus 787 to -405 | 4 | 7 | 1 | 2 | 2 | 2 |

| Reporter | Experiment | Embryo | Primary Score | Secondary Anterior Score | Secondary Posterior Score | Secondary (Average) |
|---|---|---|---|---|---|---|
| Minus 787 to -405 | 4 | 8 | 4 | 2.5 | 2.5 | 2.5 |
| Minus 787 to -405 | 4 | 13 | 2 | 1 | 1 | 1 |
| Minus 787 to -405 | 4 | 14 | 3 | 1 | 1.5 | 1.25 |
| Minus 787 to -405 | 4 | 15 | 0 | 2.5 | 2.5 | 2.5 |
| Minus 787 to -405 | 4 | 16 | 2 | 2.5 | 3 | 2.75 |
| Minus 787 to -405 | 5 | 2 | 1 | 1 | 2.5 | 1.75 |
| Minus 787 to -405 | 5 | 8 | 0 | 2 | 4 | 3 |
| Minus 787 to -405 | 5 | 9 | 1 | 1 | 1 | 1 |
| Minus 787 to -405 | 5 | 10 | 1.5 | 2.5 | 2.5 | 2.5 |
| Minus 787 to -405 | 5 | 11 | 2.5 | 0 | 3 | 1.5 |
| Minus 787 to -405 | 5 | 12 | 1 | 1 | 1 | 1 |
| Minus 787 to -405 | 5 | 14 | 4 | 4 | 3 | 3.5 |
| Minus 787 to -405 | 5 | 15 | 4 | 4 | 2 | 3 |
| Minus 787 to -405 | 6 | 5 | 4 | 2 | 2 | 2 |
| Minus 787 to -405 | 6 | 7 | 1 | 4 | 4 | 4 |
| Minus 787 to -405 | 6 | 8 | 0 | 4 | 4 | 4 |
| Minus 787 to -405 | 6 | 9 | 1.5 | 2 | 2.5 | 2.25 |
| Minus 787 to -405 | 6 | 11 | 1 | 4 | 4 | 4 |
| Minus 787 to -405 | 6 | 13 | 2 | 2 | 2 | 2 |
| Minus 787 to -405 | 6 | 14 | 0 | 2 | 2 | 2 |
| Minus 488 to +13 | 4 | 3 | 1 | 2 | 3 | 2.5 |
| Minus 488 to +13 | 4 | 4 | 1 | 2 | 3 | 2.5 |
| Minus 488 to +13 | 4 | 7 | 1.5 | 4 | 4 | 4 |
| Minus 488 to +13 | 4 | 8 | 2 | 3 | 3 | 3 |
| Minus 488 to +13 | 4 | 9 | 1 | 1.5 | 2.5 | 2 |
| Minus 488 to +13 | 4 | 11 | 2.5 | 1.5 | 1 | 1.25 |
| Minus 488 to +13 | 5 | 2 | 0 | 0 | 1 | 0.5 |
| Minus 488 to +13 | 5 | 3 | 0 | 1 | 2 | 1.5 |
| Minus 488 to +13 | 5 | 4 | 2 | 3 | 2.5 | 2.75 |
| Minus 488 to +13 | 5 | 5 | 1 | 2 | 2.5 | 2.25 |
| Minus 488 to +13 | 5 | 7 | 3 | 2 | 2.5 | 2.25 |
| Minus 488 to +13 | 5 | 8 | 1 | 2 | 2 | 2 |
| Minus 488 to +13 | 5 | 9 | 1 | 1 | 1 | 1 |
| Minus 488 to +13 | 5 | 10 | 2.5 | 2.5 | 2 | 2.25 |
| Minus 488 to +13 | 5 | 11 | 1 | 1 | 3 | 2 |
| Minus 488 to +13 | 5 | 12 | 1.5 | 1 | 2.5 | 1.75 |
| Minus 488 to +13 | 5 | 13 | 1 | 2.5 | 2.5 | 2.5 |
| Minus 488 to +13 | 5 | 14 | 1 | 1.5 | 1 | 1.25 |
| Minus 488 to +13 | 5 | 15 | 2.5 | 2.5 | 2 | 2.25 |
| Minus 488 to +13 | 6 | 3 | 2.5 | 2 | 1 | 1.5 |

| Reporter | Experiment | Embryo | Primary Score | Secondary Anterior Score | Secondary Posterior Score | Secondary (Average) |
|---|---|---|---|---|---|---|
| Minus 488 to +13 | 6 | 4 | 2.5 | 4 | 4 | 4 |
| Minus 488 to +13 | 6 | 5 | 2 | 4 | 4 | 4 |
| Minus 488 to +13 | 6 | 8 | 2.5 | 0 | 1 | 0.5 |
| Minus 488 to +13 | 6 | 9 | 0 | 1 | 2.5 | 1.75 |
| Minus 488 to +13 | 6 | 10 | 1 | 1 | 2.5 | 1.75 |
| Minus 488 to +13 | 6 | 11 | 3 | 2 | 4 | 3 |
| Minus 488 to +13 | 6 | 13 | 1 | 2 | 2 | 2 |
| Minus 488 to +13 | 6 | 14 | 1 | 2 | 4 | 3 |
| Minus 488 to +13 | 9 | 1 | 1 | 2 | 2.5 | 2.25 |
| Minus 488 to +13 | 9 | 2 | 0 | 1 | 2.5 | 1.75 |
| Minus 488 to +13 | 9 | 4 | 1 | 1 | 1.5 | 1.25 |
| Minus 488 to +13 | 9 | 5 | 1 | 1.5 | 2 | 1.75 |
| Minus 488 to +13 | 9 | 6 | 1 | 2.5 | 2.5 | 2.5 |
| Minus 488 to +13 | 9 | 7 | 0 | 1.5 | 1.5 | 1.5 |
| Minus 488 to +13 | 9 | 8 | 0 | 1 | 1 | 1 |
| Minus 488 to +13 | 9 | 9 | 2 | 1 | 1 | 1 |
| Minus 488 to +13 | 9 | 10 | 0 | 1 | 1.5 | 1.25 |
| Minus 488 to +13 | 9 | 11 | 4 | 1 | 1 | 1 |
| Minus 488 to +13 | 10 | 1 | 2.5 | 2.5 | 2.5 | 2.5 |
| Minus 488 to +13 | 10 | 2 | 3 | 0 | 1 | 0.5 |
| Minus 488 to +13 | 10 | 3 | 1.5 | 0 | 0 | 0 |
| Minus 488 to +13 | 10 | 4 | 0 | 1 | 1 | 1 |
| Minus 488 to +13 | 10 | 5 | 2 | 3 | 1 | 2 |
| Minus 488 to +13 | 10 | 6 | 0 | 2.5 | 2 | 2.25 |
| Minus 488 to +13 | 10 | 7 | 1 | 1.5 | 1.5 | 1.5 |
| Minus 488 to +13 | 10 | 8 | 1 | 1.5 | 1.5 | 1.5 |
| Minus 488 to +13 | 11 | 1 | 1.5 | 1 | 1 | 1 |
| Minus 488 to +13 | 11 | 2 | 3 | 2.5 | 3 | 2.75 |
| Minus 488 to +13 | 11 | 3 | 1 | 1 | 1 | 1 |
| Minus 488 to +13 | 11 | 4 | 2 | 2 | 3 | 2.5 |
| Minus 488 to +13 | 11 | 6 | 1.5 | 1 | 1 | 1 |
| Minus 488 to +13 | 11 | 7 | 2 | 3 | 3 | 3 |
| Minus 488 to +13 | 14 | 1 | 1 | 0 | 2.5 | 1.25 |
| Minus 488 to +13 | 14 | 2 | 1 | 2.5 | 3 | 2.75 |
| Minus 488 to +13 | 14 | 3 | 1 | 2.5 | 1.5 | 2 |
| Minus 488 to +13 | 14 | 4 | 2.5 | 2 | 2 | 2 |
| Minus 488 to +13 | 14 | 5 | 1.5 | 1.5 | 1.5 | 1.5 |
| Minus 488 to +13 | 14 | 6 | 1.5 | 1 | 1 | 1 |
| Minus 488 to +13 | 14 | 7 | 0 | 2.5 | 2.5 | 2.5 |
| Minus 488 to +13 | 14 | 8 | 0 | 1 | 2.5 | 1.75 |

| Reporter | Experiment | Embryo | Primary Score | Secondary Anterior Score | Secondary Posterior Score | Secondary (Average) |
|---|---|---|---|---|---|---|
| Minus 488 to +13 | 15 | 1 | 0 | 1 | 0 | 0.5 |
| Minus 488 to +13 | 15 | 2 | 0 | 4 | 4 | 4 |
| Minus 488 to +13 | 15 | 3 | 0 | 1.5 | 2.5 | 2 |
| Minus 488 to +13 | 15 | 4 | 1 | 0 | 1 | 0.5 |
| Minus 488 to +13 | 15 | 5 | 0 | 1 | 1 | 1 |
| Minus 488 to +13 | 15 | 6 | 0 | 1.5 | 1.5 | 1.5 |
| Minus 488 to +13 | 15 | 7 | 0 | 1 | 1 | 1 |
| Minus 488 to +13 | 15 | 8 | 1.5 | 1 | 1 | 1 |
| Minus 488 to +13 | 16 | 1 | 2 | 2 | 3 | 2.5 |
| Minus 488 to +13 | 16 | 2 | 1 | 1.5 | 1.5 | 1.5 |
| Minus 488 to +13 | 16 | 3 | 3 | 2 | 2 | 2 |
| Minus 488 to +13 | 16 | 4 | 0 | 2 | 1.5 | 1.75 |
| Minus 488 to +13 | 16 | 5 | 0 | 1.5 | 1 | 1.25 |
| Minus 488 to -165 | 9 | 1 | 3 | 3 | 3 | 3 |
| Minus 488 to -165 | 9 | 2 | 2 | 1.5 | 1.5 | 1.5 |
| Minus 488 to -165 | 9 | 3 | 2.5 | 2 | 2.5 | 2.25 |
| Minus 488 to -165 | 9 | 4 | 2 | 1 | 1 | 1 |
| Minus 488 to -165 | 9 | 5 | 1 | 2 | 2 | 2 |
| Minus 488 to -165 | 9 | 6 | 4 | 4 | 4 | 4 |
| Minus 488 to -165 | 9 | 7 | 3 | 1.5 | 2 | 1.75 |
| Minus 488 to -165 | 9 | 8 | 2.5 | 2 | 2.5 | 2.25 |
| Minus 488 to -165 | 10 | 1 | 4 | 3 | 3 | 3 |
| Minus 488 to -165 | 10 | 2 | 4 | 2 | 2 | 2 |
| Minus 488 to -165 | 11 | 1 | 2.5 | 2.5 | 2.5 | 2.5 |
| Minus 488 to -165 | 11 | 3 | 2.5 | 4 | 4 | 4 |
| Minus 488 to -165 | 11 | 5 | 4 | 1 | 1 | 1 |
| Minus 488 to -165 | 11 | 6 | 2 | 2 | 2.5 | 2.25 |
| Minus 488 to -165 | 11 | 7 | 4 | 3 | 4 | 3.5 |
| Minus 488 to -245 | 9 | 1 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 2 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 3 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 4 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 5 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 6 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 7 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 8 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 9 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 10 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 11 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 12 | 0 | 0 | 0 | 0 |

| Reporter | Experiment | Embryo | Primary Score | Secondary Anterior Score | Secondary Posterior Score | Secondary (Average) |
|----------|-----------|--------|---------------|--------------------------|---------------------------|---------------------|
| Minus 488 to -245 | 9 | 13 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 1 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 2 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 3 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 4 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 5 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 6 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 7 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 8 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 9 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 1 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 2 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 3 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 4 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 5 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 6 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 7 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 8 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 9 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 10 | 0 | 0 | 0 | 0 |
| Minus 245 to +13 | 14 | 1 | 0 | 1 | 1 | 1 |
| Minus 245 to +13 | 14 | 2 | 1.5 | 3 | 3 | 3 |
| Minus 245 to +13 | 14 | 3 | 0 | 2 | 2 | 2 |
| Minus 245 to +13 | 14 | 4 | 0 | 1 | 1 | 1 |
| Minus 245 to +13 | 14 | 5 | 0 | 2.5 | 2.5 | 2.5 |
| Minus 245 to +13 | 14 | 6 | 1 | 0 | 0 | 0 |
| Minus 245 to +13 | 15 | 1 | 3 | 1.5 | 1 | 1.25 |
| Minus 245 to +13 | 15 | 2 | 0 | 1 | 1 | 1 |
| Minus 245 to +13 | 15 | 3 | 2 | 1 | 1 | 1 |
| Minus 245 to +13 | 15 | 4 | 2 | 1 | 1 | 1 |
| Minus 245 to +13 | 15 | 5 | 0 | 1 | 1 | 1 |
| Minus 245 to +13 | 15 | 6 | 2 | 2 | 2 | 2 |
| Minus 245 to +13 | 15 | 7 | 0 | 1 | 1 | 1 |
| Minus 245 to +13 | 16 | 1 | 1 | 3 | 3 | 3 |
| Minus 245 to +13 | 16 | 2 | 2 | 1 | 1 | 1 |
| Minus 322 to +13 | 9 | 1 | 0 | 3 | 3 | 3 |
| Minus 322 to +13 | 9 | 2 | 1 | 2 | 2 | 2 |
| Minus 322 to +13 | 9 | 3 | 1 | 3 | 3 | 3 |
| Minus 322 to +13 | 9 | 4 | 1.5 | 1.5 | 1.5 | 1.5 |
| Minus 322 to +13 | 9 | 5 | 1 | 2 | 2 | 2 |

| Reporter | Experiment | Embryo | Primary Score | Secondary Anterior Score | Secondary Posterior Score | Secondary (Average) |
|---|---|---|---|---|---|---|
| Minus 322 to +13 | 9 | 6 | 0 | 3 | 3 | 3 |
| Minus 322 to +13 | 9 | 7 | 1 | 2.5 | 2.5 | 2.5 |
| Minus 322 to +13 | 9 | 8 | 1 | 2 | 2 | 2 |
| Minus 322 to +13 | 9 | 9 | 4 | 1 | 1 | 1 |
| Minus 322 to +13 | 9 | 10 | 1 | 2.5 | 2.5 | 2.5 |
| Minus 322 to +13 | 9 | 11 | 1 | 2 | 2 | 2 |
| Minus 322 to +13 | 10 | 1 | 1 | 4 | 4 | 4 |
| Minus 322 to +13 | 10 | 2 | 0 | 2.5 | 3 | 2.75 |
| Minus 322 to +13 | 10 | 3 | 1 | 3 | 3 | 3 |
| Minus 322 to +13 | 10 | 4 | 1 | 3 | 4 | 3.5 |
| Minus 322 to +13 | 11 | 1 | 0 | 2 | 2 | 2 |
| Minus 322 to +13 | 11 | 2 | 1.5 | 3 | 3 | 3 |
| Minus 322 to +13 | 11 | 3 | 2 | 4 | 4 | 4 |
| Minus 322 to +13 | 11 | 4 | 1 | 2.5 | 2.5 | 2.5 |
| Minus 322 to +13 | 11 | 5 | 4 | 2.5 | 3 | 2.75 |
| Minus 322 to +13 | 11 | 6 | 4 | 3 | 4 | 3.5 |
| Minus 322 to +13 | 11 | 7 | 0 | 2.5 | 2.5 | 2.5 |
| Minus 322 to +13 | 11 | 8 | 4 | 4 | 4 | 4 |

**Table A.2 Averaged Qualitative Scores**

| Reporter | Primary mean score | Secondary mean score | n |
|---|---|---|---|
| Minus 1565 to +13 | 2.0 | 2.5 | 50 |
| Minus 1565 to -758 | 0.6 | 1.1 | 54 |
| Minus 1045 to -526 | 0.5 | 1.3 | 31 |
| Minus 787 to +13 | 2.0 | 2.4 | 73 |
| Minus 787 to -245 | 1.3 | 2.2 | 25 |
| Minus 787 to -405 | 1.7 | 2.2 | 23 |
| Minus 488 to +13 | 1.2 | 1.7 | 73 |
| Minus 488 to -165 | 2.9 | 2.3 | 15 |
| Minus 488 to -245 | 0.0 | 0.0 | 32 |
| Minus 245 to +13 | 1.0 | 1.5 | 15 |
| Minus 322 to +13 | 1.4 | 2.6 | 23 |

**Table A.3 Individual Ectopic Scores**

CNS: Central Nervous System, Epi: Epidermis, Mus: Muscle, End: Endoderm, Mes: Mesenchyme

| Reporter | Experiment | Embryo | CNS Score | Epi Score | Mus Score | End Score | Mes Score |
|---|---|---|---|---|---|---|---|
| Minus 1565 to +13 | 32 | 1 | 2 | 1.5 | 0 | 1 | 1 |
| Minus 1565 to +13 | 32 | 2 | 0 | 0 | 0 | 0 | 0 |
| Minus 1565 to +13 | 32 | 3 | 2 | 1 | 1 | 2.5 | 1 |
| Minus 1565 to +13 | 32 | 5 | 1.5 | 0 | 0 | 1 | 1 |
| Minus 1565 to +13 | 32 | 6 | 0 | 0 | 2 | 0 | 2 |
| Minus 1565 to +13 | 32 | 7 | 1 | 0 | 0 | 0 | 0 |
| Minus 1565 to +13 | 33 | 1 | 2.5 | 0 | 0 | 2 | 2 |
| Minus 1565 to +13 | 33 | 2 | 0 | 0 | 1 | 3 | 3 |
| Minus 1565 to +13 | 33 | 3 | 2 | 0 | 1 | 2.5 | 1.5 |
| Minus 1565 to +13 | 33 | 4 | 2 | 0 | 2 | 3 | 4 |
| Minus 1565 to +13 | 33 | 5 | 1.5 | 1 | 1 | 1 | 2 |
| Minus 1565 to +13 | 33 | 6 | 2.5 | 2 | 3 | 2.5 | 1.5 |
| Minus 1565 to +13 | 34 | 1 | 2 | 2 | 1 | 0 | 2 |
| Minus 1565 to +13 | 34 | 5 | 0 | 0 | 0 | 1.5 | 1.5 |
| Minus 1565 to -758 | 32 | 3 | 0 | 0 | 0 | 1 | 1.5 |
| Minus 1565 to -758 | 32 | 4 | 0 | 0 | 0 | 0 | 0 |
| Minus 1565 to -758 | 32 | 5 | 0 | 0 | 1 | 0 | 2 |
| Minus 1565 to -758 | 32 | 6 | 1 | 0 | 1 | 1 | 2 |
| Minus 1565 to -758 | 33 | 1 | 0 | 0 | 0 | 1 | 2 |
| Minus 1565 to -758 | 33 | 3 | 2 | 1 | 0 | 1 | 0 |
| Minus 1565 to -758 | 33 | 6 | 0 | 0 | 0 | 1 | 1 |
| Minus 1565 to -758 | 34 | 4 | 1 | 1 | 1 | 1 | 1.5 |
| Minus 1565 to -758 | 34 | 5 | 0 | 0 | 1 | 1 | 2 |
| Minus 1045 to -526 | 32 | 1 | 2 | 2 | 1 | 2 | 1.5 |
| Minus 1045 to -526 | 32 | 2 | 2 | 2 | 2 | 1 | 2 |
| Minus 1045 to -526 | 32 | 4 | 2.5 | 3 | 2 | 1 | 1.5 |
| Minus 1045 to -526 | 32 | 5 | 2 | 2 | 2.5 | 2 | 1.5 |
| Minus 1045 to -526 | 32 | 6 | 1 | 0 | 2.5 | 1 | 1 |
| Minus 1045 to -526 | 32 | 7 | 2 | 2 | 2 | 1 | 1 |
| Minus 1045 to -526 | 33 | 1 | 2 | 2 | 2 | 0 | 2 |
| Minus 1045 to -526 | 33 | 2 | 1 | 2 | 2 | 1.5 | 1 |
| Minus 1045 to -526 | 33 | 3 | 1.5 | 1.5 | 2 | 2.5 | 1 |
| Minus 1045 to -526 | 34 | 2 | 1 | 1.5 | 1.5 | 1.5 | 2.5 |
| Minus 1045 to -526 | 34 | 4 | 1 | 1 | 0 | 1 | 1 |
| Minus 1045 to -526 | 34 | 5 | 2.5 | 2.5 | 2 | 3 | 2 |
| Minus 1045 to -526 | 34 | 7 | 0 | 0 | 1.5 | 2.5 | 1.5 |

| Reporter | Experiment | Embryo | CNS Score | Epi Score | Mus Score | End Score | Mes Score |
|----------|------------|--------|-----------|-----------|-----------|-----------|-----------|
| Minus 787 to +13 | 4 | 1 | 0 | 1 | 2 | 1 | 1 |
| Minus 787 to +13 | 4 | 2 | 2.5 | 2 | 2 | 1.5 | 2 |
| Minus 787 to +13 | 4 | 3 | 0 | 2 | 3 | 1 | 2.5 |
| Minus 787 to +13 | 4 | 5 | 3 | 2 | 3 | 3 | 2 |
| Minus 787 to +13 | 4 | 6 | 2.5 | 2.5 | 3 | 1.5 | 1.5 |
| Minus 787 to +13 | 4 | 8 | 2 | 1.5 | 2 | 1 | 1.5 |
| Minus 787 to +13 | 4 | 10 | 1.5 | 2 | 2 | 1 | 1 |
| Minus 787 to +13 | 4 | 11 | 1.5 | 1.5 | 2 | 3 | 2 |
| Minus 787 to +13 | 5 | 1 | 1 | 1 | 1 | 2 | 2 |
| Minus 787 to +13 | 5 | 2 | 1.5 | 1 | 1.5 | 1 | 2 |
| Minus 787 to +13 | 5 | 3 | 3 | 3 | 3 | 1 | 1 |
| Minus 787 to +13 | 5 | 6 | 1 | 1 | 1 | 1 | 1 |
| Minus 787 to +13 | 5 | 7 | 2 | 2 | 2 | 1 | 1 |
| Minus 787 to +13 | 5 | 8 | 1 | 1 | 2 | 1 | 1 |
| Minus 787 to +13 | 5 | 12 | 2 | 2 | 1.5 | 1 | 1 |
| Minus 787 to +13 | 5 | 15 | 1.5 | 1.5 | 1.5 | 1 | 1 |
| Minus 787 to +13 | 5 | 16 | 1 | 1 | 1 | 1 | 1 |
| Minus 787 to +13 | 6 | 3 | 2 | 2 | 3 | 1 | 2 |
| Minus 787 to +13 | 6 | 5 | 1 | 1 | 0 | 1 | 1.5 |
| Minus 787 to +13 | 6 | 6 | 1.5 | 1 | 1 | 2 | 2.5 |
| Minus 787 to +13 | 6 | 8 | 2 | 1 | 1 | 2 | 2 |
| Minus 787 to +13 | 6 | 9 | 0 | 1 | 1.5 | 2.5 | 2 |
| Minus 787 to +13 | 6 | 10 | 1 | 1.5 | 1 | 1 | 2 |
| Minus 787 to +13 | 6 | 12 | 1.5 | 1 | 2.5 | 2.5 | 2.5 |
| Minus 787 to +13 | 6 | 13 | 0 | 1 | 1 | 2 | 2 |
| Minus 787 to +13 | 32 | 1 | 2 | 1 | 2 | 2 | 2 |
| Minus 787 to +13 | 32 | 2 | 3 | 1.5 | 1 | 0 | 2 |
| Minus 787 to +13 | 32 | 4 | 1.5 | 1 | 0 | 0 | 1 |
| Minus 787 to +13 | 32 | 5 | 1.5 | 1.5 | 3 | 1.5 | 1.5 |
| Minus 787 to +13 | 32 | 6 | 0 | 0 | 0 | 1.5 | 3 |
| Minus 787 to +13 | 32 | 7 | 1 | 0 | 0 | 1 | 2 |
| Minus 787 to +13 | 33 | 1 | 2 | 0 | 1 | 1 | 2.5 |
| Minus 787 to +13 | 33 | 2 | 0 | 1 | 1 | 1.5 | 3 |
| Minus 787 to +13 | 33 | 3 | 0 | 1 | 1 | 2 | 2 |
| Minus 787 to +13 | 33 | 4 | 1 | 1 | 2 | 1 | 3 |
| Minus 787 to +13 | 33 | 5 | 0 | 3 | 1 | 2 | 1 |
| Minus 787 to +13 | 33 | 6 | 1 | 1 | 1 | 1 | 1.5 |
| Minus 787 to +13 | 34 | 1 | 1 | 1 | 1 | 2.5 | 2.5 |
| Minus 787 to +13 | 34 | 2 | 0 | 2 | 1.5 | 2 | 2 |
| Minus 787 to +13 | 34 | 3 | 1 | 2 | 1 | 1.5 | 2 |
| Minus 787 to +13 | 34 | 4 | 0 | 0 | 1 | 1 | 1.5 |

| Reporter | Experiment | Embryo | CNS Score | Epi Score | Mus Score | End Score | Mes Score |
|---|---|---|---|---|---|---|---|
| Minus 787 to +13 | 34 | 5 | 0 | 0 | 1 | 1 | 2 |
| Minus 787 to +13 | 34 | 6 | 1 | 0 | 0 | 1 | 0 |
| Minus 787 to -245 | 4 | 1 | 0 | 0 | 1 | 2 | 2 |
| Minus 787 to -245 | 4 | 2 | 1 | 1 | 1 | 1 | 2.5 |
| Minus 787 to -245 | 4 | 3 | 1 | 1 | 1 | 1 | 2 |
| Minus 787 to -245 | 4 | 9 | 1 | 1 | 1 | 1 | 2 |
| Minus 787 to -245 | 4 | 10 | 0 | 1 | 1 | 2 | 2 |
| Minus 787 to -245 | 4 | 13 | 1 | 2 | 1 | 2 | 3 |
| Minus 787 to -245 | 5 | 1 | 1.5 | 1.5 | 1 | 1 | 1.5 |
| Minus 787 to -245 | 5 | 4 | 0 | 1 | 1 | 1 | 3 |
| Minus 787 to -245 | 5 | 7 | 2 | 1.5 | 2 | 2 | 2 |
| Minus 787 to -245 | 5 | 9 | 1.5 | 0 | 1 | 1 | 1 |
| Minus 787 to -245 | 5 | 13 | 0 | 1 | 0 | 1.5 | 1 |
| Minus 787 to -245 | 5 | 15 | 1.5 | 1.5 | 1 | 2.5 | 1.5 |
| Minus 787 to -245 | 5 | 16 | 0 | 1 | 1 | 2.5 | 1 |
| Minus 787 to -245 | 6 | 2 | 0 | 1.5 | 2 | 2.5 | 2.5 |
| Minus 787 to -245 | 6 | 5 | 0 | 1 | 3 | 2.5 | 2.5 |
| Minus 787 to -245 | 6 | 8 | 0 | 1 | 0 | 1.5 | 2.5 |
| Minus 787 to -245 | 6 | 10 | 1 | 2 | 2.5 | 2 | 3 |
| Minus 787 to -245 | 6 | 11 | 2.5 | 2.5 | 1.5 | 2.5 | 1.5 |
| Minus 787 to -245 | 6 | 13 | 0 | 0 | 0 | 1 | 1 |
| Minus 787 to -405 | 4 | 2 | 0 | 0 | 0 | 2 | 3 |
| Minus 787 to -405 | 4 | 5 | 3 | 1 | 2.5 | 2 | 3 |
| Minus 787 to -405 | 4 | 7 | 1 | 1 | 1.5 | 2.5 | 2.5 |
| Minus 787 to -405 | 4 | 8 | 1 | 1 | 1.5 | 3 | 4 |
| Minus 787 to -405 | 4 | 13 | 0 | 1 | 1 | 2 | 3 |
| Minus 787 to -405 | 4 | 15 | 0 | 0 | 0 | 2.5 | 3 |
| Minus 787 to -405 | 4 | 16 | 2 | 1 | 1.5 | 2.5 | 2.5 |
| Minus 787 to -405 | 5 | 2 | 1 | 0 | 1 | 1.5 | 1 |
| Minus 787 to -405 | 5 | 9 | 0 | 1 | 0 | 2 | 1.5 |
| Minus 787 to -405 | 5 | 10 | 1 | 2 | 1 | 1 | 2 |
| Minus 787 to -405 | 5 | 15 | 0 | 1 | 1 | 1 | 1.5 |
| Minus 787 to -405 | 6 | 5 | 0 | 0 | 0 | 2.5 | 2 |
| Minus 787 to -405 | 6 | 8 | 1 | 1 | 1.5 | 2.5 | 2.5 |
| Minus 787 to -405 | 6 | 9 | 1 | 1 | 2 | 2.5 | 3 |
| Minus 787 to -405 | 6 | 11 | 1 | 2.5 | 1 | 3 | 2.5 |
| Minus 787 to -405 | 6 | 13 | 0 | 1 | 1.5 | 2.5 | 2.5 |
| Minus 787 to -405 | 6 | 14 | 1.5 | 2 | 1 | 2.5 | 1 |
| Minus 488 to +13 | 4 | 3 | 2 | 2 | 1.5 | 1.5 | 3 |
| Minus 488 to +13 | 4 | 4 | 2 | 2 | 2.5 | 1.5 | 1 |
| Minus 488 to +13 | 4 | 8 | 0 | 1 | 1 | 2 | 2 |

| Reporter | Experiment | Embryo | CNS Score | Epi Score | Mus Score | End Score | Mes Score |
|---|---|---|---|---|---|---|---|
| Minus 488 to +13 | 4 | 9 | 1.5 | 1.5 | 1.5 | 1 | 1 |
| Minus 488 to +13 | 5 | 4 | 0 | 1 | 1 | 1 | 1.5 |
| Minus 488 to +13 | 5 | 5 | 1 | 1 | 1 | 1 | 1 |
| Minus 488 to +13 | 5 | 7 | 1 | 1 | 1 | 2 | 1 |
| Minus 488 to +13 | 5 | 8 | 1 | 1 | 1 | 2.5 | 1 |
| Minus 488 to +13 | 5 | 10 | 1 | 1 | 1 | 2.5 | 1.5 |
| Minus 488 to +13 | 5 | 12 | 1 | 1 | 1.5 | 1.5 | 1 |
| Minus 488 to +13 | 5 | 14 | 1 | 1 | 1 | 1 | 1 |
| Minus 488 to +13 | 6 | 3 | 0 | 1 | 1.5 | 2.5 | 1.5 |
| Minus 488 to +13 | 6 | 8 | 1 | 1 | 1 | 1 | 1 |
| Minus 488 to +13 | 6 | 9 | 0 | 1 | 1 | 0 | 2 |
| Minus 488 to +13 | 6 | 14 | 0 | 0 | 0 | 1.5 | 2 |
| Minus 488 to +13 | 9 | 1 | 0 | 0 | 0 | 1 | 1.5 |
| Minus 488 to +13 | 9 | 2 | 0 | 1 | 1.5 | 1.5 | 1.5 |
| Minus 488 to +13 | 9 | 5 | 0 | 0 | 0 | 1 | 1.5 |
| Minus 488 to +13 | 9 | 6 | 4 | 3 | 4 | 1 | 1.5 |
| Minus 488 to +13 | 9 | 7 | 0 | 1 | 1 | 1.5 | 2 |
| Minus 488 to +13 | 9 | 8 | 0 | 0 | 0 | 1.5 | 2 |
| Minus 488 to +13 | 9 | 9 | 3 | 0 | 2 | 2 | 3 |
| Minus 488 to +13 | 10 | 1 | 0 | 1 | 1 | 1 | 2 |
| Minus 488 to +13 | 10 | 2 | 0 | 0 | 0 | 1 | 2.5 |
| Minus 488 to +13 | 10 | 3 | 0 | 1.5 | 1.5 | 1.5 | 0 |
| Minus 488 to +13 | 10 | 5 | 2 | 2 | 1.5 | 0 | 1 |
| Minus 488 to +13 | 10 | 6 | 0 | 1 | 0 | 0 | 2 |
| Minus 488 to +13 | 10 | 7 | 0 | 1.5 | 1 | 1.5 | 2.5 |
| Minus 488 to +13 | 10 | 8 | 1.5 | 1.5 | 2.5 | 1.5 | 1 |
| Minus 488 to +13 | 11 | 1 | 2.5 | 2.5 | 2.5 | 1 | 1.5 |
| Minus 488 to +13 | 11 | 2 | 2 | 1 | 1 | 1 | 1 |
| Minus 488 to +13 | 11 | 4 | 1.5 | 1.5 | 2 | 1 | 2 |
| Minus 488 to +13 | 14 | 3 | 0 | 1 | 0 | 1 | 2 |
| Minus 488 to +13 | 14 | 4 | 0 | 0 | 0 | 0 | 2.5 |
| Minus 488 to +13 | 14 | 5 | 0 | 1 | 1 | 0 | 4 |
| Minus 488 to +13 | 14 | 6 | 1.5 | 1.5 | 0 | 0 | 1 |
| Minus 488 to +13 | 14 | 7 | 0 | 1 | 0 | 3 | 2.5 |
| Minus 488 to +13 | 14 | 8 | 1.5 | 1.5 | 1 | 0 | 1.5 |
| Minus 488 to +13 | 15 | 1 | 0 | 1 | 1 | 0 | 1.5 |
| Minus 488 to +13 | 15 | 2 | 0 | 0 | 0 | 0 | 1.5 |
| Minus 488 to +13 | 15 | 5 | 0 | 1 | 1.5 | 0 | 1.5 |
| Minus 488 to +13 | 15 | 7 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to +13 | 16 | 1 | 1.5 | 1 | 1 | 1 | 1.5 |
| Minus 488 to +13 | 16 | 2 | 1.5 | 1 | 1.5 | 1 | 2 |

| Reporter | Experiment | Embryo | CNS Score | Epi Score | Mus Score | End Score | Mes Score |
|---|---|---|---|---|---|---|---|
| Minus 488 to +13 | 16 | 3 | 1 | 1.5 | 1.5 | 1.5 | 1 |
| Minus 488 to +13 | 16 | 5 | 1 | 1 | 1 | 1 | 2 |
| Minus 488 to -165 | 9 | 1 | 2 | 2 | 2.5 | 4 | 2 |
| Minus 488 to -165 | 9 | 3 | 0 | 1 | 1 | 4 | 2 |
| Minus 488 to -165 | 9 | 6 | 2 | 2 | 2 | 3 | 3 |
| Minus 488 to -165 | 10 | 1 | 4 | 2.5 | 4 | 4 | 2 |
| Minus 488 to -165 | 11 | 1 | 3 | 2 | 3 | 4 | 2.5 |
| Minus 488 to -165 | 11 | 3 | 2.5 | 1.5 | 2.5 | 4 | 1.5 |
| Minus 488 to -165 | 11 | 5 | 2.5 | 2.5 | 2.5 | 2.5 | 1.5 |
| Minus 488 to -165 | 11 | 6 | 1.5 | 1 | 1.5 | 2.5 | 2.5 |
| Minus 488 to -165 | 11 | 7 | 1 | 1 | 1 | 4 | 1 |
| Minus 488 to -245 | 9 | 1 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 3 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 4 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 8 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 9 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 10 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 12 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 9 | 13 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 2 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 3 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 4 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 10 | 5 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 1 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 3 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 5 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 6 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 8 | 0 | 0 | 0 | 0 | 0 |
| Minus 488 to -245 | 11 | 10 | 0 | 0 | 0 | 0 | 0 |
| Minus 245 to +13 | 14 | 3 | 2 | 1 | 2 | 1 | 2.5 |
| Minus 245 to +13 | 14 | 4 | 0 | 0 | 1 | 0 | 1.5 |
| Minus 245 to +13 | 14 | 5 | 0 | 1 | 1 | 2.5 | 3 |
| Minus 245 to +13 | 14 | 6 | 1.5 | 0 | 1 | 0 | 1 |
| Minus 245 to +13 | 15 | 1 | 0 | 1 | 1 | 0 | 1.5 |
| Minus 245 to +13 | 15 | 2 | 1.5 | 1.5 | 1 | 0 | 1 |
| Minus 245 to +13 | 15 | 3 | 2.5 | 3 | 3 | 1.5 | 2 |
| Minus 245 to +13 | 15 | 4 | 2.5 | 1 | 2.5 | 1 | 1 |
| Minus 245 to +13 | 15 | 5 | 2 | 2 | 2.5 | 2 | 2.5 |
| Minus 245 to +13 | 16 | 1 | 1.5 | 2 | 3 | 1 | 2 |
| Minus 322 to +13 | 9 | 1 | 2 | 2 | 2.5 | 2 | 3 |

| Reporter | Experiment | Embryo | CNS Score | Epi Score | Mus Score | End Score | Mes Score |
|---|---|---|---|---|---|---|---|
| Minus 322 to +13 | 9 | 2 | 0 | 1 | 2 | 3 | 0 |
| Minus 322 to +13 | 9 | 3 | 1 | 0 | 3 | 3 | 4 |
| Minus 322 to +13 | 9 | 4 | 0 | 0 | 1.5 | 2.5 | 2.5 |
| Minus 322 to +13 | 9 | 5 | 0 | 0 | 2 | 2 | 3 |
| Minus 322 to +13 | 9 | 6 | 0 | 0 | 1.5 | 4 | 3 |
| Minus 322 to +13 | 9 | 7 | 0 | 0 | 2.5 | 4 | 3 |
| Minus 322 to +13 | 9 | 8 | 0 | 1 | 3 | 3 | 3 |
| Minus 322 to +13 | 9 | 10 | 2 | 2 | 1.5 | 3 | 2.5 |
| Minus 322 to +13 | 9 | 11 | 2 | 1 | 2 | 3 | 1 |
| Minus 322 to +13 | 10 | 1 | 1 | 1 | 2 | 3 | 4 |
| Minus 322 to +13 | 10 | 2 | 2 | 2 | 2 | 3 | 4 |
| Minus 322 to +13 | 10 | 4 | 1 | 1 | 2 | 3 | 4 |
| Minus 322 to +13 | 11 | 1 | 1 | 1 | 2 | 3 | 3 |
| Minus 322 to +13 | 11 | 3 | 2 | 1 | 3 | 3 | 4 |
| Minus 322 to +13 | 11 | 4 | 2.5 | 2.5 | 2 | 4 | 4 |
| Minus 322 to +13 | 11 | 5 | 1 | 1 | 2 | 2.5 | 3 |
| Minus 322 to +13 | 11 | 7 | 1 | 1 | 2 | 2.5 | 2.5 |
| Minus 322 to +13 | 11 | 8 | 3 | 1.5 | 3 | 3 | 4 |

**Table A.4 Averaged Ectopic Scores**

| Reporter | CNS | Epi | Mus | End | Mes | Average | n |
|---|---|---|---|---|---|---|---|
| Minus 1565 to +13 | 1.4 | 0.5 | 0.9 | 1.4 | 1.6 | 1.2 | 14 |
| Minus 1565 to -758 | 0.4 | 0.2 | 0.4 | 0.8 | 1.3 | 0.6 | 9 |
| Minus 1045 to -526 | 1.6 | 1.7 | 1.8 | 1.5 | 1.5 | 1.6 | 13 |
| Minus 787 to +13 | 1.2 | 1.3 | 1.5 | 1.4 | 1.8 | 1.4 | 43 |
| Minus 787 to -245 | 0.7 | 1.1 | 1.2 | 1.7 | 2.0 | 1.3 | 19 |
| Minus 787 to -405 | 0.8 | 1.0 | 1.1 | 2.2 | 2.4 | 1.5 | 17 |
| Minus 488 to +13 | 0.8 | 1.0 | 1.1 | 1.1 | 1.6 | 1.1 | 46 |
| Minus 488 to -165 | 2.1 | 1.7 | 2.2 | 3.6 | 2.0 | 2.3 | 9 |
| Minus 488 to -245 | 0 | 0 | 0 | 0 | 0 | 0 | 18 |
| Minus 245 to +13 | 1.4 | 1.3 | 1.8 | 0.9 | 1.8 | 1.4 | 10 |
| Minus 322 to +13 | 1.1 | 1.0 | 2.2 | 3.0 | 3.0 | 2.1 | 19 |

**Table A.5 Reporter Construct cloning primers**

| C11.331 enhancer region | forward primer | reverse primer | genomic coordinates |
|---|---|---|---|
| -1565 to +13 | CAGGTGCCACAAATAAACC | CCTATTTGTCCTTCTGAAATAACAG | KHC11:4060949-4062548 |
| -1565 to -758 | CAGGTGCCACAAATAAACC | ATAAAGTCCCACGATACGCAG | KHC11:4060949-4658618 |
| -1045 to -526 | ATGCATAAGTTGGCAGAGATACG | TTCTACATTTCCTCCTTACAGCG | KHC11:4061470-4062012 |
| -787 to +13 | AATTGCGAACTGCGTATCG | CCTATTTGTCCTTCTGAAATAACAG | KHC11:4061902-4062548 |
| -787 to -245 | AATTGCGAACTGCGTATCG | TACGCAGTAATTGCAGCTTTCC | KHC11:4061902-4062291 |
| -787 to -405 | AATTGCGAACTGCGTATCG | AACAGACAAATAGCAAAAACACTACC | KHC11:4061902-4062131 |
| -488 to +13 | TATTTTACTTTCGTGATAACTGC | CCTATTTGTCCTTCTGAAATAACAG | KHC11:4062051-4062548 |
| -488 to -245 | TATTTTACTTTCGTGATAACTGC | TACGCAGTAATTGCAGCTTTCC | KHC11:4062051-4062291 |
| -488 to -165 | TATTTTACTTTCGTGATAACTGC | CTAATATCCGGAAGGTTTGATG | KHC11:4062051-4062371 |
| -322 to +13 | AACAAAATTCAATTGCACGGC | CCTATTTGTCCTTCTGAAATAACAG | KHC11:4062214-4062548 |
| -322 to -92 | AACAAAATTCAATTGCACGGC | CTTTCAAATTTTAAATCACGAGACAG | KHC11:4062214-4062444 |
| -322 to -165 | AACAAAATTCAATTGCACGGC | CTAATATCCGGAAGGTTTGATG | KHC11:4062214-4062371 |
| -322 to -127 | AACAAAATTCAATTGCACGGC | AGTTTTCAATAAACCGAAATTCG | KHC11:4062214-4062409 |
| -245 to +13 | AAGGTTCAGTAGTTAGTTTAAGCG | CCTATTTGTCCTTCTGAAATAACAG | KHC11:4062291-4062548 |

**Table A.6 TFBS Mutagenesis Primers**

| TFBS to be mutated | forward primer | reverse primer |
|---|---|---|
| ETS | TCATCAAACCagagGGATATTAGTTAATTCAGATAAAAC | TCTTAATAAAACACCTCTACTC |
| GATA | AGTTAATTCActatAAACGAATTTCGGTTTATTGAAAACTTAGTTC | AATATCCGGAAGGTTTGATG |
| Myb | ATTCAGATAAggtGAATTTCGGTTTATTGAAAACTTAG | TAACTAATATCCGGAAGG |
| LMX1A | GATATTAGTTtggaCAGATAAAACGAATTTCGGTTTATTG | CGGAAGGTTTGATGATCTTAATAAAAC |

# Appendix B - Chapter 4 Supplement

**Table B.1 Complete bootstrap parameter estimates for all reporters**

| Reporter | Parameter | 95% CI Median Lower Limit | Median | 95% CI Median Upper Limit | Mean |
|---|---|---|---|---|---|
| Proximal | Top Plateau | 0.367 | 0.409 | 0.451 | 0.409 |
| Proximal | Hill Coefficient | -36.6 | -5.00 | -3.34 | -9.49 |
| Proximal | EC50 | 0.338 | 0.487 | 0.616 | 0.474 |
| Proximal | Bottom Plateau | 0.0317 | 0.0403 | 0.0667 | 0.0424 |
| Distal | Top Plateau | 0.279 | 0.340 | 0.457 | 0.345 |
| Distal | Hill Coefficient | -2.55 | -1.43 | -0.807 | -2.07 |
| Distal | EC50 | 0.0327 | 0.0686 | 0.109 | 0.0659 |
| Distal | Bottom Plateau | 0.0196 | 0.0365 | 0.0556 | 0.0369 |
| Full Length | Top Plateau | 0.871 | 1.02 | 1.48 | 1.04 |
| Full Length | Hill Coefficient A | -35.9 | -26.9 | -1.03 | -22.2 |
| Full Length | EC50 A | 0.00855 | 0.0928 | 0.102 | 0.0720 |
| Full Length | Middle Plateau | 0.469 | 0.583 | 0.701 | 0.583 |
| Full Length | Hill Coefficient B | -37.6 | -25.7 | -3.88 | -23.0 |
| Full Length | EC50 B | 0.471 | 0.613 | 0.631 | 0.599 |
| Full Length | Bottom Plateau | 0.0490 | 0.0625 | 0.0719 | 0.0620 |
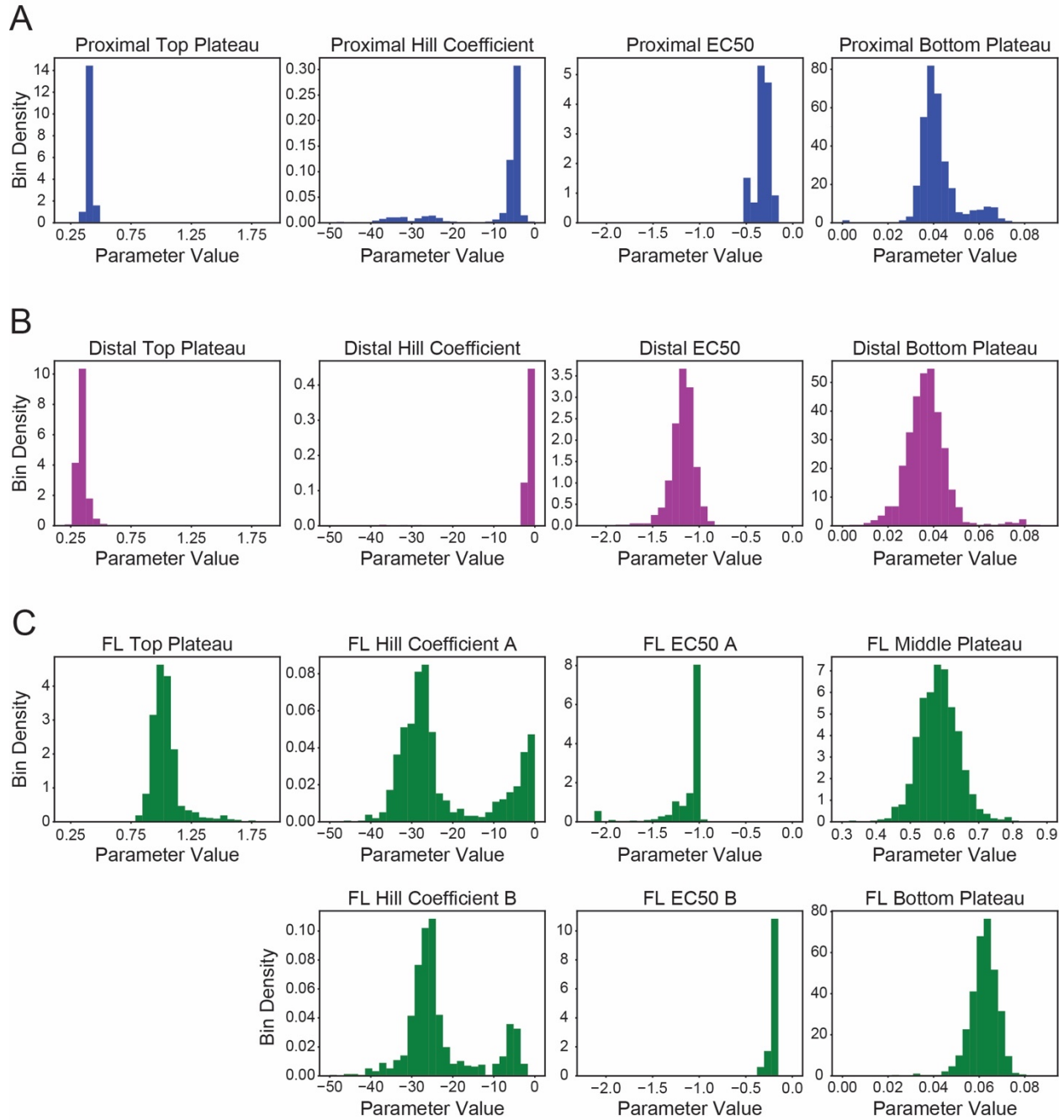
**Figure B.5.1 Parameter distributions from bootstrapped residuals.**

**(A-C)** Parameter distributions for the **(A)** Proximal, **(B)** Distal, and **(C)** Full Length reporters.