BMP-mediated cis-regulation of msh/msx genes in Drosophila and vertebrates

A thesis submitted in partial satisfaction of the requirements for the degree Master of Science

in

Biology

by

Erika Paige Taylor

Committee in Charge:
Professor Ethan Bier, Chair
Professor Bill McGinnis
Professor Steve Wasserman

2014
The Thesis of Erika Paige Taylor is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2014
Dedication:

This thesis is dedicated to all of the people who got me through. Thank you for all of your positive influence, love and support.

Mom, thank you for always encouraging me to be the best I can be and for all of your unconditional love. Nana and DD, thank you everything. You always made sure I had a warm meal, transportation and a roof over my head. My family has played an instrumental role in helping me through graduate school and encouraging me. Thank you all for being there through thick and thin.

To my wonderful husband, Brian, thank you for all of your patience, especially when I had to come into lab on weekends and work long hours to finish my experiments. Your understanding, love and support has helped me to finish strong.

This thesis is dedicated to all of the people who got me through. Thank you for all of your positive influence, love and support.

Ultimately, I dedicate my work to the glory and honor of God, the creator of all things. I am blessed to be a part of unravelling the intricate details of His creation and unveiling His design for neural development and how it compares among different species. “For in him all things were created: things in heaven and on earth, visible and invisible, whether thrones or powers or rulers or authorities; all things have been created through Him and for Him.” Colossians 1:16 (NIV)
Table of Contents:

Signature page.................................................................iii
Dedication...........................................................................iv
Table of Contents..............................................................v
List of Figures......................................................................vi
Acknowledgements...........................................................vii
Abstract.............................................................................ix
Introduction.................................................................1
Results............................................................................8
Figures...........................................................................14
Materials and Methods....................................................19
Conclusions.......................................................................21
References.........................................................................26
List of Figures:

Figure 1: Recombination Scheme .................................................................14
Figure 2: BMP Regulation of ME.................................................................15
Figure 3: BMP Regulation of ind.................................................................16
Figure 4: Zebrafish mxsB CRM.................................................................17
Acknowledgements:

I would like to acknowledge Professor Ethan Bier, for recognizing my potential, encouraging me to apply for the MS/BS program and for providing invaluable opportunities and guidance during my time in lab. I appreciate the continuous mentorship and for being positive, especially when troubleshooting experiments.

Next, I would like to thank my mentor, Francisco Esteves. You took the time to share knowledge and to teach me technical skills I will continue to use. I will miss our afternoon coffee breaks, discussing the caveats of the project and Drosophila research in general.

To all of the members of the Bier Lab, both past and present: thank you for always having a smile and willing to help in any way you can. Annabel, Adrianne, Shannon, Valentino and Steve, your collaboration and friendship is priceless and sustained me throughout my studies.

I am privileged that my thesis work, figures and partial text, will contribute to a larger project, in preparation to be published. I would like to thank all of the authors, Francisco Esteves, Alexander Springhorn, Erika Kague, George Pyrowolakis, Shannon Fisher and Ethan Bier, for their collaboration and including me in the paper, “BMPs regulate expression of msx genes in the dorsal neuroectoderm of Drosophila and vertebrates by distinct mechanisms.” The Pyrowolakis Lab provided valuable EMSA analysis of the msh/msx CRM, from which my analysis of the SE2 site is based. The Fisher Lab is instrumental in the cross-taxa comparison. I would like to say a special thank you to Erika Kague for putting in long hours to perform injections and collect zebrafish embryos to send for in situ hybridization analysis.

The Introduction in part is currently being prepared for submission for publication of the material. Francisco Esteves, Alexander Springhorn, Erika Kague, George
Pyrowolakis, Shannon Fisher and Ethan Bier. The thesis author is the primary investigator and author of this material.

The Results in part are currently being prepared for submission for publication of the material. Francisco Esteves, Alexander Springhorn, Erika Kague, George Pyrowolakis, Shannon Fisher and Ethan Bier. The thesis author is the primary investigator and author of this material.

The Figures in part are currently being prepared for submission for publication of the material. Francisco Esteves, Alexander Springhorn, Erika Kague, George Pyrowolakis, Shannon Fisher and Ethan Bier. The thesis author is the primary investigator and author of this material.

The Materials and Methods in part is currently being prepared for submission for publication of the material. Francisco Esteves, Alexander Springhorn, Erika Kague, George Pyrowolakis, Shannon Fisher and Ethan Bier. The thesis author is the primary investigator and author of this material.

The Conclusions in part is currently being prepared for submission for publication of the material. Francisco Esteves, Alexander Springhorn, Erika Kague, George Pyrowolakis, Shannon Fisher and Ethan Bier. The thesis author is the primary investigator and author of this material.
ABSTRACT OF THE THESIS

BMP-mediated cis-regulation of msh/msx genes in Drosophila and vertebrates

by

Erika Paige Taylor

Master of Science in Biology

University of California, San Diego, 2014

Professor Ethan Bier, Chair

In a broad variety of bilaterian species the trunk central nervous system (CNS) derives from three primary rows of neuroblasts. The fates of these three rows of neuroblasts are determined in part by the expression of three conserved transcription factors: vnd/nkx2.2, ind/gsh and msh/msx in Drosophila/vertebrates, which are expressed in corresponding non-overlapping patterns along the dorsal-ventral axis.
While this conserved suite of “neural identity” gene expression strongly suggests a common ancestral origin for the patterning systems, it is unclear whether the original regulatory mechanisms establishing these patterns have been similarly conserved. In *Drosophila*, genetic evidence suggests that BMPs act in a dosage-dependent fashion to repress expression of neural identity genes. BMPs also play a dose-dependent role in patterning the dorsal and lateral regions of the vertebrate CNS, however, the mechanism by which they achieve such patterning has not yet been clearly established. In this thesis I contributed to the analysis of mechanisms by which BMPs act on cis-regulatory modules (CRMs) that control localized expression of the *Drosophila msh* and zebrafish *msxB* in the dorsal CNS. My studies helped confirm that BMPs regulate neural genes in a threshold dependent manner and support the view that BMPs act differently to regulate similar patterns of gene expression in the neuroectoderm by repressing *msh* expression in *Drosophila* while activating *msxB* expression in zebrafish. These findings suggest that the mechanisms by which the BMP gradient patterns the dorsal neuroectoderm have reversed since the divergence of these two lineages.
Introduction:

Neural development in both Drosophila and vertebrates arise from a highly conserved expression pattern of Bone Morphogenetic Proteins (BMPs) and neural identity genes, along the Dorsal/Ventral (D/V) axis. BMPs are expressed in similar relative patterns in the epidermal ectoderm, during early development, across all bilaterian species (animals with mirror symmetry across the sagittal plane) (Denes 2007) and are responsible for patterning three primary rows of neural identity gene expression, in a conserved order along the D/V axis. It is important to look at the underlying genetic mechanism behind BMP regulation of neural genes to understand the origins of this neuronal patterning. In Drosophila, BMPs repress expression of neural genes throughout development (Biehs 1996). However, in vertebrates, the predominant view is that BMPs function as neural activators and only repress neural genes during early neural induction (Bier 1997 and De Robertis 2008) and then reverse their effect during the subsequent stage of development to activate expression of neural genes in the neural tube (Lee and Jessel 1999). Close examination of BMP regulation of neural genes in both invertebrates and vertebrates will reveal if the underlying genetic mechanisms function in a conserved/ancestral manner or derived separately, from divergent linages. This study spotlights the role of BMP signaling, on the transcriptional level, or the Cis-Regulatory Module (CRM), of muscle segment homeobox (msh), in Drosophila, and the homologous vertebrate gene, msx, to gain insights into the origins of this system.

Establishment of the D/V Axis:

Tissue differentiation along the D/V axis is initiated by the maternal deposition of Dorsal (Dl), a Rel/Nf-κB family protein which forms a gradient sufficient to divide the embryo into three primary tissue types: the dorsal ectoderm, the lateral neuroectoderm
and the ventral mesoderm. Depending upon the specific target gene, DI acts as either a repressor or an activator. High doses of DI define the ventral mesoderm by activating \textit{twist} (\textit{twi}), gene responsible for activating mesoderm genes, and \textit{snail} (\textit{sna}), gene responsible for suppressing expression of neural genes. Lower levels of \textit{Dorsal}, in the lateral region, activate expression of neural genes, defining the neuroectoderm, and indirectly suppress expression of epidermal genes via activation of the BMP antagonists, \textit{short gastrulation} (\textit{sog}) (Francois 1994) and \textit{brinker} (\textit{brk}) (Jazwinska 1999). A WNT activity gradient appears to accomplish a similar role in establishing the DV axis in vertebrates.

The dorsal ectoderm is distinguished by an opposing gradient, composed of the TGF-\(\beta\)/BMP family morphogen, Decapentaplegic (Dpp), (an orthologue of BMP2/4, in vertebrates) which acts in a threshold-dependent fashion to consolidate subdivision of the DV axis into neural versus non-neural ectoderm established initially by Dorsal, then contributes to subdividing these regions into finer domains (Padgett 1987 and Arendt 1994). The high concentration of Dpp along the dorsal border defines the amnioserosa, while the second threshold, ventrally adjacent, activates epidermal genes and represses neural gene expression in the epidermis proper.

Dpp acts in conjunction with other extracellular proteins, \textit{short gastrulation} (\textit{Sog}) (the Chordin homologue), \textit{twisted gastrulation} (\textit{tsg}), \textit{tolloid} (\textit{tld}) and \textit{screw} (\textit{scw}) to pattern the dorsal ectoderm. \textit{Screw} is a TGF-\(\beta\) ligand/BMP-related protein that enhances Dpp signaling by forming heterodimers with Dpp (Arora 1994) which then bind and activate BMP-receptors that are present throughout the embryo.

The interplay between \textit{dpp} and \textit{Sog} is crucial for defining the borders between epidermal and neural tissue (Ashe 1999 and Araujo 2000). \textit{Sog} protects the neuroectoderm from Dpp auto activation and expression of epidermal genes by forming
an inhibitory complex with Dpp (Biehs 1996) and cofactor, Tsg (Mason 1994). Without Tsg, the inhibitory interaction between Dpp and Sog cannot take place (Ross 2001 and Shimmi 2003). It has been proposed that this Sog-Dpp-Tsg complex may behave as a "shuttle," carrying Dpp dorsally to maintain peak Dpp activity along the dorsal midline, where the protease, Tld, releases Dpp from the complex (Yu 2004). Sog is secreted by cells in the lateral neuroectoderm, and diffuse dorsally where it is degraded by Tld, which acts as the sink to shape the Sog gradient (Srinivasan 2002). Dpp is required as a cofactor to activate Tld, the protease responsible for cleaving Sog in the Sog-Dpp-Tsg complex. By degrading Sog, Tld then frees Dpp to bind BMP-receptors and initiate signaling (Marques 1997, as reviewed in Ashe 2002).

**BMP Signaling Pathway:**

Extracellular BMPs signal via tetrameric receptor complexes composed of type I and II serine-threonine kinase receptors. The receptor for Dpp homodimers is composed of two Thick veins (Tkv) and two Punt (Put) chains. In contrast, the Dpp-Scw receptor is comprised of one Tkv, one Saxophone (Sax) and two Put chains (Shimmi 2005). Once the BMP-receptor complexes are activated, Put, the type II chain, phosphorylates the type I chains, which in turn phosphorylate Mothers against Dpp (MAD) to generate the active form, pMAD. pMAD forms a complex with the co-SMAD Medea (Med) (homologue of Smad4) and enters the nucleus to bind DNA as a heteromeric complex. The pMad-Medea complex activates epidermal genes, including dpp (auto activation), at binding sites composed of GC-rich Mad sites separated from Med SBE (Smad Binding Element) by a variable length spacer (Gao 2005). pMAD-Medea also forms a trimeric complex with Schnurri (Shn), to form a repressor complex on SEs that inhibit neural gene expression (as reviewed in Mizutani 2008). Intracellular antagonists block Dpp signaling as well. Inhibitory MAD (I-MAD) complexes block the p-MAD-Medea complex, in the
cytoplasm (Schmierer 2007). Another transcriptional repressor protein against Dpp is Brinker (Brk) (Dunipace 2013). brk is expressed in the same fashion as the extracellular Dpp antagonist, Sog (Jazwinska 1999) and negatively regulates Dpp signaling by competing with the pMAD-Medea at DNA binding sites (Kirkpatrick 2001, Weiss 2010 and Rushlow 2001). BMPs also bind activating elements (AEs) to turn on genes such as Daughters against Dpp (Dad) and other epidermal genes with similar bipartite configurations. The spacer length in activation element (AE) sites prevents Shn binding and, instead, recruits other necessary transcriptional co-factors for activation (Gao 2005).

*Patterning of the neural identity genes:*

In order to understand how neural identity genes are expressed precisely in three mutually exclusive longitudinal columns, it is important to understand the genetic circuitry that defines the limits of these gene expression domains. The spatial subdivision of the neuroectoderm into these three domains is controlled by differential activation by Dl and by various signals impinging on the CRM. The three neural identity genes in *Drosophila* and vertebrates are: ventral nervous system defective (*vnd*)/NK2 transcription factor-related (*nkx2.2*), intermediate neuroblasts defective (*ind*)/genomic screen homeobox (*gsh*) and paired box 6 (*pax6*), and muscle segment homeobox (*msh*)/MSH homeobox (*msx*) (as reviewed in Mizutani and Bier 2008). The orthologous genes are expressed in the same sequential order, ventral to dorsal, in both *Drosophila* and vertebrates and cross-regulate each other in inhibitory fashions (Von Ohlen 2000; Cowden and Levine 2003).

In *Drosophila*, *vnd* is the first and ventral-most neural identity gene to be expressed. Expression of *vnd* is induced by a specific threshold of Dl (Markstein 2002), and *vnd* is restricted to the ventral neuroectoderm by repression from Snail (Sna) in the
mesoderm (Mellernick and Nirenberg 1995). Both activation and repression of \textit{vnd} are carried out at clustered binding sites in an intronic CRM (Stathopoulos 2002). As the DI gradient diminishes, it arrives at a threshold where it is no longer sufficient to activate \textit{vnd}. However, the lower levels of DI present in the intermediate neuroectoderm are sufficient to induce \textit{ind} expression in conjunction with Egfr signaling that is graded across the \textit{ind} domain as a consequence of the EGF-ligands Spitz and Vein diffusing dorsally from their production zone in the ventral neuroectoderm (Von Ohlen 2000 and Cornell 2000). Genetic experiments and analysis of the \textit{ind} CRM reveal that the ventral border of \textit{ind} expression results from \textit{vnd} repression (Weiss 1998; Stathopoulos 2005; Garcia 2011). The DI gradient also indirectly activates \textit{ind} by repressing expression of \textit{dpp}, a neural inhibitor. The dorsal \textit{ind} border is determined by low levels of DI/EGFR signaling (that are insufficient to activate \textit{ind}) and inhibition of Dpp expression. Dpp signaling induces Smad:Shn complex binding within the CRM to repress \textit{ind} expression. (as reviewed in Mizutani 2006; Esteves, unpublished data). Consistent with the dose-dependent model of Dpp-mediated regulation of neural genes, \textit{ind} can only tolerate low levels of Dpp (Von Ohlen 2000). DI activates neural genes; however, in the mesoderm, where the DI concentration is very high, \textit{sna} is activated and represses expression of all neural genes (Cowden and Levine 2003).

\textbf{BMP regulation of msh/msxB:}

In \textit{Drosophila}, genetic evidence suggests that BMPs act in a dosage-dependent fashion to repress expression of neural identity genes (Mizutani et al., 2006; Mizutani and Bier 2008). A 700 bp \textit{msh} CRM (henceforth referred to as Msh Element (ME)) has been identified that is directly repressed by \textit{ind} (Von Ohlen 2009). The CRM directs gene expression by recruiting specific transcription factors and signaling molecules that determine where and when the gene will be expressed. Analyzing the CRM of a gene is
important for developmental biology because we gain insight into the genetic circuits by which the developing embryo is organized (Howard 2004). The response of the ME to BMP-mediated regulation has not yet been investigated yet, however. Previous studies in the Bier lab utilized bioinformatic analysis to search the ME for known BMP binding sites and screen for the actual BMP binding sites both in vitro (EMSA analysis) and in vivo (mutant screen, in situ hybridization), however, the identified sites differ by at least one nucleotide from the known optimal binding site such as that in the ind CRM. For my thesis, I have assisted in analyzing the relative effects of optimal versus suboptimal SE sites mediating repression by BMP signaling. One of the two suboptimal SE sites in the ME was mutated to resemble the known optimal ind consensus site and I show that this results in enhanced silencing by BMPs and exclude the possibility that this mutation inadvertently interferes with activation by the ME.

BMPs also play a dose-dependent role in patterning the dorsal and lateral regions of the vertebrate CNS (Lee and Jessell 1999); (Hu 2004), however, the mechanism by which they achieve such patterning has not yet been clearly established. Previously in the Bier lab, the mechanisms by which BMPs act on cis-regulatory modules (CRMs) that control localized expression of the Drosophila msh and zebrafish msxB in the dorsal central nervous system (CNS) were analyzed. Our analysis suggests that BMPs act differently to regulate similar patterns of gene expression in the neuroectoderm by repressing msh expression in Drosophila while activating msxB expression in zebrafish. These findings suggest that the mechanisms by which the BMP gradient patterns the dorsal neuroectoderm have reversed since the divergence of these two lineages. To confirm this, in collaboration with the Fisher lab, I observed the effects of mutating the msxB ME, which confirm the role of BMP in CRM activation, rather than repression, as in the Drosophila model.
The Introduction in part is currently being prepared for submission for publication of the material. Francisco Esteves, Alexander Springhorn, Erika Kague, George Pyrowolakis, Shannon Fisher and Ethan Bier. The thesis author is the primary investigator and author of this material.
Results:

Previous experiments identified the *msh* CRM (Von Ohlen 2009) and this 700 bp element was scanned for known Smad consensus binding sites (Esteves, unpublished data). This bioinformatic analysis revealed several candidate BMP responsive sites, including: activator elements (AE), silencing elements (SE) and Brk binding sites. AE sites (GGCGCCA\(N\)\(_4\)GNCV) are known to bind Mad/Medea heteromers. SE sites (GNC\(G\)NC\(N\)\(_5\)GNCTG) bind the trimeric Mad/Medea/Shn complex. Brinker is a transcriptional repressor of SMAD-dependent BMP signaling. This repression is mediated at least in part by direct competition by Brk and SMADs for binding to common DNA sites ((T/GGCGYY) which have been found to overlap with subsets of SE elements (Weiss 2010) suggesting that the known ability of Brk to block BMP-mediated repression of *msh* expression (Mizutani 2006) could occur via competition for binding to BMP-responsive sites in the ME (see below, however).

To experimentally identify Smad binding sites within the ME, electrophoretic Mobility Shift Assays (EMSAs) were then used to comprehensively analyze all of the candidate BMP responsive sites. Protein binding was only found at two of the three putative SE sites (SE1 and SE2; Fig. 2A). These sites both assembled Mad/Medea/Shn complexes although binding to SE2 was stronger than to SE1. In contrast, no in vitro binding was observed to any of the candidate AE or Brk sites, however.

In order to better understand BMP regulation of neural genes, a ME-lacZ construct was used that recapitulates the pattern of endogenous *msh* expression. This ME-lacZ construct and mutant versions of it were inserted into the same chromosomal integration site using the PhiC31 transgenesis system (Venken 2006). Genetic analysis and mutation screening, which I participated in, confirmed that the two SE sites contribute to BMP-mediated repression of the ME. SE1 mutants showed a weak dorsal
expansion of ME-driven expression, SE2 mutants somewhat greater dorsal expression, and SE1,SE2 double mutants, yet more pronounced dorsal expansion. The levels of dorsal lacZ expression in the SE1,SE2 double mutant was still less than that observed for the intact ME element in a dpp- genetic background, however, indicating that other direct or indirect BMP-dependent inhibitory inputs must act on the ME (Esteves, unpublished data).

Since the SE2 consensus, GNTGNC(N)\_5GNCTG, differs by one nucleotide from the known optimal SE consensus, GNCGNC(N)\_5GNCTG, which is present in the ind CRM, we tested whether converting the primary SE site in the ME (SE2) to an optimal site (SE2*) might lead to greater BMP-mediated repression. Indeed, EMSA experiments confirmed that the Mad/Medea/Shn complexes bound with higher affinity to an SE2* sequence than to an SE2 oligonucleotide. Prior genetic studies also revealed that BMP signaling is more effective in repressing expression of ind than msh (Mizutani 2006). In line with SE sites in the ind and msh CRMs having differing affinities for binding Smad complexes, the SE2* mutation resulted in a significant reduction in ME-lacZ expression in embryos (Fig. 2B).

Although the greatly reduced expression of lacZ expression driven by the ME-SE2* mutant construct is consistent with the optimized SE2* site acting as a higher affinity BMP SE site, an important caveat was that this mutation might also have inadvertently eliminated a necessary activating sequence. In order to verify that the reduced expression observed in ME-SE2* embryos was indeed BMP-dependent, I recombined the SE2* mutation into a dpp- background, to ensure that no positive elements were disrupted by the SE2* mutation (Fig. 2B). The recombination required multiple crosses, to accommodate for the haploinsufficiency of dpp. First, dpp- is balanced by a maternal duplication of dpp, Dpdpp (Fig. 1, Cross 1a). I selected male
progeny with rough (SM6), red eyes (Dpdpp) and sternopleural. Concurrently, the \([w^+, SE2^*-lacZ]\) mutant is marked with FM7c on the X chromosome (Fig. 1, Cross1b). I then picked virgins with red \((w^+)\) heart shaped (FM7c) eyes and curly wings \((CyO)\). Both of these progeny are then crossed to produce flies with \([w^+, SE2^*-lacZ]\) and \(dpp^-\) on the second chromosome, balanced and marked by the maternal duplication of Dpp and FM7c (Fig. 1, Cross 2). The virgins from this stock were phenotypically identified by red, heart-shaped eyes and sternopleural. Because SE2* and \(dpp^-\) are each on the second chromosome, recombination could take place when these females are crossed to Scutoid \((Sco)/CyO23\) males \((CyO23\) contains a duplication of Dpp). Originally, sternopleural \((Sp)\) was used to track the \(dpp^-\) allele, because they are fairly close linked, with only an 18% likelihood of recombination. According to published literature, \(Sp\) is an allele of wingless \((wg)\) (http://flybase.org/reports/FBal0015984.html), cytologically mapped to 27F1. Discrepancies were found, however, with other mapping data \((Nusslein-Volhard and Wieschaus 1980)\), in which \(Sp\) was mapped at 28A1. Also, other sources claimed that \(Sp\) maps closer to the centromere than wingless \((Buratovich 1997)\). The \(att\) insertion site for the \([w^+, SE2^*-lacZ]\) complex is located at 28E. Therefore, we expected to see 2% recombination between \(Sp\) and \(SE2^*\); after extensive searching, no such recombinants were found, however, suggesting that \(Sp\) may have been mismapped. I therefore disregarded the \(Sp\) marker and selected all progeny with red \((w^+)\), heart-shaped (FM7c) eyes and curly wings \((CyO)\). Since \(dpp^-\) is located at 22F, there is a 20% chance of recombination with the \(att\) site. To generate a stock and to test possible recombinants, I crossed them back to the \(w^-; Sco/Cyo23\) stock. If the \(SE2^*\) did not recombine, I would see progeny with both curly and straight wings \((SE2^*-lacZ/Sco\) and \(SE2^*-lacZ/Cyo23)\). Because \(dpp\) is haploinsufficient, the \(dpp^-\ SE2^*-lacZ\) recombinant progeny could only have curly wings \((dpp^- SE2^*-lacZ/Cyo23)\). The
recombinant was further confirmed by cuticle prep (in collaboration with Alexa Clemmons, from the Wasserman lab). The wildtype cuticles displayed normal denticles, accordingly, dpp- mutants expressed denticles circumferentially.

Having successfully obtained the dpp- SE2*-ME recombinant, I used it to test for the predicted BMP responsiveness of this element (i.e., like the intact ME, full-level lacZ expression should be observed that expands to the dorsal midline). To visualize lacZ reporter expression of the ME, flies were raised in cages, at 25º C, and the embryos were collected and fixed in four hour intervals (see Materials and Methods). Colorimetric in situ hybridization of a lacZ antisense RNA dixogin probe was used to visualize reporter gene expression (Figure 2B) (see Materials and Methods). The wildtype ME displayed a classic msh-like expression pattern (Fig. 2B). The intact ME element when crossed into a dpp- background drove lacZ expression that expanded throughout the dorsal region (i.e., msh can be expressed dorsally in the absence of dpp repression Fig. 2B). Likewise, ME-SE2*-lacZ expression, which is reduced compared to the wildtype ME in a wild-type background, was expressed throughout the dorsal region in a dpp-background, confirming that the ind-like ME is more sensitive to BMP regulation. Importantly, the observation that in the absence of BMPs, SE2* expression expands dorsally, confirms that no positive elements were disrupted by this mutation.

I also examined the effect of making the reciprocal type of mutation in of the SE site in the ind CRM. The optimal SE site in this CRM was replaced with the SE2 site from the ME and transgenic stocks carrying this construct were isolated. I then stained for lacZ expression by in situ hybridization to determine whether this mutated ind CRM might drive altered (i.e., dorsally expanded) expression. I found that the wildtype ind and msh-like ind expression patterns were is not significantly different (Fig. 3). Since complete elimination of the ind SE element results in only modest dorsal expansion
(Stathopoulos 2002) it may be that the effects of this more modest mutation cannot be readily detected without additional sensitizing conditions (e.g., overexpression of EGF ligands to potentiate dorsal expansion of *ind*).

The above analysis of the *Drosophila msh* CRM is consistent with genetic data in this organism indicating that BMPs act by dosage sensitive repression of neural identity gene expression (Mizutani 2006). In vertebrates, however, the prevailing view is that BMP activate expression of *msh* orthologs (*msx* genes) (Lee and Jessell 1999; Hu 2004; Weis 2010; Esteves, unpublished data). To address this question, a 2.4 Kb region immediately upstream of the *msxB* coding region was identified using the powerful *Tol2* transgenesis system (Fisher 2006b) that drove reporter gene expression in stripes within the neural plate abutting the BMP expressing epidermal ectoderm. The 2.4 kb *msxB* CRM was then paired down to a 671bp minimal element that was then used for subsequent experiments for testing its response to BMPs signaling.

Echoing the approach in *Drosophila*, BMP responsive sites were sought within the minimal *msxB* CRM by first scanning bioinformatically for candidate SE or AE sites using the SMAD1/2 consensus GNCKNC and SMAD4 consensus GNC(T/V) with relaxed spacing constraints and then testing by EMAS whether they could indeed assemble *Drosophila* SMAD and/or SMAD-Shn complexes in response to BMP signaling *in vitro* in a *Drosophila* cell line (S2 cells). This analysis identified a single highly conserved site (zAE) to which BMP signal-dependent DNA binding was observed. The zAE contains candidate SMAD1/2 and SMAD4 binding sites separated by an unusually long 16 bp spacer. Further analysis of this binding motif revealed that both the SMAD1/2 and SMAD4 sites are each required for DNA binding, suggesting that the functional zAE includes both sites (Esteves, unpublished data).
Based on the identification of a single functional SMAD sites in the 671 bp msxB CRM, I then tested the effect of mutating this site in vivo. Embryos were injected with the wildtype 671bp msxB-CRM, a 36 deletion of the SMAD site (msxB-CRM^{DEL}) and a point mutant that eliminated SMAD binding in EMSA assays (msxB-CRM^{GCR}). I then stained the resulting transiently transfected embryos for gfp reporter gene expression. I visualized transient gfp expression by *in situ* hybridization, at 65ºC, with an antisense gfp RNA probe, visualized with NBT/BCIP. For both the msxB-CRM^{DEL} construct and the msxB-CRM^{GCR} construct I observed significantly reduced reporter expression *in vivo* in the transient transformant embryos as compared to embryos transfected with the intact 671 msxB-CRM construct (Fig. 4). These results indicate that a single BMP responsive site within the 671 bp zebrafish msxB-CRM is required for mediating activation of this element *in vivo*.

The Results in part are currently being prepared for submission for publication of the material. Francisco Esteves, Alexander Springhorn, Erika Kague, George Pyrowolakis, Shannon Fisher and Ethan Bier. The thesis author is the primary investigator and author of this material.
Figures:

1a) \( \frac{w^{+}; dpp^{+}Sp^{+}}{CyO23} \times \frac{Dpdpp^{+}; dpp^{+}Sp^{+}dl^{+}}{w^{+}, SM6} \rightarrow \frac{Dpdpp^{+}; dpp^{+}Sp^{+}}{SM6} \)

1b) \( \frac{w^{+}; lw^{+}, SE2^{+}-lacZ}{[w^{+}, SE2^{+}-lacZ]} \times \frac{sog brk^{+}; DTD48}{FM7c, CyO} \rightarrow \frac{w^{+}; [w^{+}, SE2^{+}-lacZ]}{FM7c, CyO} \)

2) \( \frac{Dpdpp^{+}; dpp^{+}Sp^{+}}{SM6} \times \frac{w^{+}; [w^{+}, SE2^{+}-lacZ]}{FM7c, CyO} \rightarrow \frac{Dpdpp^{+}; [w^{+}, SE2^{+}-lacZ]}{FM7c, dpp^{+}Sp^{+}} \)

3) \( \frac{w^{+}; sco}{CyO23} \times \frac{Dpdpp^{+}; [w^{+}, SE2^{+}-lacZ]}{FM7c, dpp^{+}Sp^{+}} \rightarrow \frac{FM7c; [w^{+}, SE2^{+}-lacZ]}{w^{+}, CyO23} \)

Figure 1: Genetic Recombination Scheme

Scheme of genetic crosses used for generating the \([w^{+}, SE2^{+}-lacZ]dpp\)- recombinant stock. Genotypes in black indicate stocks used. To test the recombinant flies and make a stock, recombinants were crossed back to \(w^{-}/Y; sco/CyO23\).
Figure 2: Analysis of Dosage Dependent BMP Regulation of the Drosophila ME

(A) Diagram of ME showing the location of Silencing Element (SE) 1 and 2 within highly conserved regions. Conservation score (Cons) is based on alignment of the ME region in 12 Drosophila species. (B) Lateral and dorsal views (anterior to the left) of in situ hybridization detection of lacZ expression in ME-lacZ embryos demonstrating the in vivo effects of mutating SE1 and SE2. ME: wild type embryos containing the intact ME-lacZ construct; ME dpp-: dpp- mutant embryos show complete dorsal expansion of lacZ expression; SE2*: Converting SE2 to an optimal (ind-like) SE site results in reduced lacZ expression; SE2* dpp-: SE2* in a dpp- background, shows complete dorsal expansion of lacZ.
**Figure 3: Analysis of Dosage Dependent BMP Regulation of the ind CRM**

(A) Wildtype *ind* expression, visualized by a *lacZ* reporter. (B) Point mutation of the known BMP responsive *ind* SE to resemble the *msh* ME consensus.
Figure 4: Analysis of the Role of BMPs in the Zebrafish ME

(A) Representation of the zebrafish msxB locus. Also depicted is the location of the 671bp CRM and the conservation (Cons) across vertebrates (represented by block conservation tracks). (B) Dorsal (anterior to the top) and lateral (anterior to the left) views of injected zebrafish embryos. Embryos were injected with either the 671bp msxB-CRM, or mutant version msxB-CRM^{DEL} (a 36bp deletion in the conserved region) and msxB-CRM^{GCR} constructs (point mutation disrupting the GC-rich site), driving gfp reporter expression. Transient gfp expression was detected by in situ hybridization. Both mutant constructs show reduced expression of reporter.
The Figures in part are currently being prepared for submission for publication of the material. Francisco Esteves, Alexander Springhorn, Erika Kague, George Pyrowolakis, Shannon Fisher and Ethan Bier. The thesis author is the primary investigator and author of this material.
Materials and Methods:

CRM-reporter construction and analysis:

The 700 bp *Drosophila msh* CRM is described in (Von Ohlen et al., 2009). The various *Drosophila msh*-CRM constructs were subcloned in pCR-TOPO vectors and subsequently cloned into the [P]acman vector (Venken et al., 2006) as NotI and KpnI restriction fragments. Site-directed mutagenesis PCR methods were adapted from (Hansson et al., 2008). Zebrafish constructs were cloned into pENTR-TOPO and transferred to pTol2 by Gateway Recombination and injected into zebrafish embryos as described in (Fisher et al., 2006b).

Genetic strains and crossing schemes:

The *Drosophila dpp*[^46] null allele used in this study is Flybase stock number 2061. The *schnurri* mutant allele is *shn*[^4738]. To generate the [*w+,SE2*-lacZ*] embryos, females that are *Dpdpp/*w+; *dpp[^46] wg[^sp] dl[^1]/ SM6 were crossed to w-/Y; *dpp[^46] wg[^sp]/CyO23* males. Also, females that are *Sog-brk/-FM7c; DTD48/CyO* were crossed to w-/Y; [*w+, SE2*-lacZ*/ [w+, SE2*-lacZ]. Selected progeny were then crossed to generate *Dpdpp/FM7c; [w+, SE2*-lacZ] / dpp[^46] wg[^sp]* females, in which recombination takes place. Recombinants were then crossed to w-/Y; *sco/CyO23* males to test for recombination and generate a stock. The fly strain used to inject all constructs has genotype PBac{yellow[+]-attP-3BVK00002 and injections were outsourced to BestGeneInc (http://www.thebestgene.com/).

*Drosophila embryo collections:*

All embryos were collected after 4 hours at 25ºC, from agar-grape juice plates, and they were fixed with 37% Formaldehyde. Embryos were then stored in MeOH, -20ºC.
**In situ hybridization:**

The colorimetric *in-situ* hybridization methods used were performed, according to O’Neill and Bier, 1994, in *Drosophila* embryos and adapted to zebrafish embryos by increasing the hybridization temperature: 55°C in *Drosophila* to 65°C in zebrafish embryos. Antibodies used: anti-digoxigenin (Roche). We also used colorimetric staining methods performed according to The DNA template used to generate the *msx*B probe was a generous gift from the Riley lab. Histochemical stain images were acquired using a Nikon optical microscope. Images were adjusted for color, brightness and contrast using Adobe Photoshop software. The antisense RNA probes were generated from plasmid DNA template, linearized, purified and synthesized using digoxin NTPs.

**Zebrafish Tol2 transgenesis:**

Tol2 is a DNA transposon tagged with targets for site-specific recombination. Once the gene of interest, *msx*B-*gfp* reporter gene, is cloned into the Tol2 expression vector, it is amplified by PCR and injected into 1-2 somite zebrafish embryos (Fisher 2006b). For this experiment, the embryos were aged to 5-6 somite stage then fixed in formaldehyde for *in vivo* analysis. A small, highly conserved 671bp fragment was identified as the site of BMP regulation, by observing the GFP-reporter expression in neuroectodermal progenitor cells (Esteves, unpublished data).

The Materials and Methods in part is currently being prepared for submission for publication of the material. Francisco Esteves, Alexander Springhorn, Erika Kague, George Pyrowolakis, Shannon Fisher and Ethan Bier. The thesis author is the primary investigator and author of this material.
Conclusions:

BMPs play a highly conserved role in neural induction and also contribute to establishment of D/V polarity within the CNS. In the latter case, however, it has not been established whether they act through common or distinct mechanisms to effect dose-dependent patterning of neural identity genes.

Comparing BMP cis-regulation of the conserved msh/msx family genes in the dorsal neuroectoderm of both Drosophila and vertebrate embryos provides an excellent paradigm for studying whether cis-regulatory processes are maintained across distant taxa. Our analysis of the Drosophila msh CRM suggests that BMPs act in part through two SE-type binding sites that mediate repression, while AE sites do not appear to mediate responses to BMP signaling. In contrast, we identified a single SMAD binding site within an embryonic zebrafish msxB CRM that is required for BMP-dependent activation. These results suggest that the regulatory mechanisms by which BMPs act on msh/msx CRMs are opposite in these two lineages, even though they drive similar expression patterns in the dorsal neuroectoderm.

The above dissection of BMP responsive sequences within the Drosophila and zebrafish msh/msx CRMs suggests that they are under opposing forms of BMP regulation: repression in Drosophila versus activation in zebrafish. In the current case of BMP-dependent regulation of msh/msxB expression, natural selection may have operated to maintain relevant gene expression patterns, such as dorsal neuroectodermal expression, that fulfill a particular function while allowing the upstream mechanisms generating that pattern to change over time. The SE1 and SE2 sites that play a role in repressing ME activity dorsally are imperfect matches to the consensus SE sites determined by Pyrowolakis and colleagues (Pyrowolakis 2004). The ind CRM, however, which according to genetic data is more sensitive to BMP repression than msh (Mizutani 2004).
2006), contains a perfect SE site required for repressing activity of this element dorsally (Garcia and Stathopoulos 2011). When this SE2 site in the ME was mutated to similarly match the ideal SE consensus sequence (SE2*) ME expression was repressed in the dorsal ectoderm domain (Fig. 2). SE2* was also recombined into a BMP null (Dpp$^{H46}$) background to confirm the mutation did not inactivate any positive elements within the ME (Fig. 2B). Dorsal expansion of ME-lacZ expression in the SE2*dpp- recombinant verified the validity of the above conclusions.

During the recombination of SE2* into the BMP null background, Sp phenotypically marked the dpp$^{H46}$ allele, since they are closely linked, with only an 18% chance of recombination. However, difficulties were encountered when recombinants with the Sp+ phenotype were not found (Fig. 1). Upon closer examination, discrepancies in literature emerged as to where Sternopleural is cytologically mapped. It was suggested by Cohen and colleagues that Sp is an allele of wingless (Neumann and Cohen 1996). According to Nusslein-Volhard and Wieschaus (Nusslein-Volhard and Wieschaus 1980), however, Sp was mapped to 28A1-A3, and according to Flybase (http://flybase.org/reports/FBal0015984.html) Sp is located at 27F1. In addition, other studies indicate Sp is closer to the centromere than wg and is an allele of a distinct complementation group (Buratovich 1997). Given the results of my genetic scheme, Sp behaved independently of the dpp-/SE2* recombination, implying it is even closer to the centromere than previously believed (centromeric to the att site, 28E). It will be interesting to examine the location of Sp further in future studies to resolve the discrepancies in literature. Perhaps sequencing analysis to identify the causative DNA changes may help resolve this question.

In addition to analyzing the effect of increasing the affinity of the SE site in the ME to resemble an optimal ind site, we also examined the converse effect of replacing
the *ind* SE site with the lower affinity SE2 site from the ME. When the *ind* SE was mutated to resemble the ME SE2 site, we expected *ind* to be more tolerant of BMP repression and that this *msh*-like *ind* CRM mutant would exhibit dorsal expansion when compared to the wildtype *ind* CRM. However, no significant difference was found (Fig. 3). This could be due to other factors affecting *ind* activation that bear more impact on *ind* expression than BMP repression, such as EGFR signal activation and threshold dependent activation by the Dorsal gradient, or that the point mutation was not strong enough to alter expression in the *ind* region of the neuroectoderm. Other studies have shown that, when manipulated further, *ind* does expand dorsally (Garcia 2013). This last inconclusive result notwithstanding our overall findings suggest that differences in affinity of SE sites for forming Mad-Shn-Med complexes contribute to the distinct responses of the two CRMs to BMP-mediated threshold-dependent repression.

The same methods used for identifying the SE in the *Drosophila* ME were mirrored in zebrafish to assess the role of BMPs in regulating expression of vertebrate *msx* genes. Analysis of a minimal 671bp *msxB*-CRM identified a single conserved SMAD binding site. To confirm BMP activation of neural genes, we injected embryos and observed that GFP reporter gene expression was lost in transient transformant embryos (Fig. 4B). Reduced transient *gfp* expression indicates that a single BMP responsive site within the 671 bp zebrafish *msxB* CRM is required for mediating activation of this element *in vivo*, not repression, as in the *Drosophila* paradigm. Also, in additional studies that I contributed to, *msxB* expression and *msxB*-reporter gene expression were lost upon inhibition of BMP signaling via heat induction of a HS-BMP construct. Thus, both mutational analysis and *in vivo* testing suggest opposing mechanisms for BMP-dependent regulation of the *msh* and *msxB* genes in the early neuroectoderm (Esteves, unpublished data). These experiments also revealed lower level ectopic *msxB*
expression in response to BMPs, however, suggesting that additional BMP-responsive elements may be present within the msxB locus.

As summarized above, our analysis strongly suggests that BMPs pattern the neuroectoderm primarily via repression in Drosophila, while in zebrafish, BMPs function, at least in part, to activate the orthologous msxB gene. Genetic studies and exogenous BMP treatment in zebrafish suggest that msx gene expression may also be repressed by high levels of BMP signaling. Whether the BMP-responsive site in the 671 bp msxB CRM together with other potential BMP-responsive elements mediate such a biphasic response will be interesting to address in future experiments. One can imagine various scenarios under which BMP-mediated regulation of msh/msxB genes could have progressed and switched its effect between vertebrates and invertebrates. In vertebrates, BMP targets frequently contain Drosophila SEs, which activate rather than repress transcription. Since msh is more tolerant of BMP repression, relative to ind, while msx genes require high neuroectodermal levels of BMPs to activate msx genes, an intermediate CRM state may have existed in which BMPs both weakly activated msx gene expression within the neuroectoderm at moderate levels while repressing gene expression at the peak BMP levels present in the adjacent epidermis. Indeed the intact zebrafish msxB gene may represent such a bifunctional condition since prior in vivo studies indicate that high levels of BMPs can inhibit msxB expression (Esteves, unpublished data). It remains to be determined whether such proposed positive and negative inputs might be mediated by a single or multiple independent CRM element(s). Such biphasic responses might have then been rendered monophasic in opposing directions, in divergent lineages, to account for the observed differences in the Drosophila versus vertebrate msx CRM.
In future analyses it will also be important to examine BMP-mediated regulation of additional neural identity genes expressed along the dorsal-ventral axis including the Gsh/ind and Nnx2.2/vnd genes, as the CRMs controlling expression of each of these genes will have undergone an independent progression by trajectories. Since there is evidence that laterally ventrally expressed genes in vertebrates are inhibited by BMPs (Furuta 1997; Golden 1999; Hartley 2001; Liem 2000; Pierani 1999) and because the more ventrally expressed ind gene in Drosophila is more sensitive to BMP-mediated repression than msh (Mizutani 2006), one might expect to find similar, and perhaps conserved ancestral modes, of BMP-mediated regulation of these genes across bilateria. It will also be interesting to understand how flexible the ancestral metazoan state was by investigating the relationship between BMPs and msx genes in basal metazoans such as jellyfish. In these diploblastic animals, although the BMP-msx relationship has not been tested, BMP2/4 (Reber-Müller 2006) and msx (Takahashi 2008) homologues are expressed in adjacent regions during development, as is the case in the majority of triploblastic animals.

The Conclusions in part is currently being prepared for submission for publication of the material. Francisco Esteves, Alexander Springhorn, Erika Kague, George Pyrowolakis, Shannon Fisher and Ethan Bier. The thesis author is the primary investigator and author of this material.
References:


