Published in final edited form as:

Dev Dyn. 2009 November ; 238(11): 2735–2744. doi:10.1002/dvdy.22096.

Ind represses *msh* expression in the intermediate column of the *Drosophila* neuroectoderm, through direct interaction with upstream regulatory DNA

Tonia Von Ohlen*, Cade Moses, and Will Poulson

Introduction

In order to establish and maintain the dorso-ventrally (DV) or anterior-posteriorly (AP) restricted stripes of homeodomain gene expression in the developing embryo it is necessary to coordinate multiple signaling pathways and the assembly of transcription factors on the regulatory DNA of each gene (Skeath, 1998; Von Ohlen and Doe, 2000) (Yagi et al., 1998; Miller et al., 2001). Regulation of the DV restricted homeodomain proteins Ventral nervous system defective (Vnd) and Intermediate neuroblasts defective (Ind) in the embryonic neuroectoderm has been investigated extensively. For both of these genes, multiple enhancer elements have been identified that control different aspects of expression. For example, initiation of *vnd* expression is controlled by a 348 bp element located in the first intron of the gene, called the *vnd* NeuroEctodermal Enhancer (NEE) (Markstein et al., 2004). This element contains binding sites for the transcription factors Twist a mesoderm and ventral neuroectoderm specific transcription factor, and Dorsal a global regulator of DV patterning. Both Twist and Dorsal are known to activate *vnd* in the ventral neuroectoderm (Mellerick and Nirenberg, 1995; Stathopoulous et al., 2002; Markstein et al., 2004). Later aspects of *vnd* expression, including expression in neuroblasts, are controlled by regulatory elements located upstream of the coding sequence (Saunders et al., 1998; Estes et al., 2001). These elements control activation in neuroblasts as well as repression in the midline (Saunders et al., 1998; Estes et al., 2001; Shao et al., 2002). Likewise, Ind expression is also controlled by multiple regulatory elements. Initiation of *ind* is associated with an element located downstream of the coding sequence (Weiss et al., 1998; Cowden and Levine, 2003; Stathopoulous and Levine, 2005). This element contains binding sites for Dorsal, known to be essential for activation of Ind and Vnd. Also present are Vnd binding sites, which is the mechanism of Vnd repression of *ind* in the ventral neuroectoderm (Weiss et al., 1998; Von Ohlen and Doe, 2000; Stathopoulous and Levine, 2005). An additional Ind regulatory element located upstream of the coding sequence participates in Ind-dependent maintenance of Ind expression (Von Ohlen et al., 2007b). Despite extensive knowledge of how expression of Ind and Vnd are regulated, very little is known about the regulation of *muscle segment homeobox (msh)* expression and until now, no regulatory DNA has been identified. Interestingly, the regulatory DNA necessary for initiation of both *vnd* and *ind* appears to contain elements necessary for both activation of expression in the appropriate domains and repression in inappropriate domains. Because the enhancer and silencer elements for these genes and presumably *msh* are inseparable, we will refer to these as regulatory elements and not enhancer elements for the remainder of this manuscript.

* Author for correspondence Division of Biology Kansas State University 104 Ackert Hall Manhattan, KS 66506 (785) 532-6149 tonia@ksu.edu.

Our work has been extensively focused on the role of the homeodomain transcription factor Ind in patterning and cell fate specification of the *Drosophila* embryonic nervous system. In this work we wanted to identify *msh* regulatory DNA capable of driving reporter gene expression in a pattern consistent with *msh* expression in the embryonic neuroectoderm, and to demonstrate the requirement for Ind in limiting the ventral boundary of *msh* expression. We have identified essential regulatory DNA that controls *msh* expression in the early embryonic neuroectoderm. This regulatory element is not only essential for activation of *msh* in the lateral neuroectoderm, but it also contains sequences essential for repression by Vnd and more specifically Ind in the ventral and intermediate neuroectoderm. Thus, as with Vnd and Ind, the regulation of *msh* involves an element that controls both activation in the lateral column and repression in more ventral regions of the neuroectoderm. We also identified an additional element responsible for expression in the lateral mesodermal cells. We believe that this element may be responding to regulation by the mesoderm specific transcription factor Tinman (tin). Finally, we demonstrate that, unlike Vnd and Ind, Msh does not appear to be essential for initiation or maintenance of its own expression.

Results

Identification of *msh* regulatory DNA

We used a variety of approaches to identify potentially important *msh* regulatory elements. Initially we use Evoprinter to identify five regions of highly conserved sequence among several Drosophilid species (Figure 1A, green boxes and Figure S1). (Odenwald et al., 2005) We refer to these regions as *mshI-4* from furthest upstream of the coding sequences, and *mshI* for the piece located in the intron (Figure 1A). We used Cis-analyst and other search programs to search for binding sites of genes known to regulate Msh expression (Berman et al., 2001). Specifically, we looked for Vnd, Ind and Mothers against Dpp (Mad) binding sites, the best-known early regulators of Msh expression (D'Alessio and Frasch, 1996; Von Ohlen and Doe, 2000). The combination of these approaches identified two regions of interest. The first region we chose to focus on, named *msh2*, is located approximately 6.4 to 8.5kb upstream of the coding sequences. The second region named *mshI*, is located within an intron closely adjacent to the first exon.

Recent identification of transcriptional targets of mesoderm specific factors and DV patterning transcription factors has allowed us to narrow down the potential regions of *msh* regulatory DNA (Sandmann et al., 2007) (Zeitlinger et al., 2007). The chromatin-IP experiments done by these groups identified a region corresponding to our region 2. One group also identified the *msh* intron region as a region bound by Twist and other mesoderm specific factors (Sandmann et al., 2007). It is known that *msh* is expressed in and required for muscle development (Lord et al., 1995; Nose et al., 1998). Since Twist is a major regulator of mesoderm formation, it is entirely possible that the regions identified by these groups are muscle specific enhancers for *msh*. In order to confirm that potential regions are true regulatory elements we generated reporter constructs with the *msh2* and *mshI* regions in a *lacZ* reporter vector. Because the other regions did not meet both our criteria we did not investigate them further. Our experiments indicate that the *msh2* region contains the regulatory sequences controlling the early aspects of Msh expression (Figure 1B-E). Specifically, we first saw stripes of expression in a pattern similar to the earliest *msh* expression around stage six (Figure 1B). From stages 7 to 10 the *lacZ* message was detected in stripes in the presumptive neuroectoderm in a pattern that closely resembles the endogenous *msh* mRNA pattern (Figure 1B-C). By late stage 11, the *LacZ* message had largely faded from the neuroectoderm, and expression remained in the salivary gland placodes (Figure 1D). Thus, these results indicate that we have identified a regulatory element controlling early DV restricted Msh expression.

Based on the presence of predicted Vnd binding sites, the intron piece (*mshI*) was also tested for its ability to drive reporter gene expression in the embryo. We found that this reporter construct also expressed in the developing embryo. However, not in the neurectoderm, instead, we observed expression in segmentally repeated spots of cells located laterally in stage 10-11 embryos (Figure 2A). These cells could be the lateral mesodermal precursors, a tissue known to express Msh (D'Alessio and Frasch, 1996). To test this we performed double labels with anti-Msh and anti- β gal antibodies. We found that the expression of Msh and β gal were only partially over-lapping (Figure 2C-E). These data suggested that there might be additional regulatory information required to precisely duplicate the endogenous Msh pattern in the lateral cells. Because Msh is also expressed in the larval imaginal discs during development, we also examined both the *Msh2-lacZ* and *MshI-lacZ* reporter lines for expression during larval stages and did not detect expression in third instar larval wing or leg discs for any of the lines tested (Data not shown). This suggests we have successfully identified regulatory DNA controlling expression of *msh* in the neurectoderm and also an element that partially functions in expression in the lateral mesodermal precursors.

The *msh2* regulatory element contains information for repression ventrally by Ind and Vnd

Ind is a direct regulator of *msh* expression. Thus, we predicted that the *msh2* element would contain sequences necessary for repression by Ind. To test this hypothesis, we first performed double label experiments for *lacZ* message and Ind protein. We found that the domains of *lacZ* and Ind were mutually exclusive just as the endogenous patterns are (Figure 3C-D). In order to examine the effect of loss of *ind* on reporter gene expression, we recombined the *Msh2-lacZ* transgene onto an *ind* mutant chromosome. This experiment revealed that like the endogenous *msh* message, the expression of *lacZ* was expanded ventrally in the *ind* mutant embryos (Compare Figure 3B to F and G to H). We also tested the ability of ectopically expressed Ind to repress reporter expression in the lateral column. To do this, we built recombinant chromosomes in which the reporter transgene was recombined onto the chromosomes carrying both the *Krüppel Gal4* driver transgene and the *UAS Ind* transgene. Again we found that ectopic expression of Ind efficiently repressed *Msh2-lacZ* expression (Figure 3I). Taken together, these results strongly support the hypothesis that the region of genomic DNA contained in the *Msh2-lacZ* reporter construct, not only contains sequences essential for initiation of *msh* in the lateral neurectoderm, but also the sequences essential for repression of *msh* in the intermediate column.

Vnd can directly repress Msh expression (Mc Donald et al., 1998). To determine if the *Msh2* element we identified is directly repressed by Vnd, we also built *vnd; indMsh2-lacZ* double mutants to produce double mutant embryos. Msh protein expands ventrally in *vnd; ind* embryos (Von Ohlen and Doe, 2000). Therefore, we anticipated a similar result if the *Msh2* regulatory element also contained sequences necessary for repression by Vnd. We also predicted that ectopic expression of Vnd should be able to repress transcription of the reporter construct. Indeed, in *vnd; ind Msh2-lacZ* embryos *lacZ* message was observed throughout the neurectoderm (Figure 3J). Also, ectopic expression of Vnd across the DV axis was sufficient to repress transcription of the reporter construct (Figure 3K). These data suggest that the *Msh2* regulatory DNA contains the information necessary for repression by both Vnd and Ind in ventral neurectodermal tissue.

Identification of a minimal regulatory element

We wanted to further characterize the action of Ind on the *msh* regulatory element. To do this, we needed to identify the minimal sequence required for expression in the early neurectoderm. Four additional reporter constructs were built with partially overlapping pieces of the original *Msh2* element (Figure 4A). These experiments revealed that the *Msh2a* fragment positioned furthest from the coding sequences produced a pattern of reporter gene

expression largely consistent with endogenous *msh* expression (Figure 4B-E). The *Msh2b* element drove reporter gene expression in the dorsal epidermis in a pattern reminiscent of highly responsive Dpp targets (Figure 4F). We failed to detect expression in embryos from elements *Msh2c* or *Msh2d* (data not shown). These results demonstrate that we have narrowed the region essential for initiation of *msh* expression in the lateral neurectoderm and repression in the ventral and intermediate neurectoderm down to the 699bp fragment encompassed by the *Msh2a* element. In this region we found four putative Ind binding sites and four putative Vnd binding sites. The presence of putative binding sites for Vnd and Ind suggests that the *msh2a* element contains the sequences necessary for repression of *msh* in the ventral neuroectoderm.

Ind binds directly to the *msh2a* element in a sequence specific manner

If Ind is acting to directly repress *msh* through the *msh2a* regulatory element, then we predicted that Ind would bind to sequences in the *msh2a* fragment in a sequence specific manner. To test this hypothesis, we identified four potential Ind/homeodomain binding sites within the sequence of the *Msh2a* region. For each of these there were pairs of two sites within 20 bp of each other. We designed probes for DNA binding experiments that contained each of the pairs of sites (Figure 5A & D). We tested the ability of Ind to bind to a fragment containing sites one and two. For these experiments, we used the IndHD protein described in (Von Ohlen et al., 2007b), to assess binding. We found that when sites one and two were intact, IndHD protein was able to bind to and alter the mobility of the fragment in a non-denaturing gel (Figure 5B). When we mutated the sequence of each of these sites to sequences that no longer allow Ind binding, we found that mutation of site two resulted in loss of ability of IndHD to bind to the fragment (Figure 5C). These data suggested, that Ind bound specifically to site two but not site one. Next, we examined the ability of Ind to bind sites three and four in the same manner. We found that Ind bound to and altered the mobility of a fragment containing both sites (Figure 5E). However, when we mutated the sequence of the sites, we found that Ind bound to the fragment mutated for site three but not for site four (Figure 5F). This demonstrates that binding is dependent on an intact site four. We conclude that Ind-mediated regulation of *msh* is via at least two sites in the *Msh2a* regulatory DNA.

An alignment of the *msh2a* sequences from six of the *Drosophila* species used to generate the evo-print reveals interesting details about predicted versus actual Ind binding sites (See supplemental Figure S2). We predicted Ind binding site based on a five out of six base pair match to the homeodomain recognition sequence TAATGG, described in (Zhao et al., 2007). When we compared sequences from the different species, we found that the sequences for site 2, 3 and 4 are highly conserved in at least four of the six sequences. Site one is only conserved in the two most closely related species and does not conserve the TAAT core consensus sequence that is typical for homeodomain proteins. Because site one lacks the TAAT sequence it is not surprising that Ind failed to bind this sequence. Both sites 3 and 4 are conserved in all six species and retain the TAAT core consensus sequence. It is unclear what the difference between these two sites but we hypothesize that additional flanking sequences may play a role in Ind binding. What these analyses lead us to conclude is that Ind binding requires an exact match to the TAAT or ATTA core sequence within the binding site and that additional flanking sequence may be important for efficient binding.

Intact Ind binding sites in the *Msh2a* element are necessary for repression in the intermediate column

Once we had identified the sites essential for Ind binding to the *msh2a* element, we wanted to confirm that these were the functional Ind binding sites in vivo. Therefore, we mutated sites two and four in the *Msh2a* element to sequences that would no longer support Ind binding, and generated reporter transgenic lines to test the effect on expression of the

reporter gene *lacZ*. We then examined expression of the reporter gene with either *lacZ* antisense probe or anti- β gal antibody in conjunction with anti-Ind antibody. We found that the expression of *lacZ* in the mutant reporter construct was slightly wider and expanded ventrally compared with expression of the wild-type reporter (Figure 6; compare D & E to A & B). In the embryos carrying the wild-type *Msh2a-lacZ* transgene we occasionally observed some over-lap between the expression of β gal and Ind (not shown), suggesting that there are additional Ind binding sites located outside the *msh2a* region that regulate more complete repression in the intermediate column. However, when we compared the *msh2a-mutant lacZ* expression pattern to the wild-type reporter, we found a higher degree of overlap between the Ind protein pattern and β gal expression (Figure 6F). These results strongly support the hypothesis that, the Ind binding sites identified above (Figure 5) are at least partially responsible for Ind mediated repression of *msh2a* reporter gene expression.

Msh is not essential for maintaining its own expression

Both Vnd and Ind are known to be essential for maintaining their own expression (Saunders et al., 1998; Uhler et al., 2007; Von Ohlen et al., 2007b). Thus, we tested the hypothesis that the Msh homeodomain protein served a similar role in maintaining *msh* expression. To do this, we took two approaches: First we tested for the presence or absence of *msh* message in embryos that were homozygous for loss of function alleles of *msh*. For these experiments we used the *msh^{SOD54}* allele described by (Mozer, 2001). We chose this allele because it was generated by EMS mutagenesis and had the potential to produce transcript. Other described alleles were generated by P-element excision and are known to take out the transcription start site (Isshiki et al., 1997). We found that *msh* mRNA was detectable in *msh^{SOD54}* homozygous mutant embryos in the normal pattern throughout the early stages of neurogenesis (Figure 7D-F). Second, we recombined the *Msh2-lacZ* reporter transgene onto the *msh^{Δ68}* mutant chromosome. When we examined reporter gene expression we found that *lacZ* message was expressed in the embryos in a pattern similar to the wild type expression pattern (Figure 7G-I). Thus, Msh activity was not exclusively necessary for the maintenance of either endogenous message or reporter gene expression. Other genes, perhaps those required for initiation of *msh* expression, are also essential for maintenance of the pattern. Furthermore, additional sequences not identified here may also be required to confer *msh* dependent expression.

Discussion

In *Drosophila* and in vertebrates expression of proteins in precisely delimited stripes is critical for proper development of the embryo and differentiation of specific cell types. Although these stripes can be in different orientations and tissues we have focused on the formation of DV restricted stripes in the embryonic neuroectoderm. The mechanism of regulation of *msh* expression in the neuroectoderm of *Drosophila* embryos has not been studied at the molecular level.

This article represents the first description of regulatory DNA essential for expression of *msh* in the early neuroectoderm of the embryo and direct repression by Ind. We identified two regions of genomic DNA that potentially could function as neuroectodermal regulatory DNA. One of these regions (*msh1*) is located in the intron between the first and second exons. This element appears to be partially able to reproduce *msh* expression in the lateral mesodermal precursors. The second region (*msh2*) located 6.5 to 8.7kb upstream of the coding sequences, is capable of driving reporter gene expression in a pattern consistent with the endogenous *msh* pattern in stripe in the early embryonic neuroectoderm. Further analysis of this element identified a 699bp fragment (*msh2a*) that functions as a minimal regulatory element. This element was directly bound by the Ind homeodomain in a sequence specific manner. We identified four putative binding sites for Ind in the minimal element. We found

that, Ind efficiently bound two of these sites and that this interaction was required for maximal repression of a reporter gene in the intermediate column. Furthermore, the identified regulatory elements are each required for specific subsets of the embryonic expression pattern. Thus far, we have examined expression of reporter constructs for these two conserved regions of potential regulatory DNA, in the embryo. We also tested for expression in larval tissues such as imaginal discs and brain, and failed to detect any reporter gene expression in these tissues. However, the pattern of *msh* expression in the embryo and throughout development is quite dynamic. Thus, the other regions of conserved sequence may include regulatory information essential for other aspects of *msh* expression including expression in the muscle precursors and larval tissues. We also demonstrated that, Msh function was not required for maintenance of *msh* expression in the embryonic neuroectoderm. This is interesting because both Ind and Vnd are required for maintenance of their own expression (Shao et al., 2002; Von Ohlen et al., 2007b). Our results point to a different mechanism of maintenance for *msh* that does not require an auto-regulatory interaction.

msh regulatory DNA was identified based on multiple criteria including; conservation of sequence among Drosophilid species and presence of binding sites for known regulators of *msh* expression. The *msh2* regulatory region and the *msh1* regulatory region were both found to contain exact matches to the known Vnd binding site (Gruschus et al., 1997). Although the putative Vnd binding sites in the *msh1* region meet the criteria for Vnd binding sites they are more likely to be Tin-binding sites instead. Tin and Vnd have identical or highly similar DNA binding specificities (Chen and Schwartz, 1995; Gruschus et al., 1997; Weiss et al., 1998). Although we have not tested this hypothesis directly, we believe that because the *msh1-lacZ* reporter expresses in the mesodermal precursors, and Tin is a master regulator of mesoderm specification, that the sites we identified as Vnd sites in the *msh1* region, are more likely Tin sites.

Vnd and Ind interact with additional co-factors such as sox domain proteins Dichaete and SoxNeuro (soxN) (Buescher et al., 2002; Overton et al., 2002) (Zhao and Skeath, 2002) (Zhao et al., 2007). For example, Ind and Dichaete interact on the *achaete* regulatory DNA, and this interaction is necessary for effective repression of *achaete* in the intermediate column (Zhao et al., 2007). Because we believe these types of interactions are highly variable and dependent on which regulatory DNA the homeodomain proteins are interacting with, we explored how *msh* regulatory DNA is affected by these interactions. Dichaete interaction has also been implicated in Ind and Vnd mediated repression of Msh (Zhao and Skeath, 2002). However, our data suggest that interaction with Dichaete is not essential for regulation of Msh by Ind. Our reasoning for this is multi-fold, first in the absence of Dichaete function, *msh* expression appears completely normal and not partially expanded ventrally (Buescher et al., 2002; Overton et al., 2002). Second, we have carefully examined the sequence of the *msh2a* regulatory element defined here, and cannot identify any potential Sox domain binding sites. Finally, ectopically expressed Ind was an efficient repressor of *msh* in the lateral column, despite of the fact that Dichaete is only expressed in the ventral and intermediate columns (Von Ohlen et al., 2007a). Thus, association between Ind and Dichaete is not essential for Ind to repress *msh* expression.

Materials and methods

Fly stocks

UASInd (Von Ohlen et al., 2007b), KrGal4, *yw*, *tin*³⁴⁶/TM3ftzlacZ was obtained from Rolf Bodmer (San Diego); *ind*^{16.2}/TM3ftzlacZ (Weiss et al., 1998); *msh*^{SOD54}/TM3ftzlacZ was obtained from Brian Mozer (NIH) (Mozer, 2001); *msh*⁴⁶⁸/TM3ftzlacZ (Isshiki et al., 1997); UASvnd and *vnd*⁶/FM7 (Skeath et al., 1994; Mc Donald et al., 1998).

Generation of *msh lacZ* reporter constructs—An *msh* evoprint was generated with the help of Thomas Brody at the NIH (Odenwald et al., 2005). *msh2* and *msh1* fragments were amplified by PCR from genomic DNA isolated from *yw* flies, using the following primers: *msh2F*: agttgccccgaggtctggagtt and *msh2R*: gctggcttcgttggtttctgc. *Msh1F*: ccgctcgagatttagacctaatac and *msh1R*: ccgctcgaggttaaattatgaagcac.

Msh2a-d constructs were generated by the same approach using the following primer sets:

GTTGCCCCGAGGTCTGGAGTTAGT *msh2a* upper

CACGCCAGATCGAAAAGAAGGAGA *msh2a* lower

TACCTCAGTTACACTGG *msh2 b* upper

TATGAAAATCCGCCAG *msh2b* lower

CGAAACGCTGGCGGCTGTGT *msh2c* upper

CCGCGTATCCATATCCCTGTGCTC *msh2c* lower

GCGGCGATATGCGGACTC *msh2d* upper

ACCCCTGCTGACTCTTGTGGACT *msh2 d* lower

In all cases PCR products were cloned using the pCR-TOPO4 cloning kit from Invitrogen, and subsequently cloned into the pCaSpeR Hsp43lacZ transformation vector at the *EcoRI* site. Transformation was done at the Non-mammalian model systems unit at Duke University. For each reporter constructs several independent insertion lines were tested. Figures show representative examples of the expression patterns observed.

DNA binding assays

Electrophoretic mobility shift assays were performed using the Histagged Ind Homeodomain protein (His6-IndHD) as described in (Von Ohlen et al., 2007b). WT and mutant oligos were annealed to form dsDNA molecules. EMSA assay was performed using the Dig non-radioactive Gel shift kit (second generation, Roche). The DNA/protein complexes were resolved on a 5% non-denaturing polyacrylamide gel. The probes used are as follows: Fragment 1

Mut 1(TTATGG) Mut 2 (GCATTA)

5'CATTCTGGTGCACACACATTCGAGCATTCTGGCACCTGGCACACAGGAGCACCA
CCAGGG3'

Fragment 2: Mut 1(TTATGG) Mut 2 (GCATTA)

5'CCTTTTGTGTTGGGCCACGCGGTTCTTTACTTTAAATACGCGCGCATAACTGCA
GCTGAA3'

Mutated sequence is show in *italics* actual sequence for each site is shown in parenthesis above.

Site directed mutagenesis

Putative Ind binding sites in the GN4 enhancer element we mutagenized using the Quick Change Multi-site directed mutagenesis kit (Stratagene). The following mutagenic primers were used to convert the core ATTA sequence to sequences that no longer support Ind binding. The Mutagenic oligos used are:

Site 2: AGCATTCTGGCACCTGGGCGCGGAGCACCACCAGGGTGTG

Site 4:

GGGATTCTTTACTTTAAATACCAAGGGATAACTGCAGCTGAAGCTGGAA

Italics indicate altered base pairs.

Antibody stains and insitu hybridizations

In situ hybridizations were performed according to standard procedures as described in: (Tautz and Pfeiffle, 1989). In all experiments shown we used a digoxigenin labeled anti-sense *lacZ* probe.

The antibody stains used rabbit anti-Ind (Von Ohlen and Moses, 2009) at a 1:2000 dillution and mouse anti-βGal at a 1:500 dillution (Promega). Secondary antibodies were biotin labeled anti-rabbit (Vector labs) was used in conjunction with the Vectastain Elite streptavidin-Horse radish peroxidase kit to produce a brown precipitate. Alkaline Phosphatase-conjugated anti-mouse (Cappel) was used to produce purple precipitate.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Dr. S. Keith Chapes for his help editing the manuscript. We are also grateful to Dr Tom Brody for help with generating and interpreting the Evoprint. This publication was in part, made possible by an Innovative Research Award, to TVO, from the Terry C. Johnson Center for Basic Cancer Research at Kansas State University. Additional funding was provided by: Grant Number P20 RR016475, to TVO, from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH). The contents of this publication are solely the responsibility of the author and do not necessarily represent the official views of NCRR or NIH.

Literature Cited

- Berman BP, Nibu Y, Pfeiffer BD, Tomancak P, Celniker SE, Levine M, Rubin GM, Eisen MB. Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the *Drosophila* genome. PNAS. 2001; 99:757–762. [PubMed: 11805330]
- Buescher M, Hing FS, Chia W. Formation of Neuroblasts in the embryonic central nervous system of *Drosophila melanogaster* is controlled by *SoxNeuro*. Development. 2002; 129:4193–4203. [PubMed: 12183372]
- Chen CY, Schwartz RJ. Identification of novel DNA binding targets and regulatory domains of a murine tinman homeodomain factor, *nkx-2.5*. J Biol Chem. 1995; 270:15628–15633. [PubMed: 7797561]
- Cowden J, Levine M. Ventral dominance governs sequential patterns of gene expression across the dorsal-ventral axis of the neuroectoderm in the *Drosophila* embryo. Dev Bio. 2003; 262:335–349. [PubMed: 14550796]
- D'Alessio M, Frasch M. *msh* may play a conserved role in dorsoventral patterning of the neuroectoderm and mesoderm. Mech Dev. 1996; 58:217–231. [PubMed: 8887329]
- Estes P, Mosher J, Crews ST. *Drosophila* single-minded represses gene transcription by activating the expression of repressive factors. Dev Biol. 2001; 232:157–175. [PubMed: 11254355]
- Gruschus JM, Tsao DHH, Wang L-H, Nirenberg M, Ferretti JA. Interactions of the *vnd/NK-2* homeodomain with DNA by nuclear magnetic resonance spectroscopy: Basis of binding specificity. Biochem. 1997; 36:5372–5380. [PubMed: 9154919]
- Isshiki T, Takeichi M, Nose A. The role of the *msh* homeobox gene during *Drosophila* neurogenesis: implication for the dorsoventral specification of the neuroectoderm. Development. 1997; 124:3099–3109. [PubMed: 9272951]

- Lord PC, Lin MH, Hales KH, Storti RV. Normal expression and the effects of ectopic expression of the *Drosophila* muscle segment homeobox (*msh*) gene suggest a role in differentiation and patterning of embryonic muscles. *Dev Biol.* 1995; 171:627–640. [PubMed: 7556942]
- Markstein M, Zinzen R, Markstein P, Yee K-P, Erives A, Stathopoulous A, Levine M. A regulatory code for neurogenic gene expression in the *Drosophila* embryo. *Development.* 2004; 131:2387–2364. [PubMed: 15128669]
- Mc Donald JA, Holbrook S, Isshiki T, Weiss J, Doe CQ, Mellerick DM. Dorso-ventral patterning in the *Drosophila* CNS: the *vnd* homeobox gene specifies ventral column identity. *Genes Dev.* 1998; 12:3606–3612.
- Mellerick DM, Nirenberg M. Dorsal-Ventral patterning genes restrict NK-2 homeobox gene expression to the ventral half of the central nervous system of *Drosophila* embryos. *Developmental Biology.* 1995; 171:306–316. [PubMed: 7556915]
- Miller DF, Holtzman SL, Kalkbrenner A, Kaufman TC. Homeotic Complex (Hox) gene regulation and homeosis in the mesoderm of the *Drosophila melanogaster* embryo: the roles of signal transduction and cell autonomous regulation. *Mech Dev.* 2001; 102:17–32. [PubMed: 11287178]
- Mozer BA. Dominant Drop mutants are gain-of-function alleles of the muscle segment homeobox gene (*msh*) whose overexpression leads to the arrest of eye development. *Dev Biol.* 2001; 233:380–393. [PubMed: 11336502]
- Nose A, Isshiki T, Takeichi M. Regional specification of muscle progenitors in *Drosophila*: the role of the *msh* homeobox gene. *Development.* 1998; 125:215–223. [PubMed: 9486795]
- Odenwald WF, Rasband W, Kuzin A, Brody T. EVOPRINTER, a multigenomic comparative tool for rapid identification of functionally important DNA. *Proc Natl Acad Sci U S A.* 2005; 102:14700–14705. [PubMed: 16203978]
- Overton PM, Meadows LA, Urban J, Russell S. Evidence for differential and redundant function of the *Sox* genes *Dichaete* and *SoxN* during CNS development in *Drosophila*. *Development.* 2002; 129:4219–4228. [PubMed: 12183374]
- Sandmann T, Girardot C, Brehme M, Tongprasit W, Stolz V, Furlong EE. A core transcriptional network for early mesoderm development in *Drosophila melanogaster*. *Genes Dev.* 2007; 21:436–449. [PubMed: 17322403]
- Saunders H-MH, Koizumi K, Odenwald W, Nirenberg M. Neuroblast pattern formation: regulatory DNA that confers the *vnd*/NK2 homobox gene pattern on a reporter gene in transgenic lines of *Drosophila*. *PNAS.* 1998; 95:8316–8321. [PubMed: 9653184]
- Shao X, Koizumi K, Nosworthy N, Tan DP, Odenwald W, Nirenberg M. Regulatory DNA required for *vnd*/NK-2 homeobox gene expression pattern in neuroblasts. *Proc Natl Acad Sci U S A.* 2002; 99:113–117. [PubMed: 11752402]
- Skeath JB. The *Drosophila* EGF receptor controls the formation and specification of neuroblasts along the dorsal-ventral axis of the *Drosophila* embryo. *Development.* 1998; 125:3301–3312. [PubMed: 9693134]
- Skeath JB, Panganiban GF, Carroll SB. The ventral *nervous system defective* gene controls proneural gene expression at two distinct steps during neuroblast formation in *Drosophila*. *Development.* 1994; 120:1517–1524. [PubMed: 8050360]
- Stathopoulous A, Levine M. Localized repressors delineate the neurogenic ectoderm in the early *Drosophila* embryo. *Dev Bio.* 2005; 280:482–493. [PubMed: 15882587]
- Stathopoulous A, Van Drenth M, Erives A, Markstein M, Levine M. Whole-genome analysis of Dorsal-ventral patterning in the *Drosophila* embryo. *Cell.* 2002; 111:687–701. [PubMed: 12464180]
- Tautz D, Pfeiffle C. A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma.* 1989; 98:81–85. [PubMed: 2476281]
- Uhler J, Zhang H, Syu LJ, Mellerick DM. The Nk-2 box of the *Drosophila* homeodomain protein, *Vnd*, contributes to its repression activity in a Groucho-dependent manner. *Mech Dev.* 2007; 124:1–10. [PubMed: 17070676]
- Von Ohlen T, Doe CQ. Convergence of Dorsal, Dpp and Egfr signaling pathways subdivides the *Drosophila* neuroectoderm into three dorsal-ventral columns. *Dev Bio.* 2000; 224:362–372. [PubMed: 10926773]

- Von Ohlen T, Syu LJ, Mellerick DM. Conserved properties of the *Drosophila* homeodomain protein, Ind. *Mech Dev.* 2007a; 124:925–934. [PubMed: 17900877]
- Von Ohlen TL, Harvey C, Panda M. Identification of an upstream regulatory element reveals a novel requirement for Ind activity in maintaining ind expression. *Mech Dev.* 2007b; 124:230–236. [PubMed: 17224261]
- Von Ohlen TL, Moses C. Identification of Ind transcription activation and repression domains required for dorsoventral patterning of the CNS. *Mech Dev.* 2009
- Weiss JB, Von Ohlen T, Mellerick D, Dressler G, Doe CQ, Scott MP. Dorsoventral patterning in the *Drosophila* Central Nervous System: The intermediate neuroblasts defective Homeobox gene specifies intermediate column identity. *Genes Dev.* 1998; 12:3591–3602. [PubMed: 9832510]
- Yagi Y, Suzuki T, Hayashi S. Interaction between *Drosophila* EGF receptor and *vnd* determines three dorsoventral domains of the neuroectoderm. *Development.* 1998; 125:3625–3633. [PubMed: 9716528]
- Zeitlinger J, Zinzen RP, Stark A, Kellis M, Zhang H, Young RA, Levine M. Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev.* 2007; 21:385–390. [PubMed: 17322397]
- Zhao G, Boekhoff-Falk G, Wilson BA, Skeath JB. Linking pattern formation to cell-type specification: Dichaete and Ind directly repress achaete gene expression in the *Drosophila* CNS. *Proc Natl Acad Sci U S A.* 2007; 104:3847–3852. [PubMed: 17360441]
- Zhao G, Skeath JB. The Sox-domain containing gene *Diachaete/fish-hook* acts in concert with *vnd* and *ind* to regulate cell fate in the *Drosophila* neuroectoderm. *Development.* 2002; 129:1165–1174. [PubMed: 11874912]

Figure 1.

Identification of Msh regulatory DNA. Top Line A) Diagram of the genomic region upstream of, and including *msh* coding sequences. Green boxes indicate regions of conserved sequence among *Drosophila* species. Approximate distances from the coding sequences are indicated for the end of each conserved block. Red boxes indicate Twist binding regions identified by ChIP experiments (Sandmann et al., 2007) (Zeitlinger et al., 2007). Blue boxes are a schematic of *msh* exons. Black scale bar represents approximately 1kb. B-E) *lacZ* mRNA in *msh2-lacZ* embryos at stages of increasing age. B) Stage 6 C) Stage 9 D) Stage 11 E) stage 14. All embryos are ventro-lateral views, anterior is up.

Figure 2.

Sequences in the first intron of *msh* control expression in mesodermal precursors. A) stage 11 *mshI-lacZ* embryo, *lacZ* mRNA. B) WT embryo stage 11, *msh* mRNA. Arrows indicate the positions of lateral mesodermal precursor cells. C-E) Doubly labeled, stage 11 *mshI-lacZ* embryo, β -gal protein brown and Msh protein in blue. D-E) High magnification images of double label embryos arrows show that domains are partially overlapping. F) late stage *mshI-lacZ* embryo, *lacZ* mRNA. Expression is detected in glial cells. A-C) Embryos are lateral views, anterior is up. D-E) Embryos are ventral views, anterior is up.

Figure 3.

Msh2lacZ contains sequences necessary for repression by Ind in the intermediate column. A-D) Wild type embryos. A) *Msh2-lacZ* mRNA Stage 6 B) *Msh2-lacZ* mRNA Stage 9 C) *Msh2-lacZ* mRNA (blue), Ind protein (brown) stage 9. D) *Msh2-lacZ* mRNA (blue), Ind protein (brown) stage 11. E) *ind^{16.2}* embryo *Msh2-lacZ* mRNA stage 6. F) *ind^{16.2}* embryo *Msh2-lacZ* mRNA stage 9. G-H) High magnification images of stage 9/10 WT (G) and *ind^{16.2}* (H) embryos arrows emphasize change in distance between *msh2lacZ* mRNA stripes. I) *Msh2-lacZ* mRNA in KrGal4x UAS ind embryo stage 10. J) *Msh2-lacZ* mRNA in *vnd; ind* *Msh2-lacZ* embryo stage 10. K) *Msh2-lacZ* mRNA in KrGal4x UAS Vnd embryo stage 9/10. All embryos are ventral views except A and E which are lateral views and I and K which are ventro-lateral views. Anterior is up in all cases.

Figure 4.

A 699bp fragment contains the minimal element necessary for neurectoderm expression. A) Schematic illustrating the subdivision of the *msh2* element to further define the minimal element. Top shows sizes of fragments a, b, c, and d. bottom shows approximate position of *msh2* element relative to the coding sequences. B) *Msh2a-lacZ* mRNA stage 6. C) *Msh2a-lacZ* mRNA stage 9. D) *Msh2a-lacZ* mRNA stage 10. E) *Msh2a-lacZ* mRNA stage 11. F) *Msh2b-lacZ* mRNA stage 6. All embryos are positioned anterior up.

Figure 5.

The Ind homeodomain binds to the msh2a element in a sequence specific manner. A) sequence of msh2a region use as probe for DNA binding in panels B and C. Possible Ind sites are highlighted in grey. B) IndHD binds to a fragment containing the first two putative Ind binding sites. C) Ind binds to and shifts DNA containing mutated site 1 (lanes 4-6) but not mutated site 2 (lanes 1-3), or both (lanes 7-9). D) sequence of msh2a region use as probe for DNA binding in panels E and F. Possible Ind sites are highlighted in grey. E) IndHD binds to a fragment containing putative Ind binding sites three and four. F) Ind binds to and shifts DNA containing mutated site 3 (lanes 1-3) but not mutated site 4 (lanes 4-6), or both (lanes 7-9).

Figure 6.

Intact Ind bindings sites in the *msh* regulatory DNA are essential for repression of *msh2a-lacZ* in the intermediate column. A-B) expression of the WT *msh2a-lacZ* message in stage 9 embryos. A) lateral view B) ventral view. C) WT *msh2a* embryo labeled for Ind protein and βGal protein. D-E) Expression of the *msh2a mut-lacZ* message in stage 9 embryos. D) lateral view E) ventral view. Arrows are added to emphasize the change in distance between the stripes of *lacZ* expression. F) *msh2a mut-lacZ* embryo labeled for Ind protein and βGal protein. Arrows point to cells that appear doubly labeled for both Ind and βGal, illustrating that there are regions of over-lap between the Ind and βGal patterns.

Figure 7.

Msh activity is not required for initiation or maintenance of *msh* expression. A-C) Wild type embryos, *msh* mRNA. A) stage 6. B) stage 9. C) stage 10. D-F) *msh*^{SOD54} embryos, *msh* mRNA. D) stage 6. E) stage 9. F) stage 10. G-I) *msh*^{Δ68} embryos *msh2-lacZ* mRNA. G) stage 6. H) stage 9. I) stage 10. All embryos are ventrolateral views, anterior is up.