Developmental Biology 445 (2019) 226–236

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/locate/developmentalbiology

# The *cis*-regulatory logic underlying abdominal Hox-mediated repression versus activation of regulatory elements in *Drosophila*

Arya Zandvakili[a,b], Juli D. Uhl[a,c], Ian Campbell[d], Joseph Salomone[a,b], Yuntao Charlie Song[a], Brian Gebelein[e,f,*]

[a] Graduate Program in Molecular and Developmental Biology, Cincinnati Children's Hospital Research Foundation, Cincinnati, OH 45229, USA
[b] Medical-Scientist Training Program, University of Cincinnati College of Medicine, Cincinnati, OH 45229, USA
[c] Department of Biological Sciences, University of Cincinnati, Cincinnati, OH, USA
[d] Department of Biomedical Engineering, University of Cincinnati, Cincinnati, OH, USA
[e] Division of Developmental Biology, Cincinnati Children's Hospital, 3333 Burnet Ave, MLC 7007, Cincinnati, OH 45229, USA
[f] Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA

ARTICLE INFO

*Keywords:*
Hox
Extradenticle
Homothorax
cis-regulatory module
Transcription factor
Gene regulation

ABSTRACT

During development diverse transcription factor inputs are integrated by *cis*-regulatory modules (CRMs) to yield cell-specific gene expression. Defining how CRMs recruit the appropriate combinations of factors to either activate or repress gene expression remains a challenge. In this study, we compare and contrast the ability of two CRMs within the *Drosophila* embryo to recruit functional Hox transcription factor complexes. The *DCRE* CRM recruits Ultrabithorax (Ubx) and Abdominal-A (Abd-A) Hox complexes that include the Extradenticle (Exd) and Homothorax (Hth) transcription factors to repress the *Distal-less* leg selector gene, whereas the *RhoA* CRM selectively recruits Abd-A/Exd/Hth complexes to activate *rhomboid* and stimulate Epidermal Growth Factor secretion in sensory cell precursors. By swapping binding sites between these elements, we found that the *RhoA* Exd/Hth/Hox site configuration that mediates Abd-A specific activation can convey transcriptional repression by both Ubx and Abd-A when placed into the *DCRE*. We further show that the orientation and spacing of Hox sites relative to additional binding sites within the *RhoA* and *DCRE* is critical to mediate cell- and segment-specific output. These results indicate that the configuration of Exd, Hth, and Hox site within *RhoA* is neither Abd-A specific nor activation specific. Instead Hox specific output is largely dependent upon the presence of appropriately spaced and oriented binding sites for additional TF inputs. Taken together, these studies provide insight into the *cis*-regulatory logic used to generate cell-specific outputs via recruiting Hox transcription factor complexes.

## 1. Introduction

The generation of distinct cell types within the metazoan body plan requires the accurate regulation of gene products needed to make specialized cell types. At the transcriptional level, *cis*-regulatory modules (CRMs) use transcription factor binding sites (TFBSs) to recruit and integrate diverse transcriptional inputs that modulate the cell specific production of mRNA (Spitz and Furlong, 2012). Recent genomic studies indicate most genes have numerous distinct CRMs, each of which contributes to the overall gene transcriptional output (Kieffer-Kwon et al., 2013). Much like we can read the primary genomic DNA sequence to predict intron/exon boundaries and protein coding regions, we would also like to read CRM sequences to predict both the relevant TFBSs and the transcriptional output mediated by the CRM. However, there are several challenges in predicting CRM function from primary sequence. First, metazoan genomes encode for hundreds to thousands of different sequence-specific transcription factors, many of which are members of protein families that bind highly similar sequences in a degenerate manner (Weirauch et al., 2014). Second, the relatively few CRMs that have been thoroughly characterized reveal differences in the importance of TFBS organization in mediating robust outcomes. Some CRMs are able to use flexible TFBS arrangements (i.e. the billboard model) whereas others require precise TFBS organization to mediate cooperative TF complexes (i.e. the enhanceosome model) (Spitz and Furlong, 2012). In this paper, we dissect how distinct TFBS arrangements within two conserved Hox-regulated CRMs contribute to opposing transcriptional outcomes in the *Drosophila* ectoderm.

* Corresponding author at: Division of Developmental Biology, Cincinnati Children's Hospital, 3333 Burnet Ave, MLC 7007, Cincinnati, OH 45229, USA.
*E-mail address:* brian.gebelein@cchmc.org (B. Gebelein).

https://doi.org/10.1016/j.ydbio.2018.11.006
Received 26 October 2018; Accepted 12 November 2018
Available online 20 November 2018
0012-1606/ 

*Hox* genes encode homeodomain transcription factors that specify distinct cell fates along the developing anterior-posterior (A-P) axis of metazoans (McGinnis and Krumlauf, 1992). Most animals contain at least five Hox genes that are often clustered in the genome (Pearson et al., 2005). *Drosophila melanogaster* encodes a single set of eight Hox genes that are split into two clusters (five in the *Antennapedia* complex and three in the *Bithorax* complex), whereas mammalian genomes have undergone Hox cluster and gene duplication resulting in four clusters that encode a total of 39 Hox genes (Pearson et al., 2005). While the number of Hox genes varies between animals, Hox genes share the property of instructing cells to adopt a "regional" (or "segment") identity within the organism by regulating the expression of target genes (Mann et al., 2009). Genomic studies have found that Hox factors affect the expression of hundreds of downstream target genes (Hueber et al., 2007; Slattery et al., 2011a; Prasad et al., 2016; Zandvakili and Gebelein, 2016). Since each segment under the control of a Hox factor is composed of many cell- and tissue-types, these findings present two challenges in understanding how Hox genes sculpt the body plan: First, what makes one Hox factor different from another to specify distinct embryonic regions during development? Second, how can a regionally expressed Hox factor regulate target genes in a cell- or tissue-specific manner?

Much of the focus on how Hox factors regulate distinct cell fates has been to define the mechanisms underlying DNA binding specificity. Comparative studies between Hox factors revealed each binds similar AT-rich DNA sequences (Noyes et al., 2008; Berger et al., 2008; Affolter et al., 2008). These findings raise a paradox: how do proteins that bind similar DNA sequences in vitro regulate distinct target genes and cell fates in vivo? A partial explanation for this phenomenon is that Hox factors form complexes with additional transcription factors. The Extradenticle (Exd, *Drosophila*)/Pbx (vertebrate) and Homothorax (Hth, *Drosophila*)/Meis (vertebrate) homeodomain proteins represent the best characterized Hox co-factors (Moens and Selleri, 2006; Mann and Chan, 1996; Mann and Affolter, 1998; Merabet and Mann, 2016; Ortiz-Lombardia et al., 2017). Exd/Pbx and Hth/Meis are widely expressed during development and cooperatively bind DNA with Hox factors. As each protein in the complex interacts with DNA in a sequence-specific manner, Hox/Exd/Hth complexes enhance both target affinity and specificity over Hox binding alone (Chan et al., 1994; Chang et al., 1996, 1997). Moreover, a selection assay revealed that Hox factors gain discriminatory power when binding DNA with Exd, a concept called latent specificity (Slattery et al., 2011b). The best example of latent specificity is the activation of *forkhead* (*fkh*) by the Sex combs reduced (Scr) Hox factor during salivary gland development (Joshi et al., 2007; Abe et al., 2015). The *fkh* Exd/Hox binding site has a narrow minor groove that is only bound by Scr when in complex with Exd, and changing this site to match a generic Exd/Hox consensus (*Fkhcon*) resulted in loss of Hox specificity as evidenced by high affinity binding by other Hox factors (Ryoo and Mann, 1999; Joshi et al., 2010). However, not all the Hox factors that bound the *Fkhcon* sequence similarly mediated activation as Scr, as a subset instead repressed transcription through unknown mechanisms (Ryoo and Mann, 1999). Thus, different Exd/Hox sites can discriminate between Hox factors, and once bound, the Hox/Exd complexes can differ in mediating distinct regulatory outcomes.

To better understand the mechanisms of how Hox binding sites mediate distinct outcomes, we have focused on defining how the Abdominal-A (Abd-A) Hox factor contributes to both activation and repression when in complex with the Exd and Hth proteins. The *RhoBAD* CRM contains a highly conserved sequence (*RhoA*) encoding an adjacent set of Exd/Hth/Hox sites that recruits an Abd-A complex to mediate *rhomboid* (*rho*) activation in a subset of sensory organ precursor (SOP) cells (Li-Kroeger et al., 2008, 2012; Gebelein, 2008; Witt et al., 2010). The activation of *rho*, which encodes a serine protease that triggers the release of an EGF ligand, results in the induction of neighboring cells to form an essential set of hepatocyte-like cells known as oenocytes (Brodu et al., 2002; Elstob et al., 2001; Rusten et al., 2001; Gutzwiller et al., 2010). In contrast, the *Distal-less Conserved Regulatory Element* (*DCRE*) contains three Hox/co-factor binding sites that recruit Abd-A/Exd/Hth complexes to repress *Distal-less* (*Dll*) gene expression in the abdominal ectoderm (Gebelein et al., 2002, 2004; Uhl et al., 2016). *Dll*, an appendage selector gene that promotes leg formation in thoracic segments, is thereby restricted from the abdomen to block appendage formation in these segments (Vachon et al., 1992). Intriguingly, the *RhoA* and *DCRE* elements differ in their ability to discriminate between functional Hox complexes, as the *DCRE* is regulated by both Abd-A and Ultrabithorax (Ubx), whereas *RhoA* is regulated by only Abd-A.

What determines if a CRM is activated or repressed when bound by a specific Hox factor? Current models suggest that the *RhoA* and *DCRE* CRMs integrate additional transcription factors that help dictate the sign of transcription. For example, *RhoA* requires a nearby Pax2 binding site to mediate activation, whereas the DCRE contains a nearby FoxG binding site (*Drosophila* express two FoxG homologues, Sloppy-paired 1 (Slp1) and Slp2, which are largely redundant) to mediate repression (Li-Kroeger et al., 2012; Gebelein et al., 2004). How these additional factors are integrated with the Hox transcription factor complexes and the role TFBS organization plays in mediating each cell-specific output is unclear. In this study, we use a series of quantitative reporter assays to define the underlying *cis*-regulatory logic and mechanisms of Hox regulatory specificity by comparing and contrasting the ability of abdominal Hox factors to affect the activity of the *DCRE* and *RhoA* CRMs in conjunction with FoxG and Pax2. Our findings provide new insights into how the organization of Hox, Exd, and Hth binding sites within CRMs contribute to achieving both Hox specificity and positive versus negative regulatory specificity.

## 2. Materials and methods

### 2.1. Transgenic reporter assays

Oligonucleotides for *DCRE* sequence variants were ordered from Integrated DNA Technologies and cloned into the *pAttB-LacZ* plasmid containing *3xGBE*. *RhoAAA* sequences were similarly ordered and cloned into the *pAttB-LacZ* plasmid. DNA sequences for each site are found in Supplemental data. The DNA spacer sequence was generated by PCR amplification of a portion of the *kanamycin* gene as previously described (Swanson et al., 2010). All plasmids were sequence-confirmed prior to injection. Transgenic flies were created using the ϕ-C31 system with each construct inserted into the same locus (51C) (Bischof et al., 2007). Injections were conducted by Rainbow Transgenics Inc.

*Drosophila* stocks carrying *lacZ* transgenes were made homozygous for reporter constructs, and embryos were collected and stained using standard procedures at 25 °C. The *UAS-HA-Ubx* and *UAS-HA-AbdA* lines were a kind gift from Richard Mann, and the *PrdG4;UAS-HA-Ubx* and *PrdG4;UAS-HA-AbdA* experiments were performed at 25 °C. Embryos were immunostained using the following primary antibodies: chicken anti-β-gal (1:1000) (Abcam), guinea-pig anti-Abd-A (1:500) (Li-Kroeger et al., 2012), rat anti-Slp2 (1:500) (Uhl et al., 2016), mouse anti-Ubx (DSHB, 1:50), rabbit anti-Salm (1:2000) (Xie et al., 2007), and rat anti-HA (1:1000). Immunostains were detected using fluorescent secondary antibodies (Jackson Immunoresearch Inc and AlexaFluor).

All samples were imaged using a Zeiss Axio Imager upright microscope with an Apotome filter for optical sectioning. A AxioCam MRm digital camera was used to capture the images. Images were quantified manually using NIH ImageJ software. For each embryo, a region-of-interest (ROI) was placed over Slp2+ cells in embryonic segments T2 through A4 (See Fig. S1 for an example). An additional ROI was created to measure background fluorescence. The area and shape of the ROIs were the same for all measurements within a comparison set. For each embryo, the mean β-gal intensity of the

background ROI was subtracted from the mean β-gal intensity of each foreground ROI. The resultant background-subtracted β-gal intensities were normalized to the background-subtracted β-gal intensity of the T3 (third thoracic segment) ROI. To statistically compare the abdominal activity between reporter genotypes, the resultant normalized β-gal intensities of abdominal segments were mean averaged per embryo and compared using the statistical test indicated in each Figure legend. Statistical analyses and plotting were conducted using R and the ggplot2 package.

### 2.2. DNA binding assays

A His-tagged Slp1 construct was made by cloning the N-terminus through the DNA binding domain of Slp1 (amino acids #1 through 216) into a pET14b vector. His-tagged Exd-Hth heterodimers, Abd-A, and Slp1 protein were purified from BL21 using Ni+ beads, as described previously (Gebelein et al., 2004; Uhl et al., 2010). SDS-PAGE and Coomassie blue staining was used to confirm purification of the protein of interest. For EMSA, fluorescent DNA probes (Integrated DNA Technologies) were mixed with proteins as indicated in Figure legends, incubated for 10 min prior to running on polyacrylamide gels, and imaged using an Odyssey LiCOR cLX scanner as previously described (Uhl et al., 2016; Wang et al., 2017).

## 3. Results

### 3.1. The DCRE mediates short-range transcriptional repression

*Dll* encodes an appendage selector gene that is directly bound and repressed by both the Ultrabithorax (Ubx) and Abdominal-A (Abd-A) Hox factors to suppress abdominal leg development in *Drosophila* (Vachon et al., 1992; Agelopoulos et al., 2012). *Dll* expression in the embryonic leg primordia is mediated by the *DMX cis*-regulatory module that can be divided into two parts: the *DMEact* conveys activation in thoracic and abdominal segments, and the *DCRE* recruits Ubx and Abd-A to repress transcription in a compartment-specific manner in abdominal segments (Gebelein et al., 2002; Agelopoulos et al., 2012). In anterior compartment cells, Ubx and Abd-A mediate repression with the Slp FoxG factors, whereas in posterior compartment cells Ubx and Abd-A cooperatively bind and repress the DCRE with the Engrailed (En) repressor protein (Gebelein et al., 2004). However, recent findings revealed the abdominal Hox factors also repress the *DMEact* via unknown DCRE-independent mechanisms (Uhl et al., 2016). Hence, to study the *cis*-regulatory logic utilized by the *DCRE* to mediate abdominal repression, we used an assay that isolates the *DCRE* from the *DMEact* by placing it adjacent to three copies of the Grainyhead (Grh) binding element (*3xGBE*). *3xGBE* sites are sufficient to activate gene expression throughout the ectoderm of the *Drosophila* embryo (Uhl et al., 2016; Uv et al., 1997). As previously reported, comparisons between *3xGBE-lacZ* (*G-lacZ*) and *3xGBE-DCRE-lacZ* (*GD-lacZ*) transgenes inserted into identical loci revealed the *DCRE* mediates robust repression in anterior compartment abdominal cells that co-express a *Drosophila* FoxG (Slp2) factor (Fig. 1**A-C**), but not in En-positive posterior compartment cells (Uhl et al., 2016). To quantify abdominal repression in anterior compartment cells, we measured β-gal intensity in Slp2+ cells (See Methods and Fig S1) and found that *G-lacZ* drives equivalent reporter levels in thoracic and abdominal segments whereas *GD-lacZ* embryos had ~70% less activity in Slp2-positive cells of abdominal segments relative to thoracic segments (Fig. 1**G**). Hence, the *GD-lacZ* assay provides a means to isolate and study *DCRE*-mediated repression by abdominal Hox factors and Slp FoxG proteins independent from the more complex *DMX* element.

To define the range of *DCRE* repression activity, we engineered a series of constructs that alter the location of the *DCRE* relative to the *3xGBE* (Fig. 1**A**). First, we swapped the order of the *DCRE* and *3xGBE* (*DCRE-3xGBE-lacZ*, *DG-lacZ*) and found that in this configuration the *DCRE* mediates transcriptional repression, albeit weaker and predominantly in a subset of Slp2+ abdominal cells (Fig. 1**D, G**). Next, we moved the *DCRE* further from the *3xGBE* (*DCRE-sp-3xGBE-lacZ, DspG-lacZ*) by inserting a 736 bp sequence from the *kanamycin* gene which was previously found to be transcriptionally inert (Swanson et al., 2010). Consistent with this spacer DNA not having significant transcriptional activity, we found that inserting it adjacent to the *3xGBE* did not significantly alter reporter expression (Fig S2). Importantly, the *DCRE* was unable to convey abdominal repression when it was separated from the *3xGBE* sites by the spacer DNA sequence (Fig. 1**E, G**). However, moving the *3xGBE* adjacent to the distant *DCRE* (*3xGBE-DCRE-sp-LacZ, GDsp-lacZ*) rescued repression (Fig. 1**F-G**). Thus, these findings are consistent with the *DCRE* functioning as a relatively short-range element that represses transcription when placed adjacent to activation elements.

### 3.2. Hox specificity and the cis-regulatory logic of RhoA activation versus DCRE repression

The *Ubx* and *abd-A* Hox genes encode nearly identical homeodomains, bind highly similar DNA sequences as monomers, and form similar transcription factor complexes with the Exd and Hth Hox cofactor proteins on DNA in vitro (Noyes et al., 2008; Slattery et al., 2011b). Consistent with these findings, prior studies demonstrated that Ubx and Abd-A repress *Dll* via the *DCRE (*Gebelein et al., 2002; Gebelein et al., 2004). To determine if both Ubx and Abd-A repress the *DCRE* in the *GD-lacZ* assay, we first compared *GD-lacZ* activity in Slp2+ anterior compartment cells of abdominal A1 segments that only express Ubx versus subsequent abdominal segments that express both Ubx and Abd-A, and observed a similar degree of repression in each abdominal segment (Fig. 2**A-B**, see quantified segment data in Fig. 1**G**). These results are consistent with Ubx being sufficient to repress gene expression via the *DCRE* in the abdominal A1 segment. Moreover, genetic removal of *Ubx* function resulted in a specific loss of *GD-lacZ* repression in A1 segments, whereas *abd-A* mutant embryos maintained significant repression in all abdominal segments (Fig. 2**C-E**). However, embryos mutant for both *Ubx* and *abd-A* were previously found to de-repress *GD-lacZ* activity in all abdominal segments (Uhl et al., 2016). Thus, both Ubx and Abd-A can repress *GD-lacZ* activity in Slp2+ cells, and analysis of *GD-lacZ* reporter activity in the A1 abdominal segment spcifically tests for Ubx-dependent transcriptional repression.

Unlike *Dll* and the *DCRE*, which are repressed by Abd-A, *rhomboid* (*rho*) is activated by Abd-A in a subset of abdominal sensory organ precursor cells (SOPs) via a highly conserved Exd/Hth/Hox binding site within *RhoBAD* (Li-Kroeger et al., 2008, 2012). *rho* encodes a serine protease that triggers the release of an EGF ligand and neighboring cells that receive the EGF signal are specified to form larval oenocytes (Elstob et al., 2001; Rusten et al., 2001). Previous studies showed that the loss of *abd-A*, but not *Ubx*, resulted in a failure to induce oenocytes, suggesting that only Abd-A is required for activating *rho* in abdominal SOP cells (Brodu et al., 2002). However, Ubx is weakly expressed in these abdominal SOP cells. To test whether Ubx is capable of activating *RhoBAD*, we used the *PrdG4* driver to ectopically express high levels of Ubx in the thorax and found that neither *RhoBAD-lacZ* nor oenocytes (marked by high Spalt-major (Salm) expression) were substantially induced in thoracic segments (Fig. 2**G**). In contrast, *PrdG4;UAS-Abd-A* embryos induced both *RhoBAD-lacZ* activity and oenocytes in the thorax (Fig. 2**F**). Thus, while both Abd-A and Ubx can repress the *DCRE* to inhibit leg development, only Abd-A activates *RhoBAD* to induce abdominal oenocyte cells.

A notable difference between the *DCRE* and *RhoA* sequences is the organization of the Hox, Exd, and Hth sites. *RhoA* contains a single set of contiguous Exd/Hth/Hox sites, whereas the *DCRE* has multiple Hox

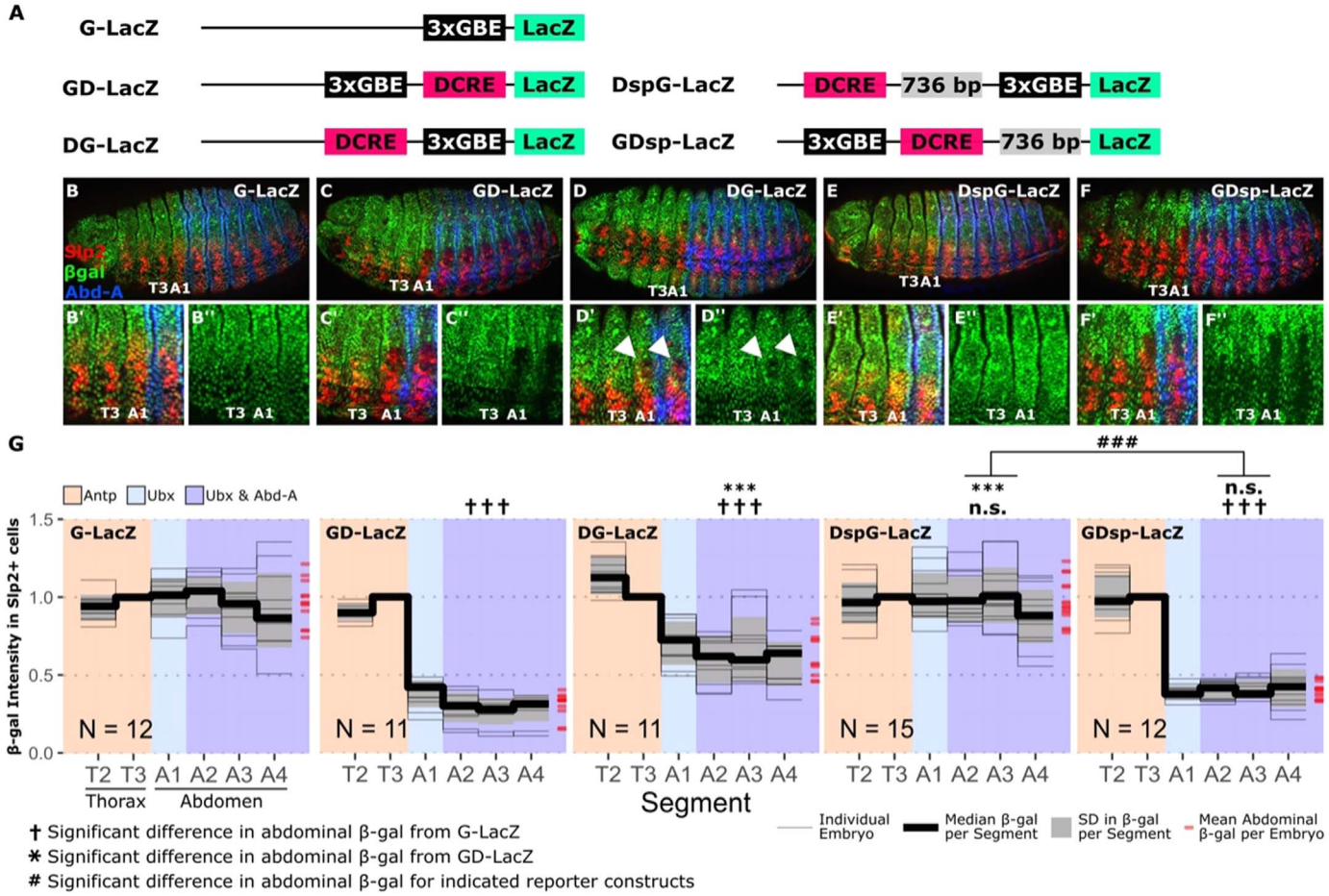


**Fig. 1. The *DCRE* is a short-range transcriptional repressor. (A)** *LacZ* reporter constructs used to test DCRE activity. The 3xGBE is a pan-ectodermal enhancer. The 736 bp spacer is transcriptionally inert (Fig S1). **(B-F)** Lateral view of Stage 15 embryos expressing the indicated *LacZ* reporter constructs. Embryos were immunostained for β-gal (green), Slp2 (red), and Abd-A (blue). **(B′-F′)** Close-up view of the T2-A2 segments from panels B-F. **(B″-F″)** Same as B′-F′ but only showing β-gal stain. **(G)** Quantification of β-gal immunostain intensity among Slp2+ cells in segments T2-A4 relative to intensity in T3 segment. Background color indicates the Hox expression domain as shown by the legend (top-left). Light gray lines are intensity values from individual embryos, while the dark black line displays median value at each segment. The grey ribbon displays the standard deviation at each segment. Tick marks on the right border of the plots indicate mean abdominal β-gal value per embryo (as shown by legend, bottom-right). Abdominal reporter expression was statistically compared between genotypes using ANOVA followed by post-hoc Tukey's test. Prior to ANOVA, abdominal reporter expression in each individual embryo was mean averaged. Crosses mark statistically significant difference from *G-LacZ* ("n.s." $p >= 0.05$, "†" $p < 0.05$, "††" $p < 0.01$, "†††" $p < 0.001$). Asterisks mark statistically significant difference from *GD-LacZ* ("n.s." $p >= 0.05$, "*" $p < 0.05$, "**" $p < 0.01$, "***" $p < 0.001$). Additional pairwise comparisons are indicated by bars and hash-signs ("n.s." $p >= 0.05$, "#" $p < 0.05$, "##" $p < 0.01$, "###" $p < 0.001$).

sites that are each coupled to either an adjacent Exd or Hth binding site (Fig. 3**A**). To determine if the organization of Hox, Exd, and Hth sites contributes to Hox specificity (Abd-A and not Ubx regulation via *RhoA* binding sites), we generated transgenic *GD-lacZ* lines in which the core Exd/Hth/Hox sites of *RhoA* replaced the Hox/Exd-Hth/Hox sequences within the *DCRE* (Fig. 3**B**). Intriguingly, we found that the *RhoA* Exd/Hth/Hox sites can function in the *DCRE* to repress gene expression, but only in one orientation. For example, in the arbitrarily assigned forward direction (AF) the Exd/Hth/Hox sites failed to repress whereas in the reverse orientation (AR) these same sequences mediated significant repression (Fig. 3**E-F, I**). Electromobility shift assays (EMSAs) using purified Exd/Hth heterodimers and Abd-A revealed no significant differences in binding patterns using probes of the *DCRE-AF* versus the *DCRE-AR*, suggesting that differences in Hox DNA binding activity cannot explain the failure of the *DCRE-AF* to mediate repression (Fig S3). Moreover, segment specific analysis of *GD-AR-LacZ* activity in Slp2+ cells revealed that the A1 segment cells that express only Ubx exhibit a similar degree of repression as the abdominal segments that also express Abd-A (Fig. 3**E-F, I**). These data suggest that Ubx can bind to and mediate repression on the *RhoA* Exd/Hth/Hox configuration of sites in the context of the *DCRE*, but it cannot mediate activation in the context of the *RhoBAD* element (Fig. 2**G**). Consistent with this idea, EMSA analysis revealed Ubx binds the *RhoA* probe with Exd/Hth (Fig S4). In addition, point mutations within the *RhoA* Hox site that were previously shown to disrupt Abd-A mediated activation of the *RhoBAD* CRM (Li-Kroeger et al., 2008), resulted in a dramatic loss of abdominal repression activity in all abdominal segments of *GD-AR-LacZ*, including in the Ubx-specific A1 abdominal segment (Fig. 3**F-G, I**). In total, these data show that the *RhoA* Exd/Hth/Hox sites are not strictly Abd-A specific and Ubx can utilize this configuration of sites to mediate repression when placed into the *DCRE*.

### *3.3. The orientation and spacing of Hox sites relative to FoxG sites is critical for Abdominal Hox-mediated repression of the DCRE*

The findings that the *RhoA* Exd/Hth/Hox sites can mediate abdominal repression in Slp+ cells when inserted into the *DCRE* in only one direction (DCRE-AR) suggests that Hox binding site orientation relative to the FoxG sites may be a critical factor in conveying output. Consistent with this idea, the Hox/Exd sites in the *DCRE-AR* (but not in the *DCRE-AF*) are in a similar orientation and spacing relative to the FoxG sites as in the wild type *DCRE* (Fig. 3**B**). Moreover, we found that when the entire Hox/Exd-Hth/Hox sequence within the

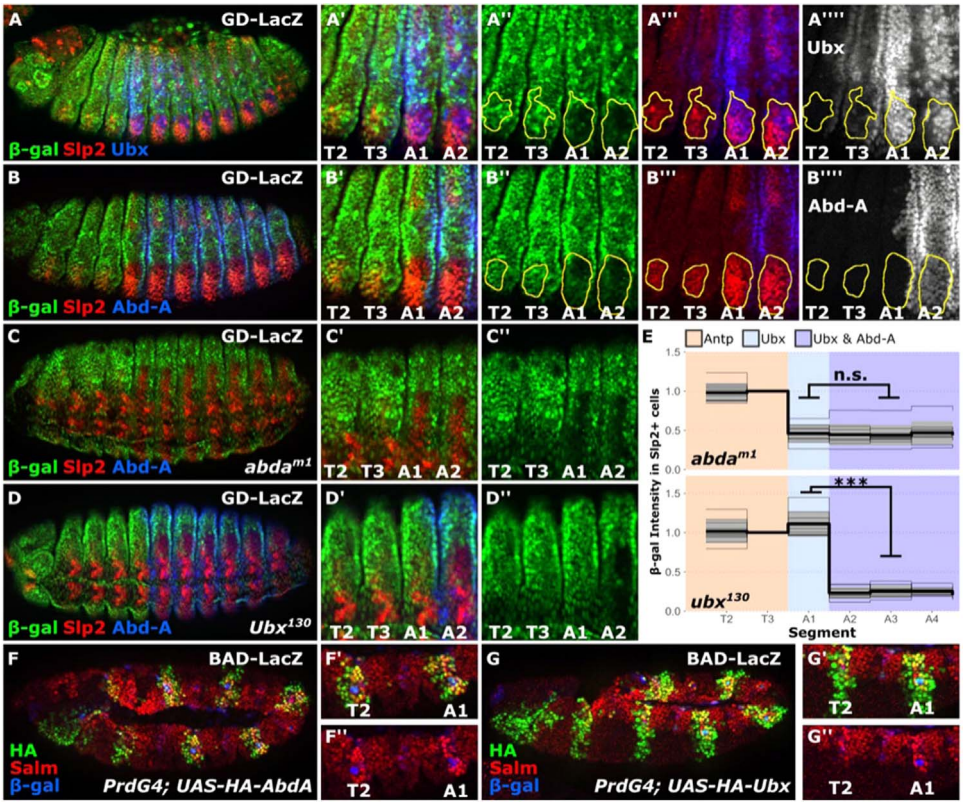


**Fig. 2. Abd-A and Ubx both repress the *DCRE* but only Abd-A activates *RhoA*.** **(A-B)** *3xGBE-DCRE-LacZ* (*GD-LacZ*) reporter embryo immunostained for β-gal (green), Slp2 (red), and Ubx (A) or Abd-A (B) (blue). **(A′-B′)** Close-up view of the T2-A2 segments shown in panels A-B. **(A′′-B′′)** Same as A′-B′ but only showing β-gal. Yellow outline indicates position of Slp2+ cells. **(A′′′-B′′′)** Same as A′′-B′′ showing only Slp2 stain with either Ubx (A′′′) or Abd-A (B′′′) stain. **(A′′′′-B′′′′)** Same as A′′-B′′ showing only Ubx (A′′′′) or Abd-A (B′′′′) stain only. Note, only Ubx is expressed in the Slp2+ cells in the A1 segment whereas both Ubx and Abd-A are expressed in the Slp2+ cells in the A2 through A4 segmeents. **(C-D)** *GD-LacZ* reporter embryos immunostained for β-gal (green), Slp2 (red), Abd-A (blue) in *abda* mutant (C) and *Ubx* mutant (D) embryos. **(E)** Quantification of β-gal immunostain intensity of *GD-LacZ* in *abd-A* and *Ubx* mutant backgrounds. Statistical comparison of β-gal intensity between A1 and mean A2-A4 abdominal segments was conducted using a *t*-test ("n.s." p > = 0.05, "*" p < 0.05, "**" p < 0.01, "***" p < 0.001). **(F-G)** *RhoBAD-lacZ;PrdG4* embryos (Stage 11) expressing either an HA-tagged AbdA (F) or HA-tagged Ubx (G) protein. Embryos were immunostained for β-gal (blue), HA (green), and Spalt-major (Salm, red). Note, Abd-A increases β-gal levels and induces additional Salm+ cells in the T2 segment whereas Ubx does not.

*DCRE* was "flipped" over in the opposite orientation (*DCRE-HF*), it placed the Hox/Hth site in a similar orientation and spacing relative to the FoxG site as the original Hox/Exd site and repressed Slp2+ abdominal gene expression as well as the wild type *DCRE* (Fig. 3**H-I**).

To further test the importance of spacing between the *DCRE* FoxG and Hox sites, we inserted short DNA sequences between these sites. Care was used to ensure the inserted sequences did not code for additional Hox, Exd, Hth, or FoxG binding sites (Fig. 4**A**). Five nucleotide intervals were used to systematically alter the DNA phasing of binding sites along the DNA helix (10 nucleotides = ~1 turn of the DNA helix). Intriguingly, we found that inserting +5 nucleotides (a half phase of the DNA helix) resulted in a complete loss of repression even though all three Hox sites and the FoxG sites are present (compare Fig. 4**B** with **4C**, quantified in **4G**). In contrast, inserting +10 nucleotides partially rescued abdominal repression in Slp2+ cells, and *GD-lacZ* reporters with +15 or +20 nucleotide insertions were also able to mediate repression as well or even better than the +10 spacer (Fig. 4**D-G**). Since the +5 bp insertion resulted in a complete loss of repression, we used EMSA analysis to compare DNA binding activity to wild type *DCRE*, *DCRE+5* and *DCRE+10* probes and found no significant difference in Abd-A and Exd/Hth binding (Fig S5). These findings suggest that the repression activity mediated by Hox factors is constrained to a specific spacing/orientation when in close proximity to the nearby FoxG sites within the *DCRE*, but when the distance between these sites is increased, binding site phasing is less critical in mediating transcriptional repression.

### *3.4. Altering the RhoA Hox and Pax2 site configuration disrupts gene activation in abdominal SOP cells*

Our studies revealed that the three Hox sites within the *DCRE* that are coupled to either a Exd or Hth site can recruit abdominal Hox complexes to mediate repression (Uhl et al., 2016). In contrast, the *RhoA* element contains a single contiguous set of Exd/Hth/Hox sites that mediate activation with a nearby Pax2 site (Fig. 3**A**). To determine if the Hox, Exd, and Hth sites from the *DCRE* can similarly mediate *RhoA* activation in abdominal SOP cells, we used a previously established transgenic reporter assay based on three copies of the *RhoA* element (*RhoAAA-lacZ*) being sufficient to mediate activation in abdominal SOP cells (Li-Kroeger et al., 2012; Witt et al., 2010) (Fig. 5**B**). To do so, we made *RhoAAA-DF* (forward) and *RhoAAA-DR* (reverse) constructs in which the Pax2 site was maintained, but the *RhoA* Exd/Hth/Hox sites were replaced with the *DCRE* Hox/Exd-Hth/Hox sites in the "forward" and "reverse" orientations (Fig. 5**A**). Analysis of embryos carrying these transgenes revealed neither *RhoAAA-DR-lacZ* nor *RhoAAA-DF-lacZ* were capable of activating transcription in abdominal SOPs (Fig. 5**C-D**). Given that inserting the *DCRE* Hox/Exd-Hth/Hox sites into *RhoA* alters the spacing between Hox and Pax2 sites, we next tested how spacing between the Pax2 and *RhoA* Exd/Hth/Hox sites affects Abd-A mediated activation. For this purpose, we inserted +5 or +10 nucleotide sequences between the Pax2 and Exd/Hth/Hox sites (Fig. 5**E**). Since *RhoA* also encodes an overlapping Senseless (Sens) binding site that can repress thoracic gene expression, care was taken to ensure that a low-affinity Sens site was maintained and that no new Pax2, Exd, Hth, or Hox sites were

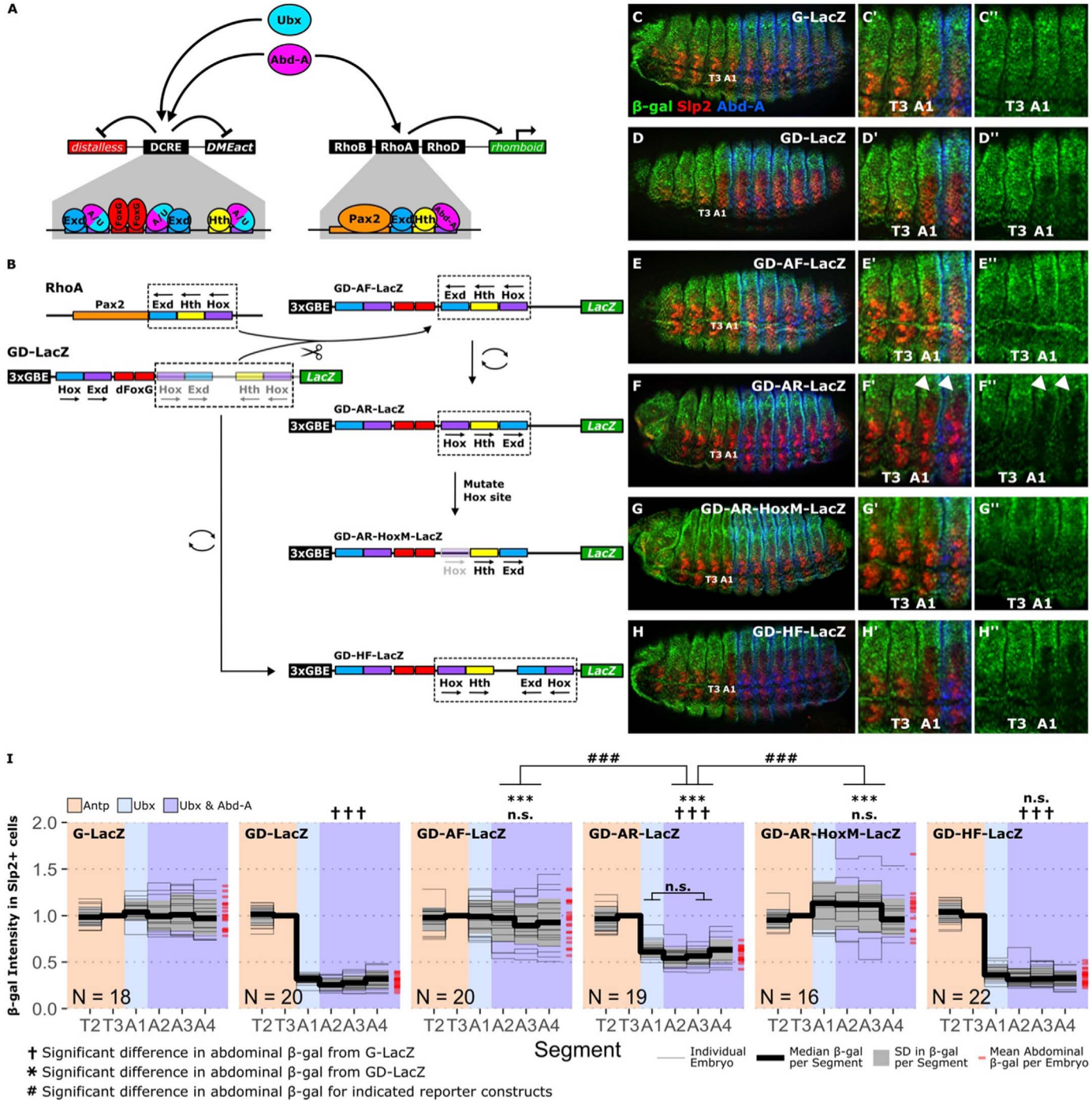


**Fig. 3. Hox, Exd, and Hth binding site arrangement does not define Abd-A vs Ubx specificity of RhoA and DCRE (A)** A schematic of Ubx and Abd-A action on the *DCRE* and *RhoA* elements. Note, both Ubx and Abd-A repress the *DCRE*, while only Abd-A can activate the *RhoA*. **(B)** Schematic demonstrating how the Hox/Exd-Hth/Hox binding sites in *DCRE* were modified in the *3xGBE-DCRE-LacZ* (GD-LacZ) constructs. For GD-AF the Hox/Exd-Hth/Hox binding sites in *DCRE* were replaced with the Exd/Hth/Hox binding sites from *RhoA*, whereas they were inserted in the reverse complement direction to create *GD-AR-LacZ*. The RhoA Hox binding site in *GD-AR-LacZ* was mutated to create *GD-AR-HoxM-LacZ*. The Hox/Exd-Hth/Hox binding sites from wildtype *DCRE* were placed in the reverse complement orientation to create the *GD-HF-LacZ* construct. **(C-H)** Lateral views of Stage 15 embryos expressing the indicated *LacZ* reporter constructs. Embryos were immunostained for β-gal (green), Slp2 (red), and Abd-A (blue). **(C′-H′)** Close-up view of T2-A2 segments from panels C-G. **(C′′-H′′)** Same as C′-G′ but only showing β-gal stain. **(I)** Quantification of β-gal immunostain intensity among Slp2+ cells in segments T2-A4 relative to intensity in T3 segment. Abdominal reporter expression was statistically compared between genotypes using ANOVA followed by post-hoc Tukey's test. Prior to ANOVA, abdominal reporter expression in each individual embryo was mean averaged. Crosses mark statistically significant difference from *G-LacZ* ("n.s." $p >= 0.05$, "†" $p < 0.05$, "††" $p < 0.01$, "†††" $p < 0.001$). Asterisks mark statistically significant difference from *GD-LacZ* ("n.s." $p >= 0.05$, "*" $p < 0.05$, "**" $p < 0.01$, "***" $p < 0.001$). Additional pairwise comparisons are indicated by bars and hash-signs ("n.s." $p >= 0.05$, "#" $p < 0.05$, "##" $p < 0.01$, "###" $p < 0.001$).

created within the *RhoA* element (Zandvakili et al., 2018). In fact, EMSA analysis revealed that adding the +5 or +10 sequences into the *RhoA* probe did not significantly alter Abd-A/Exd/Hth binding (Fig S6). Like the *DCRE*, we found that insertion of 5 bp sequences between the Pax2 and Exd/Hth/Hox sites resulted in a loss of Hox mediated activation (Fig. 5**G**). However, unlike the *DCRE*, the *RhoA* element did not regain activity when a full helical phase of DNA sequence (+10 bp) was inserted between these sites (Fig. 5**H**). These findings are

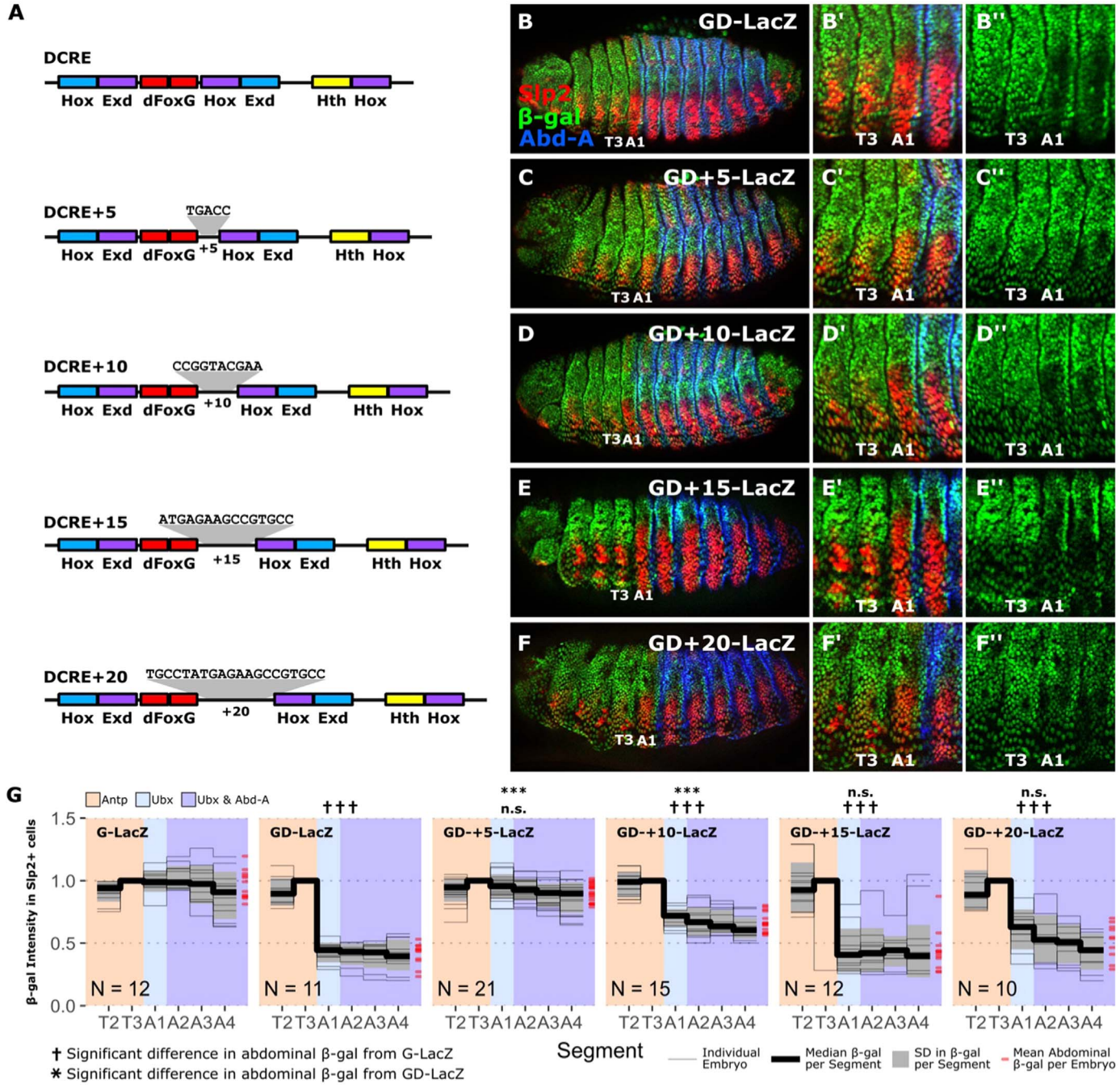


**Fig. 4. Spacing between Hox and dFoxG sites is critical for Hox-mediated repression of the *DCRE*.** **(A)** Schematics of *DCRE* variants with +5, +10, +15, or +20 bp inserted between the dFoxG and Hox binding sites. **(B-F)** Lateral views of Stage 15 embryos expressing the indicated *LacZ* reporter constructs. Embryos were immunostained for β-gal (green), Slp2 (red), and Abd-A (blue). **(B′-F′)** Close-up view of T2-A2 segments from panels B-F. **(B″-F″)** Same as B′-F′ but only showing β-gal stain**. (G)** Quantification of β-gal immunostain intensity *among* Slp2+ cells in segments T2-A4 relative to intensity in T3 segment of the indicated LacZ reporter embryos. Abdominal reporter expression was statistically compared between genotypes using ANOVA followed by post-hoc Tukey's test. Prior to ANOVA, abdominal reporter expression in each individual embryo was mean averaged. Crosses mark statistically significant difference from *G-LacZ* ("n.s." $p >= 0.05$, "†" $p < 0.05$, "††" $p < 0.01$, "†††" $p < 0.001$). Asterisks mark statistically significant difference from *GD-LacZ* ("n.s." $p >= 0.05$, "*" $p < 0.05$, "**" $p < 0.01$, "***" $p < 0.001$).

consistent with two interpretations: First, the Pax2 and Exd/Hth/Hox sites may be highly constrained in order to mediate transcriptional activation in abdominal SOPs. Second, the re-engineered *RhoA* elements disrupt additional unknown binding sites that are required for abdominal SOP activation. Currently, we favor the former as a previous scanning mutagenesis assay only uncovered the Pax2 binding site (Li-Kroeger et al., 2012). However, additional experiments are needed to conclusively distinguish between these two possibilities.

### *3.5. The orientation of the FoxG sites is critical for Abdominal Hox-mediated repression of the DCRE*

The ability to re-engineer functional *DCRE* repression elements using distinct Hox binding site configurations in the *GD-lacZ* assay also provides an opportunity to assess if altering the orientation of FoxG sites similarly affects the ability of the *DCRE* to mediate transcriptional repression. Previous studies identified two other experimentally confirmed CRMs that are repressed by the *Drosophila* FoxG factors (Slp1 and Slp2): an *even-skipped* (*eve*) enhancer in the early embryonic

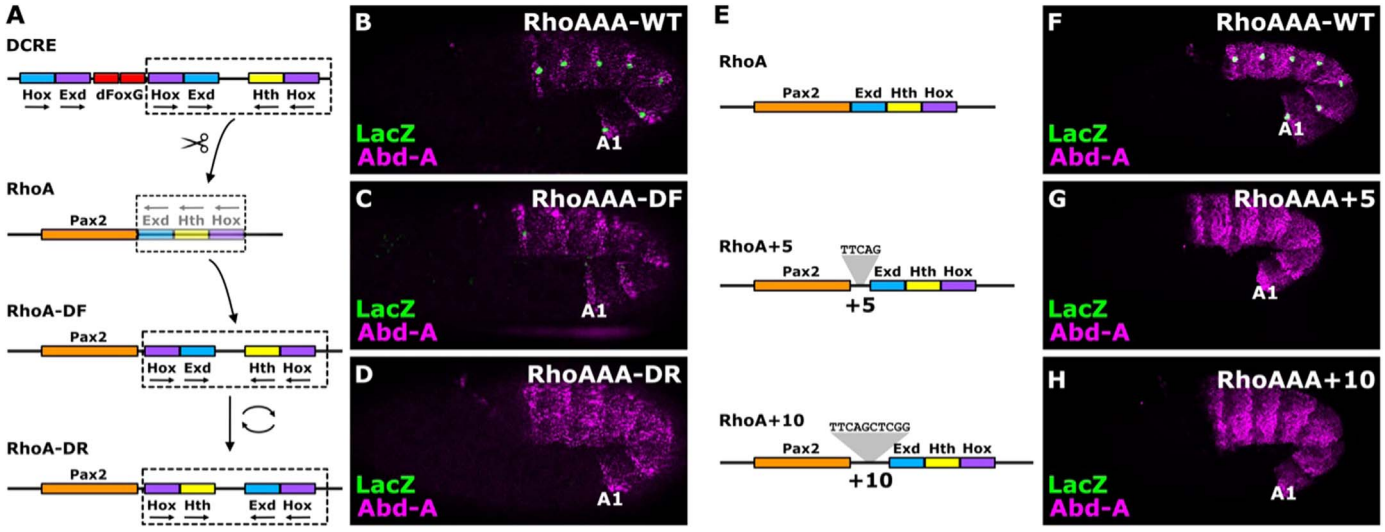


**Fig. 5. Altering the arrangement and spacing of the Hox, Exd, and Hth binding sites disrupts *RhoA* activity in abdominal SOP cells. (A)** Schematic showing how the Exd/Hth/Hox binding sites in *RhoA* were replaced with the Hox/Exd-Hth/Hox binding sites from the *DCRE* to create the *RhoA-DF* and *RhoA-DR* constructs. **(B-D)** Lateral view of Stage 10 embryos expressing the indicated *LacZ* reporter construct. Embryos were immunostained for β-gal (green), and Abd-A (magenta). **(E)** Schematic of *RhoA* variants with an additional 5 or 10 bp between the Pax2 and Exd-Hth-Hox binding sites. **(F-H)** Lateral view of Stage 10 embryos expressing the indicated *LacZ* reporter construct. Embryos were immunostained for β-gal (green) and Abd-A (magenta).

ectoderm and a *bagpipe* (*bap*) enhancer in the embryonic visceral mesoderm (Andrioli et al., 2004, 2002; Lee and Frasch, 2005). Unlike the *DCRE*, both the *eve* and *bap* enhancers are unlikely to directly integrate Hox inputs, as Slp represses *eve* prior to Hox gene expression and Slp represses the *bap* enhancer in all segments (Andrioli et al., 2004, 2002; Lee and Frasch, 2005). Sequence comparisons between the *DCRE*, *eve*, and *bap* elements reveals each contains at least two FoxG binding sites, but in distinct orientations. The *DCRE* FoxG sites are in a head-to-head (HH) orientation, the *eve* FoxG sites are in a head-to-tail (HT) orientation, and the *bap* FoxG sites are in a tail-to-tail (TT) orientation (Fig. 6**A**). To determine if the *eve* and *bap* FoxG binding sites can mediate repression in the *3xGBE* assay, we replaced the native *DCRE* FoxG sites with those from *bap* and *eve* (Fig. 6**A**). Comparative EMSAs using *DCRE* probes with either the *bap* or *eve* FoxG sites revealed purified Slp1 protein binds both sequences as well or better than the wild type FoxG sites (Fig. 6**B**). In contrast, point mutations within the wild type FoxG sites (SlpM) weakens Slp1 binding to the *DCRE* (Fig. 6**B**). Surprisingly, expression analysis of *GD-lacZ* embryos containing either the *eve* or *bap* FoxG sites revealed a significant loss of repression that was comparable to the *GD-SlpM-lacZ* embryos (Fig. 6**C-E** and **H**). Similar results were also seen when the FoxG sites from the *eve* enhancer were inserted in the reverse complement orientation of the *DCRE* relative to the Hox sites (EveRC, Fig. 6**F** and **H**). Importantly, comparative EMSA assays revealed that *DCRE* probes containing each FoxG binding site configuration retained the ability to bind the Abd-A/Exd/Hth complex, indicating that a loss of Hox binding cannot explain the loss of repression activity (Fig S7). Given that one of the differences between the FoxG sites in the *DCRE* versus the *eve* and *bap* enhancers is the orientation of binding sites, we re-engineered the *bap* FoxG sites into a Head-to-Head orientation within the *DCRE* (BapC) and found that it repressed as well as the wild type *DCRE* (Fig. 6**G-H**). In all cases, similar behaviors were observed in the A1 segment that only expresses Ubx as compared to abdominal segments that express both Ubx and Abd-A. Taken together, these findings indicate that the orientation and spacing of the Hox and FoxG sites are critical for mediating robust abdominal transcriptional repression by Ubx and Abd-A in Slp+ cells.

## 4. Discussion

In this study, we interrogated how the Hox, Exd, and Hth binding site configurations within the *DCRE* and *RhoA* regulatory elements contribute to specific transcriptional outcomes. Our results indicate that: A) The *DCRE* mediates transcriptional repression over a relatively short distance; B) The Exd/Hth/Hox configuration of binding sites from *RhoA*, which recruit Abd-A/Hth/Exd complexes to mediate activation in abdominal SOP cells, can also recruit Abd-A Hox repression complexes when placed into the context of the *DCRE* – with the outcome dependent upon the presence of nearby TFBSs for either the Pax2 (activation) or FoxG factors (repression). C) While the Ubx Hox factor does not regulate activation via the *RhoA* Exd/Hth/Hox sites in the context of the *RhoBAD* enhancer, a Ubx/Exd/Hth Hox complex can bind to and repress the *RhoA* configuration of binding sites when inserted into the *DCRE*; D) The orientation and spacing of FoxG binding sites relative to the Hox sites is critical to mediate transcriptional repression in Slp2+ abdominal cells. In total, these findings reveal new insights into the TFBS grammar underlying how CRMs yield cell- and segment-specific outputs by recruiting specific Hox transcriptional complexes.

### *4.1. cis-regulatory grammar: The role of orientation and spacing between TFBSs in mediating proper transcriptional output*

*cis*-regulatory modules (CRMs) integrate transcriptional inputs to mediate cell-specific output. Current models of CRM function include the billboard, enhanceosome, and transcription factor (TF) collective (Spitz and Furlong, 2012). The billboard model simply requires binding sites to be present within the CRM and their spacing/orientation/order has little effect on transcriptional outcomes (Arnosti and Kulkarni, 2005). Hence, transcription output is largely additive and the loss of any one binding site often has a modest impact on overall transcription levels and pattern. In contrast, the enhanceosome requires precisely spaced and oriented sites to mediate cooperative complex formation. Hence, the loss of any one site can disrupt both complex formation and synergistic output (Merika and Thanos, 2001). Lastly, the TF collective model similarly stipulates cooperative complex formation on DNA, but the arrangement of TFBSs needed to mediate cooperativity isn't highly constrained because protein-protein interactions between transcription factors can compensate for changes in DNA sequence (Junion et al., 2012). CRMs have been identified that are representative of each model; the billboard model by the *even-skipped* enhancers in *Drosophila* (Arnosti and Kulkarni, 2005), the

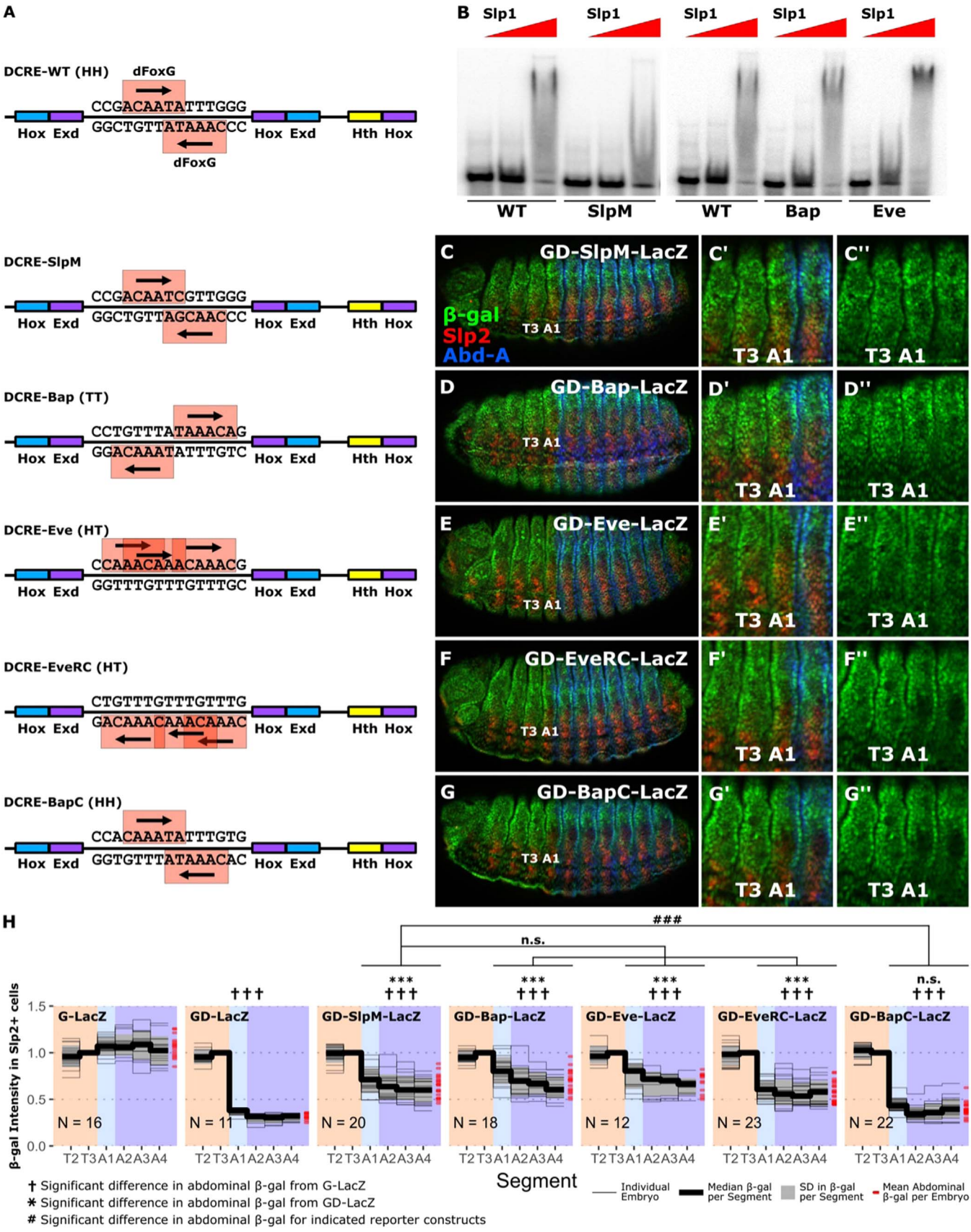
A
DCRE-WT (HH)
dFoxG
CCGACAATATTTGGG
GGCTGTTATAAACCC
Hox Exd Hox Exd Hth Hox
dFoxG
DCRE-SlpM
CCGACAATCGTTGGG
GGCTGTTAGCAACCC
Hox Exd Hox Exd Hth Hox
DCRE-Bap (TT)
CCTGTTTATAAACAG
GGACAAATATTTGTC
Hox Exd Hox Exd Hth Hox
DCRE-Eve (HT)
CCAAACAAACAAACG
GGTTTGTTTGTTTGC
Hox Exd Hox Exd Hth Hox
DCRE-EveRC (HT)
CTGTTTGTTTGTTTG
GACAAACAAACAAAC
Hox Exd Hox Exd Hth Hox
DCRE-BapC (HH)
CCACAAATATTTGTG
GGTGTTTATAAACAC
Hox Exd Hox Exd Hth Hox
B
Slp1
WT
SlpM
WT
Bap
Eve
C GD-SlpM-LacZ
β-gal
Slp2
Abd-A
T3 A1
C' C''
D GD-Bap-LacZ
D' D''
E GD-Eve-LacZ
E' E''
F GD-EveRC-LacZ
F' F''
G GD-BapC-LacZ
G' G''
H
###
n.s.
***
†††
n.s.
G-LacZ
GD-LacZ
GD-SlpM-LacZ
GD-Bap-LacZ
GD-Eve-LacZ
GD-EveRC-LacZ
GD-BapC-LacZ
N = 16
N = 11
N = 20
N = 18
N = 12
N = 23
N = 22
β-gal Intensity in Slp2+ cells
1.5
1.0
0.5
0.0
T2 T3 A1 A2 A3 A4
Segment
† Significant difference in abdominal β-gal from G-LacZ
* Significant difference in abdominal β-gal from GD-LacZ
# Significant difference in abdominal β-gal for indicated reporter constructs
Individual Embryo
Median β-gal per Segment
SD in β-gal per Segment
Mean Abdominal β-gal per Embryo

**Fig. 6. Orientation of dFoxG sites is critical for Hox-mediated repression of the *DCRE*.** **(A)** Schematics of *DCRE* variants: wildtype (*DCRE-WT*), weakened dFoxG sites (*DCRE-SlpM*), or differing dFoxG site compositions (*DCRE-Eve*, *DCRE-Bap*, *DCRE-EveRC, DCRE-BapC*). Orientations of the dFoxG binding sites are noted as: HH = head to head sites; TT = tail to tail sites; HT = head to tail sites. **(B)** EMSAs comparing two concentrations of Slp1 protein (50ngs and 250ngs) binding to *DCRE-WT* vs *DCRE-SlpM* as well as *DCRE-WT* vs *DCRE-Eve* and *DCRE-Bap*. **(C-G)** Lateral view of Stage 15 embryos expressing the indicated *LacZ* reporter constructs. Embryos were immunostained for β-gal (green), Slp2 (red), and Abd-A (blue). **(C′-G′)** Close-up view of T2-A2 segments from embryos in panels C-G. **(C′′-G′′)** Same as C′-G′ but only showing the β-gal stain. **(H)** Quantification of β-gal immunostain intensity among Slp2+ cells in segments T2-A4 relative to intensity in T3 segment of the indicated *LacZ* reporter embryos. Abdominal reporter expression was statistically compared between genotypes using ANOVA followed by post-hoc Tukey's test. Prior to ANOVA, abdominal reporter expression in each individual embryo was mean averaged. Crosses mark statistically significant difference from *G-LacZ* ("n.s." $p >= 0.05$, "†" $p < 0.05$, "††" $p < 0.01$, "†††" $p < 0.001$). Asterisks mark statistically significant difference from *GD-LacZ* ("n.s." $p >= 0.05$, "*" $p < 0.05$, "**" $p < 0.01$, "***" $p < 0.001$). Additional pairwise comparisons are indicated by bars and hash-signs ("n.s." $p >= 0.05$, "#" $p < 0.05$, "##" $p < 0.01$, "###" $p < 0.001$).

enhanceosome model by the interfernon-β enhancer in mammalian cells (Merika and Thanos, 2001), and the TF collective model by a set of *Drosophila* heart CRMs and neural enhancers in *C elegans* (Junion et al., 2012; Lloret-Fernandez et al., 2018). Of course, these models are not mutually exclusive and CRMs can have aspects consistent with the flexibility associated with the billboard/TF collective models and the constraints associated with the enhanceosome model (Swanson et al., 2010). Here, we highlight how our current understanding of the integration of transcriptional inputs by the *DCRE* and *RhoA* elements reveals new insight into how TFBS organization affects Hox-mediated outputs in light of these three CRM models.

Prior studies revealed the *DCRE* encodes two Hox/Exd sites and one Hox/Hth site that contribute to abdominal Hox mediated repression via a mechanism largely consistent with the TF collective model of CRM function; i.e. numerous protein-protein and protein-DNA contacts between Exd, Hth, abdominal Hox factors, and the *DCRE* ensure robust complex formation and transcriptional outcomes (Uhl et al., 2016). For example, point mutations in any one Hox, Exd, or Hth site had only a modest effect on both DNA binding and transcriptional repression. Moreover, a Hth isoform that completely lacks its DNA binding domain can still mediate cooperative protein-protein interactions to form abdominal Hox complexes that repress the *DCRE*. Altogether, these findings indicate that multiple distinct Hox TF complexes can yield functional repression (Noro et al., 2006). Consistent with these results, we found that the *RhoA* configuration of Exd/Hth/Hox sites can replace a set of Hox/Exd and Hth/Hox sites in the *DCRE* to mediate cooperative complex formation and robust transcriptional repression within the context of the *DCRE*. However, they do so in only one orientation relative to nearby FoxG binding sites. Moreover, we found that the spacing between the Hox and FoxG sites as well as the orientation of the FoxG sites contribute to transcriptional repression. Altogether, these data are congruent with the *DCRE* having aspects of both the TF collective and enhanceosome models of CRM function: while some different combinations of nearby Hox, Exd, Hth, and FoxG sites can mediate robust complex formation and appropriate outputs, not all combinations are functional indicating that the binding sites are at least somewhat constrained. This later point is also in agreement with the high degree of DCRE sequence and TFBS conservation observed between different *Drosophilid* species (Uhl et al., 2016).

Comparative studies on the highly conserved *RhoA* element similarly revealed constrained TFBS configurations as well as additional TFBS affinity and competition features that are required for accurate cell- and segment specific output. Five transcription factors have been shown to directly impact *RhoA* output: Pax2 and Exd/Hth/Hox sites promote activation whereas an overlapping binding site for the Sens transcription factor represses gene expression in thoracic segments (Li-Kroeger et al., 2008, 2012). We recently defined two key TFBS properties required for proper abdominal SOP output. First, the Pax2 and Sens binding sites are required to be low affinity, as a high affinity Pax2 site resulted in ectopic *RhoA* activity in additional SOP cells, whereas a high affinity Sens site resulted in repressed *RhoA* activity in all SOP cells (Zandvakili et al., 2018). Second, uncoupling the Sens and Pax2 sites so that they no longer overlap and compete for binding disrupted segment-specific *RhoA* output (Zandvakili et al., 2018). Both of these features provide insights into additional constraints on TFBS organization that have not been formally incorporated into the billboard, enhanceosome, or TF collective models. In this study, we further show that changing the spacing of the Pax2 site relative to the Exd/Hth/Hox sites disrupts activity in abdominal SOP cells; data that suggests enhanceosome-like activity. Like the DCRE, this idea is supported by the high degree of *RhoA* sequence conservation observed across numerous *Drosophilid* species in terms of both TFBS sequence and organization (Zandvakili et al., 2018). However, it should be pointed out that while Pax2 and Abd-A can interact in cells, cooperativity has not been detected between these factors on the *RhoA* element (Li-Kroeger et al., 2012). In addition, since all of the tested *RhoA* TFBS manipulations failed to reconstitute transcriptional activation in abdominal SOP cells, it is possible that the introduced sequence changes disrupt additional, unknown binding sites within the *RhoA* element required to mediate activation. Hence, future studies are needed to determine if the constraints on the Hox and Pax2 binding sites represent an enhancesome-like activity or are indicative of constraints consistent with other TFBS features such as overlapping and/or low affinity sites of additional transcription factor inputs.

### 4.2. Hox specificity: target activation vs target repression

While genome-wide ChIP assays define where a transcription factor binds, they do not determine if the binding event mediates a change in regulatory activity. Hence, a fundamental question is: What defines whether a specific Hox/Exd/Hth complex activates, represses, or fails to regulate transcription once bound to DNA? Our comparative studies of the *DCRE* and *RhoA* elements show that the presence of additional TFBSs is critical to mediate appropriate output. For example, the *RhoA* CRM contains both a set of Exd/Hth/Hox binding sites and a Pax2 binding site to mediate Abd-A specific gene activation in abdominal SOP cells. However, if the *RhoA* Exd/Hth/Hox sites are moved into the *DCRE*, they can mediate abdominal repression by not only Abd-A but also the Ubx Hox factor in a manner dependent upon nearby FoxG binding sites. Hence, the *RhoA* configuration of Exd/Hth/Hox sites is neither an intrinsically activating configuration of sites nor is it strictly an Abd-A specific configuration of binding sites. Instead, Hox regulatory specificity is encoded and dependent upon the presence of adjacent binding sites within the CRM.

Given that both Pax2 and the Slp FoxG factors are expressed in all *Drosophila* segments and not exclusively in the abdomen, these tissue-specific factors must be selectively integrated with specific abdominal Hox factors via nearby DNA binding sites on the RhoA and DCRE elements, respectively. Consistent with this idea, previous studies demonstrated that Abd-A and Pax2 could be co-immunoprecipitated in cell culture, whereas a thoracic Hox factor (Antennapedia, Antp) that fails to activate *RhoA* failed to form complexes with Pax2 (Li-Kroeger et al., 2012). Moreover, the vertebrate Hox11 proteins were also found to form complexes with Pax2 to regulate target gene expression in the mammalian kidney (Gong et al., 2007). While less is known about how Hox factors interact with Slp/FoxG, recent bimolecular-fluorescence (BiFC) assays in *Drosophila* found that Slp2 interacts with both Abd-A and Ubx in embryos, and Abd-A does so in a manner dependent on its ability to bind DNA (Baeza et al., 2015). Moreover, the thoracic Hox

factor, Antp, failed to interact with Slp2 in BiFC assays, and we previously found that instead of mediating repression Antp stimulates the *DMX* leg enhancer in a *DCRE*-dependent manner via unknown mechanisms (Uhl et al., 2016). Altogether, these findings are consistent with the Pax2 factor selectively working with an Abd-A Hox complex to mediate gene activation and FoxG/Slp factors working with Ubx and Abd-A on the *DCRE* to mediate abdominal repression.

## Funding

This work was supported by a National Institutes of Health Grant (GM079428) to BG and a National Instititutes of Health Grant (GM063483) to AZ. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ydbio.2018.11.006.

## References

Abe, N., et al., 2015. Deconvolving the recognition of DNA shape from sequence. Cell 161 (2), 307–318.
Affolter, M., Slattery, M., Mann, R.S., 2008. A lexicon for homeodomain-DNA recognition. Cell 133 (7), 1133–1135.
Agelopoulos, M., McKay, D.J., Mann, R.S., 2012. Developmental regulation of chromatin conformation by Hox proteins in Drosophila. Cell Rep. 1 (4), 350–359.
Andrioli, L.P., et al., 2002. Anterior repression of a Drosophila stripe enhancer requires three position-specific mechanisms. Development 129 (21), 4931–4940.
Andrioli, L.P., et al., 2004. Groucho-dependent repression by sloppy-paired 1 differentially positions anterior pair-rule stripes in the Drosophila embryo. Dev. Biol. 276 (2), 541–551.
Arnosti, D.N., Kulkarni, M.M., 2005. Transcriptional enhancers: intelligent enhanceosomes or flexible billboards? J. Cell. Biochem. 94 (5), 890–898.
Baeza, M., et al., 2015. Inhibitory activities of short linear motifs underlie Hox interactome specificity in vivo. Elife 4.
Berger, M.F., et al., 2008. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. Cell 133 (7), 1266–1276.
Bischof, J., et al., 2007. An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc. Natl. Acad. Sci. USA 104 (9), 3312–3317.
Brodu, V., Elstob, P.R., Gould, A.P., 2002. abdominal A specifies one cell type in Drosophila by regulating one principal target gene. Development 129 (12), 2957–2963.
Chan, S.K., et al., 1994. The DNA binding specificity of Ultrabithorax is modulated by cooperative interactions with extradenticle, another homeoprotein. Cell 78 (4), 603–615.
Chang, C.P., et al., 1996. Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. Mol. Cell. Biol. 16 (4), 1734–1745.
Chang, C.P., et al., 1997. Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins. Mol. Cell. Biol. 17 (10), 5679–5687.
Elstob, P.R., Brodu, V., Gould, A.P., 2001. spalt-dependent switching between two cell fates that are induced by the Drosophila EGF receptor. Development 128 (5), 723–732.
Gebelein, B., et al., 2002. Specificity of Distalless repression and limb primordia development by abdominal Hox proteins. Dev. Cell 3 (4), 487–498.
Gebelein, B., 2008. The control of EGF signaling and cell fate in the Drosophila abdomen. Fly 2 (5), 257–258.
Gebelein, B., McKay, D.J., Mann, R.S., 2004. Direct integration of Hox and segmentation gene inputs during Drosophila development. Nature 431 (7009), 653–659.
Gong, K.Q., et al., 2007. A Hox-Eya-Pax complex regulates early kidney developmental gene expression. Mol. Cell. Biol. 27 (21), 7661–7668.
Gutzwiller, L.M., et al., 2010. Proneural and abdominal Hox inputs synergize to promote sensory organ formation in the Drosophila abdomen. Dev. Biol. 348 (2), 231–243.
Hueber, S.D., et al., 2007. Comparative analysis of Hox downstream genes in Drosophila. Development 134 (2), 381–392.
Joshi, R., et al., 2007. Functional specificity of a Hox protein mediated by the recognition of minor groove structure. Cell 131 (3), 530–543.
Joshi, R., Sun, L., Mann, R., 2010. Dissecting the functional specificities of two Hox proteins. Genes Dev. 24 (14), 1533–1545.
Junion, G., et al., 2012. A transcription factor collective defines cardiac cell fate and reflects lineage history. Cell 148 (3), 473–486.
Kieffer-Kwon, K.R., et al., 2013. Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. Cell 155 (7), 1507–1520.
Lee, H.H., Frasch, M., 2005. Nuclear integration of positive Dpp signals, antagonistic Wg inputs and mesodermal competence factors during Drosophila visceral mesoderm induction. Development 132 (6), 1429–1442.
Li-Kroeger, D., et al., 2008. Hox and senseless antagonism functions as a molecular switch to regulate EGF secretion in the Drosophila PNS. Dev. Cell 15 (2), 298–308.
Li-Kroeger, D., Cook, T.A., Gebelein, B., 2012. Integration of an abdominal Hox complex with Pax2 yields cell-specific EGF secretion from Drosophila sensory precursor cells. Development 139 (9), 1611–1619.
Lloret-Fernandez, C., et al., 2018. A transcription factor collective defines the HSN serotonergic neuron regulatory landscape. Elife 7.
Mann, R.S., Affolter, M., 1998. Hox proteins meet more partners. Curr. Opin. Genet. Dev. 8 (4), 423–429.
Mann, R.S., Chan, S.K., 1996. Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. Trends Genet. 12 (7), 258–262.
Mann, R.S., Lelli, K.M., Joshi, R., 2009. Hox specificity unique roles for cofactors and collaborators. Curr. Top. Dev. Biol. 88, 63–101.
McGinnis, W., Krumlauf, R., 1992. Homeobox genes and axial patterning. Cell 68 (2), 283–302.
Merabet, S., Mann, R.S., 2016. To be specific or not: the critical relationship between Hox and TALE proteins. Trends Genet. 32 (6), 334–347.
Merika, M., Thanos, D., 2001. Enhanceosomes. Curr. Opin. Genet. Dev. 11 (2), 205–208.
Moens, C.B., Selleri, L., 2006. Hox cofactors in vertebrate development. Dev. Biol. 291 (2), 193–206.
Noro, B., et al., 2006. Distinct functions of homeodomain-containing and homeodomain-less isoforms encoded by homothorax. Genes Dev. 20 (12), 1636–1650.
Noyes, M.B., et al., 2008. Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. Cell 133 (7), 1277–1289.
Ortiz-Lombardia, M., et al., 2017. Hox functional diversity: novel insights from flexible motif folding and plastic protein interaction. Bioessays 39 (4).
Pearson, J.C., Lemons, D., McGinnis, W., 2005. Modulating Hox gene functions during animal body patterning. Nat. Rev. Genet. 6 (12), 893–904.
Prasad, N., et al., 2016. A comparative genomic analysis of targets of Hox protein Ultrabithorax amongst distant insect species. Sci. Rep. 6, 27885.
Rusten, T.E., et al., 2001. Spalt modifies EGFR-mediated induction of chordotonal precursors in the embryonic PNS of Drosophila promoting the development of oenocytes. Development 128 (5), 711–722.
Ryoo, H.D., Mann, R.S., 1999. The control of trunk Hox specificity and activity by Extradenticle. Genes Dev. 13 (13), 1704–1716.
Slattery, M., et al., 2011a. Genome-wide tissue-specific occupancy of the Hox protein Ultrabithorax and Hox cofactor Homothorax in Drosophila. PLoS One 6 (4), e14686.
Slattery, M., et al., 2011b. Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. Cell 147 (6), 1270–1282.
Spitz, F., Furlong, E.E., 2012. Transcription factors: from enhancer binding to developmental control. Nat. Rev. Genet. 13 (9), 613–626.
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 (3), 359–370.
Uhl, J.D., Cook, T.A., Gebelein, B., 2010. Comparing anterior and posterior Hox complex formation reveals guidelines for predicting cis-regulatory elements. Dev. Biol. 343 (1–2), 154–166.
Uhl, J.D., Zandvakili, A., Gebelein, B., 2016. A Hox transcription factor collective binds a highly conserved distal-less cis-regulatory module to generate Robust transcriptional outcomes. PLoS Genet. 12 (4), e1005981.
Uv, A.E., Harrison, E.J., Bray, S.J., 1997. Tissue-specific splicing and functions of the Drosophila transcription factor Grainyhead. Mol. Cell Biol. 17 (11), 6727–6735.
Vachon, G., et al., 1992. Homeotic genes of the Bithorax complex repress limb development in the abdomen of the Drosophila embryo through the target gene Distal-less. Cell 71 (3), 437–450.
Wang, G., et al., 2017. A Hox complex activates and potentiates the Epidermal growth factor signaling pathway to specify Drosophila oenocytes. PLoS Genet. 13 (7), e1006910.
Weirauch, M.T., et al., 2014. Determination and inference of eukaryotic transcription factor sequence specificity. Cell 158 (6), 1431–1443.
Witt, L.M., et al., 2010. Atonal, senseless, and abdominal-A regulate rhomboid enhancer activity in abdominal sensory organ precursors. Dev. Biol. 344, 1060–1070.
Xie, B., et al., 2007. Senseless functions as a molecular switch for color photoreceptor differentiation in Drosophila. Development 134 (23), 4243–4253.
Zandvakili, A., et al., 2018. Degenerate Pax2 and Senseless binding motifs improve detection of low-affinity sites required for enhancer specificity. PLoS Genet. 14 (4), e1007289.
Zandvakili, A., Gebelein, B., 2016. Mechanisms of specificity for hox factor activity. J. Dev. Biol. 4 (2).