Low Affinity Binding Sites in an Activating CRM Mediate Negative Autoregulation of the *Drosophila* Hox Gene *Ultrabithorax*

Rebecca K Delker[1¶], Vikram Ranade[3¶], Ryan Loker[3], Roumen Voutev[1], Richard S Mann[1,2*]

1. Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY, USA

2. Department of Neuroscience, Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA

3. Department of Genetics, Columbia University, New York, NY, USA

* Corresponding author

Email: rsm10@columbia.edu

¶ These authors contributed equally to this work

Abstract

Specification of cell identity and the proper functioning of a mature cell depend on precise regulation of gene expression. Both binary ON/OFF regulation of transcription, as well as more fine-tuned control of transcription levels in the ON state, are required to define cell types. The *Drosophila melanogaster* Hox gene, *Ultrabithorax* (*Ubx*), exhibits both of these modes of control during development. While ON/OFF regulation is needed to specify the fate of the developing wing (*Ubx* OFF) and haltere (*Ubx* ON), the levels of *Ubx* within the haltere differ between compartments along the proximal-distal axis. Here, we identify and molecularly dissect the novel contribution of a previously identified *Ubx* cis-regulatory module (CRM), *anterobithorax* (*abx*), to a negative auto-regulatory loop that maintains decreased *Ubx* expression in the proximal compartment of the haltere as compared to the distal compartment. We find that Ubx, in complex with the known Hox cofactors, Homothorax (Hth) and Extradenticle (Exd), acts through low-affinity Ubx-Exd binding sites to reduce the levels of *Ubx* transcription in the proximal compartment. Importantly, we also reveal that Ubx-Exd-binding site mutations sufficient to result in de-repression of *abx* activity in the proximal haltere in a transgenic context are not sufficient to de-repress *Ubx* expression when mutated at the endogenous locus, suggesting the presence of multiple mechanisms through which Ubx-mediated repression occurs. Our results underscore the complementary nature of CRM analysis through transgenic reporter assays and genome modification of the endogenous locus; but, they also highlight the increasing need to understand gene regulation within the native context to capture the potential input of multiple genomic elements on gene control.

41

42 ***Author Summary***

43

44 One of the most fundamental questions in biology is how information encoded in
45 the DNA is translated into the diversity of cell-types that exist within a multicellular
46 organism, each with the same genome. Regulation at the transcriptional level, mediated
47 through the activity of transcription factors bound to *cis*-regulatory modules (CRMs),
48 plays a key role in this process. While we typically distinguish cell-type by the specific
49 subset of genes that are transcriptionally ON or OFF, it is also important to consider the
50 more fine-tuned transcriptional control of gene expression level. We focus on the
51 regulatory logic of the Hox developmental regulator, *Ultrabithorax* (*Ubx*), in fruit flies,
52 which exhibits both forms of transcriptional control. While ON/OFF control of *Ubx* is
53 required to define differential appendage fate in the T2 and T3 thoracic segments,
54 respectively, more fine-tuned control of transcription levels is observed in distinct
55 compartments within the T3 appendage, itself, in which all cells exhibit a *Ubx* ON state.
56 Through genetic analysis of regulatory inputs, and dissection of a *Ubx* CRM in a
57 transgenic context and at the endogenous locus, we reveal a compartment-specific
58 negative autoregulatory loop that dampens *Ubx* transcription to maintain distinct
59 transcriptional levels within a single developing tissue.

60

61 ***Introduction***

62

Though often described as the blueprint of an organism, the genome – and its underlying code – remains undecipherable unless understood within a cellular context. This is because, in addition to genomic sequence, knowledge of the subset of genes expressed throughout space and time is required. In eukaryotic cells, this transcriptional regulation is governed by the presence of non-coding cis-regulatory modules (CRMs) through which the binding of transcription factors (TFs) can either positively or negatively affect the expression of target genes. While it is common to think of cell identity as a product of a binary ON vs. OFF control of transcription, an added layer of complexity is gained by recognizing that the quantitative amount (the level/dose) of gene expression of both TFs and their downstream targets can also impact cell identity and/or function (1-3). Here, studying the Hox gene, *Ultrabithorax* (*Ubx*), in *Drosophila melanogaster* (*D. melanogaster*), we explore the question of how gene expression can be tuned once in an ON transcriptional state.

Ultrabithorax (*Ubx*), required during development at both embryonic and larval stages, is likely best known for its role in the differential development of the second and third thoracic dorsal appendages, the wing (T2) and haltere (T3), respectively (4). These serially homologous structures are derived from larval imaginal discs that differ primarily in the expression of *Ubx*: a *Ubx* OFF state is required for wing development, while a *Ubx* ON state is required for haltere development (5-8). While this demonstrates the importance of an ON/OFF binary mode of control for Ubx expression during development, it is also important to note that, although Ubx is detected in all cells of the haltere disc, distinct levels of *Ubx* expression occur along the proximal distal axis

(Figure 1A, B). While distal cells exhibit high levels of Ubx, proximal cells exhibit low levels of Ubx, resulting in a distal/proximal ratio > 1. This observation raises questions not only of the downstream effects of high versus low *Ubx* expression on cell- and tissue-fate, but also of the mechanisms by which Ubx levels are regulated in proximal versus distal compartments. While the first question remains unanswered, here, we address the latter question and provide evidence for the presence of an autoregulatory mechanism in which Ubx directly represses its own expression within the proximal compartment of the developing haltere.

The regulation of *Ubx*, like many developmental regulators that need to be expressed in very precise spatio-temporal patterns, relies on input from a number of CRMs throughout its large, ~120 kb locus (7,9-15). We find a novel role for a previously identified *Ubx* CRM, termed *anterobithorax* (*abx*), located within the large third intron of *Ubx*. Established as an enhancer necessary for activation of *Ubx* within the anterior compartment of the haltere ((9,11) and Figure 3A, B), we find that *abx* also serves a crucial role in the negative autoregulation of *Ubx* that results in distinct levels of expression in the proximal versus distal compartments of the haltere disc. Thus, interestingly, we find that a single CRM is utilized for both activating and repressive functions within the same tissue.

While this is not the first observation of *Ubx* autoregulation (13,16,17), it does offer additional insight into our understanding of autoregulatory mechanisms, as well as the potential utility of autoregulation. Previously characterized examples of *Ubx* negative

autoregulation highlight the importance of autoregulation to buffer against increases in Ubx protein level, either from ectopic pulses of transgenic Ubx or from more subtle increases in *Ubx* copy number via chromosomal aberrations containing duplications of the *Ubx* locus (13,16). In each of these scenarios, an increase in Ubx protein resulted in repression of the endogenous locus or repression of CRMs (as readout by Gal4 insertions) within the locus, respectively.

In contrast, we present a role for negative autoregulation in establishing distinct levels of *Ubx* between compartments within a single tissue during normal development. Additionally, we molecularly dissect this autoregulatory mechanism in both a transgenic context, where we can isolate the contribution of a small cluster of low affinity Ubx-Exd binding sites to decrease proximal Ubx levels, and in an endogenous context, where multiple binding sites and mechanisms are likely required to achieve autoregulation. Our results underscore the complementary nature of transgenic and endogenous modification approaches to understanding gene regulation, and highlight the increasing importance of interrogating this problem within the complexity of the native locus enabled by new technologies like CRISPR (18).

Results

Proximal/Distal (PD) Expression Bias of Ubx is Established at the Level of Transcription

131 The importance of *Ubx* expression for haltere fate was established decades ago
132 by the discovery of mutations that resulted in a loss of *Ubx* function in the developing
133 haltere and the concomitant production of four-winged flies (15,19,20). Here,
134 misregulation of the ON/OFF *Ubx* expression switch results in a complete homeotic
135 transformation of the haltere to the wing. Similarly, gain of function of *Ubx* in the
136 developing wing results in a near complete transformation of wing to haltere (8,21). In
137 addition, more subtle changes in *Ubx* expression level have also been shown to result
138 in malformations of the adult appendage: decreases in *Ubx* dose lead to a slightly
139 enlarged haltere (16) and increases in *Ubx* dose result in decreased adult haltere size –
140 though size is buffered to a certain extent against changes in *Ubx* dose (16,22). While
141 this establishes the importance of maintaining appropriate levels of *Ubx* expression in a
142 given tissue context, and the presence of autoregulatory responses to genetic
143 perturbation to maintain levels within the appropriate window, it does not address
144 mechanisms present during normal development to establish differential expression
145 levels between compartments in a single tissue.
146
147 In third instar larval haltere imaginal discs, *Ubx* is expressed at higher levels in
148 the distal region than in the proximal region (Figure 1A, B) offering a system to probe
149 how different expression levels are established. This differential *Ubx* expression pattern
150 can be seen at both the level of total cellular protein (Figure 1B) – quantified to reveal
151 an average distal/proximal ratio of 1.67 (Figure 1B', Figure 1 Supplemental Figure 1) –
152 and at the level of nascent RNA production (Figure 1C), providing evidence that the
153 proximal-distal bias is established, at least in part, due to transcriptional control. These

findings are additionally corroborated by the presence of a proximal-distal bias in the activity of several Gal4 enhancer-trap lines, including *Ubx-lacZ*^{HC166D} (Figure 1D, left) (23), which presumably captures input from multiple CRMs required for the establishment of the observed proximal-distal bias. Interestingly, RNAi-mediated clonal reduction of proximal Ubx elevates levels of LacZ driven by *Ubx-lacZ*^{HC166D} to distal levels (Figure 1D, right). Taken together, these results suggest that the observed proximal-distal bias is established through the active repression of proximal Ubx levels at the level of transcriptional control and that this occurs, either directly or indirectly, through *Ubx* activity.

Proximal autorepression of Ubx relies on Hox cofactors Hth/Exd

The establishment of the PD bias through proximally-restricted repression suggests the existence of a factor that is similarly restricted to the proximal compartment. We therefore investigated the potential contribution of the well-established Hox cofactors, Homothorax (Hth) and Extradenticle (Exd). *hth* expression is restricted to the proximal compartment, with the exception of expression in the dorsal, distal hinge (Figure 2A, asterisk), thus largely mirroring the pattern of low *Ubx* (Figure 2A). To test if Hth represses *Ubx* transcription in the proximal haltere, we generated *hth* null (*hth*^{P2}) clones and performed immunostaining against endogenous Ubx protein. In the absence of Hth, we observed a de-repression of proximal Ubx, indicating the involvement of Hth in the downregulation of *Ubx* in the proximal haltere (Figure 2B).

Given the involvement of Hth, we next tested the role of *exd* – which encodes an obligate binding partner of Hth. Hth interacts directly with Exd and is required to translocate Exd from the cytoplasm of the nucleus (24). In the distal haltere disc, where Hth is absent, Exd is cytoplasmic and seemingly non-functional. To test if Exd works along with Hth to repress *Ubx* transcription in the proximal haltere, we generated *exd* null clones with two different alleles (exd^{1} and exd^{2}). As we saw for hth^{P2} clones, loss of Exd in the proximal (but not distal) haltere resulted in de-repression of *Ubx* (Figure 2D, Figure 2 Supplemental Figure 1A). This result contrasts with a previous report showing loss of Ubx protein in exd^{2} mutant clones, implying the positive regulation of *Ubx* by Exd (25). To resolve this discrepancy, we further tested the role of Exd by knocking it down in clones expressing *exd* RNAi. In these clones, and in agreement with our *exd* null clones, proximal Ubx levels are elevated (Figure 2 Supplemental 1B).

Three known isoforms of Hth exist due to alternative splicing of the locus: a full-length isoform that contains a homeodomain and is thus able to bind DNA (Hth^{FL}) and two homeodomain-less isoforms that cannot directly bind DNA and are thus reliant on complex formation for recruitment to DNA (collectively referred to as Hth^{HM}) (26). To test if Hth requires its homeodomain for proximal *Ubx* repression, we generated clones of the $hth^{100\text{-}1}$ allele, which only produces Hth^{HM} due to a premature stop codon before the homeodomain (27). In these clones, and thus in the absence of Hth^{FL}, Ubx levels are not elevated and the PD bias remains intact (Figure 2C). Thus, although Hth is necessary for proximal *Ubx* repression, direct binding of Hth to DNA is dispensable; HD-less isoforms of Hth are sufficient for the bias in Ubx levels along the PD axis.

Given the evidence for the repressive effects of Hth on *Ubx* expression, we sought to understand if there was mutual repression between the two factors. We generated *Ubx* null (Ubx^{9-22}) mitotic clones and performed immunostaining against Hth protein. In the absence of Ubx, Hth protein levels were unaffected suggesting that, while Hth represses *Ubx*, Ubx does not repress *hth* (Figure 2F). These results were confirmed by immunostaining for Exd protein in *Ubx* null clones (Figure 2E). As Hth is required for the translocation of Exd to the nucleus, the subcellular localization of Exd can serve as a readout for the presence or absence of Hth. In the presence of Hth, Exd is nuclear, but if Hth is lost upon induction of *Ubx* null clones, Exd should become cytosolic. However, contrary to this prediction, the levels of Exd and subcellular localization of Exd are unaffected in *Ubx* null clones (Figure 2E).

A Ubx*-regulated Module within the Intronic abx CRM Mediates autorepression*

Previous work on the *cis*-regulatory logic of *Ubx* identified two large spatially distinct genomic regulatory regions: a region upstream of the TSS (Upstream Control Region – UCR) that regulates *Ubx* in parasegment 6 (PS6) and the posterior haltere, and a region within the large third intron of *Ubx* (Downstream Control Region – DCR) that regulates *Ubx* in PS5 and anterior haltere (11,14). While a 35 kb UCR fragment (35UZ) drives reporter expression in the posterior haltere, various fragments derived from the DCR, termed the *anterobithorax* (*abx*) CRM, drive reporter expression throughout the haltere that recapitulates *Ubx* expression, including the presence of a

PD bias (11,13). In addition, classically defined alleles of the *Ubx* locus affecting anterior expression in the haltere are located within the *abx* region (12)(7). Smaller fragments of *abx* of 6.8 kb (*abx6.8*) and 3.2 kb (*abx3.2*) were sufficient to recapitulate the *Ubx* expression pattern, including the PD bias, in the haltere (Figure 3A) (11). These fragments also drive reporter expression in other imaginal discs where *Ubx* expression is absent, presumably due to the absence of a necessary repressive element, such as a Polycomb Response Element (PRE) (Figure 3C).

That the *abx* region functions as a *cis*-regulatory module is further corroborated by published FAIRE-seq data that shows a broad region of open chromatin particularly within the domains of the two small *abx* fragments, *abx6.8* and *abx3.2* – hereafter called *abxFAIRE* (Figure 3A). Not only is this region specifically accessible in the haltere as compared to the wing (Figure 3A, track 1 versus track 2), but it also displays both Ubx and Hth binding, as assayed by previous ChIP-chip experiments (Figure 3A, track 3 and 4) (28,29). Consistent with these earlier results, clonal deletion of *abxFAIRE* results in a decrease in Ubx levels – ranging from a slight reduction to complete loss – in both proximal and distal cells throughout the anterior of the haltere disc (Figure 3B, Figure 3 Supplemental 1B, B') (9). Despite this activating activity, *abx*-driven reporter expression maintains the PD bias, suggesting that this element may also contain sequences required for down-regulation of *Ubx* in the proximal haltere.

Within *abxFAIRE* we focused on a ~1.4 kb fragment (*abxF*) that included the highest enrichment for both Ubx and Hth binding in ChIP experiments (Figure 3A). As

246 expected based on prior reports, *abxF-lacZ* reporter constructs recapitulate *Ubx*
247 expression throughout the haltere, including the PD bias (Figure 3C) (11). Similar to
248 native *Ubx* levels, reporter expression driven by *abxF* is elevated in the absence of
249 either Hth or Exd, as determined by examining mitotic clones null for each gene (Figure
250 3D, 3E). Further, as with native *Ubx* levels, proximal repression of *abxF* activity does
251 not require the full-length isoform of Hth (Figure 3F), confirming that direct DNA binding
252 of Hth is unnecessary for the PD bias. Finally, *Ubx* null mitotic clones ($Ubx^{9\text{-}22}$) similarly
253 results in a de-repression of proximal *abxF* activity to levels comparable to that in the
254 distal pouch (Figure 3G). Taken together, these results confirm that the *abxF* fragment
255 is able to capture the input of Ubx-Hth-Exd in proximal repression and the concomitant
256 establishment of the PD bias, providing evidence for a dual role for *abx* in mediating
257 both activation (distally and proximally) and partial repression (proximally). Thus, *abxF*
258 provides a platform to more finely dissect the mechanism of auto-repression by Ubx and
259 its cofactors.
260
261 *Ubx directly binds abx to auto-repress in the proximal compartment*
262
263 Given that the *abxF* CRM recapitulates *Ubx* expression bias along the PD axis,
264 and that ChIP-chip results indicate that Ubx is bound at this location, we attempted to
265 identify a Ubx binding site in *abxF* (28). As a first step, we minimized the 1.4 kb *abxF*
266 fragment and identified a 531 bp minimal autoregulatory CRM, *abxN*, which
267 recapitulates the expression pattern of the larger *abxF* fragment, and *Ubx* itself (Figure
268 4A) when driving a *lacZ* reporter. Using the recently developed versatile maximum

likelihood framework, *No Read Left Behind* (*NRLB*), to predict Ubx-Exd dimer binding, we identified two clusters of low affinity Ubx-Exd binding sites (Relative Affinity > 10^{-3}, Figure 4B) (30). Mutation of each binding site cluster to abrogate Ubx binding (Figure 4B') followed by immunostaining of the *lacZ* reporter revealed a decrease in PD bias, though this was more severe following mutation of Cluster 1 (Figure 4C, quantified in 4C'). In fact, mutation of both clusters together exhibited a PD bias defect comparable to that of Cluster 1 alone (Figure 4C, quantified in 4C'). Using both *in vitro* and *in vivo* techniques, we next established that Ubx can bind directly to *abxN* together with its cofactors. *In vitro* EMSAs revealed the binding of Ubx-Exd-Hth^{HM} to a probe containing the Cluster 1 binding sites. Importantly, this binding was lost upon mutation of these sites (Figure 4 Supplemental 1B); and, *in vivo* ChIP-qPCR was used to demonstrate Ubx binding to both the endogenous *abxN* CRM (recapitulating published ChIP-chip data) and the transgenic CRM-*lacZ* reporter gene (28). Again, mutation of Cluster 1 binding sites resulted in decreased occupancy of Ubx at transgenic *abxN* (Figure 4 Supplemental 1C). Further, *Ubx* null clones generated in the background of a Cluster 1-mutated *abxN-lacZ* transgene revealed no further de-repression (Figure 4 Supplement 1A), providing additional evidence that the repressive activity of Ubx occurs through direct binding of *abxN* via Cluster 1 binding sites.

It is important to note that while it is clear that this reduction in PD bias results in part from an increase in proximal reporter expression, a decrease in distal reporter expression, particularly within the distal hinge region, is also apparent as a consequence of mutating Cluster 1.These results could suggest an additional

292 autoregulatory mechanism within the distal compartment that functions in an inverse
293 manner; however, this was not supported by any of our other genetic analyses, and
294 thus further experiments would be necessary to support this conclusion.
295
296 All of the predicted binding sites within *abxN*, including those found within Cluster
297 1, are low affinity binding sites, defined as less than 0.1% of the maximal binding site
298 affinity in the genome (31). Interestingly, when we mutate Cluster 1 to contain a high
299 affinity binding site as determined by *in vitro* SELEX experiments and *NRLB* predictions
300 (Figure 4B''), levels of the *lacZ* reporter within the proximal compartment are decreased
301 and the PD bias is increased accordingly (Figure 4C and 4C') (31). Not only does this
302 result support the autoregulatory role of Ubx in controlling the PD bias – increasing the
303 binding affinity results in greater proximal repression – but it also suggests the
304 functional importance of low affinity binding sites (32,33). In this case, low affinity
305 binding sites may allow for a more tunable state of transcription that is responsive to
306 changes in TF levels. Here, abx low affinity binding sites appear to be optimized to
307 produce the correct amount of *Ubx* repression.
308
309 *Mutation of abxN at the native locus is not sufficient to alter the Proximal-Distal bias*
310
311 Having established the proximally-restricted repressive role of Ubx on *abxN* to
312 form the PD bias in a transgenic context, we sought to understand this phenomenon at
313 the endogenous locus. To this end we devised a two-step strategy that utilizes
314 CRISPR/Cas9 genome engineering along with PhiC31 based recombinase mediated

cassette exchange (PhiC31 RMCE) to replace the native, wildtype *abx* CRM with Ubx-binding site mutant varieties (Figure 5B). Two replacements were performed: one that encompasses the entirety of *abxFAIRE,* including *abxF* and *abxN* (*Targeted Region*^{4kb}, ~4 kb), and a second that encompasses *abxN* (*Targeted Region*^{2kb}, ~2.2 kb, Figure 5A). These "*abx*-replacement platforms," verified by both PCR and Southern Blot analysis, were then used to re-insert *abx* sequences (via RMCE) with and without the desired Ubx-Exd binding site mutants (Figure 5B). Because this is not a scarless editing strategy, it was important to confirm that the sites remaining as a product of attP/attB recombination did not affect *Ubx* expression. To test this, we used RMCE to reinsert the wildtype sequence of each targeted region (*Targeted Region*^{2kb} in Figure 5 and Figure 5 Supplemental 1; *Targeted Region*^{4kb} in Figure 6 and Figure 6 Supplemental 1) such that the only difference between the engineered allele and the native allele is the presence of the recombinase sites. The wildtype *abx* replacement alleles were able to homozygose, and homozygous flies developed normally with no noticeable defect. In addition, the generation of clones that are homozygous for the wildtype *abx* replacement alleles showed no differences in *Ubx* expression compared to neighboring wild type cells, as assayed at the protein level by immunostaining (Figure 5 Supplemental 1A, top panel and Figure 6 Supplemental 1B, top panel).

Having established that these RMCE platforms do not affect *Ubx* expression, we next sought to understand the effect of Ubx binding site mutants within the context of the *Ubx* locus on proximal repression and the formation of the PD bias. Because the mutation of Cluster 1 binding sites showed a greater reduction in the PD bias in our

transgenic studies, we used our 2 kb replacement platform (Targeted Region^{2kb}) to generate *abxN* replacement alleles containing mutations to either abrogate Ubx-Exd binding (*abxN-Cluster1^{Kill}*) or enhance Ubx-Exd binding (*abxN-Cluster1^{Hi}*). While each of these mutations altered the PD bias in the context of the transgenic *lacZ* reporter, neither perturbed the PD bias of Ubx, as assayed at the protein level. This was true whether we assayed *Ubx* expression in haltere discs homozygous for the *abxN-Cluster1^{Kill}* allele (Figure 5C, quantified in 5C') or through the generation of homozygous *abxN-Cluster1^{Kill}* clones (Figure 5 Supplemental 1A). Thus, mutation of a single cluster of low affinity Ubx-Exd binding sites within the >100 kb *Ubx* locus is insufficient to impact Ubx expression.

Multiple low affinity Ubx-Exd binding sites are required for PD bias formation

The disparity between the effect of Cluster 1 mutations on the PD bias of the CRM *lacZ* reporter transgene and the absence of an effect of the same mutations at the endogenous locus suggested the existence of additional regulatory inputs; as mentioned above, the *Ubx* locus itself is >100 kb, *abxFAIRE* is ~4 kb, and *abxF* is ~1.4 kb, whereas *abxN* contains only 531 bp of *abx* sequence. Thus, the difference could suggest the involvement of multiple Ubx-Exd binding sites in Ubx-mediated proximal repression that are present in the greater *abx* region, but not within the *abxN* transgene. In fact, we can see some evidence of this even within the *abxN* transgenic CRM, itself. Even though the abrogation of binding to Cluster 1 resulted in the greatest amount of proximal de-repression, mutation of Cluster 2 also showed some effect (Figure 4C, C').

Thus, we sought to expand our window outside of *abxN* to search for additional predicted Ubx-Exd binding sites. Using *NRLB* (30), we identified the top twenty predicted Ubx-Exd binding sites within *abxFAIRE*, ranked by relative affinity (Figure 6A). Very few high affinity sites are present (colored in green, Figure 6A), and these binding sites exist near the border of the FAIRE peak, suggesting that they may not even be accessible *in vivo* (Figure 6A-Inset); the remaining predicted Ubx-Exd binding sites are low affinity sites, on par with those found within the smaller *abxN* fragment.

To streamline our mutagenesis efforts, we focused our attention on 1) mutating the small set of high affinity binding sites (green, Figure 6A), and 2) a larger set of low affinity binding sites all residing within *abxF* (red, Figure 6A). Mutations were made to abrogate Ubx-Exd binding within each of these sets independently and together, resulting in three mutant genotypes (Figure 6 Supplemental 1A). In order to generate *abxFAIRE* mutants at the endogenous locus we utilized our 4 kb replacement platform (Targeted Region^{4kb}, Figure 5A, B) to perform RMCE. As above, we confirmed with an *abxFAIRE*^{WT} replacement that the presence of the RMCE scars does not affect *Ubx* expression (Figure 6 Supplemental 1B, top panel).

A comparison of the PD bias in flies homozygous for the *abxFAIRE* replacement alleles revealed no defect in proximal repression upon mutation of either the set of high, low, or combined high and low binding sites (Figure 6B, quantification shown along with *abxN* replacements in B'). Thus, just as with the *abxN*-Cluster 1 mutants, mutation of up to fourteen binding sites (Figure 6 Supplemental 1A) within *abxFAIRE* was unable to

384 perturb the PD bias. This result was confirmed by generating homozygous *abxFAIRE*
385 mutant mitotic clones in developing halteres; again, levels of proximal *Ubx* expression
386 remained comparable to wildtype (Figure 6 Supplemental 1B).
387
388 Because not all low affinity binding sites within *abxFAIRE* were mutated –
389 including those with relative affinities within the top 20 selected sites (Figure 6A, shown
390 in black) and those below the relative affinity cutoff (not shown) – it remained possible
391 that the lack of effect observed upon mutation of endogenous *abxFAIRE* was due to the
392 presence of additional, unmutated binding sites. To address this, we first sought to
393 understand the effect of additional *abx* mutations in a transgenic CRM reporter gene.
394 We focused our attention on the set of low affinity *abxF* binding sites (red, Figure 6A)
395 and returned to our *abxF-lacZ* reporter transgene to conduct an analysis of PD bias as a
396 result of increasing mutation load on the ~1.4 kb *abxF* CRM. While mutation of the
397 Cluster 1 or Cluster 2 binding sites alone did not result in proximal de-repression in this
398 context, abrogation of binding to the set of all low affinity binding sites (red, Figure 6A)
399 resulted in a slight, but statistically significant, de-repression and a concomitant
400 reduction in the PD bias (Figure 6C, quantified in C'). The requirement for many
401 mutations to begin to observe de-repression in the context of a longer *abx* fragment
402 supports our hypothesis that multiple low affinity binding sites are necessary to achieve
403 *Ubx* negative autoregulation. This also helps resolve the disparity between our results
404 with the transgenic *abx-lacZ* reporters and at the endogenous *Ubx* locus. The presence
405 of even more binding sites outside of the *abx* region could also contribute to PD bias

406 formation, masking the de-repression effect of mutations made on only a subset of
407 binding sites.
408
409 To further assess the existence of binding sites outside of *abxN*, we replaced the
410 endogenous ~4 kb *abxFAIRE* region with the much shorter (531 bp) *abxN* with and
411 without Cluster 1 binding sites mutated, effectively deleting ~3.5 kb of *abxFAIRE* (Figure
412 6 Supplemental Figure 2A, B). A reduction of the total number of potential Ubx-Exd
413 binding sites within the *abxFAIRE* region through this strategy holds the potential to
414 distinguish between two conclusions: (1) if an impact of mutating Cluster 1 on the PD
415 bias is observed, it would argue for the sufficiency of binding sites within *abxFAIRE* for
416 PD bias formation, or (2) if there is no impact of mutating Cluster 1, it is likely that
417 additional Ubx-Exd binding sites outside of the *abxFAIRE* region are involved in PD bias
418 formation. Two phenotypes are observable upon replacement of *abxFAIRE* with the
419 shorter *abxN*. First, for both the WT *abxN* and Cluster 1^{kill} replacements, large patches
420 of tissue exhibit no *Ubx* expression (Figure 6 Supplemental Figure 2B). Thus,
421 minimization of the *abx* CRM results in a seemingly stochastic loss of *Ubx* activation,
422 underscoring the activating role of *abx*. While this phenotype makes it more challenging
423 to address the effects of the *abxFAIRE-abxN* replacements on PD bias, in cells where
424 *Ubx* expression was present, the PD bias remained intact even upon mutation of *abxN*
425 Cluster 1 (Figure 6 Supplemental Figure 2B). This result confirms the likely involvement
426 of binding sites elsewhere in the locus (outside of *abxFAIRE*) in *Ubx* negative
427 autoregulation.
428

Depletion of Ubx Protein de-represses proximal Ubx transcription at the native locus

While we were able to establish the requirement for the cofactors, Hth and Exd, for proximal repression of the native *Ubx* locus through the generation of mutant mitotic clones (Figure 2B, D), evidence for the involvement of Ubx protein, itself, was restricted to the use of *abx-lacZ* transgenes (Figures 3 and 4) and the *lacZ*^{HC166D} enhancer trap (Figure 1). To address the role of Ubx protein at the native *Ubx* locus, we made use of the nanobody-based deGradFP system designed to direct depletion of a GFP fusion protein using a genomically encoded and spatially restricted anti-GFP nanobody coupled to a ubiquitin E3 ligase that leads to proteasomal degradation of the targeted fusion protein (34,35). In parallel, we generated two *Ubx* knockin alleles through a two-step CRISPR/RMCE targeting strategy that replaced *Ubx exon 1* with either a *GFPUbx exon 1* protein fusion (GFPUbx) or a *GFP-t2a-Ubx Exon 1* allele (*GFP-t2a-Ubx*) that produces both GFP and Ubx protein in a 1:1 ratio (Figure 7 Supplemental 1). By combining the deGradFP system with these GFPUbx alleles, along with smRNA FISH to assay Ubx transcription, we could directly interrogate whether a reduction of Ubx protein results in proximal de-repression of the endogenous *Ubx* locus.

Restriction of the expression of the deGradFP nanobody (Nslmb-vhhGFP), and thus depletion of GFPUbx, was achieved both spatially using the Gal4/UAS system, and temporally by addition of a temperature-sensitive Gal80 repressor. For our purposes, degradation of Ubx was restricted to either the distal compartment (using *nubbin-Gal4*) or to the proximal compartment (using *teashirt-Gal4*), and allowed to occur for 24 hours

prior to dissection through a temperature shift from 18°C to 30°C (Figure 7A). smRNA FISH of homozygous *GFPUbx* haltere discs reveals the expected PD bias with proximal transcription levels lower than distal. This can also be observed by assessing native GFP fluorescence from the fusion gene (Figure 7B, top). Degradation of GFPUbx induced distally did not affect *Ubx* transcription as assayed by smRNA FISH (Figure 7B, middle left), further corroborated by GFP intensity in these discs. Distal GFPUbx protein levels decreased almost to proximal levels because, though transcription remains the same, the resulting protein is constantly degraded by the nanobody (Figure 7B middle right). In contrast, degradation of GFPUbx proximally resulted in increased proximal transcription as assayed by smRNA FISH (Figure 7B, left). Thus, consistent with our findings using transgenic *abx* reporters, Ubx is necessary for proximal repression at the endogenous locus; depletion of Ubx protein results in *Ubx* de-repression. This negative autoregulatory loop, though, appears restricted to the proximal compartment as loss of Ubx distally does not result in increased distal transcription levels. Finally, to ensure that these effects were due to the loss of Ubx, we repeated the same experiments on the *GFP-t2a-Ubx* allele. Here, degradation of GFP occurs independent of Ubx (Figure 7A); and, in fact, though clear loss of GFP intensity can be observed with expression of the deGradFP nanobody, no effect on transcription from the *GFP-t2a-Ubx* allele was seen either distally or proximally (Figure 7C).

Discussion

The proper regulation of gene expression levels is critical for proper cell function; this is particularly true for developmental genes that must be tightly regulated both in space and time. In this study, we have combined transgenic assays with genome engineering of a native CRM to characterize the mechanism of quantitative transcriptional tuning of a single developmental gene, *Ubx.* Using transgenic CRM reporter assays we have revealed a novel negative autoregulatory mechanism to partially repress proximal Ubx levels relative to distal, which functions, at least in part, through the intronic CRM, *abx*. Ubx in complex with its cofactors, Hth and Exd, binds a cluster of low affinity binding sites within *abx* to dampen transcription. In truncated *abx* transgenes, the number of binding sites that need to be mutated in order to observe de-repression increases as the length of the CRM – and thus the number of potential low affinity binding sites – increases. However, in our most truncated version (*abxN)*, in which only two clusters of binding sites are predicted, mutation of a single cluster (Cluster 1) that abrogates Ubx binding is sufficient to observe de-repression; and mutation of the same cluster to a high affinity binding site enhances proximal repression – thus establishing a connection between the strength of Ubx-Exd binding to the CRM and the strength of repression. While previous reports demonstrated the utility of clusters of low affinity binding sites to balance specificity and robustness (32,36), which is likely at play here, low affinity binding sites may also make transcription more tunable by allowing CRMs to be more responsive to changes in local TF concentration (37). Further, in this case it appears that the affinity of the Ubx-Exd-Hth binding sites is tuned to create the right amount, and not too much, repression.

497 By performing transgenic *abx* reporter assays and genome engineering of the
498 native *abx* CRM in parallel, we were able to reveal disparities in the impact of binding
499 site mutations in each context. While we established the role of Ubx (through targeted
500 protein degradation) and Hth-Exd (through the use of genetic null alleles) in mediating
501 proximal repression at the endogenous locus, mutation of 14 predicted binding sites
502 within the native *abx* CRM did not result in de-repression of Ubx, despite having this
503 effect in a transgenic reporter context. This disparity, along with our understanding of
504 the involvement of multiple low affinity binding sites throughout the *abx* region, suggests
505 a model in which proximal repression and the formation/maintenance of the PD bias
506 relies on the contribution of sequence information beyond the *abx* region (Figure 7D).
507 We focused in on a ~4 kb region within *abx* that both exhibits the highest signal in
508 FAIRE accessibility assays, shows significant enrichment of Ubx and Hth binding, and is
509 able to recapitulate the PD bias when paired with a reporter. Despite this, additional
510 regions of chromatin accessibility and Ubx-Hth binding are observed within the locus,
511 and the contribution they may play to PD bias formation cannot be ruled out (Figure 3A).
512 Further, the UCR, as captured by a 35 kb region upstream of the TSS fused to *lacZ*
513 (35UZ), also exhibits a PD bias – though only within the posterior domain where the
514 UCR is active (13). Thus, it appears that generation of the PD bias through proximal
515 repression is not established by a single regulatory element, but rather by modifying the
516 activity of many (or all) regulatory elements within the locus.
517
518 This idea raises interesting questions regarding how proximal repression is
519 achieved upon binding of Ubx to *abx*. Given our evidence that *abx* mediates both

activation of Ubx transcription (both proximally and distally) and repression of Ubx transcription (proximally), it becomes interesting to ask if Ubx-mediated repression in this context is a dampening of the inherent activating potential of *abx.* This becomes particularly pertinent when considering the absence of Ubx repression in the distal hinge even though Ubx, Hth, and Exd are present (Figure 2A). The contribution of other factors (either proximally or distally) could potentially modulate the repressive effect of Ubx at *abx.*

Author Contributions

All authors contributed to the design of this study; VR, RKD, RL, RV conducted experiments. RKD, VR, and RSM wrote the manuscript.

Acknowledgements

The authors thank members of the Mann lab for comments and suggestions throughout this study and Chaitanya Rastogi and Harmen Bussemaker for help with binding site analyses.

Methods

Fly Strains

All wildtype strains used are *yw*. Mutant alleles used are as follows: exd^{1}, exd^{2}, hth^{P2}, $hth^{100\text{-}1}$, $Ubx^{9\text{-}22}$, Dcr^{2}. UAS-*ExdRNAi* was obtained from the Vienna Stock Center (VDRC), and UAS-*UbxRNAi* was generated in our lab using the following primer sets – UbxF-NheI: GATCGCTAGCAACTCGTACTTTGAACAGGCC; Ubx-F-AvrII: GATCCCTAGGAACTCGTACTTTGAACAGGCC; Ubx-R-XbaI: GATCTCTAGAGCTGACACTCACATTACCGC; Ubx-R-EcoRI: GATCGAATTCGCTGACACTCACATTACCGC. $Ubx\text{-}lacZ^{HC166D}$ (also referred to as $Ubx\text{-}lacz^{166}$ (16)) was a gift from Welcome Bender and is described in Bender *et al.* 2000 (23). UAS-GFPdegrade flies (UAS-NSlmb-vhhGFP4) are from Bloomington (34). The following CRISPR alleles (described below) were made for and used in this study: Targeted Region$^{2kb\text{-}ubiDsRed}$ Replacement Platform, $abxN^{WT}$, $abxN^{Cluster1Kill}$, $abxN^{Cluster1Hi}$, *Targeted Region*$^{4kb\text{-}ubiDsRed}$ Replacement Platform, $abxFAIRE^{WT}$, $\Delta abxFAIRE$ (which is: abxFAIRE^{MCS}), $abxFAIRE^{Low\text{-}Mutant}$, $abxFAIRE^{High\text{-}Mutant}$, , $abxFAIRE^{Low+High\text{-}Mutant}$ $abxFAIRE\text{-}abxN^{WTreplace}$, $abxFAIRE\text{-}abxN^{Cluster1Killreplace}$, $UbxExon1^{P3RFP}$ Replacement Platform, $UbxExon1^{WT}$, *GFPUbx* Fusion, *GFP-t2a-Ubx.*

CRISPR Targeting

Two regions within *abx* and one region encompassing *Ubx* Exon1 were targeted with CRISPR/Cas9. For each targeting event, two gRNAs were designed flanking the region of interest. gRNA sequences for each of the regions are as follows – *abx-Targeted Region*^{2kb}*:* GAGATGCTTTTGAATTCTCG and GGCAGATCGGATTGGATCTT; *abx-Targeted Region*^{4kb}*:* GGCTTTGCAACTAATTGAAA and

565 GTAAATGTTGGCTATTCAAAA; *Ubx* Exon1: GAATTCGAAGAAAATTAG and
566 GTAAGACATATGAAAGC. gRNAs were cloned into the pCFD4 dual gRNA vector
567 (http://www.crisprflydesign.org/, Port *et al.* 2014 (38)). Homemade donor vectors were
568 made containing either a ubiDsRED or P3-RFP fluorescent selection marker flanked by
569 inverted PhiC31 attP recognition sequences – the ubiDsRED cassette contains a full
570 attP sequence
571 (GTACTGACGGACACACCGAAGCCCCGGCGGCAACCCTCAGCGGATGCCCCGGG
572 GCTTCACGTTTTCCCAGGTCAGAAGCGGTTTTCGGGAGTAGTGCCCCAACTGGGG
573 TAACCTTTGAGTTCTCTCAGTTGGGGGCGTAGGGTCGCCGACATGACACAAGGGG
574 TTGTGACCGGGGTGGACACGTACGCGGGTGCTTACGACCGTCAGTCGCGCGAGC
575 GCGA), whereas the P3-RFP cassette contains a minimal attP sequence
576 (CCCCAACTGGGGTAACCTTTGAGTTCTCTCAGTTGGGGG) from Voutev *et al.* 2018
577 (39). ~1.5 kb homology arms were cloned on either side of the inverted attP sites.
578 Primers used to clone the homology arms are as follows: *abx-Targeted Region^{2kb}:* (Left
579 Arm) GCCAGAAGCTGCAAATTCAAG and CTTTGGGTTCTGTTCCACAGC, (Right
580 Arm) GAATTCAAAAGCATCTCCGCATAAAG and GCCAACCGCAGACTGTGCGA;
581 *abx-Targeted Region^{4kb}:* (Left Arm) GATGTAGGCCATGGTTTCGGC and
582 TGAATAGCCAACATTTACTGACTCG, (Right Arm)
583 AAACGGTAAAACTTGAGATTTTCTTATT and CGGAGAATCCGTATGAATCG; *Ubx*
584 *Exon1:* (Left Arm) GCTCAACTGTAGTTTTCTGTTCG and
585 ATTTTCTTCGAATTCTTATATGCTAT, (Right Arm) AGCAGGCAGAACAGACCTT and
586 CTCGCAGAGATTGTCTGACAC. The gRNA template (pCFD4) and donor template
587 were injected into a germline-expressing Cas9 strain (nanos-Cas9, Kondo *et al.* 2013

588 (40)) at a concentration of 250 ng/µL and 500 ng/µL, respectively. Selection of positive
589 CRISPR events was done by screening for the presence of ubiDsRED or P3-RFP.
590 Positive fly lines were validated by PCR and Southern Blot analysis.
591
592 ***Recombinase Mediated Cassette Exchange (RMCE)***
593
594 PhiC31-mediated RMCE was used to replace the ubiDsRED/P3-RFP selection markers
595 inserted using CRISPR/Cas9 into *abx/Ubx Exon1*. A homemade vector was used,
596 containing inverted PhiC31 attB recognition sequences
597 (CGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTAC) flanking a multiple
598 cloning site for insertion of sequences used for replacement alleles. Replacement of
599 wildtype sequence from *abx-Targeted Region^{2kb}, abx-Targeted Region^{4kb},* and *Ubx*
600 Exon1 was performed by amplifying regions of interest from the yw genome. The
601 following primers were used – *abx-Targeted Region^{2kb}:* ATCCAATCCGATCTGCCCAG
602 and TCGAGGAGTGAGTAAGAGATTGATAAAG*; abx-Targeted Region^{4kb}:*
603 AAACGGGAGGCTTTTGCTG and CAATTAGTTGCAAAGCCGTTTTTC; *Ubx Exon 1*:
604 TAGAGGTTGTATTGTTTTATTAATAAAAAACCTATTG and
605 TTCATATGTCTTACATTACAAGTTGTTATCTGTTTTTCC. Mutations of replacement
606 regions were made either by site directed mutagenesis or through synthesis and ligation
607 of mutated restriction fragments within the region of interest. In the generation of *abx-*
608 *Targeted Region^{4kb}* mutant replacements, the ~4 kb fragment with mutations was
609 stitched together from synthesizing restriction fragments based on reference sequence
610 with engineered mutations. In doing this, additional natural variants (SNPs, small indels)

611 between the reference strain and our wildtype strain (yw) were introduced into the
612 *abxFAIRE*^{LowAffinityMut} and *abxFAIRE*^{LowandHighMut}, and thus differ from the wildtype
613 *abxFAIRE*^{WT} replacement allele in these additional locations. These natural variants do
614 not reside within predicted Ubx/Exd binding sites, do not create new predicted Ubx/Exd
615 binding sites, nor do our results show that they have an effect on PD bias formation.
616 The *abxFAIRE* deletion allele was generated by using an attB donor plasmid containing
617 a multiple cloning site
618 (gaagcttcctaggaggcctagatctgcggccgcttaattaaacgcgtgaatgggcgcgccgctagccatatgggtaccg
619 gatcc). The replacement of *abxFAIRE* with *abxN* was done using the following primers:
620 GAACACAAAGGAGTCTGGTG and AACGTCGGAGGATGTAGG. The GFPUbx fusion
621 and GFP-t2a-Ubx replacement alleles were cloned using overlap PCR and inserted at
622 the transcription start site of *Ubx Exon 1.* The GFPUbx fusion contained the linker,
623 GGSGGSG; the GFP-t2a-Ubx allele contained the t2a sequence,
624 ggttctggagagggccgcggcagcctgctgacctgcggcgatgtggaggagaaccccgggccc. Replacement
625 donor plasmids were injected into flies with the attP platform cassette either at *abx* or at
626 *Ubx Exon 1*. The necessary recombinase enzyme, PhiC31, was either injected as
627 plasmid along with the donor cassette (*abx* replacements) or was expressed from a
628 genomic insertion on the X chromosome of nanos-PhiC31 (*Ubx Exon 1* replacements).
629 Progeny from injected flies were screened for the loss of the fluorescent selection
630 marker (UbiDsRED, P3-RFP). Because the attP/attB reaction does not provide
631 directionality, replacements can be inserted in the forward or reverse direction.
632 Southern blot was performed to ensure the correct directionality of the replacement.
633

634 ***Antibodies***

635

636 Antibodies used are as follows – anti-Ubx (mouse, 1:10 FP3.38 from Developmental
637 Studies Hybridoma Bank (DSHB) in supernatant or ascites form); anti-ß-gal (Rabbit,
638 1:5000, MP Biochemicals 559762); anti-Hth^{HD} (guinea pig, 1:500, Noro and Mann 2006
639 (26)); anti-Hth (guinea pig, 1: 5000, Ryoo and Mann 1999 (41)); anti-Exd (rabbit, Mann
640 and Abu-Shaar 1996 (42)); where GFP signal is shown, native GFP fluorescence was
641 acquired.

642

643 ***FAIRE and ChIP-CHIP Data Accession***

644

645 FAIRE data shown was from McKay *et al*. (29), and downloaded from NCBI GEO
646 database with accession number: GSE38727. ChIP data for Ubx and Hth was obtained
647 from Slattery *et al.* (28), and downloaded from NCBI GEO database with accession
648 number: GSE26793.

649

650 ***Immunostaining of Imaginal Discs***

651

652 Wandering third instar larvae were collected and dissected in PBS to invert the head
653 region and expose attached imaginal discs to solution. Inverted heads were fixed in Fix
654 Solution (PBS/4% Paraformaldehyde/.1% TritonX/.1% Sodium Deoxycholate) for 25
655 minutes at RT. Fix solution was removed and replaced with Staining Solution (PBS/.3%
656 TritonX/1% BSA). Inverted heads were washed 2X with Staining Solution for 20 minutes

at RT and stained overnight with desired antibody in Staining Solution at 4C. The following day antibody solution was removed and inverted heads washed 4X with Staining Solution followed by a 1.5hr incubation with secondary antibody and DAPI (1:1000) in Staining Solution at RT. This was followed by two washes with Staining Solution, dissection of discs from the inverted heads in PBS and mounting of the discs in Vectashield. Imaging of discs was conducted on the following microscopes: Zeiss Apotome.2 Microscope, Leica SP5 Confocal Microscope, and Zeiss LSM 800 Confocal Microscope.

Calculation of Distal/Proximal Ratio

All analysis was done in Fiji/ImageJ. Confocal Z-stacks of discs were imported into Fiji and slices near the edge of the stack that contained peripodial membrane were manually removed. All pixels surrounding the haltere disc of interest were cleared by drawing an ROI around the disc itself. ROIs for the distal compartment and proximal compartment were drawn based on a single Z-slice in which the distal pouch was clearly demarcated (using either the antibody stain of interest (Ubx/LacZ) or DAPI). These ROIs were propagated to all slices of the image and the average intensity of the stain of interest was acquired, excluding black pixels. For each disc, the mean of these single-slice average intensities was computed for the whole disc, distal compartment, and proximal compartment. Multiple discs were analyzed to produce the scatter plots shown. Each point is representative of a single disc. A one-way ANOVA analysis

679 followed by Tukey's multiple comparisons test with alpha = .05 was used for statistical
680 analysis. This process is depicted in a schematic in Supplemental Figure 1.
681
682 ***Imaging of Adult Haltere Structure***
683
684 Whole adult flies were submerged in 70% ethanol, followed by three washes in
685 PBS/.3% TritonX. The head and abdomen were removed, leaving the thorax with
686 appendages attached. Thoraces were fixed overnight at 4C in PBS/4%
687 Paraformaldehyde. The following day, thoraces were washed 5X with PBS/.3% TritonX,
688 followed by dissection and mounting of halteres in Vectashield. Imaging was conducted
689 on a Leica SP5 confocal microscope, acquiring autofluorescence of the cuticle with the
690 488 laser. Images shown are MAX projections of the stacks acquired.
691
692 ***Generation of lacZ Reporter Constructs***
693
694 All CRM lacZ reporter constructs were generated by cloning the regulatory DNA of
695 interest (*abxN, abxF)* into pRVV54-lacZ (43) using the NotI and HindIII restriction
696 enzyme sites. Primers (5' to 3') used to amplify each regulatory region are as follows –
697 *abxN*: GAACACAAAGGAGTCTGGTG and AACGTCGGAGGATGTAGG; *abxF:*
698 GAACACAAAGGAGTCTGGTGAG and GTTAAGCATTTTGGGTGCGAG. All lacZ
699 reporter constructs were inserted into the attp40 landing site on chr2. Mutations were
700 made either through site-directed mutagenesis or through synthesis of mutated

701 restriction fragments within the regulatory region of interest, which were then ligated
702 back into the pRVV54-lacZ backbone.
703
704 ***NRLB Binding Site Prediction***
705
706 Binding sites for Exd/UbxIVa were predicted using the NRLBtools package in R (30).
707
708 ***Generation of RNAi Flip-out Clones***
709
710 Flip-out clones were generated by crossing *act<y<Gal4, UAS-GFP* to different *hs-flp*;
711 *UAS* lines and heat-shocking larvae for between 8-10 minutes at 37C.
712
713 ***Generation of Mitotic Clones***
714
715 Mutant alleles of interest were recombined with standard FRT lines. Flies with mutant
716 recombined alleles were crossed to either FRT ubiGFP or ubiDsRED (to mark the
717 clones) and progeny of this cross were heatshocked at 37C for 40min-1hr 48 hours after
718 egg laying (AEL). Wandering third instar larvae were collected 72 hours after heatshock,
719 dissected, and subjected to the immunostain protocol as described.
720
721 ***ChIP-qPCR***
722

723 Wandering third instar larvae from two different genotypes (*abxN WT-lacZ* and *abxN*
724 *Cluster 1^{kill}*-lacZ) were dissected and haltere imaginal discs were collected in PBS on
725 ice. Discs were fixed with 1.8% formaldehyde, crosslinked chromatin was sonicated,
726 and chromatin preparation and immunoprecipitation was performed as described in
727 Estella *et al.* 2008 (44). The IP was done with rabbit anti-Ubx (Ubx1, generated by
728 modENCODE) at a final concentration of 1.5 µg/mL for each IP. Rabbit IgG (Sigma)
729 was used for the control IP. The following primer pairs were used for qPCR –
730 Endogenous *abx*: TGGAGCTCCAAATGAAACGC and CGCTCAACATTGTTAGTGGC;
731 Transgenic *abxN:* CAGTGCTGGCTGCATTTGCT and ACAACTGATGCTCTCAGCCA;
732 Intergenic Control: CCGAACATGAGAGATGGAAAA and
733 AAAGTGCCGACAATGCAGTTA. qPCR was done on an Applied Biosystems 7300
734 machine and calculations were done using the 2-ΔΔCt method in MS Excel. IPs were
735 done in triplicate.
736
737 ***Protein Purification and Electrophoretic Mobility Shift Assays (EMSAs)***
738
739 Ubx protein was His-tagged and purified from *E. coli* (BL21 or BL21pLysS; Agilent) 4
740 hours of induction with isopropyl-B-D-thiogalactopyranoside (IPTG) using Co-
741 chromatography. Exd (in pET9a) and Hth^{HM} (in pET21b) were co-expressed and co-
742 purified, through the His-tag attached to Hth^{HM}, in *E. coli* (BL21) in the same way as Ubx
743 recombinant protein and used as a complex for all EMSAs. Protein concentrations were
744 determined by the Bradford assay and then confirmed by SDS/PAGE and Blue
745 Coomassie analysis (SimplyBlue SafeStain, Invitrogen). EMSAs were carried out as

746 previously described ((45)). UbxIa was used at 250 ng/lane (high concentration) and
747 200 ng/lane (low concentration) and Exd/HthHM was used at 150 ng/lane. Sequences
748 for probes used are as follows – *abxWT*:
749 CCTCGTCCCACAGCTcgaatatattttataacccggagcaaatgcagcca; *abxCluster1^{Kill}:*
750 CCTCGTCCCACAGCTcgaatccccttccccacccggagcaaatgcagcca. Uppercase is linker
751 sequence. DNA binding was observed using phosphoimaging as detected by a
752 Typhoon Scanner (Amersham).
753
754 ***smRNA FISH***
755
756 A probe library containing 48 20-nt Stellaris FISH probes was designed to target the first
757 2 kb of Ubx Intron 1. Libraries were ordered from Biosearch Technologies and labelled
758 with Quasar 670. Wandering third instar larvae were collected and dissected in PBS to
759 invert heads and expose discs to solution. Inverted heads were washed in PBSM
760 (PBS/5mM MgCl2) 1X at RT, followed by fixation in PBSM/4% PFA for 10 min at RT.
761 Discs were permeabilized with PBS/.5% TritonX for 10 min at RT and washed once with
762 PBSM for 10 min at RT. Inverted heads were washed 1X with Pre-Hyb (10% deionized
763 formamide in 2X SSC) for 10 min at RT prior to hybridization. Hybridization was
764 performed overnight in a thermoshaker at 37C (~600RPM) covered in foil. Hybridization
765 buffer contains: 2X SSC, .2 mg/mL BSA, 50% Dextran Sulfate, 10% deionized
766 formamide, 50 µg/mL *E. coli* tRNA, 50 µg/mL salmon sperm ssDNA, and 125 nM Ubx
767 Intron Probe. 100 µL of hybridization buffer was used for each sample. The following
768 day, hybridization buffer was removed and heads were washed with Pre-Hyb buffer for

769 20 minutes at 37C and 20 min at RT. Inverted heads were washed with PBS for 10 min
770 at RT, stained with DAPI (PBS/DAPI (1:1000 dilution) for 30 min at RT and
771 resuspended in PBS. Discs were dissected from inverted heads in PBS/1%BSA and
772 mounted in Vectashield prior to imaging.
773
774 ***GFP Degrade Assay***
775
776 deGRADFP flies, UAS-NSlmb-vhhGFP4, were obtained from Caussinus *et al.* 2011 (34)
777 and paired with tubGal80ts, and either nubbin (nub)-Gal4 or teashirt (tsh)-Gal4. Flies
778 were crossed to GFPUbx fusion or GFP-t2a-Ubx knockin flies. Progeny were kept at
779 18C (Gal80 inactive) until early third instar stage and then shifted to 30C for 24 hours
780 prior to dissection. Wandering third instar larvae were collected, dissected, and
781 subjected to smRNA FISH analysis. Genotypes are as follows: “No Degrade:” yw;
782 gal80ts, UAS-GFPdegrade, gal80ts/UAS-GFPdegrade, gal80ts; GFPUbx (or GFP-t2a-
783 Ubx)/GFPUbx (or GFP-t2a-Ubx); “Distal Degrade:” yw; nubGal4/gal80ts, UAS-
784 GFPdegrade; GFPUbx (or GFP-t2a-Ubx)/ GFPUbx (or GFP-t2a-Ubx); “Proximal
785 Degrade:” yw; tshGal4/gal80ts, UAS-GFPdegrade; GFPUbx (or GFP-t2a-Ubx)/ GFPUbx
786 (or GFP-t2a-Ubx).
787
788
789 ***Figure Captions***
790

Figure 1. Proximal/Distal Expression Bias of *Ubx* is Established at the Level of Transcription.

(A) (Left) A schematic of the 3rd Instar haltere imaginal disc, denoting regions of high versus low *Ubx* expression along with regions fated to become the notum (yellow), the proximal hinge (red), the distal hinge (green) and the pouch (blue). Domains of expression of key proximal (*teashirt* (*tsh*), *homothorax* (*hth*)) and distal (*nubbin* (*nub*)) specifying genes are shown. (Right) A confocal image (MAX projection) of an adult haltere. Dorsal view is shown. Structures along the proximal-distal axis are specified. **(B)** Ubx immunostain (along with DAPI nuclear stain) of a 3rd Instar haltere imaginal disc showing proximal-distal expression bias. Haltere disc is outlined in yellow. MAX projection of 4 slices shown. **(B')** Quantification of average Ubx immunostain intensity in the whole disc ("Disc"), proximal compartment, and Distal compartment. Each colored point represents a single disc (N=11). Dotted line represents the mean for each compartment. Reported D/P ratio is computed by dividing the mean distal intensity by the mean proximal intensity. **(C)** Single-molecule RNA FISH (smRNA FISH) against the first intron of *Ubx* (along with DAPI nuclear stain) showing proximal-distal bias at the level of transcription. MAX projection of all slices shown. **(D)** (Left) LacZ immunostain driven by the enhancer trap insertion, $lacZ^{HC166D}$. (Right) LacZ immunostain driven by $lacZ^{HC166D}$ upon generation of *UbxRNAi* clones. Ubx and GFP immunostains are also shown. GFP+ tissue marks cells expressing *UbxRNAi*. All scale bars shown are 50 micron in size.

812 **Figure 2. Proximal autorepression of *Ubx* Relies on Hox cofactors Hth/Exd. (A)** Immunostain
813 of Ubx (Left) and Hth (Middle), along with a merge of the two (Right) in 3rd instar haltere
814 imaginal discs. The distal hinge is marked with an asterisk to denote a Ubx-high region that
815 also expresses *hth*. **(B)** Ubx immunostain in haltere discs in which *hth* null clones (*hth*^{P2})
816 have been induced. Hth and GFP immunostains are also shown. Clones are Hth- and GFP-
817 negative. Yellow arrows point to the clone. **(C)** Ubx immunostain in haltere discs in which
818 Hth^{HM}-only clones (*hth*$^{100\text{-}1}$), lacking the C-terminal homeodomain) have been induced. An
819 immunostain for the C-terminus of Hth (Hth^{HD}) and GFP are shown. Clones are Hth^{HD}- and
820 GFP-negative. **(D)** Ubx Immunostain in haltere discs in which *exd* null clones (*exd*^{1}) have
821 been induced. An Exd and GFP immunostain are also shown. Clones are Exd- and GFP-
822 negative. Yellow arrows point to clones. **(E)** Exd immunostain in haltere discs in which *Ubx*
823 null clones (*Ubx*$^{9\text{-}22}$) have been induced. Ubx and GFP immunostains are also shown. Clones
824 are Ubx- and GFP-negative, and outlined in yellow. **(F)** Hth immunostain in haltere discs in
825 which *Ubx*-null clones (*Ubx*$^{9\text{-}22}$) have been induced. A Ubx and GFP immunostain are also
826 shown. Clones are Ubx- and GFP-negative, and outlined in yellow. All scale bars shown are
827 50 micron in size.
828
829 **Figure 3. A Ubx-regulated Module within the Intronic *abx* CRM Mediates autorepression.**
830 **(A)** A screenshot of the *Ubx* locus within the Integrative Genomics Viewer (IGV) is shown.
831 FAIRE accessibility peaks from McKay *et al.* ((29)) are shown for 3rd instar wing and haltere
832 imaginal discs (top two tracks). Ubx and Hth ChIP-chip in 3rd instar haltere discs from

833 Slattery *et al.* ((28)) are shown (bottom two tracks). The intronic *anterobithorax* (*abx*)
834 regulatory region is boxed. The location of two *abx* fragments that drive reporter expression
835 (*abx3.2, abx6.8)* are shown in red, along with the annotated location of the *abx* Polycomb
836 repressive element (PRE). Regions of interest (*abxN*, *abxF*, *abxFAIRE*) in this paper are
837 shown beneath the annotated *Ubx* gene. **(B)** Ubx Immunostain in haltere discs containing
838 Δ*abxFAIRE* clones. A GFP immunostain, in addition to a merge of GFP/Ubx, is also shown.
839 Clones are marked by yellow arrows and outlined in yellow. **(C)** LacZ immunostain in 3rd
840 instar imaginal discs expressing *lacZ* from an *abxF-lacZ* transgene. A wing, T3 leg and haltere
841 are shown. A Ubx immunostain, in addition to a merge of LacZ/Ubx, is also shown. **(D)** LacZ
842 immunostain in *abxF-lacZ* transgenic haltere discs in which *hth* null clones (hth^{P2}) have
843 been induced. GFP and Hth immunostains are also shown. Clones are GFP- and Hth-
844 negative, and marked with a yellow arrow. **(E)** LacZ immunostain in *abxF-lacZ* transgenic
845 haltere discs in which *exd* null clones (exd^{1}) have been induced. GFP and Exd immunostains
846 are also shown. Clones are GFP- and Hth-negative, and marked with a yellow arrow. **(F)**
847 LacZ immunostain in *abxF-lacZ* transgenic haltere discs in which Hth^{HM}-only clones (hth^{100-1}, lacking the C-terminal homeodomain) have been induced. An immunostain for the C-
849 terminus of Hth (Hth^{HD}) and GFP are also shown. Clones are Hth^{HD}- and GFP-negative, and
850 outlined in yellow. **(G)** LacZ immunostain in *abxF-lacZ* transgenic haltere discs in which
851 *Ubx* null clones (Ubx^{9-22}) have been induced. A GFP immunostain, in addition to a merge of
852 GFP/LacZ, is shown. Clones are GFP-negative and marked with a yellow arrow.
853

Figure 4. Ubx Directly Binds *abxN* to Autorepress in the Proximal Compartment. (A) LacZ immunostain in *abxN-lacZ* haltere discs. A DAPI nuclear stain and a merge of LacZ/DAPI are shown. **(B)** *NRLB* predicted Ubx-Exd binding sites within the 531 bp *abxN*, along with their relative affinities. Strandedness of the predicted binding sites is depicted by color: black bars above the axis denote the start of a site in the forward direction and red bars below the axis denote start of a site in the reverse direction. The boxed region is focused on in B'. **(B')** 49 bp zoomed in region from **B**. Binding sites are divided into Cluster 1 and Cluster 2. **(B")** Schematic of mutations of Cluster 1 and 2 made in this paper. Cluster 1 was mutated to abrogate Ubx/Exd binding (Kill) and enhancer Ubx/Exd binding (Hi). Cluster 2 was mutated to abrogate Ubx/Exd binding (Kill). **(C)** LacZ immunostains in *abxN-lacZ* transgenic haltere discs. Exemplary images are shown for discs transgenic for wildtype and mutant forms of the transgenic *abxN.* **(C')** Quantification of the Distal/Proximal ratio in discs described in **C**. Colors indicate discs from a single experiment. Dotted red bar signifies the mean value. Significance values are derived from one-way ANOVA analysis followed by Tukey's multiple comparisons test with alpha = .05. Multiplicity adjusted p-values for each level of significance are: <.0001 for ****, .0005 for ***, and .0143 for *. All scale bars shown are 50 micron in size.

Figure 5. Mutation of *abxN* at the Native Locus is not Sufficient to Alter the Proximal-Distal Bias. (A) (Top) IGV Genome Browser Snapshot displaying FAIRE accessibility peaks in the wing and haltere (from McKay *et al.* (29)), and Ubx and Hth ChIP-chip peaks in the haltere (from Slattery *et al.* (28)). The bottom conservation track was downloaded from UCSC

875 Genome Browser and denotes the evolutionary conservation of *D. melanogaster* sequence
876 with twelve other *Drosophila* species, mosquito, honeybee and red flour beetle. Regions
877 devoid of red lack conservation. (Bottom) Regions targeted with CRISPR/Cas9. Black bars are
878 the regions that were targeted for deletion and/or replacement. Yellow triangles denote
879 regions of low conservation that were targeted by the Cas9/gRNA complex. These are also
880 the points of insertion of the integrase recognition sequence (*attP*). **(B)** A schematic of the
881 two-step CRM replacement strategy. gRNA sites were chosen in non-conserved regions
882 surrounding the *abx* sequence of interest. A dual-gRNA expressing plasmid was injected
883 along with a donor cassette containing an attP-flanked fluorescent selection marker (*ubi-*
884 *dsRED*) into *nanos-Cas9* flies. The resulting "*abx* Replacement Platform" serves as a means to
885 delete or insert modified versions of *abx* using PhiC31-based RMCE. **(C)** Ubx immunostain in
886 haltere discs homozygous for *abxN* replacement alleles. WT and Cluster 1 mutants are shown
887 alongside a yw control. **(C')** (Top) A schematic of the Targeted Region^{2kb} replacement
888 boundaries for discs shown in **C**. (Bottom) Quantification of the Distal/Proximal ratio in discs
889 from **C**. Each point is an individual disc. Points are color-coded by experiment. Dotted red
890 line signifies mean value. Significance was tested using a one-way ANOVA analysis followed
891 by Tukey's multiple comparisons test with alpha = .05. All scale bars shown are 50 micron in
892 size.
893
894 **Figure 6. Multiple Low Affinity Ubx/Exd Binding Sites are Required for PD Bias Formation.**
895 **(A)** The top twenty *NRLB* predicted Ubx-Exd binding sites within the 4057 bp *abxFAIRE*,

along with their relative affinities. Sites are segregated into high affinity (green), low affinity (red), both of which are mutated. Sites in black were not mutated. Locations of *abxN* and *abxF* are boxed. The *inset* places the mutated sites in the context of haltere *abx* accessibility. **(B)** Ubx immunostain in haltere discs homozygous for *abxFAIRE* replacement alleles. WT, Low Affinity Mutants, High Affinity Mutants, and the combination are shown. **(B')** (Top) A schematic of the Targeted Region^{4kb} replacement boundaries for discs shown in **B**, along with a Ubx immunostain in haltere discs from *yw* flies for control. (Bottom) Quantification of the Distal/Proximal (D/P) ratio in discs from **B** combined with *abxN* replacements from Figure 5. Each point is an individual disc. Points are color-coded by experiment. Dotted red line signifies mean value for each genotype. Significant differences between each of the replacement alleles was tested using one-way ANOVA analysis followed by Tukey's multiple comparisons test with alpha = .05. All scale bars shown are 50 micron in size. **(C)** LacZ immunostain in haltere discs with a wildtype or mutant *abxF-lacZ* reporter transgene. Haltere discs are outlined in yellow. **(C')** Quantification of Distal/Proximal ratios in *abxF-lacZ* transgenic discs. Each point denotes an individual disc. Dotted red line denotes mean value for each genotype. Significance values are derived from one-way ANOVA analysis followed by Tukey's multiple comparisons test with alpha = .05. Multiplicity adjusted p-values for each level of significance are: <.0001 for ****.

Figure 7. Depletion of Ubx Protein Derepresses Proximal *Ubx* Transcription at the Native Locus. **(A)** Schematic of the GFP-degrade system used. A *GFPUbx* fusion and a *GFP-t2a-Ubx*

knockin were generated by CRISPR. These *Ubx* alleles were coupled with a UAS-driven nanobody-based GFP degrade system (from (34,35), compartment restricted Gal4 drivers, and a temperature-sensitive Gal80 transgene to restrict GFP degradation to a short window. Larvae were shifted from 18°C (Gal80 active) to 30°C (Gal80 inactive) to allow for GFP degradation 24hr prior to dissection. **(B)** (Left) smRNA FISH against the first intron of *Ubx* in *GFPUbx* fusion knockin haltere discs. (Middle) Cropped image of smRNA FISH from left. (Right) Acquisition of GFP fluorescence. "No Degrade" signifies the absence of a Gal4 to drive the degrade system. "Distal Degrade" denotes the use of *nubbin-Gal4* (*nub-Gal4*), and "Proximal Degrade" denotes the use of *teashirt-Gal4* (*tsh-Gal4*). **(C)** The same experiment as in **B** for the *GFP-t2a-Ubx* allele. **(D)** A model of Ubx negative autoregulation in the proximal haltere. Multiple low affinity Ubx-Exd binding sites within *abx* are necessary for proximal repression, but we cannot rule out the involvement of additional sequence juxtaposed to *abx* or elsewhere in the gene. Because *abx* is necessary for the activation of *Ubx* expression (both distally and proximally), we suggest the possibility that the proximal repression effect results from a modification of this inherent activating role. All scale bars are 50 micron in size.

References

1. Driever W, Nüsslein-Volhard C. The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. Cell. 1988 Jul

1;54(1):95–104.

2. Grueber WB, Jan LY, Jan Y-N. Different levels of the homeodomain protein cut regulate distinct dendrite branching patterns of Drosophila multidendritic neurons. Cell. 2003 Mar 21;112(6):805–18.

3. DeKoter RP, Singh H. Regulation of B lymphocyte and macrophage development by graded expression of PU.1. Science. 2000 May 26;288(5470):1439–41.

4. Tomoyasu Y. ScienceDirectUltrabithorax and the evolution of insect forewing/hindwing differentiation. Current Opinion in Insect Science. Elsevier Inc; 2017 Feb 1;19(C):8–15.

5. White RA, Wilcox M. Distribution of Ultrabithorax proteins in Drosophila. EMBO J. 1985 Aug;4(8):2035–43.

6. White RA, Wilcox M. Protein products of the bithorax complex in Drosophila. Cell. 1984 Nov;39(1):163–71.

7. Bender W, Akam M, Karch F, Beachy PA, Peifer M, Spierer P, et al. Molecular Genetics of the Bithorax Complex in Drosophila melanogaster. Science. 1983 Jul 1;221(4605):23–9.

8. White RAH, Akam ME. Contrabithorax mutations cause inappropriate expression of Ultrabithorax products in Drosophila. Nature. 1985 Dec 12;318(6046):567–9.

9. Little JW, Byrd CA, Brower DL. Effect of abx, bx and pbx mutations on expression of homeotic genes in Drosophila larvae. Genetics. 1990 Apr;124(4):899–908.

10. Müller J, Bienz M. Long range repression conferring boundaries of Ultrabithorax expression in the Drosophila embryo. EMBO J. 1991 Nov;10(11):3147–55.

11. Simon J, Peifer M, Bender W, O'Connor M. Regulatory elements of the bithorax complex that control expression along the anterior-posterior axis. EMBO J. 1990 Dec;9(12):3945–56.

12. Peifer M, Bender W. The anterobithorax and bithorax mutations of the bithorax complex. EMBO J. 1986 Sep;5(9):2293–303.

13. Irvine KD, Botas J, Jha S, Mann RS, Hogness DS. Negative autoregulation by Ultrabithorax controls the level and pattern of its expression. Development. 1993 Jan;117(1):387–99.

14. Irvine KD, Helfand SL, Hogness DS. The large upstream control region of the Drosophila homeotic gene Ultrabithorax. Development. 1991.

15. Lewis EB. A gene complex controlling segmentation in Drosophila. Nature. 1978 Dec 7;276(5688):565–70.

16. Crickmore MA, Ranade V, Mann RS. Regulation of Ubx expression by epigenetic enhancer silencing in response to Ubx levels and genetic variation. PLoS Genet. 2009 Sep;5(9):e1000633.

17. Bienz M, Tremml G. Domain of Ultrabithorax expression in Drosophila visceral mesoderm from autoregulation and exclusion. Nature. 1988 Jun 9;333(6173):576–8.

18. Delker RK, Mann RS. From Reductionism to Holism: Toward a More Complete View of Development Through Genome Engineering. In: Tsang SH, editor. Precision Medicine, CRISPR, and Genome Engineering: Moving from Association to Biology and Therapeutics. Cham: Springer International Publishing; 2017. pp. 45–74. (Precision Medicine, CRISPR, and Genome Engineering: Moving from Association to Biology and Therapeutics).

19. Lewis EB. Genes and Developmental Pathways. Am Zool. 1963;3(1):33–56.

20. Duncan I, Montgomery G. Anecdotal, historical and critical commentaries on genetics - E. B. Lewis and the bithorax complex: Part I. Genetics. 2002 Apr;160(4):1265–72.

21. González-Gaitán MA, Micol JL, Garcia-Bellido A. Developmental genetic analysis of Contrabithorax mutations in Drosophila melanogaster. Genetics. 1990 Sep;126(1):139–55.

22. Smolik-Utlaut SM. Dosage requirements of Ultrabithorax and bithoraxoid in the determination of segment identity in Drosophila melanogaster. Genetics. 1990 Feb;124(2):357–66.

23. Hudson WBAA. P element homing to the Drosophila bithorax complex. 2000 Aug 21;:1–12.

24. Rieckhof GE, Casares F, Ryoo HD, Abu-Shaar M, Mann RS. Nuclear translocation of extradenticle requires homothorax, which encodes an extradenticle-related homeodomain protein. Cell. 1997 Oct 17;91(2):171–83.

25. Azpiazu N, Morata G. Functional and regulatory interactions between Hox and extradenticle genes. Genes Dev. 1998 Jan 15;12(2):261–73.

26. Noro B, Culi J, McKay DJ, Zhang W, Mann RS. Distinct functions of homeodomain-containing and homeodomain-less isoforms encoded by homothorax. Genes Dev. 2006 Jun 15;20(12):1636–50.

27. Kurant E, Eytan D, Salzberg A. Mutational analysis of the Drosophila homothorax gene. Genetics. 2001 Feb;157(2):689–98.

28. Slattery M, Ma L, Négre N, White KP, Mann RS. Genome-wide tissue-specific occupancy of the Hox protein Ultrabithorax and Hox cofactor Homothorax in

Drosophila. PLoS ONE. 2011;6(4):e14686.

29. McKay DJ, Lieb JD. A common set of DNA regulatory elements shapes Drosophila appendages. Dev Cell. 2013 Nov 11;27(3):306–18.

30. Rastogi C, Rube HT, Kribelbauer JF, Crocker J, Loker RE, Martini GD, et al. Accurate and sensitive quantification of protein-DNA binding affinity. Proceedings of the National Academy of Sciences. 2018 Apr 17;115(16):E3692–701.

31. Slattery M, Riley T, Liu P, Abe N, Gomez-Alcala P, Dror I, et al. Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. Cell. 2011 Dec 9;147(6):1270–82.

32. Crocker J, Abe N, Rinaldi L, McGregor AP, Frankel N, Wang S, et al. Low Affinity Binding Site Clusters Confer Hox Specificityand Regulatory Robustness. Cell. Elsevier Inc; 2015 Jan 15;160(1-2):191–203.

33. Liu Z, Tjian R. Visualizing transcription factor dynamics in living cells. J Cell Biol. 2018 Jan 29.

34. Caussinus E, Kanca O, Affolter M. Fluorescent fusion protein knockout mediated by anti-GFP nanobody. Nat Struct Mol Biol. 2011 Dec 11;19(1):117–21.

35. Caussinus E, Kanca O, Affolter M. Protein knockouts in living eukaryotes using deGradFP and green fluorescent protein fusion targets. Curr Protoc Protein Sci. 2013 Sep 24;73:Unit30.2.

36. Crocker J, Noon EP-B, Stern DL. The Soft Touch: Low-Affinity Transcription Factor Binding Sites in Development and Evolution. 1st ed. Vol. 117, Essays on Developmental Biology Part B. Elsevier Inc; 2016. 15 p.

37. Liu Z, Tjian R. Visualizing transcription factor dynamics in living cells. J Cell Biol. 2018 Apr 2;217(4):1181–91.

38. Port F, Chen HM, Lee T, Bullock SL. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. Proceedings of the National Academy of Sciences. 2014 Jul 7.

39. Voutev R, Mann RS. Robust ΦC31-Mediated Genome Engineering in Drosophila melanogasterUsing Minimal attP/attB Phage Sites. G3 (Bethesda). 2018 May 4;8(5):1399–402.

40. Kondo S, Ueda R. Highly Improved Gene Targeting by Germline-Specific Cas9 Expression in Drosophila. Genetics. 2013 Sep 3.

41. Ryoo HD, Mann RS. The control of trunk Hox specificity and activity by Extradenticle. Genes Dev. 1999 Jul 1;13(13):1704–16.

42. Mann RS, Abu-Shaar M. Nuclear import of the homeodomain protein extradenticle in response to Wg and Dpp signalling. Nature. 1996 Oct 17;383(6601):630–3.

43. Slattery M, Voutev R, Ma L, Négre N, White KP, Mann RS. Divergent Transcriptional Regulatory Logic at the Intersection of Tissue Growth and Developmental Patterning. Desplan C, editor. PLoS Genet. 2013 Sep 5;9(9):e1003753.

44. Estella C, McKay DJ, Mann RS. Molecular integration of wingless, decapentaplegic, and autoregulatory inputs into Distalless during Drosophila leg development. Dev Cell. 2008 Jan;14(1):86–96.

45. Gebelein B, Culi J, Ryoo HD, Zhang W, Mann RS. Specificity of Distalless repression and limb primordia development by abdominal Hox proteins. Dev Cell. 2002 Oct;3(4):487–98.

Figure 1

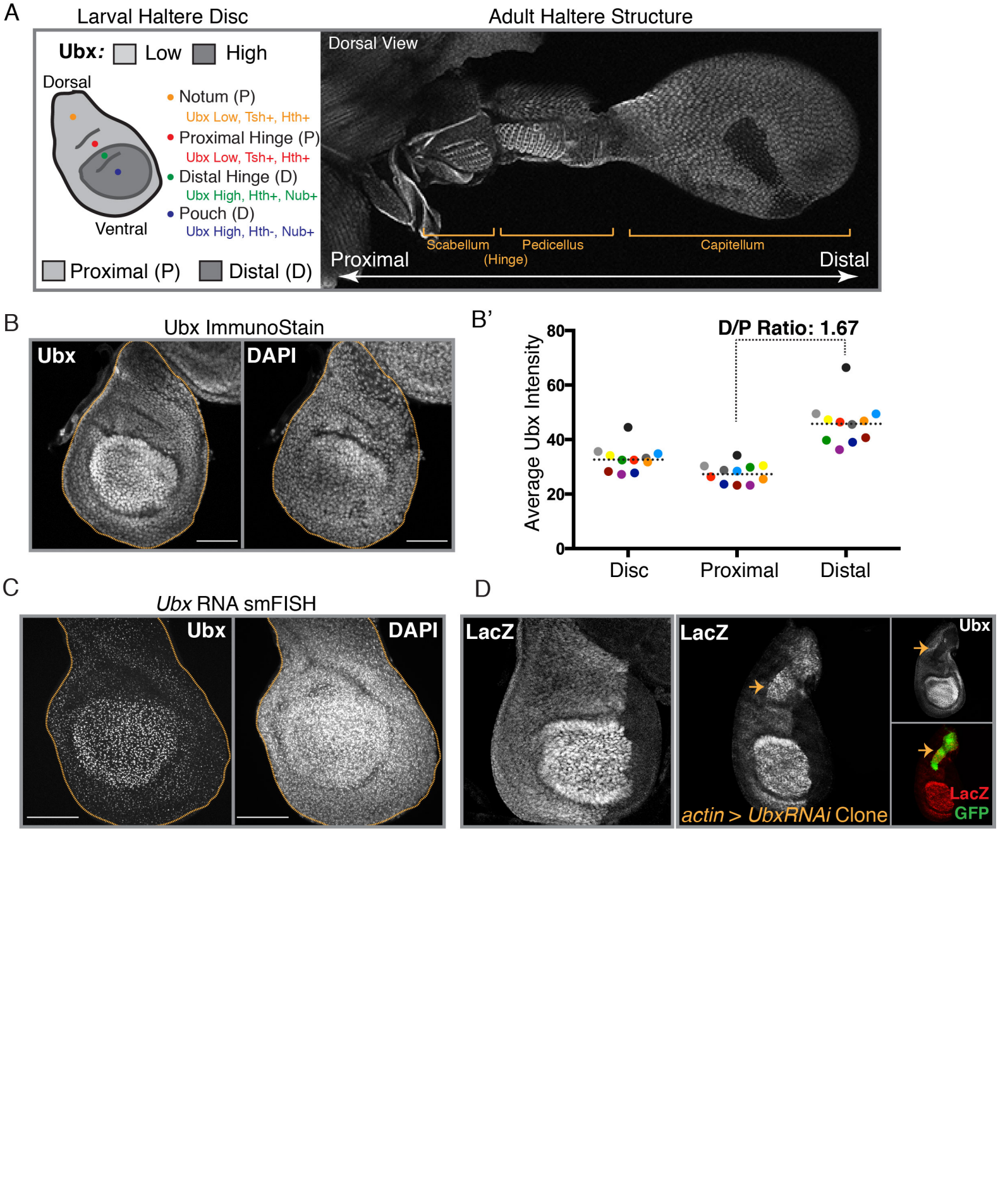


Figure 2

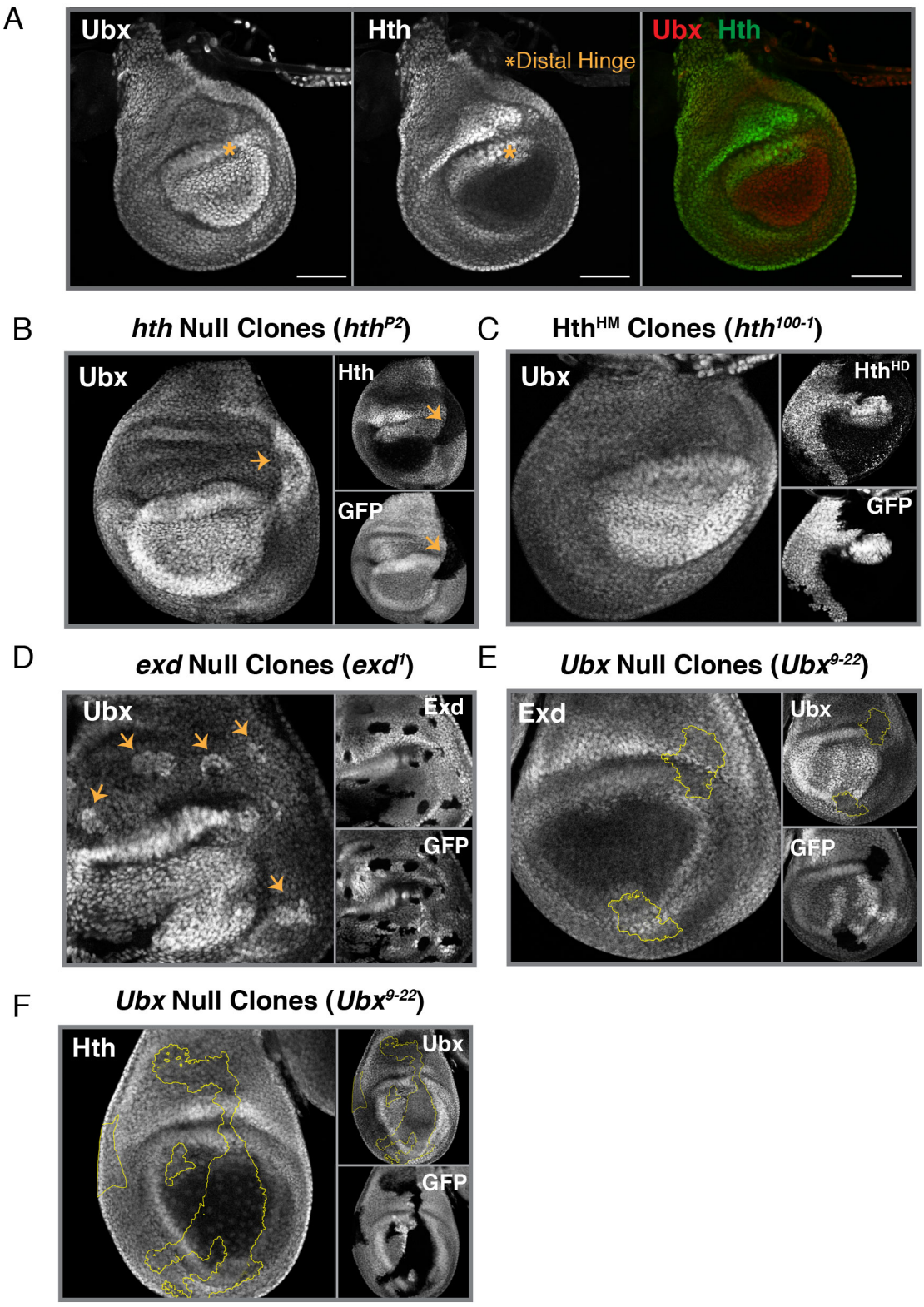


Figure 3

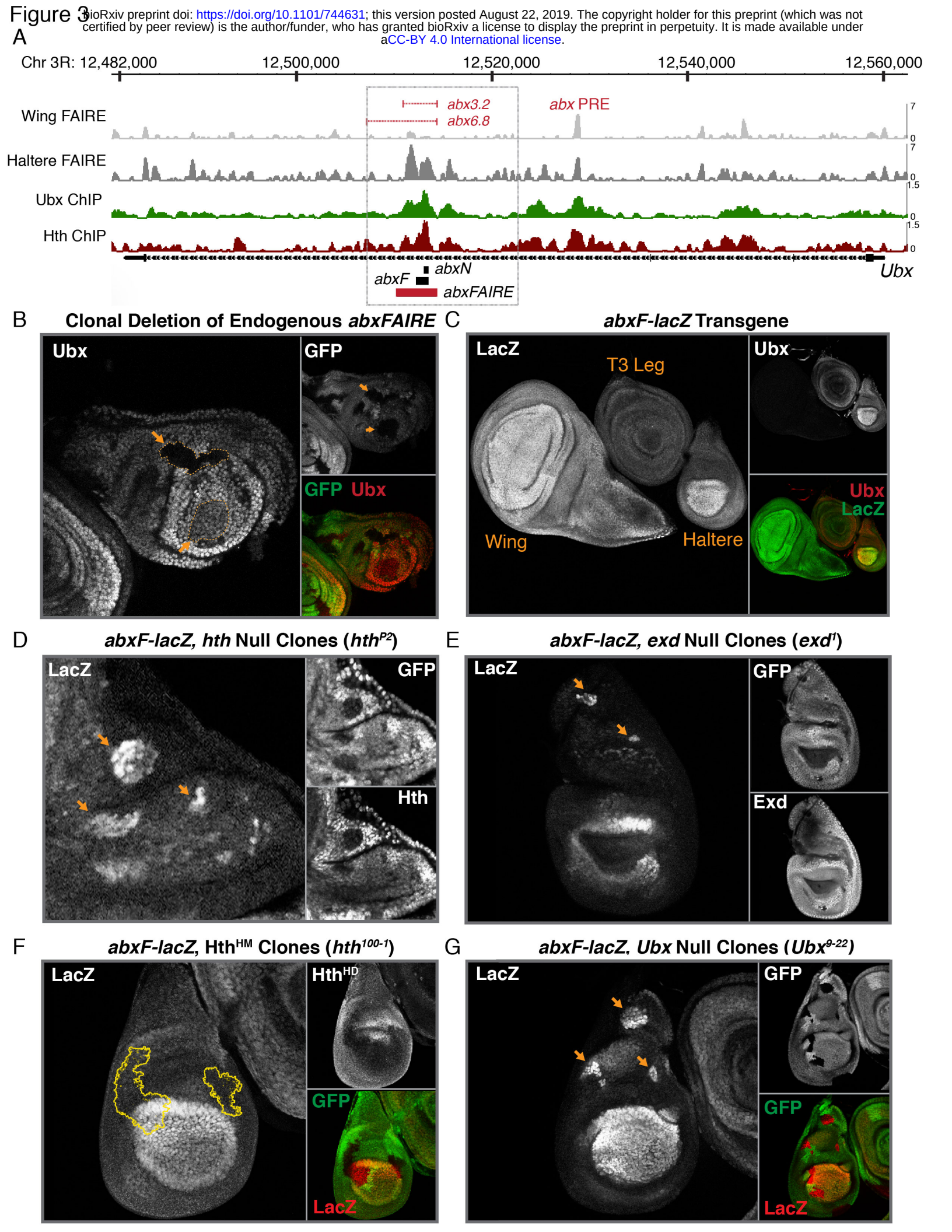


Figure 4

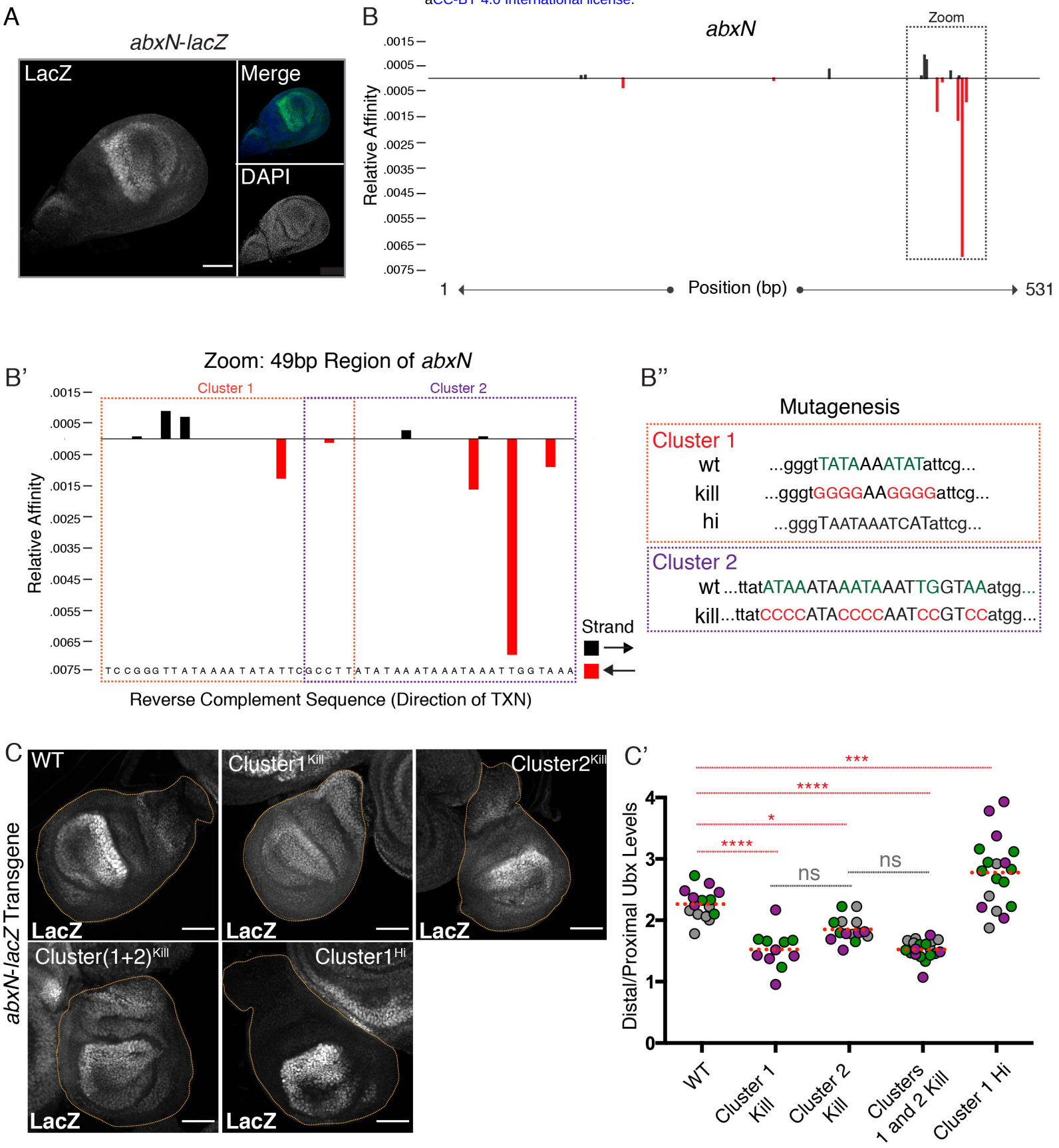


Figure 5

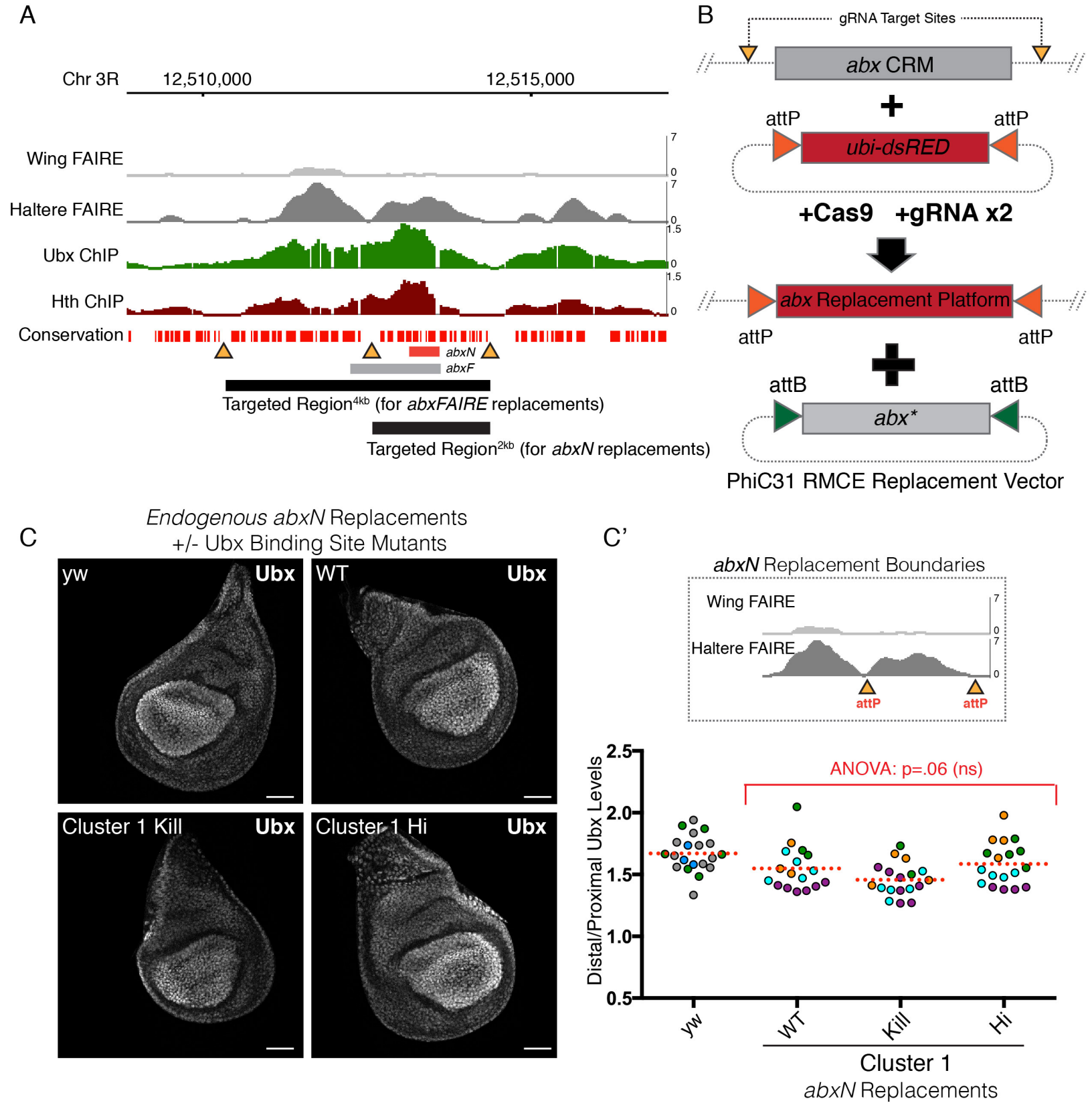
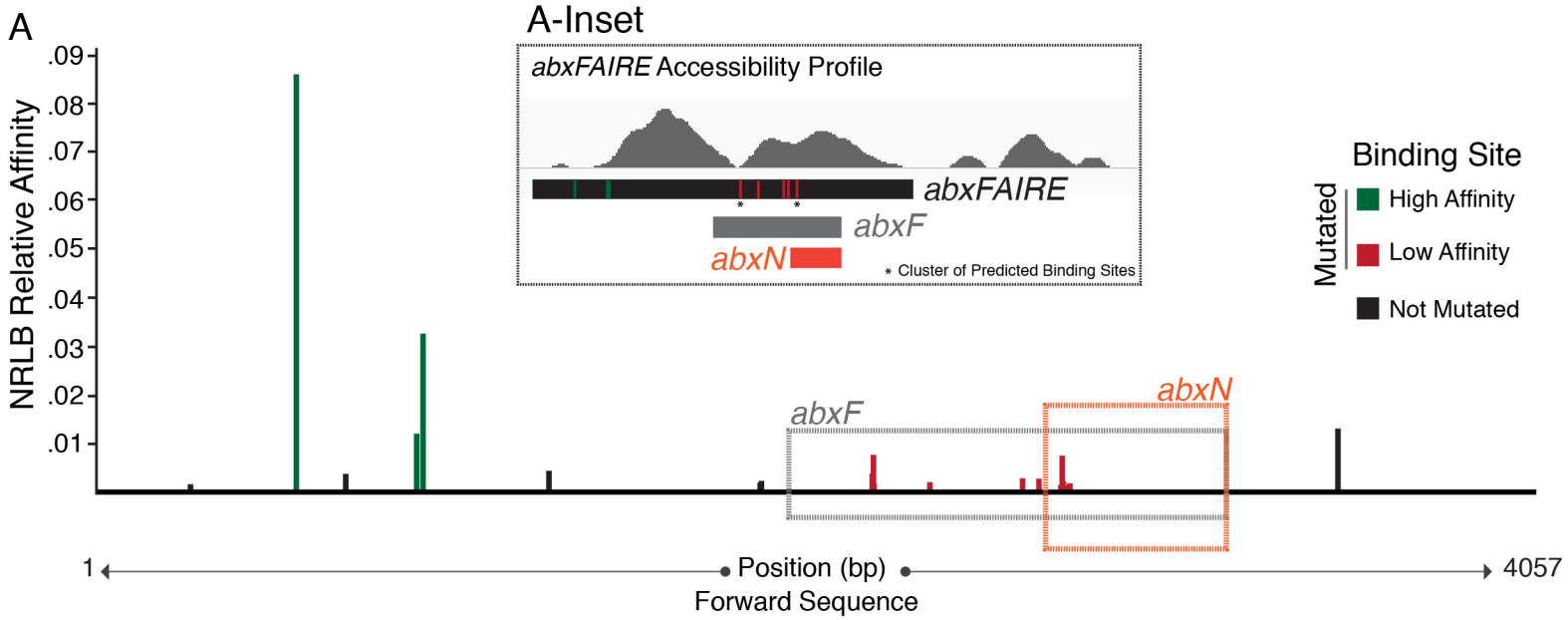
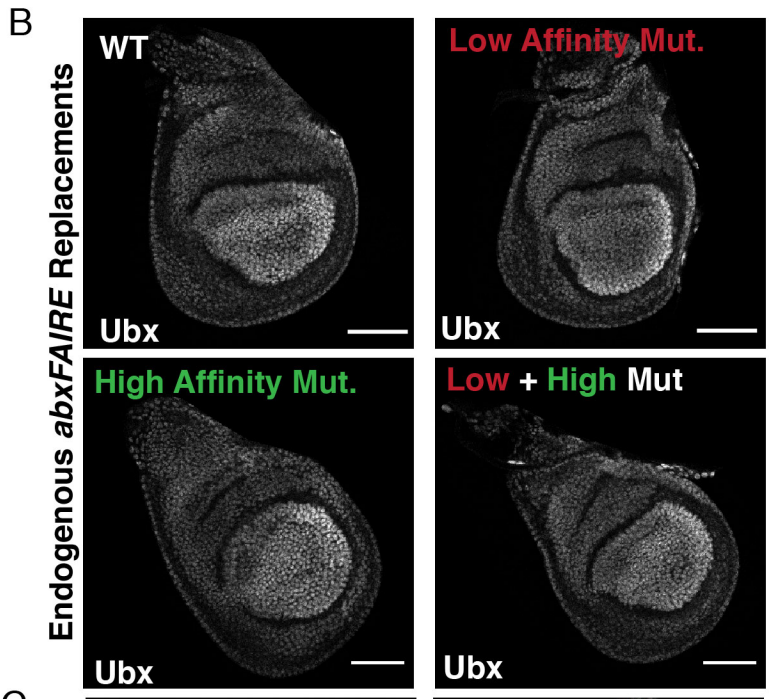
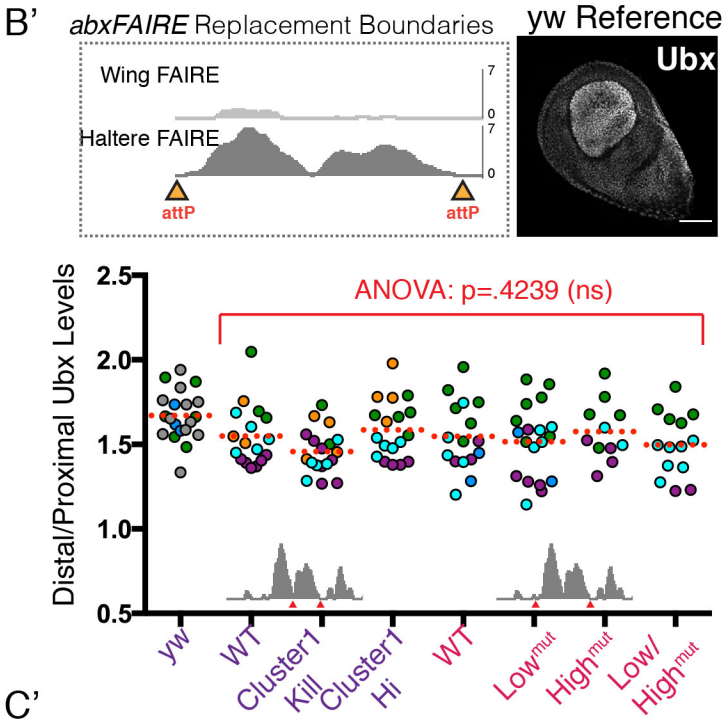
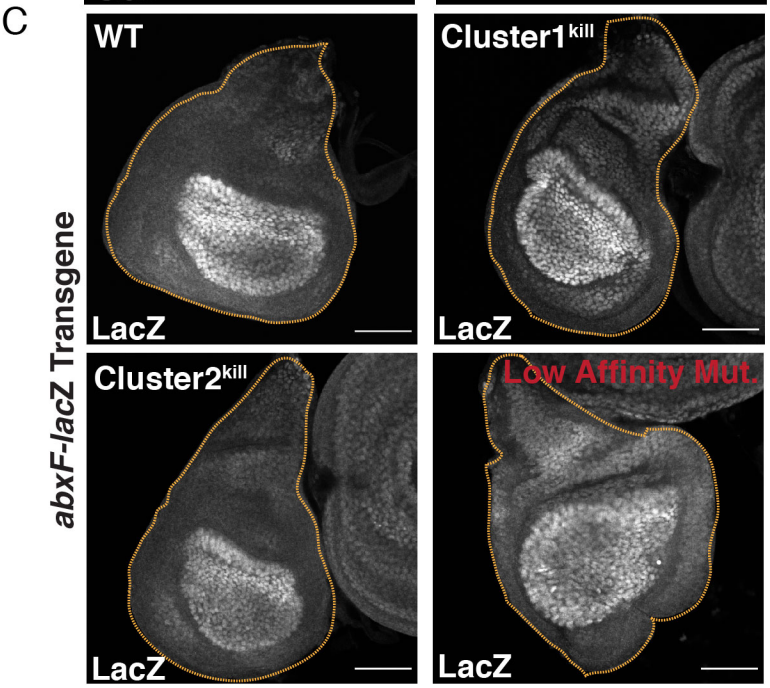


Figure 6









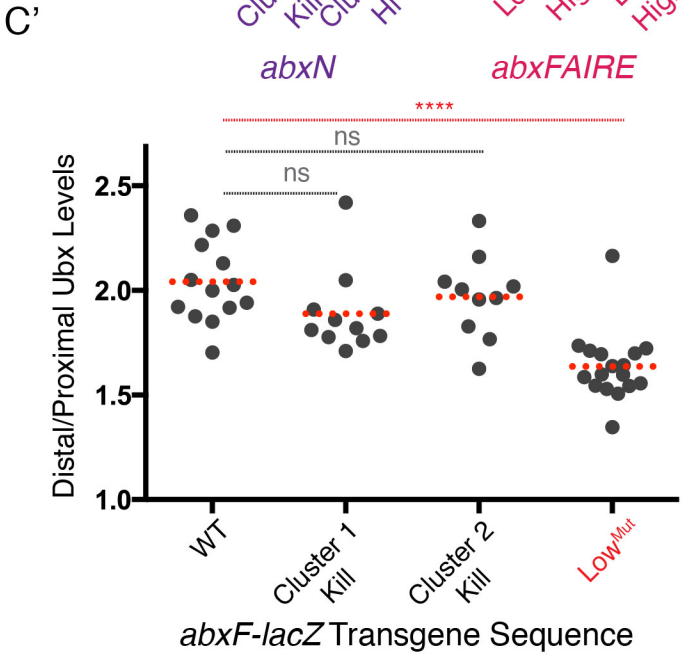


Figure 7

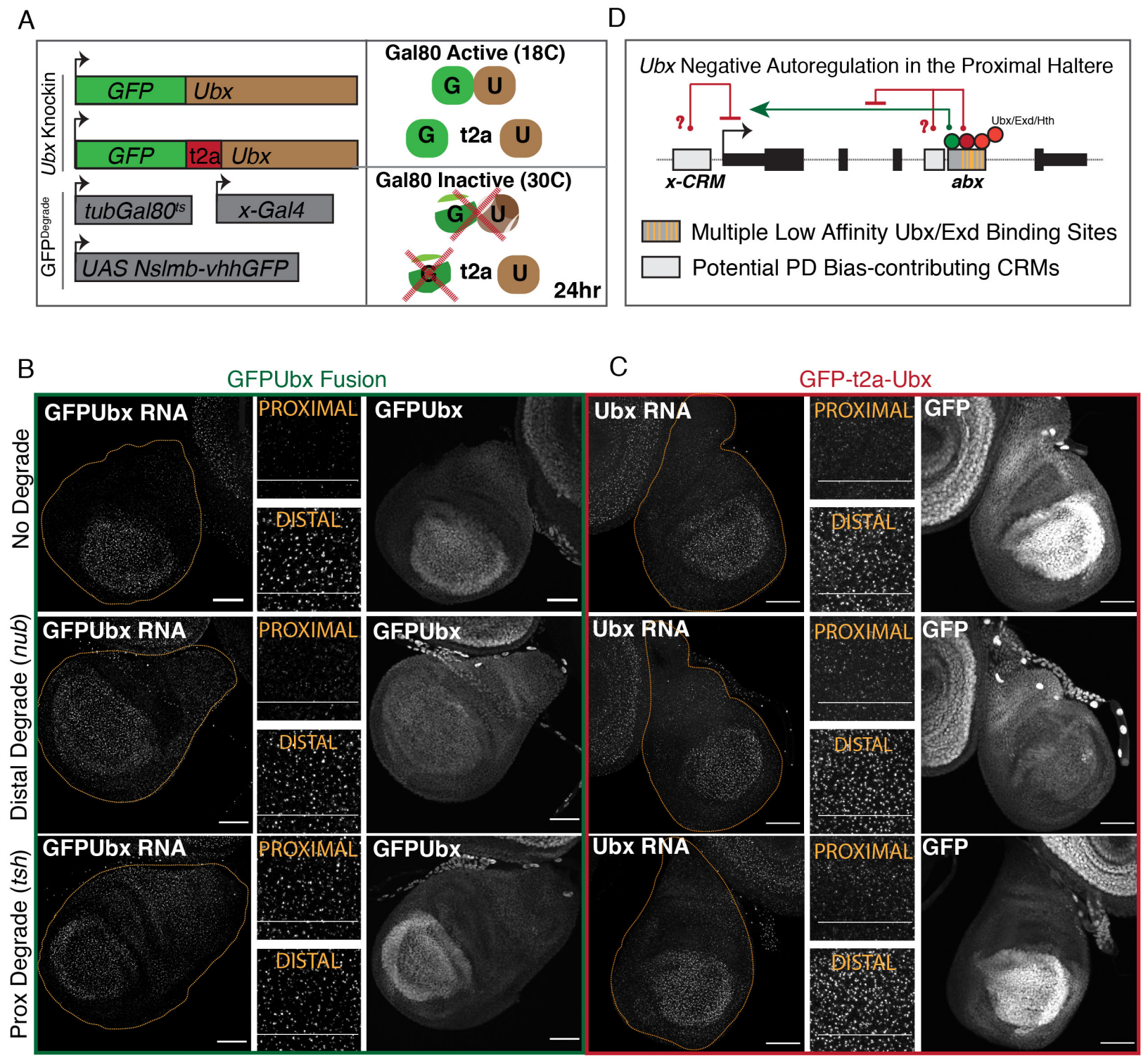
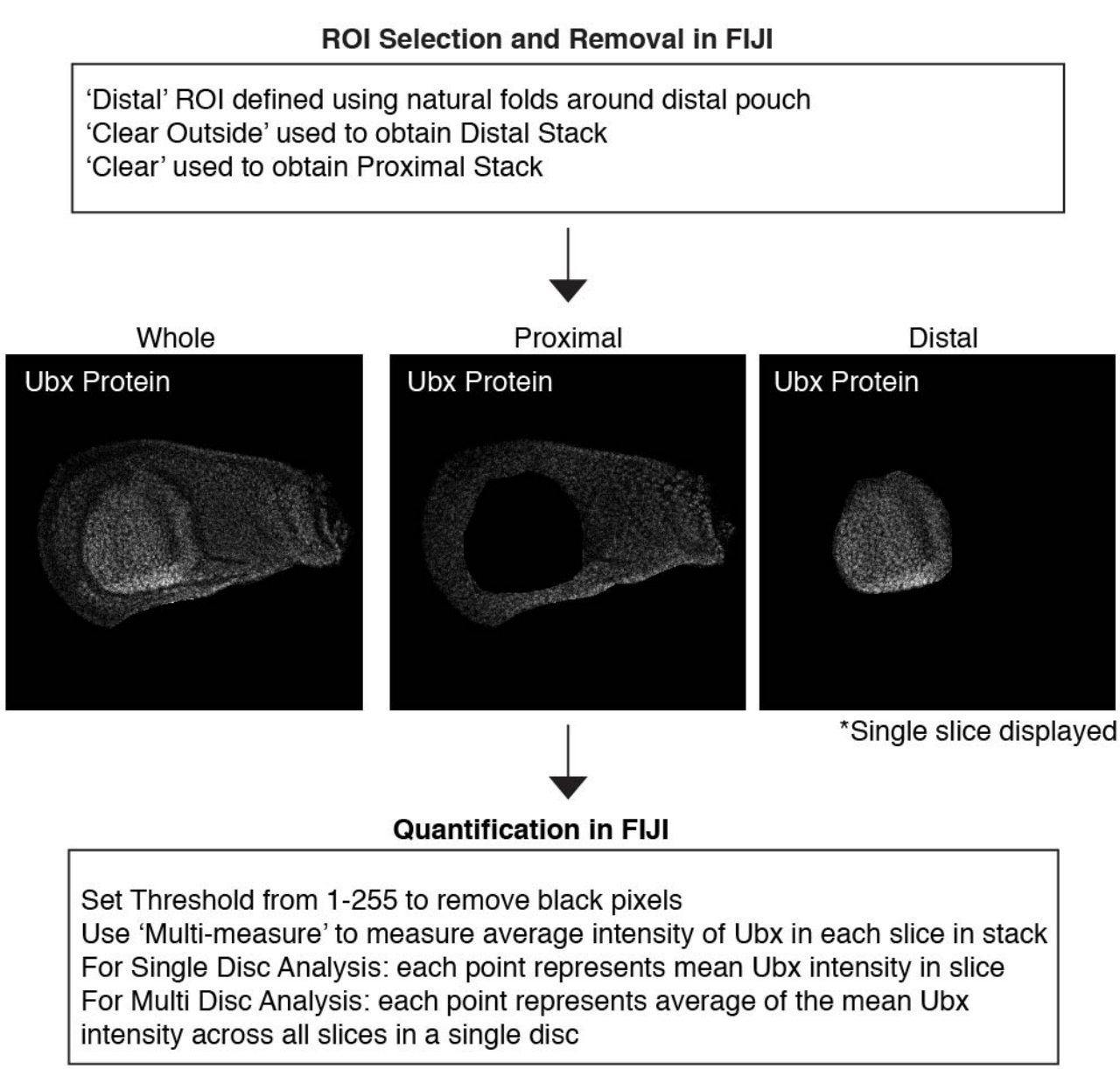
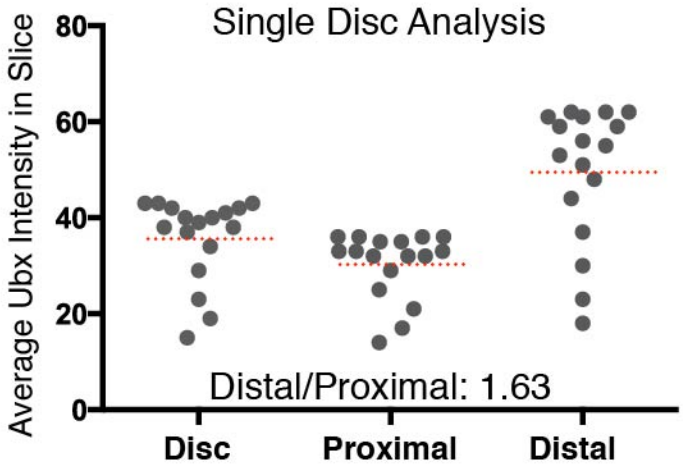


Figure 1-Supplemental 1

A



B



C

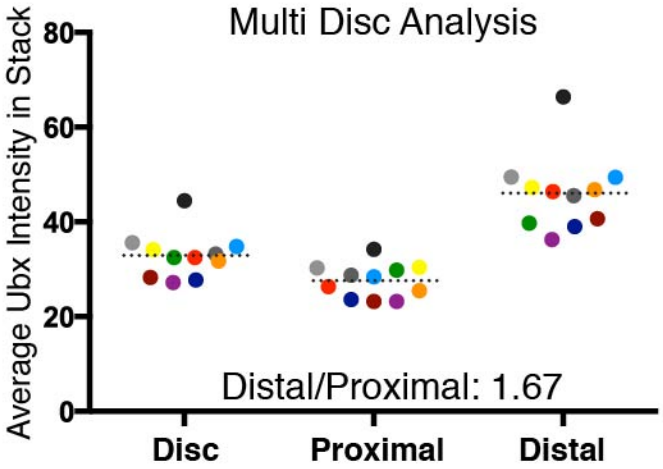


Figure 2 Supplemental 1

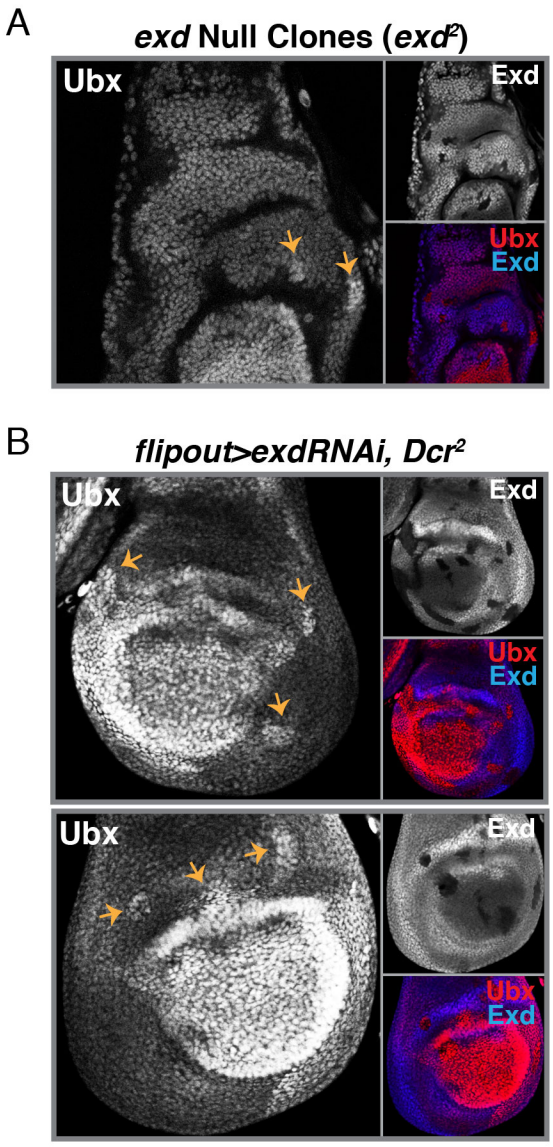


Figure 3 Supplemental 1

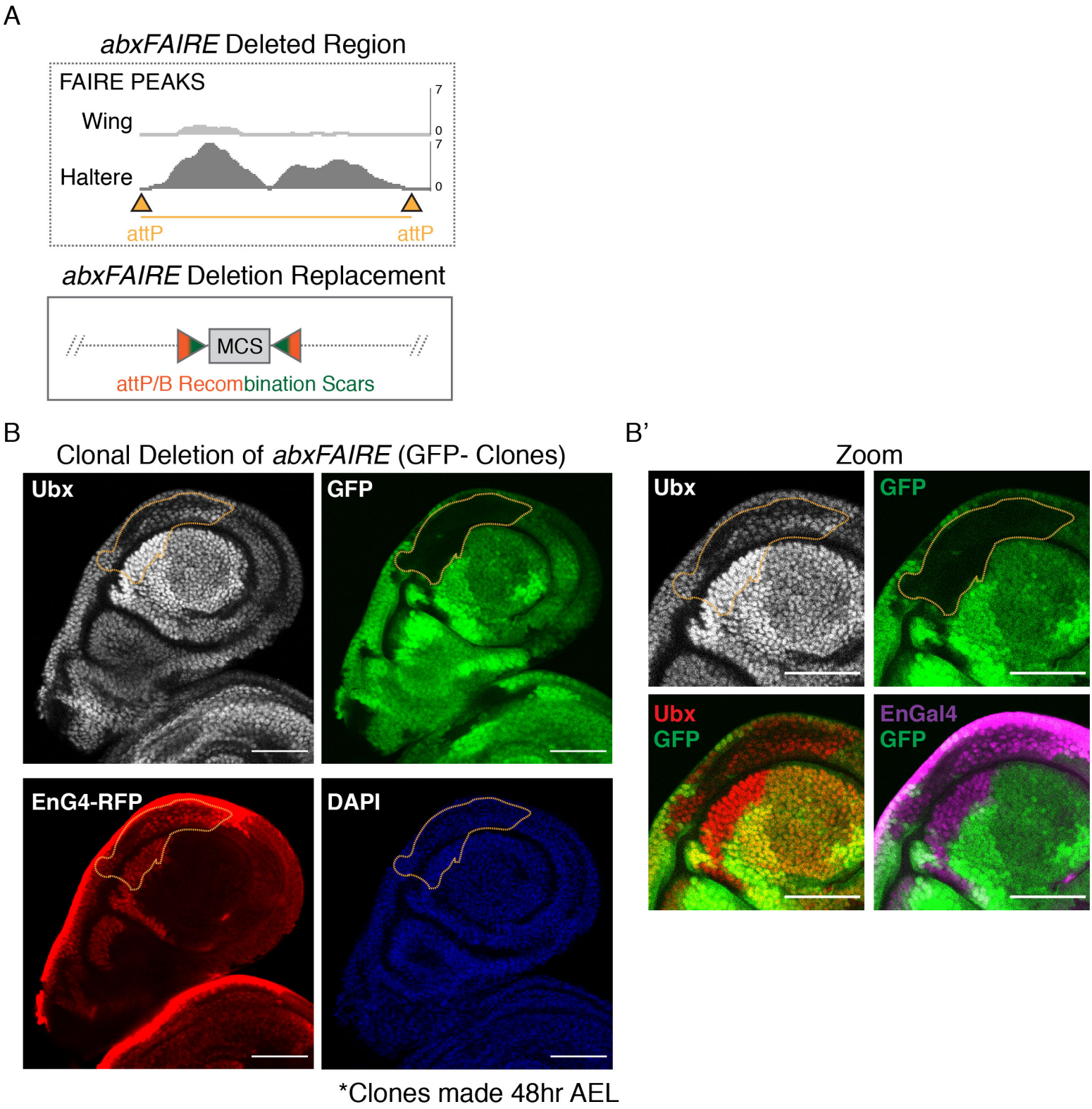
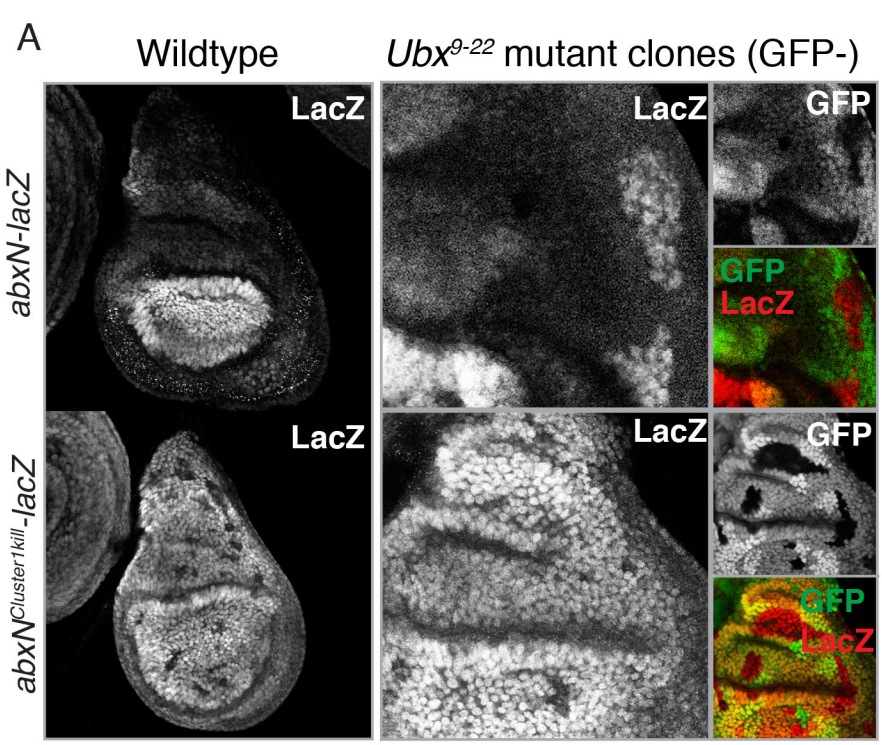
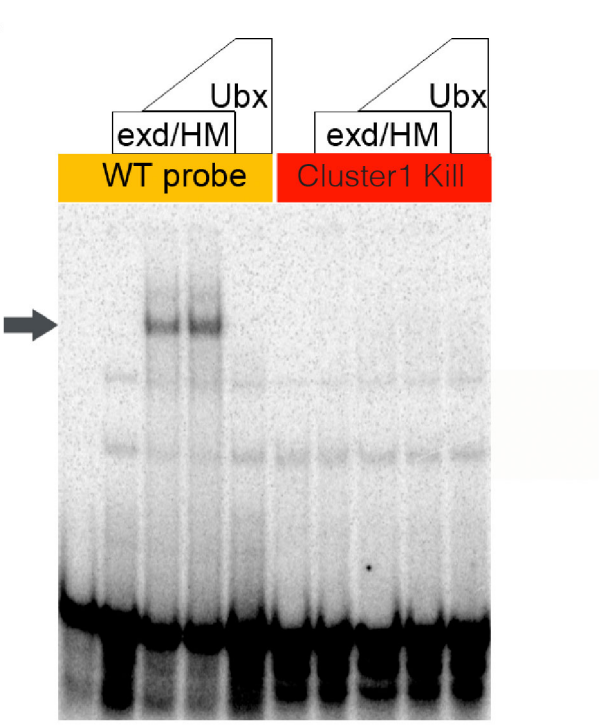


Figure 4 - Supplemental 1





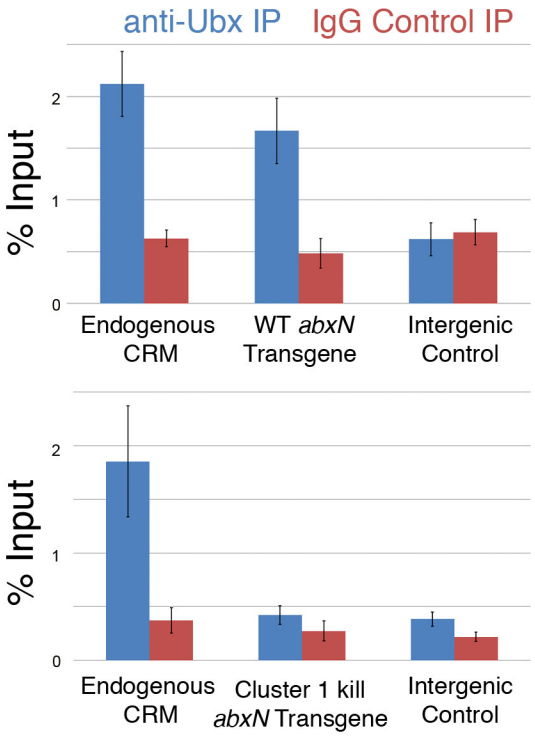


Figure 5 Supplemental 1

A

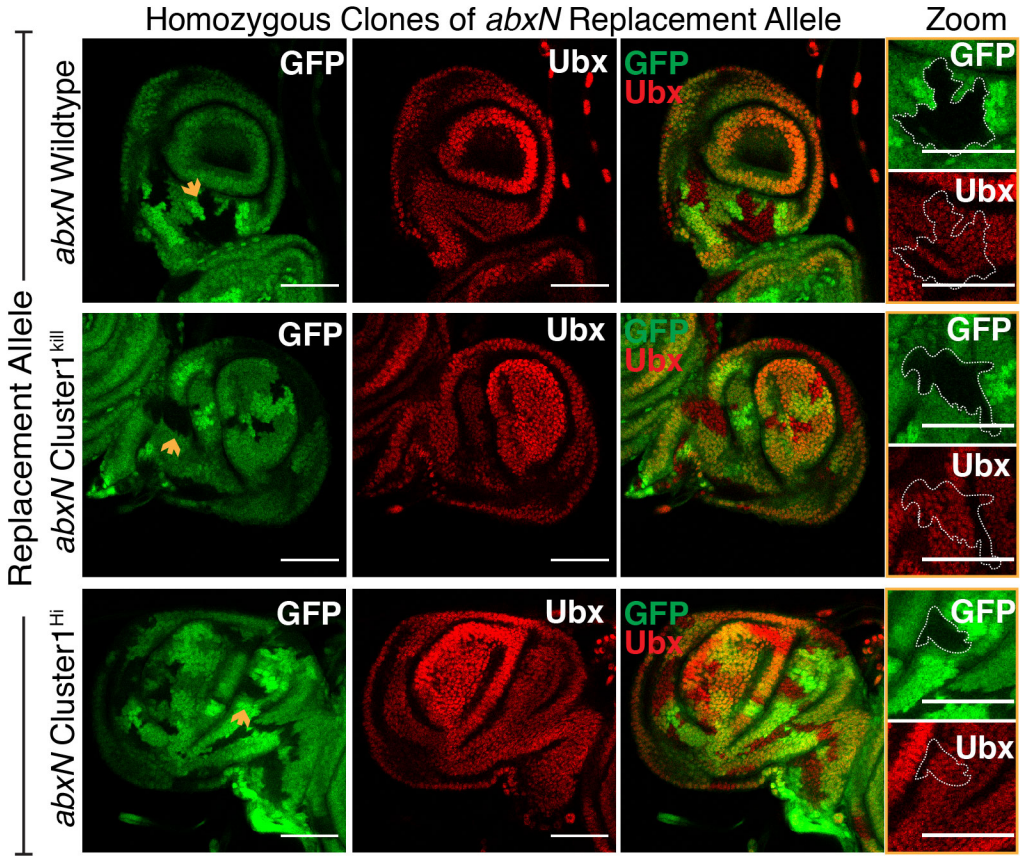


Figure 6 - Supplemental

A

Name	# Predicted BS	Minus Strand Sequence WT Sequence	Minus Strand Sequence Mutant Sequence	NRLB Relative Affinities	Category
C1	2	CCGGGTTatAAaaTAtaTTCGCC	CCGGGTTggAAggTAggTTCGCC	8.96E-4, 1.28E-3	Low
C2	3	GCCTTatATAAATAaaTAaaTTGGTA	GCCTTggATggATAggTAggTTGGTA	9.03E-4, 1.6E-3, 6.9E-3	
BS3	1	CGATGAtaaATGCGACAA	CGATGAgggATGCGACAA	2.2E-3	
BS4	1	ATTGAatttATTTGCTCG	ATTGAggggATTTGCTCG	2.3E-3	
BS5	1	GGGAAttgTCAATCATCG	GGGAAcccTCAATCATCG	1.5E-3	
C6	3	ATATtaTTAatAAAtaTTtCATTT	ATATggTTAggAAAggTTgCATTT	1.2E-3, 3.3E-3, 7E-3	
BS7	1	CTGTGatttATGGCTGGA	CTGTGccccATGGCTGGA	8.5E-2	High
C8	2	TCGAGatttATGGTCCATG CGGCAGTTGTaaatCAAAT	TCGAGccccATGGTCCATG CGGCAGTTGTggggCAAAT	1.1E-2, 3.2E-2	

"C"= Cluster, "BS"= Binding Site

B

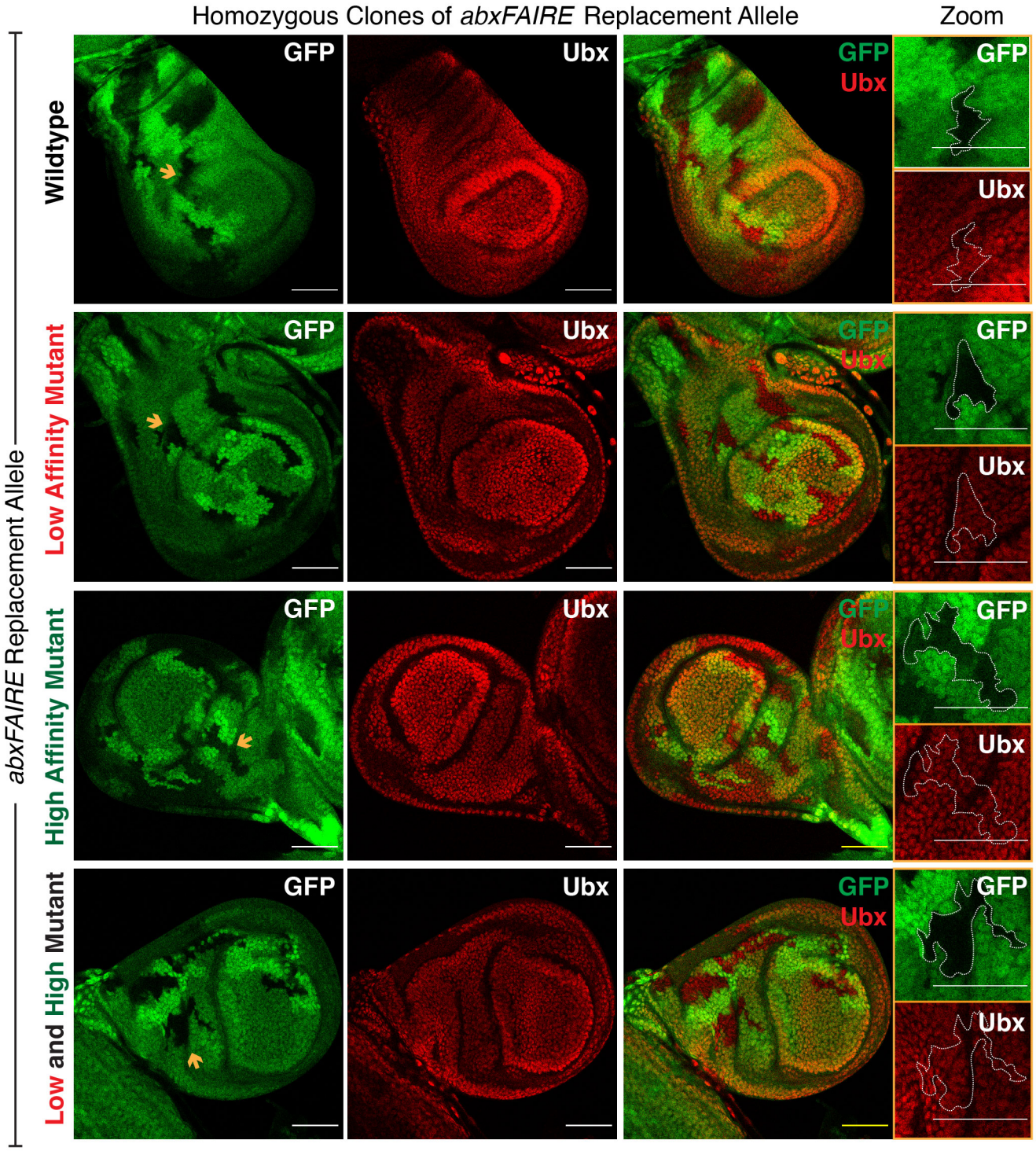


Figure 6 - Supplemental 2

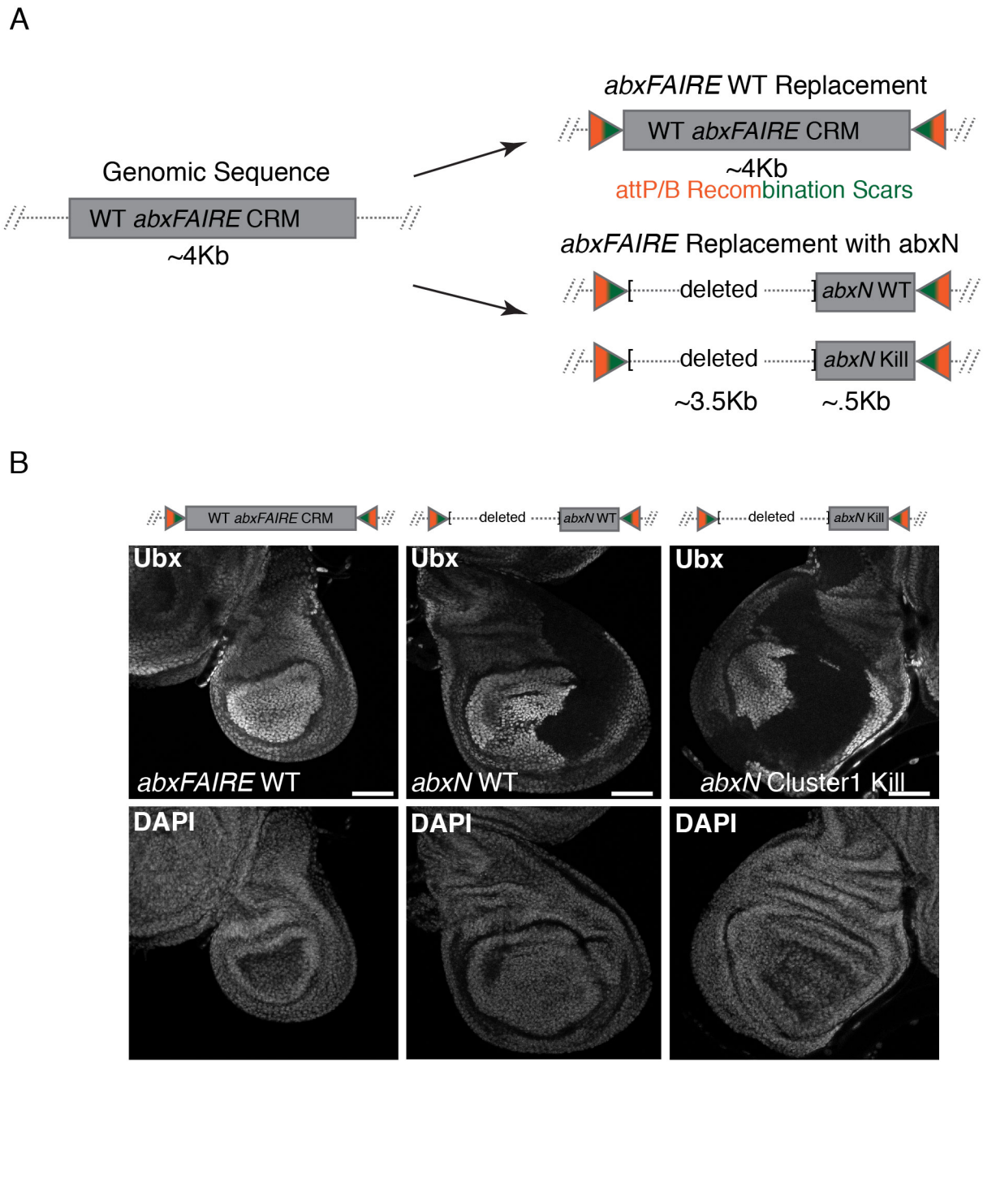
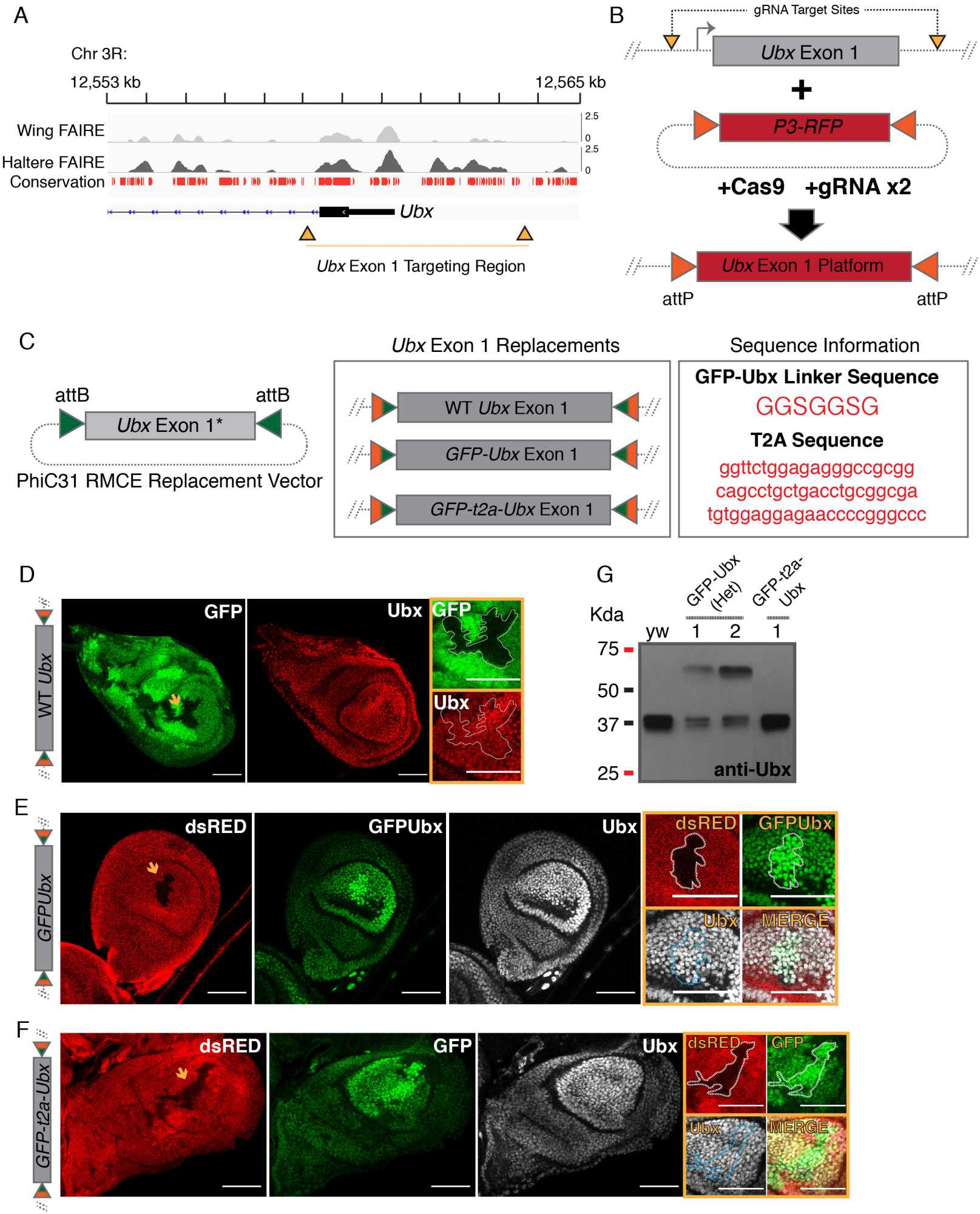


Figure 7-Supplemental 1



Supplemental Figure Legends

Figure 1 Supplemental 1. Overview of Strategy to Quantify the Proximal/Distal Bias of Ubx Expression. (A) Analysis pipeline conducted within FIJI is shown. Selection of whole disc region of interest (ROI), Distal ROI, and Proximal ROI is followed by the measurement of the average intensity of Ubx within that ROI for each slice in the stack. Black pixels are removed from the analysis to prevent skewing of the average from differences in sizes of the ROI. **(B)** (Left) An analysis of a single haltere disc is shown. Each dot represents the average intensity for each slice in the stack in the whole disc ("Disc"), the proximal ROI, and the distal ROI. Dotted red line represents the mean for each compartment; reported distal/proximal ratio is the distal mean/proximal mean. **(C)** As shown in Figure 1. The mean average Ubx intensity for each compartment from the single disc analysis is reported as a single point in the multi-disc analysis. Individual discs are color-coded such that the compartment-specific averages for each disc can be compared to one another. Dotted line is the mean average intensity for each compartment.

Figure 2 Supplemental 1. Additional *exd* mutant Clones. (A) Ubx Immunostain in haltere discs in which *exd* null clones (exd^{2}) have been induced. An Exd immunostain, in addition to a merge of Ubx/Exd, is shown. Clones are marked by a yellow arrow. **(B)** Ubx immunostain in two haltere discs in which *exdRNAi* clones have been induced in the background of a *Dcr2* mutant. An Exd immunostain, in addition to a merge of Ubx/Exd, is shown. Clones are marked by a yellow arrow.

Figure 3 Supplemental 1. Clonal Deletion of *abxFAIRE*. (A) (Top) A schematic of the *abxFAIRE* region targeted and deleted. FAIRE accessibility peaks from McKay *et al.* (29) are shown in the wing and haltere. (Bottom) The Δ*abxFAIRE* allele was generated by replacing the *abxFAIRE* sequence with a minimal cloning site (MCS) sequence using PhiC31-based RMCE. Through this method scars are left on either side of the replacement. **(B)** Mitotic clones homozygous for Δ*abxFAIRE* (GFP-) were induced at 48hr after egg-laying (AEL). Clones in the posterior compartment (*EnGal4*-RFP+) do not show a defect in Ubx expression. Clone is outlined in yellow. **(B')** A zoomed in image of the clone from B. Merges of Ubx/GFP and *EnGal4*-RFP/GFP are shown. Scale bars are 50 micron in size.

Figure 4 Supplemental 1. Loss of Cluster 1 binding sites Results in Loss of Ubx Binding. (A) (Left) LacZ immunostain of haltere discs containing either an *abxN-lacZ*^{WT} transgene (top) or an *abxN-lacZ*$^{Cluster1\text{-}Kill}$ transgene. (Right) LacZ immunostain of discs of wildtype and mutant *abxN-lacZ* genotype upon induction of *Ubx* null clones (*Ubx*$^{9\text{-}22}$). A GFP immunostain, in addition to a merge of GFP/LacZ, is shown. Clones are GFP-negative. **(B)** EMSA assay of Ubx-Exd-Hth^{HM} *in vitro* binding to a probe containing the Cluster 1 binding sites (left) and the probe mutated to abrogate the Cluster 1 binding sites (right). Gray arrow points to the shifted trimer band. **(C)** Chromatin Immunoprecipitation (ChIP-qPCR) was performed on haltere discs from transgenic flies containing either the *abxN-lacZ*^{WT} reporter (top) or the *abxN-lacZ*$^{Cluster1\text{-}Kill}$ reporter (bottom). % Input for an anti-Ubx IP and an IgG Isotype control IP are shown for three

genomic regions: (1) the endogenous *abx* CRM, (2) the transgenic CRM, and (3) an intergenic region on chromosome 2 that serves as a negative control. Averages and standard deviation from three independent IPs are reported.

Figure 5 Supplemental 1. Generation of Clones Homozygous for Wildtype and Mutant *abxN* replacement alleles. (A) GFP and Ubx immunostains in haltere discs in which clones homozygous for *abxN* replacement alleles were induced 48hr AEL. Clones are GFP-negative and denoted with a yellow arrow. Zoomed images of single clones (outlined) are shown to the right. All scale bars shown are 50 micron in size.

Figure 6 Supplemental 1. Multiple Low Affinity Ubx/Exd Binding Sites are Required for PD Bias Formation. (A) Table of *NRLB*-predicted Ubx/Exd binding sites in *abxFAIRE* that have been mutated. Many predicted sites fall within clusters so we report the number of binding sites within each cluster. Wildtype and mutated sequences are shown. Lowercase levels are bases that were mutated. Relative affinities and the category (Low, High) for each binding site or cluster are given. **(B)** GFP and Ubx immunostains in haltere discs in which clones homozygous for *abxFAIRE* replacement alleles were induced 48hr AEL. Clones are GFP-negative and denoted with a yellow arrow. Zoomed images of single clones (outlined) are shown to the right. All scale bars shown are 50 micron in size.

Figure 6 Supplemental 2. Multiple Low Affinity Ubx/Exd Binding Sites are Required for PD Bias Formation. (A) Schematic of *abxFAIRE* replacements tested in **B**. The Targeted Region^{4kb} replacement platform was used to generate either a wildtype *abxFAIRE* (+scars) allele or an allele with only the ~500 bp *abxN* wildtype or Cluster 1^{kill} sequence, deleting the remaining ~3.5 kb *abxFAIRE* sequence. **(B)** Ubx immunostain in haltere discs homozygous for the specified *abxFAIRE* replacement alleles. A DAPI nuclear stain is shown for each. All scale bars shown are 50 micron in size.

Figure 7 Supplemental 1. CRISPR Targeting of *Ubx* Exon 1. (A) Genome browser screenshot around *Ubx* exon 1. FAIRE accessibility peaks in the wing and haltere (from McKay *et al.* (29)) are shown for reference. Conservation track was downloaded from UCSC. No red denotes a lack of evolutionary conservation with twelve other *Drosophila* species, mosquito, honeybee and red flour beetle. Yellow triangles denote Cas9/gRNA target sites; and the yellow bar denotes the region replaced. **(B)** Schematic of a two-step *Ubx* exon 1 replacement strategy. gRNA sites were chosen in non-conserved regions surrounding *Ubx* exon 1. A dual-gRNA expressing plasmid was injected along with a donor cassette containing an attP-flanked fluorescent selection marker (*P3-RFP*) into *nanos-Cas9* flies. The resulting "*Ubx* Exon 1 Replacement Platform" serves as a means to insert modified versions of *Ubx* using PhiC31-based RMCE. **(C)** (Left) Schematic of RMCE replacement cassette. (Right) Schematic of replacement alleles generated and sequence information for the linker used in the GFP-Ubx fusion (amino acids) and the T2A sequence used (DNA). **(D)** GFP native

fluorescence and a Ubx immunostain in haltere discs in which clones homozygous for the wildtype *Ubx* exon 1 replacement allele were induced. Clones are GFP-negative and marked with a yellow arrow. Cropped images of single clones (outlined) are shown. **(E)** GFPUbx native fluorescence, and dsRed and Ubx immunostains in haltere discs in which clones homozygous for the GFPUbx fusion allele were induced. Clones are dsRed-negative and marked with a yellow arrow. Cropped images of a single clone (outlined) are shown. **(F)** GFPUbx native fluorescence, and dsRed and Ubx immunostains in haltere discs in which clones homozygous for the *GFP-t2a-Ubx* allele were induced. Clones are dsRed-negative and marked with a yellow arrow. Cropped images of a single clone (outlined) are shown. **(G)** An anti-Ubx immunoblot on protein derived from the following genotypes: *yw* (lane 1), *GFP-Ubx* heterozygous (lanes 2, 3), and *GFP-t2a-Ubx* (lane 4). The lower band is Ubx and the upper band is the GFPUbx fusion. All scale bars shown are 50 micron in size.