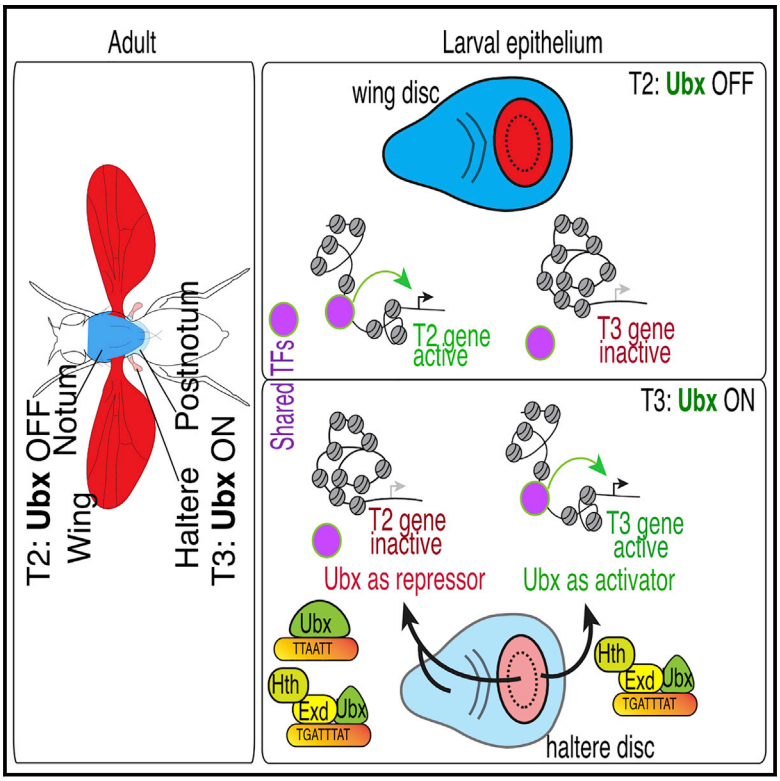
Article

Current Biology

Cell-type-specific Hox regulatory strategies orchestrate tissue identity

Graphical abstract



Highlights

- The Hox protein Ubx increases and decreases chromatin accessibility in the haltere
- These changes correlate with activation and repression of gene expression
- The direction of change depends on the position within the haltere disc
- Ubx-mediated changes in accessibility alter the binding of other shared regulators

Authors

Ryan Loker, Jordyn E. Sanner, Richard S. Mann

Correspondence

rsm10@columbia.edu

In brief

In *Drosophila*, mutations in the *Ubx* Hox gene result in a homeotic transformation, in which the second thoracic segment is duplicated in place of the third thoracic segment. Loker et al. use genome-wide approaches to understand how a single Hox factor orchestrates the dramatic transformation of an entire segment that contains multiple cell types.

Loker et al., 2021, Current Biology *31*, 1–10
October 11, 2021 © 2021 Elsevier Inc.
https://doi.org/10.1016/j.cub.2021.07.030

CellPress

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

Current Biology

CellPress

Article

Cell-type-specific Hox regulatory strategies orchestrate tissue identity

Ryan Loker,[1,2,5] Jordyn E. Sanner,[2,5] and Richard S. Mann[2,3,4,5,6,7,*]
[1]Department of Genetics and Development, Columbia University Irving Medical Center, New York, NY, USA
[2]Department of Biochemistry and Molecular Biophysics, Columbia University Irving Medical Center, New York, NY, USA
[3]Department of Neuroscience, Columbia University Irving Medical Center, New York, NY, USA
[4]Department of Systems Biology, Columbia University Irving Medical Center, New York, NY, USA
[5]Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA
[6]Twitter: @richmann10
[7]Lead contact
*Correspondence: rsm10@columbia.edu
https://doi.org/10.1016/j.cub.2021.07.030

SUMMARY

Hox proteins are homeodomain transcription factors that diversify serially homologous segments along the animal body axis, as revealed by the classic bithorax phenotype of *Drosophila melanogaster*, in which mutations in *Ultrabithorax* (*Ubx*) transform the third thoracic segment into the likeness of the second thoracic segment. To specify segment identity, we show that Ubx both increases and decreases chromatin accessibility, coinciding with its dual role as both an activator and repressor of transcription. However, the choice of transcriptional activity executed by Ubx is spatially regulated and depends on the availability of cofactors, with Ubx acting as a repressor in some populations and as an activator in others. Ubx-mediated changes to chromatin accessibility positively and negatively affect the binding of Scalloped (Sd), a transcription factor that is required for appendage development in both segments. These findings illustrate how a single Hox protein can modify complex gene regulatory networks to transform the identity of an entire tissue.

INTRODUCTION

Among the most famous mutant phenotypes in modern biology is the four-winged "bithorax" fly, in which the third thoracic (T3) segment of *Drosophila melanogaster* is transformed into a nearly complete copy of the second thoracic (T2) segment.[1,2] This dramatic homeotic transformation of segment identity is caused by loss-of-function mutations in the homeotic selector gene *Ultrabithorax* (*Ubx*), which is required to modify the "ground-state" segment identity of T2 into T3.[1–3] *Ubx* is one of eight paralogous Hox genes in *Drosophila*, all of which encode homeodomain transcription factors (TFs). Each Hox gene is expressed in a subset of segments along the anterior-posterior body axis of the fly and is responsible for determining their identities. Although the complexity in mammals is compounded by the existence of 39 Hox genes, loss-of-function mutations in the mouse establish that, as in the fly, Hox genes determine regional identities along the vertebrate body axis,[4,5] including the specification of multiple cell types.[6,7] Since their discovery, changes in Hox protein function and expression during development have been shown to be key drivers in the evolution of diversity among animal body plans.[8,9]

To modify the identity of an entire segment, a single Hox protein must function in parallel in multiple cell types, in each case by altering regulatory networks via gene activation and repression. While strides have been made to characterize how Hox proteins function as TFs, our understanding has been largely informed by analyzing individual *cis*-regulatory modules (CRMs) in disparate cell types.[10–16] A mechanistic understanding of how Hox proteins differentially modify gene regulatory networks in multiple cell populations to transform one tissue into another is lacking. A major barrier has been the technical hurdle of characterizing large sets of Hox-targeted CRMs in multiple cell types within a segment.

The dynamic interplay between TFs and nucleosome occupancy at CRMs, often analyzed through genome-wide measurements of chromatin accessibility, is an important mechanism by which cell-type-specific gene regulatory states are established.[17–21] Recent work has demonstrated that some, but not all, Hox proteins have the capacity to increase the accessibility of chromatin at target binding sites in multiple experimental systems, including Kc167 cell culture,[22,23] motor neuron induction from embryonic stem cells (ESCs),[24] and in the distal mammalian limb bud.[25] These studies suggest that the different potentials for Hox proteins to increase chromatin accessibility may contribute to their paralog-specific functions *in vivo*.

Here, we examine how a single Hox protein orchestrates the transformation of an entire segment. By comparing multiple populations of progenitors that give rise to the adult T2 and T3 segments of the fly, combined with profiling whole-genome TF binding, we directly assess the impact of Ubx on chromatin accessibility. We find that Ubx causes widespread cell-type-specific increases and decreases to chromatin accessibility that coincide with gene activation and repression, respectively. Importantly, this analysis revealed a previously unknown spatial regulation of Hox activity, wherein the function of Ubx as an

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

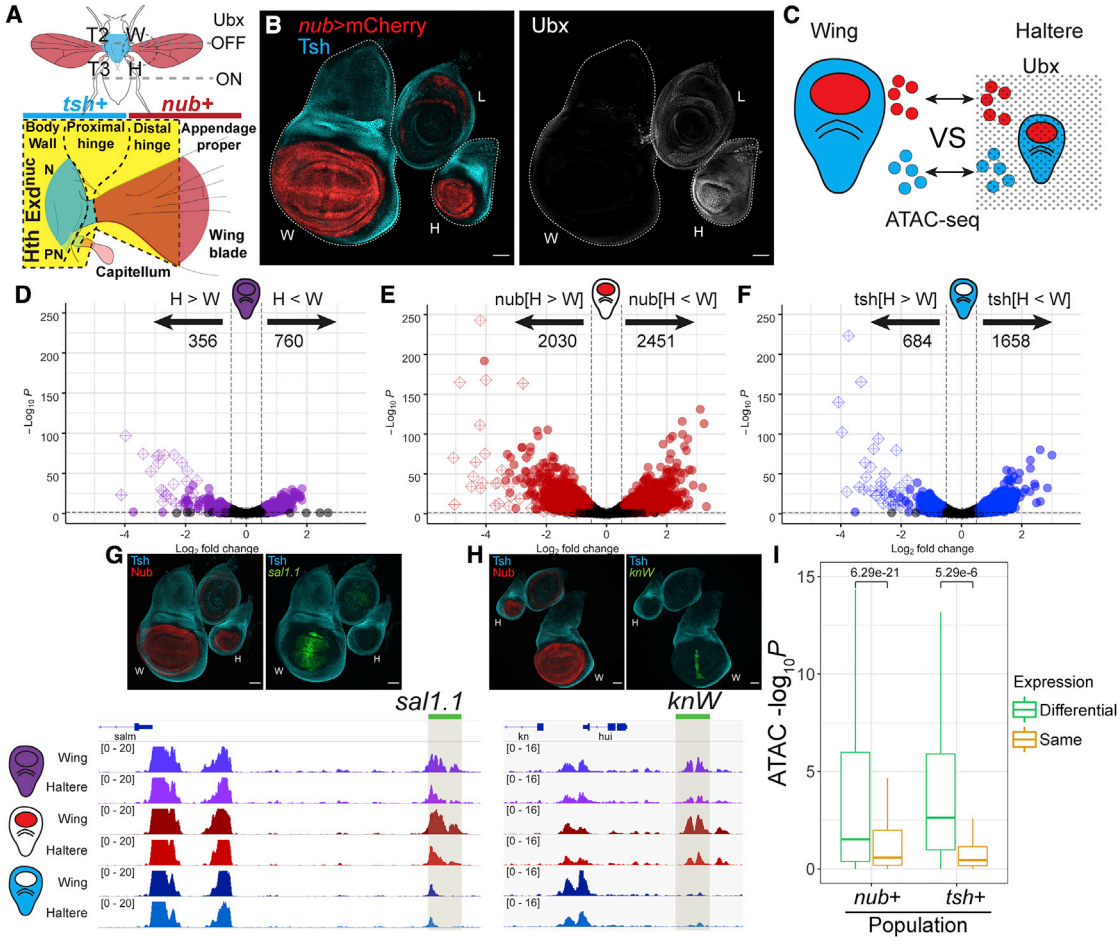


Figure 1. Segment-specific chromatin accessibility and gene expression in wing and haltere imaginal discs

(A) Schematics of an adult fly highlighting the contributions of the dorsal wing and haltere imaginal discs; the lower panel shows a magnified view of the proximal appendage (hinge) regions. For both the wing-bearing T2 and haltere-bearing T3 segments, blue marks body wall domains (notum [N] and postnotum [PN], respectively) and red marks the appendages (wing and haltere, respectively). The *tsh*$^+$ domain (blue) gives rise to the body and proximal hinge, while the *nub*$^+$ domain (red) gives rise to the distal hinge and appendage proper (wing blade and capitellum). The Hox cofactor Hth (yellow), which induces the nuclear localization of Exd (Exd^{nuc}), is expressed in the body wall, proximal hinge, and distal hinge, but is absent from the appendage proper.

(B) Left: immunostain of third larval instar wing (W) and haltere (H) imaginal discs showing distal (*nub*$^+$, red) and proximal (*tsh*$^+$, blue) populations. Also shown is the T3 leg imaginal disc (L). Right: Ubx is expressed throughout the haltere disc, and is absent from the wing disc. Scale bars for this and subsequent panels, 50 μm.

(C) Experimental scheme to compare chromatin accessibility using ATAC-seq in homologous distal (*nub*$^+$, red) and proximal (*tsh*$^+$, blue) populations of the wing and haltere imaginal discs. Dotted background indicates the presence of Ubx in all haltere imaginal disc cells.

(D–F) Genome-wide comparison of wing and haltere ATAC-seq data for whole tissue (D), *nub*$^+$ cells (E), and *tsh*$^+$ cells (F). Colored points satisfy a threshold of LFC > 0.5, adjusted $p < 0.05$. Diamond-shaped points are ATAC peaks within the *Ubx* genomic locus. A common set of 24,915 open chromatin regions, generated by merging ATAC-seq peaks in each sorted dataset, was used for comparisons.

(G and H) ATAC-seq genomic tracks at previously described Ubx target CRMs *sal1.1* (G) and *knW* (H). Cloned fragments driving reporter expression (green) above the genome tracks are indicated by the green bar.

(I) Comparison of ATAC-seq scores with transcriptome measurements from sorted *nub*$^+$ and *tsh*$^+$ cells. Differentially expressed genes (DESeq adjusted $p < 0.01$) are significantly more likely to have a differential ATAC peak (DESeq-$\log_{10}$pval) compared to genes expressed at similar levels (see STAR Methods for details). p values (above) derived from Student's t test. Median value indicated by horizontal line.

activator or repressor of transcription depends on the location in the segment and on the availability of Hox cofactors Homothorax (Hth) and Extradenticle (Exd). Finally, we show that this modified chromatin landscape alters the ability of another TF, Scalloped (Sd), to access its targets in the genome, leading to an expansion to some sites and restriction of others relative to T2. This study provides a molecular framework of how Hox proteins function to modify multiple cell types to alter the morphology of complex tissues.

RESULTS

Ubx diversifies chromatin accessibility of the homologous wing and haltere imaginal discs

The dorsal epithelium of the wing-bearing T2 and evolutionarily derived T3 segments of *Drosophila* come from the wing and haltere imaginal discs, respectively (Figures 1A and 1B). Each disc gives rise to homologous structures of the body wall, hinge, and appendage proper (listed from proximal-most to distal-most

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

position). At all positions along the proximal-distal axis, the dorsal structures of the T3 segment are highly modified relative to T2.[26–28] Notably, in both imaginal discs, the Hox cofactors Hth and Exd[29] are present only in cells that give rise to the body fates and the most proximal parts of the appendages (hinge; Figure 1A). Thus, depending on the proximal-distal position, Ubx transforms T2 into T3 both with and without these cofactors.[10,30] Although comparisons between the developing wing (T2) and haltere (T3) appendages revealed many of the transcriptional changes required to transform one appendage into another, they have not yet addressed how Ubx directly executes these extensive changes in gene expression.[31,32]

We initially performed assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) to compare the accessible chromatin profiles of the intact wing and haltere imaginal discs at the third larval instar stage. Although the overall profiles are very similar (correlation coefficient = 0.998), consistent with previous observations,[33] we find 760 sites with decreased accessibility in the haltere (H [haltere] < W [wing]) compared to the wing and 356 sites with increased accessibility (H > W) (DESeq2;[34] adjusted $p < 0.05$, log2 fold change [LFC] > 0.5; Figure 1D). Notably, approximately one-fifth of the H > W sites (n = 59) are within the *Ubx* locus, and they exhibit the highest fold difference between the wing and haltere discs (Figure 1D, open circles).

Compared to all accessible regions, the location of differentially accessible sites is biased toward introns and intergenic regions (Figure S1A). All four previously described CRMs regulated by Ubx in the haltere[35–37] are identified by this analysis, suggesting that many of the genome-wide differences we identify represent Ubx-mediated changes to the T2 gene regulatory network. The known Ubx-activated and -repressed CRMs show an increase and decrease in accessibility in the haltere relative to the wing, respectively, consistent with an inverse relationship between accessibility and repression,[38–40] and demonstrating the ability of this approach to distinguish both types of regulatory outcomes downstream of Ubx (Figure S1C).

Chromatin differences downstream of Ubx are region specific

Because *Ubx* is expressed in all haltere cells and must ultimately be responsible for all T3-specific differences, it is possible that the differences in chromatin accessibility measured above exist in all haltere cells, regardless of cell type. Alternatively, Ubx may alter accessibility differently, depending on the cell type. To discriminate between these possibilities, we repeated the ATAC-seq measurements using purified populations of nuclei from homologous distal and proximal domains from the wing and haltere imaginal discs. The distal population, marked by the expression of *nubbin* (*nub*), gives rise to the external adult appendages, including the distal hinge and appendage proper (wing blade and capitellum for the wing and haltere, respectively) (Figures 1A and 1B). The proximal population, marked by the expression of *teashirt* (*tsh*), gives rise to the non-appendage thoracic body tissue (notum and postnotum, respectively) and proximal hinge that connects the appendage to the body (Figures 1A–1C).

Comparison of the nub^+ domains yields 2,451 regions that are less accessible in the haltere compared to the wing (nub[H < W]) and 2,030 regions with increased accessibility in the haltere (nub [H > W]). In the tsh^+ domain, 1,658 regions have decreased accessibility in the haltere compared to the wing (tsh[H < W]) and 684 have increased accessibility (tsh[H > W]) (Figures 1D–1F). The majority of differentially accessible loci were specific to either the nub or tsh populations (e.g., compare columns 1, 3, and 6 in Figure S1B). As expected, most of the differential regions identified in the whole-disc comparison were also identified in the population-specific comparisons (Figure S1B). Compared to the whole-disc comparisons, the larger number of differentially accessible regions identified in the tsh^+ and nub^+ domains is likely due to greater sensitivity when comparing more homogeneous cell populations.

These data support the idea that changes in chromatin accessibility induced by Ubx are context specific. Examination of specific CRMs that are differentially expressed in only the *tsh* or *nub* populations further supports this idea. For example, the *sal1.1* and *knW* CRMs[30,36,37] are repressed by Ubx in nub^+ haltere cells, and both enhancers have less accessibility in the nub^+ domain, but no difference in the tsh^+ domain (Figures 1G and 1H).

To assess whether differences in chromatin accessibility correlate with transcriptional changes on a genome-wide scale, we performed RNA sequencing (RNA-seq) on the nub^+ and tsh^+ cell populations for both the haltere and wing imaginal discs. Differential analysis performed for nub^+ cells yielded 828 genes downregulated in the haltere and 846 genes with increased expression relative to nub^+ wing cells (Figure S1D). In the tsh^+ population, 126 genes had lower levels and 56 had higher levels in the haltere compared to tsh^+ wing cells (Figure S1D). For both populations, differentially expressed genes are more likely to have differentially accessible ATAC-seq peaks, compared to genes that are expressed at similar levels (chi-square test: p = 5.99e−47 [nub^+] and p = 3.06e−36 [tsh^+]; Figure 1I). These results suggest that tissue-specific differences in chromatin accessibility contribute to tissue-specific gene expression.

Most differences in accessibility are Ubx dependent

To confirm that the haltere-specific differences in chromatin accessibility depend on Ubx, we performed a time-sensitive knockdown of Ubx for 48 h in the nub^+ domains and repeated the ATAC-seq comparison from the nub^+ population from both the wing and haltere imaginal discs. Following knockdown, the Ubx target *salm* is derepressed in the haltere as previously described[36] (Figure 2A), and there is an increase in chromatin accessibility at the *sal1.1* CRM (Figure 2B), demonstrating that the haltere-specific chromatin accessibility at this locus is dependent on Ubx. Examining the data genome-wide, the majority of tissue-specific differences are lost after knockdown of Ubx (Figures 2C and S2A). Compared to the wing, 660 regions had less accessibility in the haltere (down from 2,451 in WT) and 237 had increased accessibility (down from 2,030 in WT) (Figure 2C). Notably, 52 sites in the *Ubx* locus remain more accessible in the haltere in the knockdown. This is expected because, with the exception of autoregulatory elements, the regulation of *Ubx* expression is upstream of Ubx activity, and therefore the accessibility of CRMs within *Ubx* should not change in response to reduced Ubx activity. The remaining tissue-specific differences may be due to incomplete knockdown, which is supported by the persistence of weak anti-Ubx antibody staining after knockdown (Figure 2A).

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

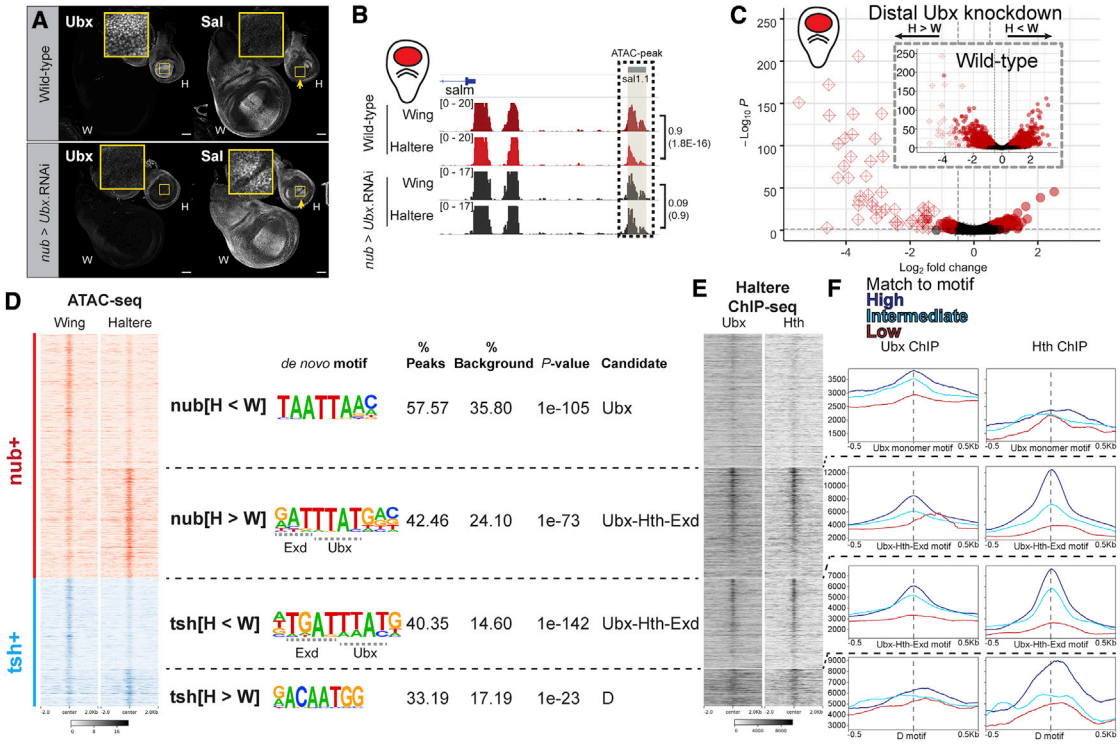


Figure 2. Ubx regulates chromatin accessibility

(A) Expression of the Ubx target *salm* in wild type and following Ubx knockdown. De-repression of *salm* in the haltere pouch is observed (arrow). Loss of Ubx expression (left) and de-repression of Spalt are magnified in the insets (yellow box).

(B) Genomic tracks showing the *salm* locus. The *sal*1.1 CRM is marked by the gray box. The region corresponding to ATAC peak generated by MACS2 and compared using DESeq2 is indicated by a dashed box. For each comparison (WT nub[W versus H] and Ubx RNAi nub[W versus H]), the LFC is indicated as the top number and adjusted p value in parentheses.

(C) Volcano plot comparing *nub*+ chromatin accessibility in wing and haltere imaginal discs following knockdown of *Ubx*. Inset: the same comparison in wild-type discs is repeated from Figure 1E for comparison. Note that the loci within the genomic region of *Ubx* (diamond shapes) remain differentially accessible, as expected given that the regulation of *Ubx* expression is upstream of Ubx activity.

(D) *De novo* motif analysis of the 4 differential ATAC-seq categories defined in Figures 1D–1F. The top-ranked motif for each category is shown. Candidate Ubx and Ubx-Hth-Exd motifs resemble motifs derived from SELEX-seq assays (Figures S2C and S2D). Heatmaps on the left show the wing and haltere ATAC-seq signals for each of the 4 categories.

(E) Heatmap showing the haltere ChIP signal for Ubx and Hth at loci within the differential ATAC-seq categories. Regions are centered around the closest match to the top-ranked *de novo* motif for that category (D) and sorted by highest-to-lowest scoring match to that motif.

(F) Plots showing distribution of average ChIP signal centered around the same motif as (E). Each category is split into thirds based on the degree of similarity of motif matching to the top-ranked *de novo* motif for that category. See STAR Methods for details.

Ubx increases and decreases chromatin accessibility depending on the region of the haltere disc

De novo searches for DNA-seq motifs can provide evidence for whether Ubx is directly responsible for changes in chromatin accessibility and whether Ubx is binding with or without its co-factors Hth and Exd. Importantly, the *nub*$^+$ population includes cells that have these cofactors (*nub*$^+$ *hth*$^+$ cells fated to become the distal hinge) and those that do not (*nub*$^+$ *hth*$^-$ pouch region fated to become the haltere capitellum), while all of the cells in the proximal *tsh*$^+$ population express these cofactors (*tsh*$^+$ *hth*$^+$; Figure 1A). Consequently, the association of specific DNA motifs with the gain or loss of accessibility also has the potential to provide spatial information about where Ubx is activating and repressing transcription.

We used an unbiased approach to look for motifs that are enriched in each differentially accessible peak set (nub[H < W], nub[H > W], tsh[H < W], tsh[H > W]; Figures 2D and S2B). Three of the four peak sets contain DNA binding motifs that are predicted to bind Hox proteins as the most enriched sequence. Interestingly, the type of motif differs between peak sets. The nub[H < W] set is highly enriched for a canonical Ubx monomer binding site, suggesting that, as with the previously described *sal*, *knot*, and *ana* targets,[30,35,37] Ubx generally represses transcription as a monomer in the *nub*$^+$ *hth*$^-$ domain. In contrast, the nub[H > W] set is enriched for a motif predicted to bind Ubx in complex with Hth and Exd (Ubx-Hth-Exd motifs[29]), suggesting that Ubx activates transcription with these cofactors in the *nub*$^+$ *hth*$^+$ domain. A Ubx-Hth-Exd motif was also enriched in the tsh[H < W] set, suggesting that Ubx represses transcription with these cofactors in the *tsh*$^+$ *hth*$^+$ domain. Equally notable is that neither of the population-specific sets are enriched for both types of Ubx motifs (Figure S2B). Furthermore, both of the discovered Ubx-related motifs match Ubx and Ubx-Hth-Exd binding sites derived from *in vitro* systematic evolution of ligands by exponential enrichment sequencing

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

(SELEX-seq) experiments, suggesting that they are bona fide Ubx monomer and Ubx-Hth-Exd binding sites, respectively[41,42] (Figure S2C). Neither type of Ubx motif is identified in the tsh[H > W] peak set.

These observations suggest that the sign of CRM regulation by Ubx differs depending on where along the proximal-distal axis Ubx functions and whether Hth and Exd are available as cofactors. A corollary to this conclusion is that in each region of the haltere disc, Ubx predominantly acts as either a repressor or an activator of transcription. Below, we provide additional evidence to support these conclusions by analyzing the activities of specific CRMs as well as the genome-wide binding of Ubx and Hth.

Ubx binds to CRMs that change chromatin accessibility in the haltere

To further examine the role of Ubx and its cofactors in regulating chromatin accessibility and CRM activity, we performed chromatin immunoprecipitation followed by sequencing (ChIP-seq) in whole haltere imaginal discs using antibodies against Ubx and Hth to directly determine their binding profiles genome-wide. The canonical Ubx-Hth-Exd complex motif is the most significantly enriched motif in Ubx ChIP-seq peaks (Figure S2E). The Ubx monomer motif is not significantly enriched in these whole-disc ChIP-seq experiments, despite the fact that several CRMs are known to bind Ubx in the absence of cofactors.[30,37] Notably, Ubx ChIP-seq experiments in other contexts also failed to identify a strongly enriched monomer motif, possibly reflecting a lower binding affinity or stability of the Ubx monomer to DNA.[22,43,44] For Hth ChIP-seq experiments, both Ubx-Hth-Exd and Hth-Exd binding site motifs are significantly enriched (Figure S2E).

The two ATAC-seq categories that are enriched for the Ubx-Hth-Exd motif (nub[H > W] and tsh[H < W]) both show strong association with Ubx and Hth binding (Figure 2E). Moreover, the strength of both the Ubx and Hth ChIP signals correlates with the *de novo* enriched Hox-Hth-Exd motif, supporting a direct interaction with these binding sites *in vivo* (Figure 2F). Although the nub[H < W] category, which is enriched for the Ubx monomer motif, shows generally low ChIP signal for both Ubx and Hth (see Discussion), the strength of the Ubx ChIP signal correlates with the presence of the *de novo* discovered Ubx momomer motif, and the region of maximum binding signal coincides with the location of the motif (Figure 2F). In contrast, the Hth ChIP signal does not show a similar correlation, supporting the conclusion that Ubx interacts directly with these regions as a monomer without Hth-Exd. These results suggest that Ubx and its cofactors directly bind to many of the sites that have haltere-specific differences in chromatin accessibility. Furthermore, the data suggest that Ubx and Hth binding are directly responsible for many of the observed segment- and region-specific differences in chromatin accessibility. The remaining differentially accessible regions that lack motif or ChIP signatures may be indirectly mediated by TFs that are downstream of Ubx.

In vivo reporters support spatial regulation of Ubx activity

The above analyses suggest that Ubx binds in nub^+ hth^- haltere pouch region as a monomer and is associated with a decrease in chromatin accessibility relative to the homologous cells in the wing disc, while Hox-Hth-Exd binding in the nub^+ hth^+ distal hinge region is associated with greater accessibility. To ask whether these categories reflect true Ubx repressed or activated targets *in vivo*, we cloned 20 putative CRMs into reporter constructs and observed their expression pattern in the wing and haltere discs. We chose loci that bind Ubx and have higher or lower accessibility in the nub^+ domain (17 from nub[H > W] and 3 from nub[H < W]) to ask whether their activity reflects the direction of change in accessibility (activating or repressing, respectively) and whether Ubx regulates them in the predicted region (distal hinge and pouch, respectively; Figure 3A). We chose candidate CRMs based solely on ATAC-seq differences and Ubx ChIP-seq signal, without taking Hth binding into consideration.

The majority of the cloned regions (13/17 nub[H > W] and 3/3 nub[H < W]) drive reporter expression in a segment-specific manner within the nub^+ domain (Figure 3B). All three candidates from the nub[H < W] category drive expression in the wing pouch and are less active in the homologous cell population in the haltere (Figures 3C and S3A–S3C). These three CRMs behave similarly to the three previously characterized CRMs in the haltere that are repressed by Ubx in this domain, leading to a total of six reporters with similar characteristics[30,35–37] (Figures 1G, 1H, and S1C).

Compared to nub[H < W] CRMs, the types of patterns driven by nub[H > W] CRMs are more varied and can be grouped into three categories: (1) expressed in the haltere distal hinge domain, but not the wing distal hinge (8/17 CRMs, e.g., Figures 3E and S3D–S3K), (2) expressed in the distal hinge domains of both tissues, but with a broader pattern in the haltere (5/17 CRMs, e.g., Figures 3F and S3L–S3P), and (3) no detectable expression in the nub^+ cells of either disc (4/17 CRMs, e.g., Figures 3D and S3Q–S3T). The third category may represent regions that change accessibility in the third larval instar stage that precede gene expression later in development. Notably, all four of these fragments are active CRMs because they drive expression in other regions of the discs (Figures 3D and S3Q–S3T). We observed no instances of repression in the distal hinge or activation in the pouch, supporting the conclusion that these Ubx activities are predominantly region specific.

The differences in reporter activity between the wing and haltere ranged from obvious to subtle. Therefore, for 11 reporters we analyzed mitotic clones of *Ubx* mutant cells in the haltere to confirm that there is a difference in activity downstream of Ubx. In all cases, loss of Ubx altered reporter activity toward the wing-like pattern, as expected (Figures 3C, 3E, and 3F, bottom panels, and S3A–S3P). These results support the conclusion that the activity of Ubx as an activator or repressor in the nub^+ cells of the haltere is spatially segregated into the nub^+ hth^+ and nub^+ hth^- domains, respectively.

Notably, even though they were chosen from the nub[H > W] and nub[H < W] sets, several reporters are also fortuitously expressed in the tsh^+ domain of the wing disc. Further supporting our conclusion that Ubx behaves as a repressor in the tsh^+ hth^+ domain, in three cases, Ubx^- clones in the haltere derepressed these reporters in that domain (Figures S3U–S3X). Added to this list of repressed targets is the autoregulatory *abx* CRM from *Ubx*, which is downregulated by Ubx-Hth-Exd predominantly in the tsh^+ domain.[10] These findings reveal that even individual loci can respond to Ubx differently, depending on the region of the imaginal disc, yet they obey the rules

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

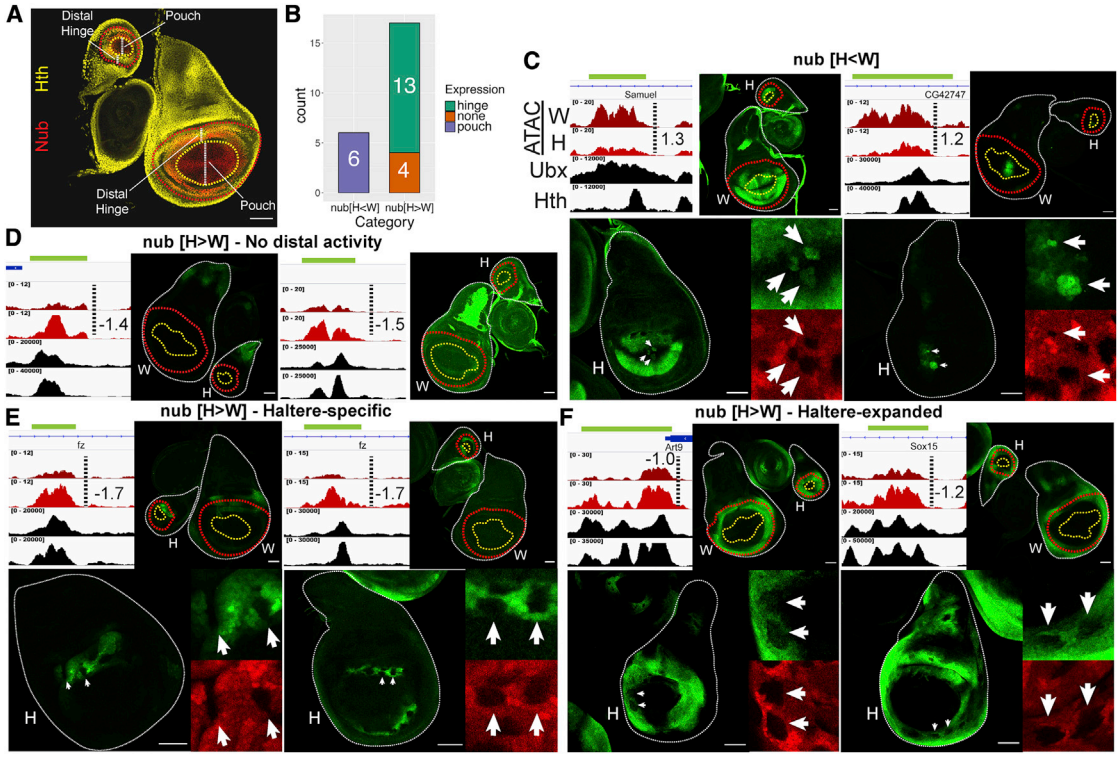


Figure 3. Analysis of Ubx-targeted CRMs

(A) Position of homologous distal hinge and pouch domains based on Hth and Nub expression in wing and haltere discs. The edges of Hth and Nub expression domains are marked with dotted yellow and red lines, respectively.

(B) Summary of CRM reporters. The nub[H < W] category includes 3 previously described CRMs: *sal1.1*, *knW*, and *ana-spot*.[30,35,37]

(C–F) Examples of nub[H < W] and nub[H > W] CRM-reporter genes (green). The upper left panels show genomic tracks for *nub*+ ATAC-seq wing, *nub*+ ATAC-seq haltere, Ubx ChIP, and Hth ChIP (LFC difference between wing and haltere indicated next to dashed line); the upper right panels show wing and haltere disc expression patterns for the reporter genes, and the bottom panels show *Ubx* null somatic clones in the haltere, with a subset of clones magnified in the insets. Clones are marked by the absence of RFP (arrows). ID for reporter examples: (C) Rep-3 (left) and Rep-2 (right); (D) Rep-18 (left) Rep-17 (right); (E) Rep-5 (left) and Rep-6 (right); and (F) Rep-14 (left) and Rep-12 (right). See Figure S3 for additional examples and Table S1 for a list of all of the reporter genes.

uncovered here, showing regional differences in Ubx-mediated gene regulation in the haltere disc.

Changes to chromatin downstream of Ubx alter the binding of another selector TF

How might Ubx-induced changes in chromatin accessibility affect CRM activity and, ultimately, transform T2 into T3? Because both tissues rely on a similar set of patterning TFs collectively referred to as selector TFs,[45] we hypothesized that Ubx may either facilitate (in the case of increased accessibility) or prevent (in the case of reduced accessibility) the binding of these shared TFs. As a test of this idea, we focused on the TF Sd because it has a similar expression pattern in wing and haltere imaginal discs and because it is required for the development of both appendages (Figure 4A).[46] Importantly, Sd is expressed in both the nub^{+} hth^{-} pouch and a subset of the nub^{+} hth^{+} hinge domains, where we hypothesize that Ubx is a repressor and an activator, respectively. Furthermore, the wing/haltere system allows the direct comparison between a Ubx^{+} state (wild-type haltere) with a Ubx^{-} state (wild-type wing), without having to analyze mutants.

To ask whether Sd binding differs in the haltere and wing imaginal discs, we performed ChIP-seq for Sd in both discs and compared the binding patterns. Although the majority of Sd binding sites are shared in the wing and haltere discs, suggesting that they are Hox independent, a subset of Sd binding sites (8.3%) are disc specific: 387 peaks show stronger binding in the haltere, while 759 peaks are stronger in the wing (Figures 4B, S4A, and S4B). *De novo* motif searches around both sets of Sd binding peaks show that, in addition to canonical Sd motifs, Ubx motifs are enriched to similar levels, suggesting that they are also targeted by Ubx (Figure 4D). However, as with the ATAC-seq data, the type of Ubx motif is distinct in peaks biased toward the different discs. H > W Sd binding events are enriched for the Ubx-Hth-Exd motif, while H < W Sd binding is associated with Ubx monomer motifs. Furthermore, Sd [H < W] and Sd [H > W] peaks overlap extensively with nub[H < W] and nub[H > W] peaks, respectively (Figures 4E and S4A). In addition, of the 589 peaks that have both tissue-specific Sd binding and differences in chromatin accessibility (51% of all tissue-specific Sd binding), 171 (29%) have a Ubx ChIP peak (Figures 4E and S4A; see Figure 4C for a specific example at the *Samuel* CRM). These data suggest that the binding of Sd is responsive to the presence of Ubx locally at the CRM, and points to a potential mechanism for how Ubx alters the output of shared TFs both positively and negatively: Ubx binding to monomer sites reduces

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

CellPress

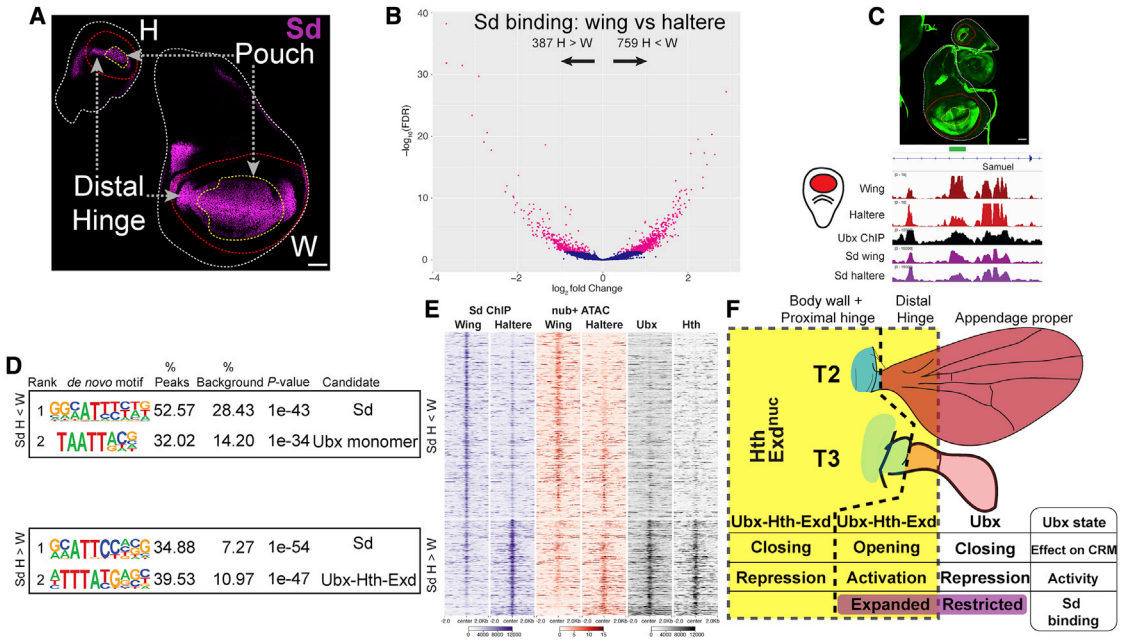


Figure 4. Ubx-mediated changes to chromatin accessibility changes where Sd binds

(A) Homologous patterns of Sd expression in the wing and haltere imaginal discs. In both tissues, Sd is expressed in the pouch, in the distal hinge, and along the dorsal-ventral compartment boundary. Boundaries of Nub (red) and Hth (yellow) expression are indicated with dotted lines as in Figure 3.

(B) Volcano plot comparing Sd binding in wing and haltere imaginal discs (Diffbind false discovery rate [FDR] < 0.05).

(C) Genomic tracks near the *Samuel* CRM (green box) and reporter expression driven by this CRM in wing and haltere discs.

(D) *De novo* motif analysis of the disc-specific Sd binding peaks for the Sd H < W and H > W categories.

(E) Heatmaps showing the ChIP signal for differential Sd binding, *nub*$^{+}$ ATAC-seq signal, Ubx ChIP signal, and Hth ChIP signal. Regions are sorted based on highest-to-lowest W:H ratio of distal ATAC-seq signal at the peak center. The top set shows the Sd H < W category and the bottom set shows the Sd H > W category, as defined in (B).

(F) Summary defining the 3 domains in T2 and T3, whether Ubx acts as a monomer or Ubx-Hth-Exd complex, whether Ubx opens or closes chromatin, and the effect on Sd binding.

accessibility and inhibits Sd binding in the pouch, while Ubx binding to Ubx-Hth-Exd sites increases accessibility and facilitates Sd binding in the distal hinge.

DISCUSSION

To transform the wing-bearing T2 segment ground state into that of the haltere-bearing T3 segment of the adult fly, we find that Ubx functions in three distinct modes that are spatially segregated in the imaginal disc: (1) Ubx reduces chromatin accessibility and represses transcription in the distal-most pouch domain as a monomer, (2) Ubx increases accessibility and activates transcription in complex with Hth-Exd in the distal hinge, and (3) Ubx reduces accessibility and represses transcription in the body wall and proximal hinge in complex with Hth-Exd.

Consistent with our findings, a subset of *Drosophila* and mammalian Hox paralogs have been shown in cultured cells, induced ESCs, and the mammalian limb bud to differ in their ability to bind and open less accessible chromatin, which has been suggested to mediate Hox paralog-specific functions *in vivo*.[22–25] Here, we tested the impact of chromatin accessibility changes by one Hox protein, Ubx, to perform the classical Hox function of serial homolog diversification. Similar to the cell culture experiments,[22,23] we find that in the T3 dorsal appendage, Ubx-Hth-Exd can increase accessibility, but the Ubx monomer does not. However, while previous studies have shown that some Hox proteins can increase accessibility, our results suggest that, at least for Ubx, Hox proteins can also decrease accessibility to diversify cell fates, both with and without Hth-Exd. Our ability to observe both increases and decreases in chromatin accessibility may be a consequence of studying the transformation of the T2 ground state into T3, which includes multiple cell fates that are modified by the presence of the same Hox protein. This is in contrast to previously studied systems, such as the vertebrate limb bud[25] or the induction of motor neuron fates from ESCs,[24] in which instead of transforming one tissue into another, Hox proteins promote the development of specific cell fates from less-differentiated progenitors.

The precise nature of how Ubx alters chromatin accessibility requires further investigation. One potential mechanism involves the recruitment of chromatin-modifying factors, several of which have been shown to interact with Hox proteins,[47,48] to modulate the compaction of the local CRM structure as suggested for the repression of *Dll* by Ubx.[49] Alternatively, Ubx may compete with the binding of activator TFs (in the case of haltere repression) or facilitate activator binding through nucleosome-mediated cooperativity[50] (in the case of haltere activation). Notably, in the cases of Ubx repression of *knot*[37] and *Dll*,[51,52] the repressive binding input of Ubx into the relevant CRMs is separable from the activating input, suggesting that for these cases, which involve

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

monomer and Ubx-Hth-Exd input, respectively, competition for binding is unlikely to be involved. Furthermore, although our ChIP data suggest that Ubx directly regulates both activated and repressed CRMs, it is notable that the Ubx ChIP signal was generally weaker when it binds as a monomer in the distal haltere compared to when it binds with Hth-Exd. This weaker ChIP signal may be a consequence of less stable and/or more transient binding by Ubx compared to a Ubx-Hth-Exd complex. Perhaps to compensate for weaker binding, repression by Ubx monomers typically requires multiple monomer binding sites, compared to fewer binding sites when repression is mediated by Ubx-Hth-Exd input.[30,37,51]

The surprising finding that Ubx predominantly acts as either an activator or repressor of transcription in a given region of the haltere disc implies that Hox proteins may exist in distinct cell-type-specific regulatory complexes that function either as dedicated activators or repressors. However, in contrast to this notion, we note that vertebrate Hox proteins can both cross-repress other Hox genes and activate downstream genes in the same domain of the spinal cord.[53] Although cross-repression by Hox genes may be a special case, future work is needed to determine whether the model proposed here extends to other Hox proteins and species.

As serially homologous tissues, the wing and haltere imaginal discs have a very similar organization of spatially restricted signaling pathways and share many of the same regionally expressed TFs. Ubx operates upon this common ground state to modify how these shared pathways and selector TFs are deployed in a T3-specific manner.[31,32,54] Ubx may alter the output from these shared systems by modifying the expression of the signaling molecules themselves, such as *wingless* repression in the posterior compartment of the haltere disc,[31] or by modifying the distribution of secreted signals, as in the case of Dpp signaling.[54] Our results reveal that Hox proteins can also modify the output of shared regulators by altering where they bind through changes to *cis*-regulatory chromatin accessibility. A pioneering role for Hox13 paralogs in the mammalian limb bud has been proposed because they permitted the binding of another Hox protein after it was ectopically expressed.[25] In light of our experiments, which assayed the binding of the essential TF, Sd, in wild-type tissues, both with and without Ubx, we suggest that not only is a pioneering role of Hox proteins during cell fate specification widespread but that Hox proteins can also function as anti-pioneers to restrict the binding of shared TFs between homologous cells.

An interesting trend emerges from comparison of the homologous adult structures that are modified by Ubx via activation or repression. In the capitellum and notum/proximal hinge, where Ubx-mediated gene repression dominates, the T3 morphology, broadly characterized, has both a reduced size and complexity. The latter can be observed through the loss of characteristic features of the T2 appendage and notum, such as highly patterned veins and large bristles (called macrochaetes), respectively. In contrast, the distal hinge of the haltere, where Ubx-mediated gene activation is the rule, develops complex T3-specific structures that are required for the haltere to provide critical sensory feedback during flight, most notably via arrays of mechanosensory neurons.[55] We speculate that diversification of tissue morphology by Hox proteins may follow a pattern wherein repressive activities contribute to simplifying the morphology of a tissue while gene activation may be required to generate novel complex cell types.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - *Drosophila* alleles and transgenes
  - Construction of enhancer reporter genes
  - Clonal analysis
  - Immunohistochemistry
  - Nuclei sorting
  - ATAC-seq library preparation and sequencing
  - RNAi knockdown
  - RNA-seq library preparation and sequencing
  - ChIP-seq library preparation and sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - ATAC-seq data processing
  - RNA-seq data processing
  - ChIP-seq data processing
  - Motif analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2021.07.030.

ACKNOWLEDGMENTS

We thank Roumen Voutev for providing reporter plasmids and contribution to Hth-Gp52 antibody production and validation, Liping Sun for contribution to Hth-Gp52 antibody production, Vikki Weake for providing the UAS-Kash.GFP strain, Chaitanya Rastogi and Harmen Bussemaker for helping to deploy SELEX-seq models genome-wide, Siqian Feng for advice on ChIP-seq, Kevin White for Ubx and GFP antibodies, Stavros Lomvardas for sharing equipment, and the Bloomington Drosophila Stock Center for fly stocks. We are also grateful to David Stern, Stavros Lomvardas, Rebecca Delker, and Nicolas Gompel for detailed comments on the manuscript. This work was supported by NIH grant R35GM118336, awarded to R.S.M.

AUTHOR CONTRIBUTIONS

Conceptualization, R.L. and R.S.M.; Methodology, R.L. and R.S.M.; Investigation, R.L. and J.E.S.; Formal Analysis, R.L.; Visualization, R.L.; Writing – Original Draft, R.L. and R.S.M.; Writing – Review & Editing, R.L. and R.S.M.; Funding Acquisition, R.S.M.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 30, 2021
Revised: June 17, 2021
Accepted: July 13, 2021
Published: August 5, 2021

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

REFERENCES

1. Lewis, E.B. (1978). A gene complex controlling segmentation in Drosophila. Nature *276*, 565–570.
2. Bridges, C.B., and Brehme, K.S. (1944). The Mutants of Drosophila melanogaster, Vol. 552 (Carnegie Institution of Washington).
3. Bender, W., Akam, M., Karch, F., Beachy, P.A., Peifer, M., Spierer, P., Lewis, E.B., and Hogness, D.S. (1983). Molecular genetics of the bithorax complex in Drosophila melanogaster. Science *221*, 23–29.
4. Mallo, M., Wellik, D.M., and Deschamps, J. (2010). Hox genes and regional patterning of the vertebrate body plan. Dev. Biol. *344*, 7–15.
5. Ghosh, P., and Sagerström, C.G. (2018). Developing roles for Hox proteins in hindbrain gene regulatory networks. Int. J. Dev. Biol. *62*, 767–774.
6. Philippidou, P., and Dasen, J.S. (2013). Hox genes: choreographers in neural development, architects of circuit organization. Neuron *80*, 12–34.
7. Arenkiel, B.R., Tvrdik, P., Gaufo, G.O., and Capecchi, M.R. (2004). Hoxb1 functions in both motoneurons and in tissues of the periphery to establish and maintain the proper neuronal circuitry. Genes Dev. *18*, 1539–1552.
8. Hrycaj, S.M., and Wellik, D.M. (2016). Hox genes and evolution. F1000Res. *5*, F1000 Faculty Rev-859.
9. Parker, H.J., Bronner, M.E., and Krumlauf, R. (2016). The vertebrate Hox gene regulatory network for hindbrain segmentation: evolution and diversification: coupling of a Hox gene regulatory network to hindbrain segmentation is an ancient trait originating at the base of vertebrates. BioEssays *38*, 526–538.
10. Delker, R.K., Ranade, V., Loker, R., Voutev, R., and Mann, R.S. (2019). Low affinity binding sites in an activating CRM mediate negative autoregulation of the Drosophila Hox gene Ultrabithorax. PLoS Genet. *15*, e1008444.
11. Zandvakili, A., Uhl, J.D., Campbell, I., Salomone, J., Song, Y.C., and Gebelein, B. (2019). The cis-regulatory logic underlying abdominal Hox-mediated repression versus activation of regulatory elements in Drosophila. Dev. Biol. *445*, 226–236.
12. Sánchez-Higueras, C., Rastogi, C., Voutev, R., Bussemaker, H.J., Mann, R.S., and Hombría, J.C.-G. (2019). In vivo Hox binding specificity revealed by systematic changes to a single cis regulatory module. Nat. Commun. *10*, 3597.
13. Bridoux, L., Zarrineh, P., Mallen, J., Phuycharoen, M., Latorre, V., Ladam, F., Losa, M., Baker, S.M., Sagerstrom, C., Mace, K.A., et al. (2020). HOX paralogs selectively convert binding of ubiquitous transcription factors into tissue-specific patterns of enhancer activation. PLoS Genet. *16*, e1009162.
14. Zheng, C., Jin, F.Q., and Chalfie, M. (2015). Hox proteins act as transcriptional guarantors to ensure terminal differentiation. Cell Rep. *13*, 1343–1352.
15. De Kumar, B., Parker, H.J., Paulson, A., Parrish, M.E., Pushel, I., Singh, N.P., Zhang, Y., Slaughter, B.D., Unruh, J.R., Florens, L., et al. (2017). HOXA1 and TALE proteins display cross-regulatory interactions and form a combinatorial binding code on HOXA1 targets. Genome Res. *27*, 1501–1512.
16. Mann, R.S., Lelli, K.M., and Joshi, R. (2009). Hox specificity unique roles for cofactors and collaborators. Curr. Top. Dev. Biol. *88*, 63–101.
17. Buenrostro, J., Wu, B., Chang, H., and Greenleaf, W. (2015). ATAC-seq: a method for assaying chromatin accessibility genome-wide. Curr. Protoc. Mol. Biol. *109*, 21.29.1–21.29.9.
18. Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the regulatory epigenome. Nat. Rev. Genet. *20*, 207–220.
19. Bozek, M., Cortini, R., Storti, A.E., Unnerstall, U., Gaul, U., and Gompel, N. (2019). ATAC-seq reveals regional differences in enhancer accessibility during the establishment of spatial coordinates in the *Drosophila* blastoderm. Genome Res. *29*, 771–783.
20. Bravo González-Blas, C., Quan, X.-J., Duran-Romaña, R., Taskiran, I.I., Koldere, D., Davie, K., Christiaens, V., Makhzami, S., Hulselmans, G., de Waegeneer, M., et al. (2020). Identification of genomic enhancers through spatial integration of single-cell transcriptomics and epigenomics. Mol. Syst. Biol. *16*, e9438.
21. Jacobs, J., Atkins, M., Davie, K., Imrichova, H., Romanelli, L., Christiaens, V., Hulselmans, G., Potier, D., Wouters, J., Taskiran, I.I., et al. (2018). The transcription factor Grainy head primes epithelial enhancers for spatiotemporal activation by displacing nucleosomes. Nat. Genet. *50*, 1011–1020.
22. Beh, C.Y., El-Sharnouby, S., Chatzipli, A., Russell, S., Choo, S.W., and White, R. (2016). Roles of cofactors and chromatin accessibility in Hox protein target specificity. Epigenetics Chromatin *9*, 1.
23. Porcelli, D., Fischer, B., Russell, S., and White, R. (2019). Chromatin accessibility plays a key role in selective targeting of Hox proteins. Genome Biol. *20*, 115.
24. Bulajić, M., Srivastava, D., Dasen, J.S., Wichterle, H., Mahony, S., and Mazzoni, E.O. (2020). Differential abilities to engage inaccessible chromatin diversify vertebrate HOX binding patterns. Development *147*, dev194761.
25. Desanlis, I., Kherdjemil, Y., Mayran, A., Bouklouch, Y., Gentile, C., Sheth, R., Zeller, R., Drouin, J., and Kmita, M. (2020). HOX13-dependent chromatin accessibility underlies the transition towards the digit development program. Nat. Commun. *11*, 2491.
26. Struhl, G. (1981). A homoeotic mutation transforming leg to antenna in Drosophila. Nature *292*, 635–638.
27. Abbott, M.K., and Kaufman, T.C. (1986). The relationship between the functional complexity and the molecular organization of the Antennapedia locus of Drosophila melanogaster. Genetics *114*, 919–942.
28. Carroll, S.B., Weatherbee, S.D., and Langeland, J.A. (1995). Homeotic genes and the regulation and evolution of insect wing number. Nature *375*, 58–61.
29. Merabet, S., and Mann, R.S. (2016). To be specific or not: the critical relationship between Hox and TALE proteins. Trends Genet. *32*, 334–347.
30. Galant, R., Walsh, C.M., and Carroll, S.B. (2002). Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. Development *129*, 3115–3126.
31. Weatherbee, S.D., Halder, G., Kim, J., Hudson, A., and Carroll, S. (1998). Ultrabithorax regulates genes at several levels of the wing-patterning hierarchy to shape the development of the Drosophila haltere. Genes Dev. *12*, 1474–1482.
32. Pavlopoulos, A., and Akam, M. (2011). Hox gene Ultrabithorax regulates distinct sets of target genes at successive stages of Drosophila haltere morphogenesis. Proc. Natl. Acad. Sci. USA *108*, 2855–2860.
33. McKay, D.J., and Lieb, J.D. (2013). A common set of DNA regulatory elements shapes Drosophila appendages. Dev. Cell *27*, 306–318.
34. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.
35. Hersh, B.M., Nelson, C.E., Stoll, S.J., Norton, J.E., Albert, T.J., and Carroll, S.B. (2007). The UBX-regulated network in the haltere imaginal disc of D. melanogaster. Dev. Biol. *302*, 717–727.
36. Walsh, C.M., and Carroll, S.B. (2007). Collaboration between Smads and a Hox protein in target gene repression. Development *134*, 3585–3592.
37. Hersh, B.M., and Carroll, S.B. (2005). Direct regulation of knot gene expression by Ultrabithorax and the evolution of cis-regulatory elements in Drosophila. Development *132*, 1567–1577.
38. Domcke, S., Hill, A.J., Daza, R.M., Cao, J., O'Day, D.R., Pliner, H.A., Aldinger, K.A., Pokholok, D., Zhang, F., Milbank, J.H., et al. (2020). A human cell atlas of fetal chromatin accessibility. Science *370*, eaba7612.
39. Berest, I., Arnold, C., Reyes-Palomares, A., Palla, G., Rasmussen, K.D., Giles, H., et al. (2019). Quantification of differential transcription factor activity and multiomics-based classification into activators and repressors: diffTF. Cell Rep. *29*, 3147–3159.e12.
40. Gray, L.T., Yao, Z., Nguyen, T.N., Kim, T.K., Zeng, H., and Tasic, B. (2017). Layer-specific chromatin accessibility landscapes reveal regulatory networks in adult mouse visual cortex. eLife *6*, e21883.

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

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

42. Rastogi, C., Rube, H.T., Kribelbauer, J.F., Crocker, J., Loker, R.E., Martini, G.D., et al. (2018). Accurate and sensitive quantification of protein-DNA binding affinity. Proc. Natl. Acad. Sci. USA *115*, E3692–E3701.

43. Slattery, M., Ma, L., Négre, N., White, K.P., and Mann, R.S. (2011). Genome-wide tissue-specific occupancy of the Hox protein Ultrabithorax and Hox cofactor Homothorax in Drosophila. PLoS ONE *6*, e14686.

44. Shlyueva, D., Meireles-Filho, A.C.A., Pagani, M., and Stark, A. (2016). Genome-wide Ultrabithorax binding analysis reveals highly targeted genomic loci at developmental regulators and a potential connection to Polycomb-mediated regulation. PLoS ONE *11*, e0161997.

45. Mann, R.S., and Carroll, S.B. (2002). Molecular mechanisms of selector gene function and evolution. Curr. Opin. Genet. Dev. *12*, 592–600.

46. Williams, J.A., Bell, J.B., and Carroll, S.B. (1991). Control of Drosophila wing and haltere development by the nuclear vestigial gene product. Genes Dev. *5*, 2481–2495.

47. Saleh, M., Rambaldi, I., Yang, X.J., and Featherstone, M.S. (2000). Cell signaling switches HOX-PBX complexes from repressors to activators of transcription mediated by histone deacetylases and histone acetyltransferases. Mol. Cell. Biol. *20*, 8623–8633.

48. Boube, M., Hudry, B., Immarigeon, C., Carrier, Y., Bernat-Fabre, S., Merabet, S., Graba, Y., Bourbon, H.M., and Cribbs, D.L. (2014). Drosophila melanogaster Hox transcription factors access the RNA polymerase II machinery through direct homeodomain binding to a conserved motif of mediator subunit Med19. PLoS Genet. *10*, e1004303.

49. Agelopoulos, M., McKay, D.J., and Mann, R.S. (2012). Developmental regulation of chromatin conformation by Hox proteins in Drosophila. Cell Rep. *1*, 350–359.

50. Mirny, L.A. (2010). Nucleosome-mediated cooperativity between transcription factors. Proc. Natl. Acad. Sci. USA *107*, 22534–22539.

51. Gebelein, B., McKay, D.J., and Mann, R.S. (2004). Direct integration of Hox and segmentation gene inputs during Drosophila development. Nature *431*, 653–659.

52. Uhl, J.D., Zandvakili, A., and 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*, e1005981.

53. Dasen, J.S., Liu, J.-P., and Jessell, T.M. (2003). Motor neuron columnar fate imposed by sequential phases of Hox-c activity. Nature *425*, 926–933.

54. Crickmore, M.A., and Mann, R.S. (2006). Hox control of organ size by regulation of morphogen production and mobility. Science *313*, 63–68.

55. Dickerson, B.H., de Souza, A.M., Huda, A., and Dickinson, M.H. (2019). Flies regulate wing motion via active control of a dual-function gyroscope. Curr. Biol. *29*, 3517–3524.e3.

56. Calleja, M., Moreno, E., Pelaz, S., and Morata, G. (1996). Visualization of gene expression in living adult Drosophila. Science *274*, 252–255.

57. Ma, J., and Weake, V.M. (2014). Affinity-based isolation of tagged nuclei from Drosophila tissues for gene expression analysis. J. Vis. Exp. (85), 51418.

58. McGuire, S.E., Mao, Z., and Davis, R.L. (2004). Spatiotemporal gene expression targeting with the TARGET and gene-switch systems in Drosophila. Sci. STKE *2004*, pl6.

59. Capdevila, M.P., and García-Bellido, A. (1981). Genes involved in the activation of the bithorax complex of Drosophila. Wilehm Roux Arch. Dev. Biol. *190*, 339–350.

60. Slattery, M., Voutev, R., Ma, L., Nègre, N., White, K.P., and Mann, R.S. (2013). Divergent transcriptional regulatory logic at the intersection of tissue growth and developmental patterning. PLoS Genet. *9*, e1003753.

61. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

62. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44* (W1), W160–W165.

63. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

64. Zhu, L.J., Gazin, C., Lawson, N.D., Pagès, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics *11*, 237.

65. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

66. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

67. Kelso, R.J., Buszczak, M., Quiñones, A.T., Castiblanco, C., Mazzalupo, S., and Cooley, L. (2004). Flytrap, a database documenting a GFP protein-trap insertion screen in Drosophila melanogaster. Nucleic Acids Res. *32*, D418–D420.

68. Barolo, S., Carver, L.A., and Posakony, J.W. (2000). GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in Drosophila. Biotechniques *29*, 726–732.

69. Bischof, J., Maeda, R.K., Hediger, M., Karch, F., and Basler, K. (2007). An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc. Natl. Acad. Sci. USA *104*, 3312–3317.

70. Xu, T., and Rubin, G.M. (1993). Analysis of genetic mosaics in developing and adult Drosophila tissues. Development *117*, 1223–1237.

71. Averof, M., and Cohen, S.M. (1997). Evolutionary origin of insect wings from ancestral gills. Nature *385*, 627–630.

72. Estella, C., McKay, D.J., and Mann, R.S. (2008). Molecular integration of wingless, decapentaplegic, and autoregulatory inputs into Distalless during Drosophila leg development. Dev. Cell *14*, 86–96.

73. Laprell, F., Finkl, K., and Müller, J. (2017). Propagation of Polycomb-repressed chromatin requires sequence-specific recruitment to DNA. Science *356*, 85–88.

74. Ghavi-Helm, Y., Zhao, B., and Furlong, E.E.M. (2016). Chromatin immunoprecipitation for analyzing transcription factor binding and histone modifications in Drosophila. Methods Mol. Biol. *1478*, 263–277.

75. Marin, E.C., Dry, K.E., Alaimo, D.R., Rudd, K.T., Cillo, A.R., Clenshaw, M.E., Negre, N., White, K.P., and Truman, J.W. (2012). Ultrabithorax confers spatial identity in a context-specific manner in the Drosophila postembryonic ventral nervous system. Neural Dev. *7*, 31.

76. Ryoo, H.D., and Mann, R.S. (1999). The control of trunk Hox specificity and activity by Extradenticle. Genes Dev. *13*, 1704–1716.

77. Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B.A., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. *46* (W1), W537–W544.

78. Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., et al. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389–393.

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-Ubx	Developmental Studies Hybridoma Bank	Cat# FP3.3858; RRID:AB_10805300
Guinea pig anti-Hth	Calico-custom production	N/A
Guinea pig anti-Tsh	Calico-custom production	N/A
Mouse anti-Nub	Developmental Studies Hybridoma Bank	Cat# 2D459; RRID:AB_2722119
Guinea Pig anti-Spalt	gift from James Hombría, CABD	N/A
Goat anti-GFP	gift from Kevin White, UChicago; ENCODE	Cat# goat-anti-GFP-UC; RRID:AB_2616146
Rabbit anti β-galactosidase	MP biomedicals	Cat# 559762, RRID:AB_2335286
Critical commercial assays		
NEB Ultra DNA library prep kit Ultra II	NEB	E7645S
Nextera DNA Library Preparation Kit	Illumina	FC-404-2005
Ovation Drosophila RNA-Seq System	Nugen	0350-32
Zymo Direct-zol RNA Microprep kit	Zymo	R2061
Deposited data		
Raw and analyzed data	This study	GEO: GSE166714
Experimental models: Organisms/strains		
Drosophila melanogaster: nub.Gal4	[56]	N/A
Drosophila melanogaster: tsh.Gal4	[56]	N/A
Drosophila melanogaster: UAS.Kash-GFP	[57]	N/A
Drosophila melanogaster: Sd-GFP	Bloomington stock center	BDSC: 50827
Drosophila melanogaster: UAS.Ubx.RNAi (chrom. 2)	[10]	N/A
Drosophila melanogaster: UAS.Ubx.RNAi (chrom. 3)	[10]	N/A
Drosophila melanogaster: UAS-mCherry.nls	Bloomington stock center	BDSC: 38425
Drosophila melanogaster: tub.Gal80^{ts}	[58]	N/A
Drosophila melanogaster: *Ubx*$^{9-22}$	[59]	N/A
Drosophila melanogaster: *yw*	Wildtype lab strain	N/A
Oligonucleotides		
Primers for reporter constructs	This study	See Table S1
Recombinant DNA		
pRVV54-LacZ	[60]	N/A
pRVV54-GFP	Gift of Roumen Voutev	N/A
Reporter constructs in pRVV54-GFP	This study	See Table S1
Software and algorithms		
Bowtie2	[61]	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
DEseq2	[34]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
Deeptools	[62]	https://deeptools.readthedocs.io/en/develop/
MACS2	[63]	https://github.com/macs3-project/MACS
ChIPPeakAnno	[64]	https://bioconductor.org/packages/release/bioc/html/ChIPpeakAnno.html
Homer	[65]	http://homer.ucsd.edu/homer/
Picard	https://broadinstitute.github.io/picard/	https://broadinstitute.github.io/picard/
Samtools	[66]	http://samtools.sourceforge.net/

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Richard Mann (rsm10@columbia.edu).

Materials availability

Fly lines and reporter constructs generated by this study will be shared by the lead contact upon request.

Data and code availability

- All ATAC-seq, RNA-seq, and ChIP-seq data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the Key resources table. Microscopy data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Experimental model for this study was the vinegar fly *Drosophila melanogaster*. A full list of strains used in the paper is included in the Key resources table. Unless otherwise described (See Method details section), flies were maintained at 25C on cornmeal food using standard laboratory techniques.

METHOD DETAILS

Drosophila alleles and transgenes

*nub.*Gal4[56]
*tsh.*Gal4[56]
UAS.Kash-GFP (Gift of Dr. Vikki Weake, Purdue Univ.)[57]
Sd-GFP (protein-trap fusion, FlyTrap, Bloomington # 50827)[67]
UAS.Ubx.RNAi (chr. II)[10]
UAS.Ubx.RNAi (chr. III)[10]
UAS-mCherry.nls (Bloomington # 38425)
tub.Gal80^{ts}[58]
Ubx$^{9-22}$[59]

Construction of enhancer reporter genes

Genomic fragments corresponding to putative CRMs were amplified from a generic laboratory *yw* stock. Regions were placed via restriction-mediated cloning into the multiple cloning site of pRVV54,[60] in which the *lacZ* ORF was replaced with the eGFP ORF (gift of Roumen Voutev, Columbia University). Coordinates of selected regions are described in Table S1. CRMs of *sal1.1* and *knW* were cloned using coordinates previously described.[30,37] *sal1.1* was inserted into pRVV54-LacZ and *knW* was synthesized as a full length fragment by Genewiz and inserted via restriction digest into a pH-Stinger[68] plasmid in which a attB sequence was inserted into the AatII restriction site. All reporters were integrated into the genome using PhiC31 system[69] at the attP40 landing site. Primers used for each reporter are listed in Table S1.

Clonal analysis

Ubx mitotic null clones were made using the Flp/FRT system[70] using the null *Ubx*$^{9-22}$ allele.[59] Larvae were heat shocked at 37°C for 40-50 minutes at the end of the 2nd instar stage and analyzed 48 hours later.

Immunohistochemistry

Wandering 3rd instar larval heads were dissected and inverted in PBS, followed by fixation in 4% PFA for 25 minutes at room temperature. Heads were then washed 2X 30 minutes in staining solution (SS: PBS, 1% BSA, 0.3% Triton-X). Primary antibodies were then added for incubation overnight in SS at 4°C. Heads were washed 4X 10 minutes in SS and incubated with fluorescent secondary antibodies for 2 hours at room temperature in dark, and washed as before. Heads were incubated overnight in Vectashield containing DAPI, and imaginal discs were subsequently dissected and mounted for imaging using a confocal microscope (Leica SP5 or Zeiss LSM 800)

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

Primary antibodies used were:

anti-Ubx (Mouse, FP3.38,[59] Developmental Studies Hybridoma Bank)
anti-Hth (Guinea Pig, Gp115, produced by Calico)
anti-Tsh (Guniea pig, Gp68, produced by Calico)
anti-Nub (Mouse, 2D4,[71] Developmental Studies Hybridoma Bank)
anti-Spalt (gift from James Hombría, CABD)
anti-GFP (Goat, gift from Kevin White, UChicago)
anti-β-galactosidase (Rabbit, Cappel)

Nuclei sorting

Nuclei were magnetically sorted from wing and haltere imaginal discs using the UAS.Kash-GFP transgene as previously described[57] with slight modifications. Briefly, imaginal discs of the genotype *nub*.G4; UAS.Kash-GFP or *tsh*.G4; UAS.Kash-GFP were isolated from larvae at the 3rd instar wandering stage by dissection in PBS with 0.01% tween-20 on ice. Dissected tissue was then washed 2X in chilled nuclei extraction buffer (NEB: 10 mM HEPES, pH = 7.5; 2.5 mM MgCl; 10 mM KCl). Nuclei were extracted in a 1 mL dounce on ice using 20 strokes of the loose pestle, followed by a 10-minute incubation, and 25 strokes of the tight pestle. Nuclei were then filtered over 30 uM cell filter, and pre-cleared for 10min with 5ul of Protein-G Dynabeads in NEB supplemented with 0.1% tween-20. Pre-clearing beads were removed with a magnet and nuclei were added to a new tube containing anti-GFP coated Dynabeads and incubated with rotation for 30 min at 4°C. Afterward bead-bound nuclei were washed 4X (5min each) with nuclei wash buffer (15 mM TRIS, pH = 7.5; 50 mM NaCl; 40 mM KCl; 2 mM MgCl2; 0.1% Tween-20). Isolated nuclei were counted on a hemocytometer, and used for ATAC-seq or RNA-seq.

ATAC-seq library preparation and sequencing

ATAC-seq was performed on 50,000 nuclei as previously described.[17] Libraries were sequenced using a 150-cycle high output (wild-type samples) or 75-cycle high output (RNAi samples) with paired end sequencing using an Illumina Nextseq. Two replicates were used for all ATAC-seq experiments, with the exception of *tsh*+ haltere experiment, for which three replicates were performed.

RNAi knockdown

Larvae of the genotype *yw; nub.G4, tub-Gal80^{ts}/UAS.Ubx.RNAi; UAS.Kash-GFP/UAS.Ubx.RNAi* were raised at 18°C until early 3rd instar and subsequently shifted to 29°C to permit expression of RNAi for 48 hours. Wandering 3rd instar larvae were collected, and subjected to ATAC-seq as described above, separately for wing and haltere discs.

RNA-seq library preparation and sequencing

RNA was extracted from sorted nuclei using TRIzol and purified using the Zymo Direct-zol RNA Microprep kit. RNA-seq libraries were prepared using total rRNA depleted RNA using Nugen Ovation *Drosophila* RNA-seq system. Libraries were sequenced using a 150-cycle high output with paired end sequencing using an Illumina Nextseq. Two replicates were used for all RNA-seq experiments.

ChIP-seq library preparation and sequencing

ChIP-seq using wing and haltere imaginal discs was performed as described previously[72] with minor modifications according to Laprell et al.[73] and Ghavi-Helm et al.[74] 3rd instar larval heads were dissected and inverted in PBS on ice. Heads were fixed for 20 minutes in 1.8% PFA in crosslinking medium (10 mM HEPES, pH = 8.0; 100 mM NaCl; 1 mM EDTA, pH = 8.0; 0.5 mM EGTA, pH = 8.0) at room-temperature with rotation, and subsequently quenched (Quench solution: 1xPBS; 125 mM glycine; 0.1% Triton X-100). Fixed-heads were then washed 2X in buffer A (10 mM HEPES, pH = 8.0; 10 mM EDTA, pH = 8.0; 0.5 mM EGTA, pH = 8.0, 0.25% Triton X-100) and 2X in buffer B (10 mM HEPES, pH = 8.0; 200 mM NaCl; 1 mM EDTA, pH = 8.0; 0.5 mM EGTA, pH = 8.0; 0.01% Triton X-100) 10 minutes each at 4°C. Wing or haltere discs were then dissected and placed in sonication buffer (10 mM HEPES, pH = 8.0 ;1 mM EDTA, pH = 8.0; 0.5 mM EGTA, pH = 8.0, 0.1% SDS). Chromatin sonication was performed using a Covaris S2 instrument at settings (105W; 2% Duty; 15 minutes).

Sonicated chromatin was brought to 1X mild-RIPA (10 mM Tris-HCl, pH = 8.0; 1 mM EDTA, pH = 8.0; 150 mM NaCl; 1% Triton X-100) concentration and pre-cleared with 40ul of Protein-G Dynabeads for 1 hour at 4°C with rotation. Pre-clearing beads were removed with magnet and antibody was added for incubation overnight, followed by the addition of 40ul of Protein-G Dynabeads with a 3 hour incubation at 4°C with rotation. Bead bound antibody-chromatin complexes were washed as follows 2X RIPA LS (10 mM Tris-HCl, pH = 8.0; 1 mM EDTA, pH = 8.0; 150 mM NaCl; 1% Triton X-100; 0.1% SDS; 0.1% DOC), 2X RIPA HS (10 mM Tris-HCl, pH = 8.0; 1 mM EDTA, pH = 8.0; 500 mM NaCl; 1% Triton X-100; 0.1% SDS; 0.1% DOC), 1X LiCl (10mM Tris-HCl, pH = 8.0; 1 mM EDTA, pH = 8.0; 250 mM LiCl; 0.5% IGEPAL CA-630; 0.5% DOC), 1X TE (10 mM Tris-HCl, pH = 8.0; 1 mM EDTA, pH = 8.0). Samples were then treated with RNase and proteinase K, and chromatin was isolated using standard phenol-chloroform extraction.

Antibodies used were anti-Ubx (7701,[75] 1:100 dilution for IP, gift from Kevin White, U. Chicago), anti-Hth (Gp52,[76] 1:300 dilution for IP), and anti-GFP (used for Sd-GFP; ab290, Abcam; 1:300 dilution for IP).

Please cite this article in press as: Loker et al., Cell-type-specific Hox regulatory strategies orchestrate tissue identity, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.07.030

ChIP-seq libraries were made following the NEBnext UltraII kit (NEB) and associated protocol.

Libraries were sequenced using a 75-cycle high output with single end sequencing using an Illumina Nextseq. Two replicates were used for all ChIP-seq experiments.

QUANTIFICATION AND STATISTICAL ANALYSIS

ATAC-seq data processing

Reads were mapped using Bowtie2[61] to the dm6 genome assembly. Mapped reads were then filtered for map quality (SAMtools[66]) and duplicates (Picard tools (broadinstitute.github.io/picard/): MarkDuplicates). The Galaxy platform[77] was used for these pre-processing steps. Genome-track files were created using Deeptools (BamCoverage; RPGC normalization). Differential analysis was performed using DESeq2[34] on a common interval of 24,915 peaks generated by merging ATAC-seq peaks called by MACS2[63](–nomodel–call-summits) from the all wild-type sorted datasets. Cut off used for calling differential accessibility was Log_2Fold change > 0.5 and adjusted p value (padj) < 0.05. Peaks within the extended *Ubx* genomic locus were defined by (Chr3R: 16655898- 16807343). Heatmaps were made using Deeptools Computematrix (options: reference-point; missingDataAsZero) and PlotHeatmap. Genomic region annotation of ATAC-seq peaks was performed using the bioconductor package ChIPpeakAnno.[64]

RNA-seq data processing

Reads were mapped using HISAT2 to the dm6 genome assembly. Mapped reads were then filtered for map quality (SAMtools[66]). Differential analysis was performed using DESeq2 with cutoff: padj < 0.01. For comparison of ATAC-seq and RNA-seq, all ATAC-seq peaks were assigned to the nearest gene that is expressed in either wing or haltere imaginal disc (count > 50). For each gene the single ATAC peak with the lowest p value determined by DESeq2 differential analysis, and the (W/H) -Log_{10}P as determined by DESeq2 (described above) was compared between peaks associated with differentially expressed versus non-differentially expressed genes using a chi-square test.

ChIP-seq data processing

Reads were mapped using Bowtie2 to the dm6 genome assembly. Mapped reads were then filtered for map quality (SAMtools[66]) and duplicates (Picard MarkDuplicates). Peaks were called using MACS2.[63] Genome-track files were created using Deeptools[62] (BamCoverage; RPKM normalization). For comparison of Sd binding in wing and haltere, differential analysis was performed using DiffBind[78] (FDR < 0.05 for significance cutoff).

Motif analysis

De novo motifs were discovered using Homer[65] (findmotifsgenome.pl). For ATAC-seq data the entire peak was used to search for enriched motifs (option: -size given) and all ATAC peaks (minus the queried group) were used to calculate background enrichment. For ChIP-seq a default 200bp window around the peak center was used. To center peaks around the best match to the de-novo motif (Figures 2E and 2F) the annotatepeaks command was used (option: -mbed) to generate the location of the motif.