PNAS RESEARCH ARTICLE | GENETICS

SpyChIP identifies cell type-specific transcription factor occupancy from complex tissues

Siqian Feng[a,b,1] and Richard S. Mann[a,b,c,1]

Edited by John Carlson, Yale University, New Haven, CT; received December 27, 2021; accepted April 18, 2022

Chromatin immunoprecipitation (ChIP) is an important technique for characterizing protein–DNA binding in vivo. One drawback of ChIP-based techniques is the lack of cell type-specificity when profiling complex tissues. To overcome this limitation, we developed SpyChIP to identify cell type-specific transcription factor (TF) binding sites in native physiological contexts without tissue dissociation or nuclei sorting. SpyChIP takes advantage of a specific covalent isopeptide bond that rapidly forms between the 15-amino acid SpyTag and the 17-kDa protein SpyCatcher. In SpyChIP, the target TF is fused with SpyTag by genome engineering, and an epitope tagged SpyCatcher is expressed in cell populations of interest, where it covalently binds to SpyTag-TF. Cell type-specific ChIP is obtained by immunoprecipitating chromatin prepared from whole tissues using antibodies directed against the epitope-tagged SpyCatcher. Using SpyChIP, we identified the genome-wide binding profiles of the Hox protein Ultrabithorax (Ubx) in two distinct cell types of the *Drosophila* haltere imaginal disc. Our results revealed extensive region-specific Ubx–DNA binding events, highlighting the significance of cell type-specific ChIP and the limitations of whole-tissue ChIP approaches. Analysis of Ubx::SpyChIP results provided insights into the relationship between chromatin accessibility and Ubx–DNA binding, as well as different mechanisms Ubx employs to regulate its downstream *cis*-regulatory modules. In addition to SpyChIP, we suggest that SpyTag–SpyCatcher technology, as well as other protein pairs that form covalent isopeptide bonds, will facilitate many additional in vivo applications that were previously impractical.

transcription factor | SpyTag | ChIP | Ubx | *Drosophila*

Significance

We have developed SpyChIP, a method that depends on covalent isopeptide bond formation between two peptides, SpyTag and SpyCatcher, to identify sites of cell type-specific transcription factor occupancy in native physiological contexts without tissue dissociation or nuclei sorting. Using SpyChIP, we characterized the genome-wide binding profiles of the Hox protein Ultrabithorax (Ubx) in two distinct cell types of the *Drosophila* haltere imaginal disc, revealing extensive cell type-specific Ubx–DNA binding events. The application of cell type-specific covalent bond formation to chromatin immunoprecipitation sets the stage for carrying out many other cell type-specific analyses and genetic manipulations in vivo that were previously impractical.

Chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) has been an important technique to query in vivo genome-wide binding profiles of transcription factors (TFs) and chromatin modifications (1). However, when assayed in whole tissues, ChIP-seq reports a mixture of TF-DNA binding signatures present in multiple cell types, making it difficult to discern a TF's cell type-specific functions. Several strategies have been developed to obtain cell type-specific TF-DNA occupancy information. Cell type-specific overexpression of tagged TFs is not an ideal solution, because nonphysiological levels or nonnative spatial and temporal expression patterns can result in false-positive or false-negative binding. An alternative is to sort cross-linked nuclei from dissociated tissues (2), but dissociation remains a significant technical challenge for many tissues, and the low yield of sorting makes this strategy only feasible for tissues that can be obtained in large quantity. Targeted DamID (TaDa), which depends on cell type-specific expression of very low level DNA adenine methyltransferase (Dam)–TF fusions, represents another powerful approach (3). However, it can be challenging to accurately control the levels of the TF–Dam fusions, and DamID-based methods have the potential to mark a mixture of past and present TF binding events, compromising the temporal resolution of the results that may be important when characterizing actively developing tissues.

To overcome the limitations of the current techniques, we developed a method based on SpyTag–SpyCatcher technology (4) that we call SpyChIP. Previous in vitro work demonstrated that the 15-amino acid SpyTag peptide spontaneously and rapidly forms a covalent isopeptide bond with a specific binding partner, a 17-kDa protein named SpyCatcher (4). We reasoned that if SpyTag and SpyCatcher were also able to form a covalent bond in the nuclei, a TF fused with SpyTag could be covalently linked to epitope tagged SpyCatcher expressed specifically in the target cell type. ChIP against the epitope on SpyCatcher would decode cell type-specific TF-DNA occupancy without tissue dissociation and nuclei sorting (Fig. 1*A*). Indeed, applying SpyChIP to the *Drosophila* Hox protein Ubx verified this approach and revealed many cell type-specific Ubx–DNA binding events in the haltere imaginal disc.

Author affiliations: [a]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10027; [b]Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY 10027; and [c]Department of Systems Biology, Columbia University, New York, NY 10027

Author contributions: S.F. designed research; S.F. performed research; S.F. and R.S.M. analyzed data; and S.F. and R.S.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: sf2607@columbia.edu or rsm10@columbia.edu.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2122900119/-/DCSupplemental.

Published June 13, 2022.

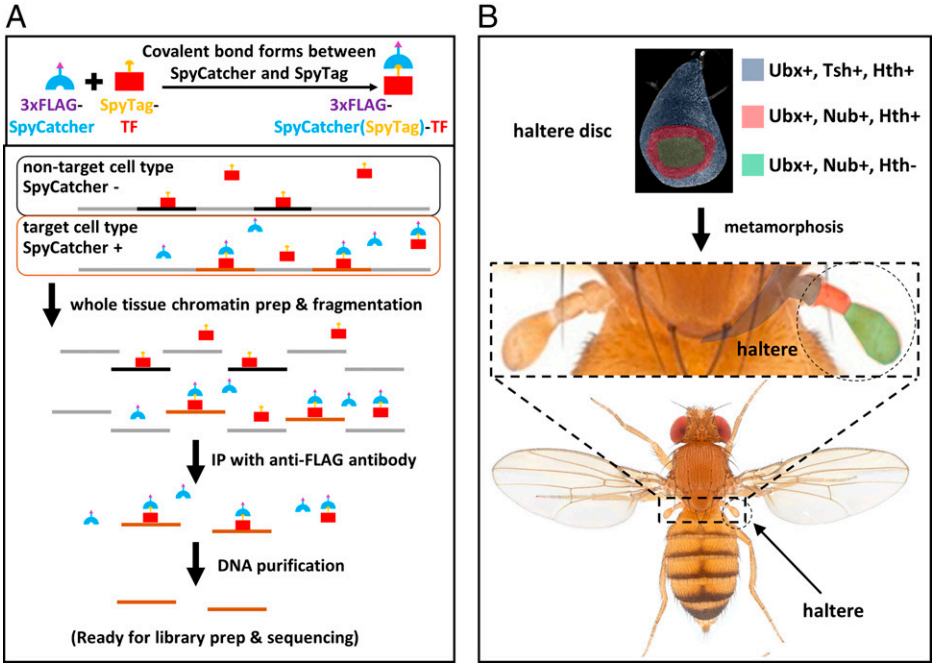

Fig. 1. Overview of SpyChIP strategy and haltere development. (*A*) A TF of interest is tagged with SpyTag by genome engineering. Upon cell-type specific expression of 3xFLAG-SpyCatcher, a covalent bond is formed between SpyTag and SpyCatcher, allowing chromatin bound by the TF to be immunoprecipitated using antibody against the 3xFLAG epitope on SpyCatcher. (*B*) Schematic of the development from larval haltere imaginal disc to adult T3 segment. During metamorphosis, the center of the haltere disc everts and becomes the distal haltere. Ubx is expressed in the entire haltere disc. The expression domains of Tsh, Nub, and Hth in the haltere disc are labeled, and the corresponding adult structures are indicated by the same colors.

Results

SpyTag and SpyCatcher Form a Covalent Isopeptide Bond In Vivo. We first tested whether SpyTag and SpyCatcher form a covalent isopeptide bond in vivo. In the nuclei of *Drosophila* embryos, we coexpressed 3xFLAG-SpyCatcher with GFP that was tagged with SpyTag at either the N or C terminus, and the V5 tag at the other end. Western blot against the 3xFLAG tag and the V5 tag was performed to follow SpyCatcher and GFP, respectively. Consistent with previous in vitro results, we detected the formation of a larger molecular weight protein that is roughly the predicted size of SpyCatcher fused to GFP (Fig. 2*A*), indicating successful covalent bond formation in *Drosophila* nuclei.

We next piloted SpyChIP by characterizing the occupancy of the Hox protein Ubx (Ultrabithorax) in different cell types in *Drosophila* haltere imaginal discs. Ubx is a selector TF that determines the identity of the third thoracic (T3) and first abdominal (A1) segments (5). We probed the genome-wide binding of Ubx in the *Drosophila* haltere imaginal disc, which gives rise to the dorsal T3 segment of the adult fly, including the haltere, an appendage critical for flight (Fig. 1*B*). Mutations in *Ubx* result in the famous four-winged *bithorax* homeotic transformation, in which the haltere-bearing T3 segment of the adult is transformed into a second copy of the wing-bearing T2 segment (5). During wild-type metamorphosis, the center of the haltere imaginal disc gives rise to most of the haltere appendage, while the periphery of the disc gives rise to the dorsal T3 body wall and the proximal haltere structures (Fig. 1*B*) (6).

We fused the SpyTag to the N terminus of Ubx at the endogenous *Ubx* locus in a scarless manner (*Materials and Methods* and *SI Appendix*, Fig S1), and expressed 3xFLAG-SpyCatcher with two cell type-specific Gal4 drivers: *tsh-Gal4*, active in the proximal haltere disc, and *nub-Gal4*, expressed in the distal haltere disc. We also used the ubiquitous driver *ubi-Gal4*, which should mimic a standard whole-tissue ChIP experiment (Fig. 1*B* and *SI Appendix*, Fig. S2). Western blotting with an anti-Ubx antibody showed that the apparent molecular weight of SpyTag-Ubx increased when SpyCatcher was expressed by all three drivers, and that the increase in size was consistent with the molecular weight of 3XFLAG-SpyCatcher (Fig. 2*B*). When *ubi-Gal4* was used to express SpyCatcher, most of the endogenous Ubx shifted to the larger molecular weight (Fig. 2*B*), indicating efficient covalent bond formation between SpyCatcher and SpyTag-Ubx in vivo in *Drosophila* nuclei. As expected, when SpyCatcher was expressed with the other two drivers, less Ubx was shifted to the larger size, consistent with their more limited expression domains within the haltere disc.

SpyChIP Faithfully Captures TF-DNA Occupancy. ChIP-seq experiments were then performed when 3xFLAG-SpyCatcher was expressed by each of the three Gal4 drivers, using chromatin prepared from whole haltere discs and anti-FLAG antibody. All Ubx::SpyChIP replicates revealed thousands of peaks, consistent with successful ChIP experiments. To assess how well SpyChIP works, we compared *ubi-Gal4* > Ubx::SpyChIP results with two independent whole haltere disc Ubx ChIP datasets. One such dataset was generated by using the same anti-FLAG antibody as we used in all Ubx::SpyChIP experiments to profile Ubx binding in 3xFLAG-Ubx flies, which was previously created by inserting the 3xFLAG tag into the endogenous *Ubx* locus in a scarless manner (7). The other whole-disc Ubx ChIP dataset was obtained by probing wild-type flies using anti-Ubx antibody (8). The average enrichment of sequencing tags in all called peaks relative to a random set of genomic regions can be used as an approximation of a ChIP's signal-to-noise ratio. We found that this enrichment is slightly higher for Ubx ChIP with anti-FLAG antibody than with anti-Ubx antibody (*SI Appendix*, Fig. S3*A*). All Ubx::SpyChIP experiments

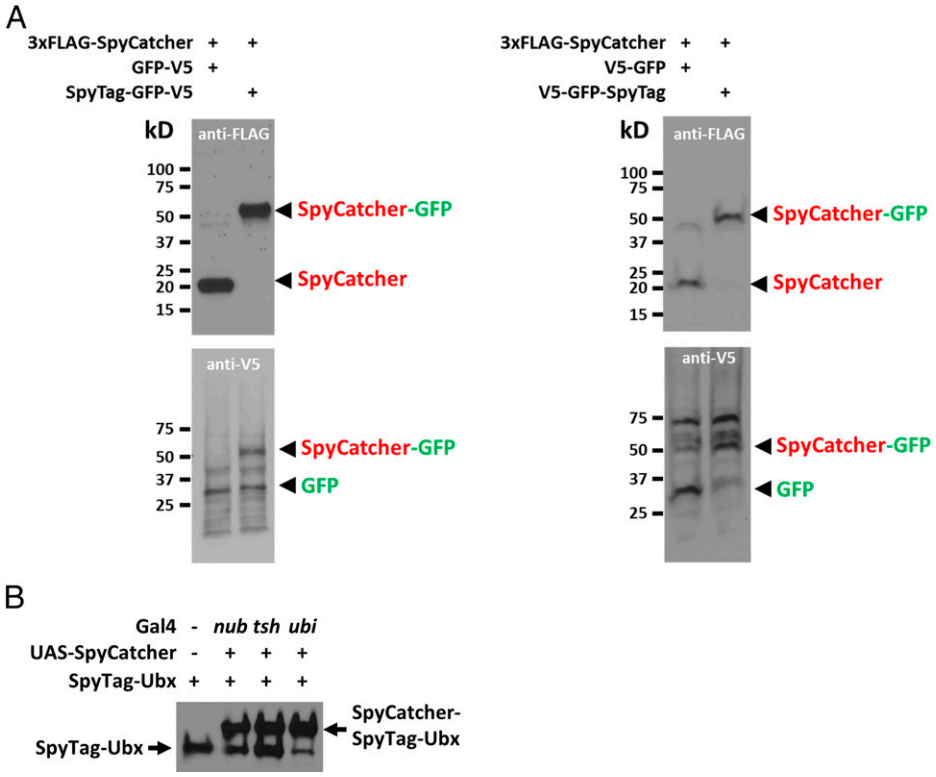

Fig. 2. SpyTag and SpyCatcher form covalent isopeptide bond in vivo. (*A*) Western blot analysis of total embryo lysates using anti-FLAG antibody and anti-V5 antibody. The embryos were F1 embryos from the following crosses: (*Left*) SpyTag at N terminus of GFP: *En-Gal4/(CyO); MKRS/TM6B* males crossed to *attP40-UAS-3xFLAG-NLS-SpyCatcher; attP2-UAS-(SpyTag)-GFP-V5* females. (*Right*) SpyTag at C terminus of GFP: *En-Gal4/(CyO)* females crossed to *attP40-UAS-3xFLAG-NLS-SpyCatcher; attP2-UAS-V5-GFP-(SpyTag)* males. In both cases, only GFP that is tagged with SpyTag shifts to a higher molecular weight after expression of 3xFLAG-SpyCatcher. (*B*) Anti-Ubx Western blot analysis of whole haltere discs. The genotypes of the lanes from left to right are: 1) *SpyTag-Ubx/SpyTag-Ubx*; 2) *nub-Gal4/+; UAS-SpyCatcher, SpyTag-Ubx/SpyTag-Ubx*; 3) *tsh-Gal4/+; UAS-SpyCatcher, SpyTag-Ubx/SpyTag-Ubx*; and 4) *ubi-Gal4/+; UAS-SpyCatcher, SpyTag-Ubx/SpyTag-Ubx*. Depending on the Gal4 driver, different amounts of SpyTag-Ubx are shifted to a higher molecular weight upon coexpression with SpyCatcher.

have similar enrichment, which is essentially the same as the enrichment of Ubx ChIP with anti-Ubx antibody, but is slightly lower than anti-FLAG Ubx ChIP (*SI Appendix*, Fig. S3*A*). We conclude that overall, the signal-to-noise ratio of SpyChIP is comparable to that of standard ChIP experiments.

In addition, pair-wise comparisons between *ubi-Gal4* > Ubx::SpyChIP and both whole haltere disc Ubx ChIPs show good agreement (Fig. 3*A* and *SI Appendix*, Fig. S3*B*). The correlation between *ubi-Gal4* > Ubx::SpyChIP and a standard Ubx ChIP is similar to the correlation between two Ubx ChIP biological replicates (*SI Appendix*, Fig. S3*B*), indicating that SpyChIP faithfully captures genome-wide Ubx occupancy.

We considered the possibility that, when SpyCatcher is expressed with *nub-Gal4* or *tsh-Gal4*, there may be a large excess of SpyCatcher compared to SpyTag-Ubx. Such an excess could result in a pool of unbound SpyCatcher that, during chromatin preparation and IP, might bind to SpyTag-Ubx from cells outside the domain targeted by Gal4, thus potentially compromising specificity. To limit this from happening, an excess of synthetic SpyTag peptide was added to quench unoccupied SpyCatcher in all experiments except for *nub-Gal4* > Ubx::SpyChIP replicate 1, which allowed us to assess the effect of quenching. The comparison between *nub-Gal4* > Ubx::SpyChIP replicates with or without quenching did not reveal significant differences (*SI Appendix*, Fig. S3*C*). This could mean that an excess of SpyCatcher does not decrease the specificity of SpyChIP or, in this case, it could be due to the fact that the endogenous Ubx levels are sufficiently high in Nub^+ cells (9) so that there is not an excess of unbound SpyCatcher.

SpyChIP Identifies Cell Type-Specific TF-DNA Binding Events. We next inspected Ubx::SpyChIP results genome-wide. Peaks shared between Tsh^+ and Nub^+ cells, as well as those specific to each cell type, could be readily identified (Fig. 3*A*). Genome-wide comparison between *tsh-Gal4* > Ubx::SpyChIP and *nub-Gal4* > Ubx::SpyChIP results identified 175 and 1,888 Ubx binding events that are specific to either the Tsh^+ domain or Nub^+ domain, respectively. In addition, there are 2,389 binding events that are shared by both datasets (Fig. 3*B*). The significant asymmetry in the numbers of Tsh^+ and Nub^+ cell-specific Ubx binding events is surprising, but is consistent with the observation that for both the wing and haltere discs, several-fold more differentially accessible loci were observed in Nub^+ cells than in Tsh^+ cells (8).

Ubx can bind to DNA either as a monomer or as a heterodimer with its cofactor Extradenticle (Exd), and the ubiquitous Exd protein is only nuclear and available as a Hox cofactor when another protein, Homothorax (Hth), is present (10). In the haltere disc, Hth is expressed in all Tsh^+ cells and some Nub^+ cells (8) (Fig. 1*B*). Consistent with the large number of Nub^+, Hth^- cells, a Ubx monomer motif is enriched in Nub^+ cell-specific Ubx-bound peaks. In contrast, an Exd-Ubx heterodimer motif is enriched in Tsh^+ cell-specific Ubx binding events, as well as in peaks shared by the two cell types (Fig. 3*B* and *SI Appendix*, Fig. S4*B*). As expected, both types of Ubx motifs are enriched in the *ubi-Gal4*>Ubx::SpyChIP peaks (*SI Appendix*, Fig. S4*A*). These results are consistent with previous results showing that Ubx binds with or without cofactors, depending on the region of the haltere disc (8), and demonstrate that SpyChIP is able to capture cell type-specific TF-DNA binding events.

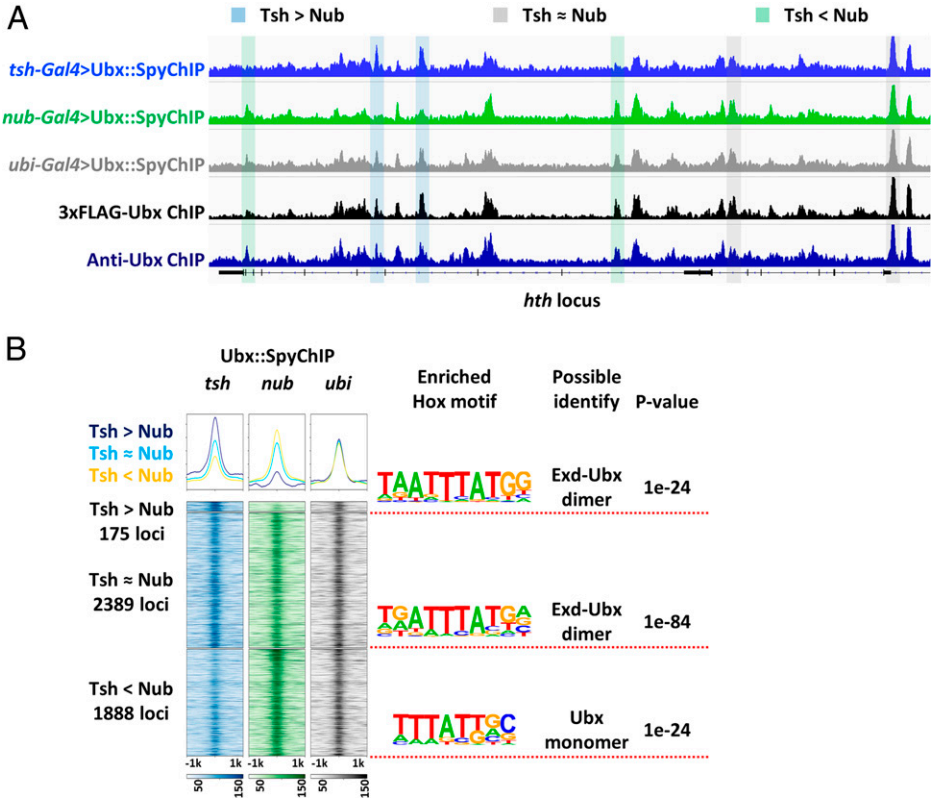

Fig. 3. SpyChIP identifies genome-wide and cell type-specific TF binding events. (*A*) SpyChIP results at the *hth* locus, which was chosen as an example. Examples of different classes of peaks are color coded: blue: Tsh Ubx::SpyChIP > Nub Ubx::SpyChIP; gray: Tsh Ubx::SpyChIP ≈ Nub Ubx::SpyChIP; and green: Tsh Ubx::SpyChIP < Nub Ubx::SpyChIP. Three SpyChIP tracks and two independent whole haltere disc Ubx ChIP tracks are shown. The 3xFLAG-Ubx (7) ChIP used the same anti-FLAG antibody as in all SpyChIP experiments. For comparison, the anti-Ubx ChIP track used an antibody directed against Ubx (8). (*B*) Heatmaps and histograms of Tsh > Nub, Tsh ≈ Nub, and Tsh < Nub Ubx::SpyChIP loci plotted for *tsh-Gal4* > Ubx::SpyChIP, *nub-Gal4* > Ubx::SpyChIP, and *ubi-Gal4* > Ubx::SpyChIP signals. Hox-related motifs significantly enriched in each class of loci are indicated. For a complete list of enriched motifs, see *SI Appendix*, Fig. S4.

The Role of Cell Type-Specific Ubx Binding. Recently, Loker et al. (8) characterized the genome-wide chromatin accessibility in Tsh$^+$ and Nub$^+$ cells of the haltere and the serially homologous wing imaginal discs. Given the cell type-specific Ubx binding data described here, we asked if there is any correlation between cell type-specific chromatin accessibility and cell type-specific Ubx binding. Notably, sites in the haltere that have Tsh > Nub Ubx binding also tend to be more accessible in Tsh$^+$ cells compared to Nub$^+$ cells, not only in the haltere disc, but also in the wing disc (Fig. 4 *A* and *B*). Since Ubx is expressed in the haltere disc but not in the wing disc, this pattern suggests that the 175 Tsh > Nub Ubx binding sites gain accessibility in Tsh$^+$ cells by a mechanism that is independent of Ubx binding. Similarly, many—but not all of the 1,888 Nub > Tsh Ubx binding sites—have biased accessibility in Nub$^+$ cells compared to Tsh$^+$ cells in both the haltere and wing (Fig. 4 *C* and *D*).

Finally, we inspected Ubx::SpyChIP patterns at selected Ubx downstream *cis*-regulatory modules (CRMs). For simplicity, we focused on CRMs that only require Ubx function in Nub$^+$ cells and also have Ubx ChIP peaks from whole haltere disc experiments, suggesting that they are direct Ubx targets. We included in our analysis *sal1.1* (11) and *kn01* (12), as well as four additional CRMs recently identified by Loker et al. (8) based on their differential accessibility in haltere Nub$^+$ cells compared to wing Nub$^+$ cells. Ubx acts as either an activator or a repressor of each CRM (Fig. 5). Among the six selected CRMs, four have Ubx binding only in Nub$^+$ cells, while the other two have Ubx binding in both Tsh$^+$ and Nub$^+$ cells. These patterns of binding and regulation are consistent with the existence of multiple modes of Ubx regulation. For example, for CRM Rep-6, which is activated by Ubx in Nub$^+$ cells, Ubx binding is observed in both Tsh$^+$ and Nub$^+$ cells and is apparently not sufficient for activation of this CRM. In contrast, Ubx only binds to CRM Rep-7 in Nub$^+$ cells, where it also acts as an activator, suggesting that the Nub$^+$ cell-specific Ubx binding pattern contributes to this CRM's restricted spatial activity in the Nub$^+$ cells.

Discussion

Characterizing cell type-specific binding is critical for understanding a TF's in vivo functions. The SpyChIP technique we describe in this study overcomes several major limitations of existing approaches. Because SpyChIP does not depend on tissue dissociation or nuclei sorting, it is especially suitable for tissues with limited availability or those that are difficult to dissociate. Contrary to the lower temporal resolution associated with DamID-based techniques, which capture a combination of past and present binding signatures, SpyChIP has as high a temporal resolution as standard ChIP and is therefore desirable in analyzing tissues undergoing dynamic rearrangements. We demonstrated the efficacy of SpyChIP by successfully obtaining cell type-specific Ubx ChIP results from the *Drosophila* haltere discs. These tiny tissues must be manually dissected and are therefore difficult to obtain in large quantity. Imaginal discs also undergo rapid cellular rearrangements during metamorphosis. In fact, to our knowledge, before our study no cell type-specific TF-DNA occupancy data have been reported

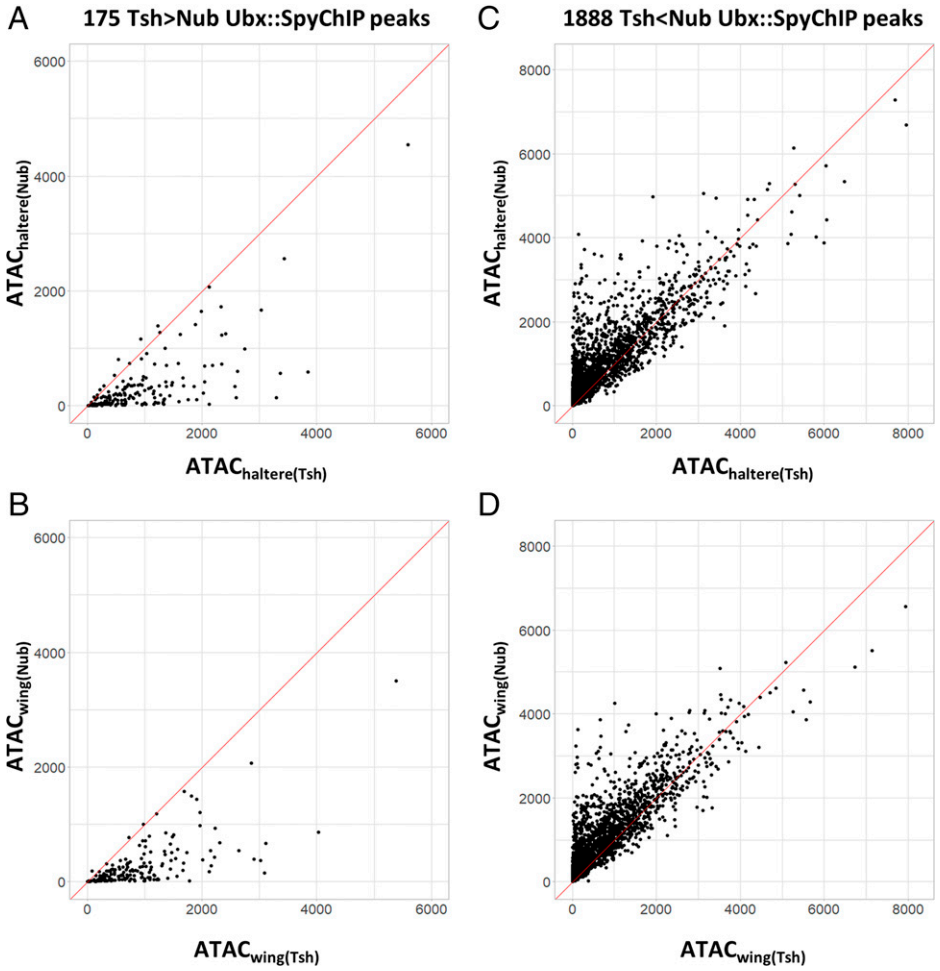

Fig. 4. Relationship between chromatin accessibility and TF-DNA binding revealed by SpyChIP. Scatter plots comparing chromatin accessibility of Tsh^+ and Nub^+ cells in 175 Tsh > Nub Ubx::SpyChIP peaks (*A* and *B*), or in 1,888 Tsh < Nub Ubx::SpyChIP peaks (*C* and *D*). The Tsh^+ vs. Nub^+ cells were compared in both the haltere disc (*A* and *C*) and the wing disc (*B* and *D*). Chromatin accessibility data are from Loker et al. (8).

from any *Drosophila* imaginal discs. The covalent bond between SpyTag and SpyCatcher is robust to diverse conditions, such as temperature and pH (4), thus SpyChIP is likely to be applicable in most tissues and in most organisms. If the target cell type represents a very small fraction in the complex tissue, SpyChIP may be combined with crude cell/nuclei sorting to partially enrich the target cells. Because the cell type-specificity of SpyChIP is genetically encoded, it is not necessary to obtain a highly pure cell population by cell sorting, which is often associated with lower yields.

Although a positive correlation is often observed between differential chromatin accessibility and differential TF binding, it is usually difficult to deduce the cause versus the consequence. With the aid of Ubx::SpyChIP, we were able to rule out that Tsh > Nub Ubx binding caused Tsh > Nub chromatin accessibility. Conversely, our results suggest that Tsh > Nub chromatin accessibility is permissive for Tsh > Nub Ubx binding pattern. It is generally believed that the same TF, especially a selective TF like Ubx, can regulate its downstream CRMs using different modes of action. However, it is not easy to demonstrate the existence of diverse mechanisms. Our Ubx::SpyChIP results show that Ubx binding is not always sufficient for CRM activation, suggesting the presence of multiple mechanisms that act in a CRM-specific manner.

Finally, we suggest that the SpyTag–SpyCatcher technology has the potential for many additional in vivo applications beyond SpyChIP. We envision that the covalent interaction between SpyTag and SpyCatcher can be combined with a variety of other techniques, such as HiChIP (13) and BioID (14), to achieve cell type-specificity without dissociation or cell/nucleus sorting. Moreover, once a factor has been fused with SpyTag by genome modification, it can be easily tagged with any peptide of interest, such as different epitopes, fluorescent proteins, or enzymes. Also noteworthy is that SpyTag and SpyCatcher are not the only pair of peptides that form a covalent bond when they interact: other orthogonal pairs have been reported to form covalent bonds in vitro (15). Therefore, there are many possibilities of in vivo applications of these covalent interacting peptide pairs.

Materials and Methods

New Fly Strains. All plasmid were generated by standard procedures, and transgenic flies were generated by integrating the plasmids into selected attP sites via phiC31 integrase mediated site-specific recombination.

The scarless *SpyTag-Ubx* allele was generated using a method we previously described (7). Briefly, a fragment of *Ubx* genomic DNA containing the *SpyTag* inserted at the N-terminal end of the *Ubx* ORF was integrated into the endogenous *Ubx* locus by phiC31 integrase mediated site-specific recombination. Double-stranded DNA breaks were then introduced to stimulate homologous recombination and repair the endogenous *Ubx* to the final scarless *SpyTag-Ubx* allele. The landing site for site-specific recombination in the *Ubx* locus has been described in detail, and the donor plasmid was generated similarly as previously described (7). The *SpyTag* sequence was inserted by overlapping extension PCR. Multiple independent *SpyTag-Ubx* alleles were generated, verified by Southern

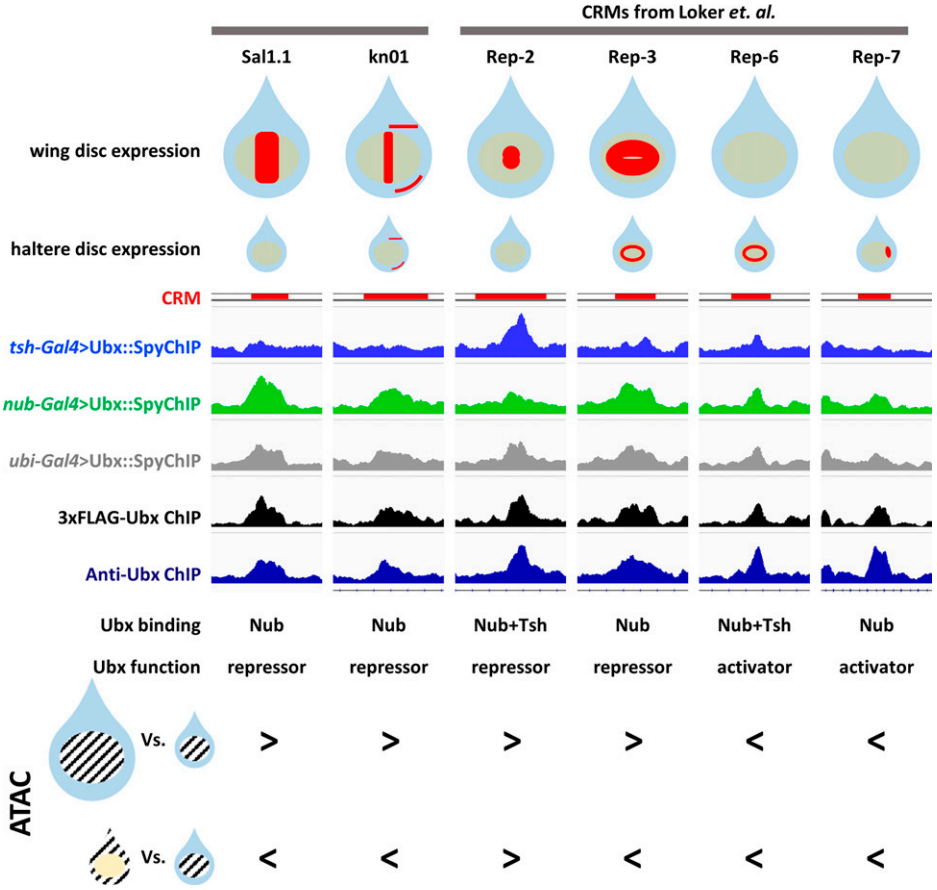

Fig. 5. SpyChIP reveals distinct Ubx regulatory strategies. Summary of Ubx binding, expression patterns and chromatin accessibility for selected Ubx-targeted CRMs. These CRMs were chosen because they bind Ubx and have been shown to require *Ubx* function, either as a repressor or activator as indicated, in Nub$^+$ cells (8). The top two rows are schematics of the CRM expression patterns in wing and haltere discs. Light blue and yellow colors mark the Tsh$^+$ and Nub$^+$ cells, respectively; red indicates CRM activity. Below are five genome browser views showing the Ubx::SpyChIP signals and whole disc Ubx ChIP signals, relative to the location of the CRMs (red bars). Whether Ubx binds in the Tsh+ or Nub$^+$ cells was determined by the MACS2 peak calling software (*Materials and Methods*). The bottom two rows compare the patterns of chromatin accessibility (8) between the Nub$^+$ cells of the wing vs. haltere (top row) and the Tsh$^+$ vs. Nub+ cells in the haltere (bottom row), for each CRM. Note that Ubx activity as a repressor or activator correlates with less or more accessibility, respectively, in haltere Nub$^+$ cells compared to wing Nub$^+$ cells. Also notable is that the four examples that have Tsh < Nub Ubx binding also have Tsh < Nub chromatin accessibility. In contrast, in the two cases where Ubx binding is observed in both Nub$^+$ and Tsh$^+$ cells, there is no correlation with accessibility differences.

blotting, and fully sequenced to make sure there were no unwanted mutations. Southern blotting was performed using DIG High Prime DNA Labeling and Detection Starter Kit II (Roche 11585614910) and DIG Wash and Block Buffer Set (Roche 11585762001) according to the manufacturer's instructions. The *Ubx* 5′ and *Ubx* 3′ probes were described previously (7). DNA Molecular Weight Marker II, DIG-labeled (Roche 11218590910) was used as the marker.

Western Blotting. Western blotting was performed using standard procedure. For embryo samples, embryos from desired crosses were collected overnight at 25 °C, and transferred to a 1-mL Wheaton homogenizer (not dechorionated). An appropriate volume of 4× SDS-PAGE loading dye (with 10% β-mercaptoethanol) was added (100 μL of the loading dye per ∼10 μL of settled embryos), and the embryos were completely homogenized. The homogenized materials were then transferred to 1.5-mL tubes. For each haltere disc sample, 35 to 55 discs were dissected in PBS+1% BSA on ice, and transferred to a 1.5-mL tube containing 0.5 mL of PBS+1% BSA. The supernatant was removed, and 100 μL of 4× SDS-PAGE loading dye (with 10% β-mercaptoethanol) was added. The haltere discs were then completely homogenized with a disposable pestle. The homogenized materials were heated at 95 °C for 6 to 7 min and chilled on ice. The samples were then spun at room temperature at maximum speed for 5 min, and the supernatant was loaded on SDS-PAGE. After SDS-PAGE, the proteins were transferred to PVDF membrane using routine procedure. The 3xFLAG epitope was detected using anti-FLAG M2-HRP (Sigma A8592; 1:10,000), and the V5 epitope was detected using mouse anti-V5 antibody (Invitrogen R96025; 1:5,000) followed by goat anti-mouse IgG-HRP (Jackson ImmunoResearch 115-035-003; 1:25,000), or with rabbit anti-V5 antibody (Abcam ab9116; 1:5,000) followed by donkey anti-rabbit IgG-HRP (Jackson ImmunoResearch 711-036-152; 1:5,000). The Ubx protein was detected using monoclonal mouse anti-Ubx FP3.38 (Developmental Studies Hybridoma Bank) at 1:100, followed by the same goat anti-mouse IgG-HRP secondary antibody at 1:10,000. SuperSigna West Pico PLUS Chemiluminescent Substrate (Thermo Scientific 34580) was used as the substrate to visualize the bands.

Chromatin Preparation. The larvae for Ubx::SpyChIP experiments were prepared by crossing *SpyTag-Ubx/(TM6B)* females to *Gal4/(CyO, GFP); attP2-UAS-3xFLAG-NLS-SpyCatcher*, *SpyTag-Ubx/(TM6B)* males. Three different *Gal4* lines–*tsh-Gal4*, *nub-Gal4*, and *ubi-Gal4*–were used. *TM6B*$^-$ and *GFP*$^-$ larvae were selected for dissection, and 100 to 150 larvae were dissected for each replicate. Homozygous *3xFLAG-Ubx* (7) larvae were also used for the whole haltere disc ChIP experiment. The larvae were pulled apart in PBS and the head parts were inverted. The inverted heat parts were cross-linked in 10 mL of cross-linking solution (10 mM Hepes pH 8.0, 100 mM NaCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, 1% formaldehyde) for 10 min at room temperature. After cross-linking, 1 mL of 2.5 M glycine was added and the samples were hand mixed for 30 s. The samples were then washed with 10 mL of quenching solution (1× PBS, 125 mM glycine, 0.1% Triton X-100) for at least 6 min at room temperature, followed by two more washes with 10 mL of ice-cold 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, with proteinase inhibitor mixture) at 4 °C, 10 min each. The gut, salivary glands, and fat bodies were then removed from all head parts in buffer A. Next,

the samples were washed twice with 10 mL of ice-cold 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, with proteinase inhibitor mixture) at 4 °C, 10 min each. The haltere discs were dissected from the head parts in buffer B, and were transferred to a 15-mL Falcon tube. The supernatant was removed, and 0.9 mL of buffer C (10 mM Hepes pH 8.0, 1 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, 1% Triton X-100, with proteinase inhibitor mixture) was added. The discs were then sonicated with Branson Sonifier 450 on ice at 15% amplitude for 12 min (15-s on/30s off). The sonicated samples were spun at maximum speed at 4 °C for 10 min, and the supernatant was transferred to new tubes, flash frozen in liquid N_2, and stored at −80 °C until the next step.

The SpyTag stock solution was prepared by dissolving synthetic SpyTag (Genscript custom peptide synthesis service) in water at a concentration of 1 mM. For replicates in which synthetic SpyTag was used to quench unoccupied SpyCatcher molecules, SpyTag was used at a final concentration of 10 μM in buffer B when dissecting haltere discs from the head parts, and in buffer C.

ChIP. ChIP was performed after all chromatin samples were prepared. The chromatin samples were thawed on ice, and to each sample, one-quarter volume of 5× chromatin dilution buffer (50 mM Tris·HCl pH 8.0, 5 mM EDTA pH 8.0, 750 mM NaCl, 1% Triton X-100) was added to adjust buffer condition, as well as appropriate volume of 100× Halt Protease Inhibitor Mixt, EDTA-Free (Thermo Scientific 87785). Next, 10 μg of normal mouse IgG was added to each sample for preclearing, and the samples were rotated at 4 °C for 1 h. Forty-microliters of protein G agarose beads suspension (Roche 11243233001) (settled beads volume 20 μL) was used for each ChIP and preclearing reaction. The beads were washed twice with 1 mL of RIPA buffer (10 mM Tris·HCl pH 8.0, 1 mM EDTA pH 8.0, 150 mM NaCl, 1% Triton X-100) for 10 min each at 4 °C with rotation, and were blocked with blocking solution (RIPA + 1.25 mg/mL BSA [Sigma A2153] + 0.25 mg/mL tRNA [Roche 10109517001]) for at least 1 h at 4 °C with rotation. The chromatin-normal IgG mixtures were added to blocked beads for preclearing, and were rotated at 4 °C for 1 h. The precleared chromatin was separated from beads by centrifugation. Next, 100 μL of each precleared chromatin was taken and stored at −80 °C as input; 12.5 μL of 100 mg/mL BSA, 25 μL of 10 mg/mL tRNA, and 10 μg of anti-FLAG M2 antibody (Sigma F1804) were added to the rest of precleared chromatin. The samples were rotated at 4 °C overnight.

In the next day, the chromatin samples were added to blocked beads, and were rotated at 4 °C for 2 h. The beads were briefly rinsed with RIPA buffer, and were subjected to the following 10-min washes at 4 °C: two washes with RIPA buffer, one wash with high-salt RIPA buffer (10 mM Tris·HCl pH 8.0, 1 mM EDTA pH 8.0, 350 mM NaCl, 1% Triton X-100), one wash with LiCl buffer (10 mM Tris·HCl pH 8.0, 1 mM EDTA pH 8.0, 250 mM LiCl, 0.1% IGEPAL CA-630), and one wash with TE buffer (10 mM Tris·HCl, 1 mM EDTA, pH 8.0, filtered). All rinses and washes were performed with 1 mL of ice-cold buffer. After the TE wash, the beads were resuspended in 500 μL of TE, and the input samples were also adjusted to 500 μL with TE buffer. Next, 5 μL of 5M NaCl, 12.5 μL of 20% SDS, and 10 μL of 1 mg/mL RNase (Sigma R5503) were added to each ChIP and input sample, and the samples were incubated at 37 °C for 30 min with rotation. Twenty microliters of 20 mg/mL proteinase K (Roche 03115836001) was then added to each sample. The samples were rotated at 55 °C for 2 to 3 h, and then at 65 °C overnight for decross-linking.

On the third day, all ChIP samples were centrifuged at room temperature at maximum speed for 5 min, and the supernatant was transferred to new tubes. Next, 100 μL of 3 M sodium acetate (pH 5.2) was added to each sample, and the samples were extracted with phenol:chloroform (1:1) and then with chloroform. One microliter of 20 mg/mL glycogen (Roche 10901393001) was then added to each sample, and the DNA was purified by isopropanol precipitation; 30 μL of 10 mM Tris buffer, pH 8.0 was used to dissolved the DNA pellet of each sample. The purified DNA was quantified using Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific Q32854).

ChIP-Seq Library Preparation and Sequencing. ChIP-seq libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs E7103) with modifications. One to 2 ng of ChIP DNA and 8 to 10 ng of input DNA was used as starting materials. No size selection was performed after adaptor ligation, and 11 PCR cycles were performed for all libraries. After PCR amplification, instead of purifying DNA using 0.9× of beads, the following purification protocol was used: 1.8× of beads was used to purify DNA from the PCR reactions. The DNA was eluted in 52 μL of elution buffer, and 50 μL was transferred to new tubes. The purified DNA was then subjected to size selection (0.65× for first bead addition, and 0.25× for second bead addition). The DNA was then eluted with 17 μL of elution buffer, and 15 μL was transferred to new tubes. The sizes of the libraries were determined by bioanalyzer, using Bioanalyzer High Sensitivity DNA Analysis (Agilent 5067-4626), and the libraries were quantified by Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific Q32854). The libraries were sequenced using Illumina Nextseq. 500 sequencer.

ChIP-Seq Data Analysis. Mapping and peak calling were performed using tools on usegalaxy.eu. The reads were mapped to *Drosophila* genome build dm6 by Bowtie2 (16) using default settings, and peak calling was performed by MACS2 (17) with the following parameters: –nomodel –extsize 200 (all other parameters were default). Differential binding analysis was performed using DiffBind (18), following default procedures. Heatmaps were generated using deeptools2 (19) (also on usegalaxy.eu), and scatter plots were generated using the R package ggplot2. de novo motif searches were performed using Homer (20), and all parameters were default except -size 80.

Data Availability. Next-generation sequencing data have been deposited in the Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/geo (accession no. GSE189554).

ACKNOWLEDGMENTS. We thank Nicolas Gompel for permission to use the fly image in Fig. 1*B*, and all members of the Mann laboratory for discussions. This study was supported by NIH Grant R35 GM118336 (to R.S.M.).

1. P. J. Park, ChIP-seq: Advantages and challenges of a maturing technology. *Nat. Rev. Genet.* **10**, 669–680 (2009).
2. S. Bonn *et al.*, Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat. Genet.* **44**, 148–156 (2012).
3. T. D. Southall *et al.*, Cell-type-specific profiling of gene expression and chromatin binding without cell isolation: Assaying RNA Pol II occupancy in neural stem cells. *Dev. Cell* **26**, 101–112 (2013).
4. B. Zakeri *et al.*, Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E690–E697 (2012).
5. E. B. Lewis, Genes and developmental pathways. *Am. Zool.* **3**, 33–56 (1963).
6. S. M. Cohen, "Imaginal disc development" in *The Development of Drosophila melanogaster*, M. Bate, A. Martinez-Arias, Eds. (Cold Spring Harbor Laboratory Press, New York, 1993), **vol. 2**, pp. 747–842.
7. S. Feng, S. Lu, W. B. Grueber, R. S. Mann, Scarless engineering of the Drosophila genome near any site-specific integration site. *Genetics* **217**, iyab012 (2021).
8. R. Loker, J. E. Sanner, R. S. Mann, Cell-type-specific Hox regulatory strategies orchestrate tissue identity. *Curr. Biol.* **31**, 4246–4255.e4 (2021).
9. R. K. Delker, V. Ranade, R. Loker, R. Voutev, R. S. Mann, Low affinity binding sites in an activating CRM mediate negative autoregulation of the Drosophila Hox gene Ultrabithorax. *PLoS Genet.* **15**, e1008444 (2019).
10. S. Merabet, R. S. Mann, To be specific or not: The critical relationship between Hox and TALE proteins. *Trends Genet.* **32**, 334–347 (2016).
11. R. Galant, C. M. Walsh, S. B. Carroll, Hox repression of a target gene: Extradenticle-independent, additive action through multiple monomer binding sites. *Development* **129**, 3115–3126 (2002).
12. D. J. McKay, J. D. Lieb, A common set of DNA regulatory elements shapes Drosophila appendages. *Dev. Cell* **27**, 306–318 (2013).
13. M. R. Mumbach *et al.*, HiChIP: Efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).
14. S. Long, K. M. Brown, L. D. Sibley, CRISPR-mediated Tagging with BirA allows proximity labeling in *Toxoplasma gondii*. *Bio Protoc.* **8**, e2768 (2018).
15. G. Veggiani *et al.*, Programmable polyproteams built using twin peptide superglues. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 1202–1207 (2016).
16. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
17. Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
18. C. S. Ross-Innes *et al.*, Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
19. F. Ramírez *et al.*, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
20. S. Heinz *et al.*, Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).