Ancestral and derived transcriptional enhancers share regulatory sequence and a pleiotropic site affecting chromatin accessibility

Yaqun Xin[a,1], Yann Le Poul[a,1], Liucong Ling[a], Mariam Museridze[a], Bettina Mühling[a], Rita Jaenichen[a], Elena Osipova[a], and Nicolas Gompel[a,2]

[a]Fakultät für Biologie, Biozentrum, Ludwig-Maximilians-Universität München, 82152 Planegg-Martinsried, Germany

Edited by William J. McGinnis, University of California San Diego, La Jolla, CA, and approved July 13, 2020 (received for review March 2, 2020)

The diversity of forms in multicellular organisms originates largely from the spatial redeployment of developmental genes [S. B. Carroll, *Cell* 134, 25–36 (2008)]. Several scenarios can explain the emergence of *cis*-regulatory elements that govern novel aspects of a gene expression pattern [M. Rebeiz, M. Tsiantis, *Curr. Opin. Genet. Dev.* 45, 115–123 (2017)]. One scenario, enhancer co-option, holds that a DNA sequence producing an ancestral regulatory activity also becomes the template for a new regulatory activity, sharing regulatory information. While enhancer co-option might fuel morphological diversification, it has rarely been documented [W. J. Glassford et al., *Dev. Cell* 34, 520–531 (2015)]. Moreover, if two regulatory activities are borne from the same sequence, their modularity, considered a defining feature of enhancers [J. Banerji, L. Olson, W. Schaffner, *Cell* 33, 729–740 (1983)], might be affected by pleiotropy. Sequence overlap may thereby play a determinant role in enhancer function and evolution. Here, we investigated this problem with two regulatory activities of the *Drosophila* gene *yellow*, the novel *spot* enhancer and the ancestral *wing blade* enhancer. We used precise and comprehensive quantification of each activity in *Drosophila* wings to systematically map their sequences along the locus. We show that the *spot* enhancer has co-opted the sequences of the *wing blade* enhancer. We also identified a pleiotropic site necessary for DNA accessibility of a shared regulatory region. While the evolutionary steps leading to the derived activity are still unknown, such pleiotropy suggests that enhancer accessibility could be one of the molecular mechanisms seeding evolutionary co-option.

transcriptional regulation | regulatory evolution | pattern formation | chromatin | enhancer

Evolutionary co-option happens when an ancestral biological object is recycled to a new function while maintaining its ancestral role. Novel *cis*-regulatory elements (transcriptional enhancers), for instance, may emerge through co-option of a preexisting element. In this case, the ancestral and the derived regulatory functions map to overlapping DNA segments, which we define as structural co-option. They may share ancestral components such as ancestral transcription factor binding sites (TFBSs), bringing co-option to a functional level but resulting in a functional dependency or pleiotropy (1–5). Because the boundaries of transcriptional enhancers are difficult to define precisely, it is most often challenging to assess sequence overlap and regulatory pleiotropy when a new regulatory activity emerges in the vicinity of an ancestral activity (6–8). An enhancer is typically defined on the basis of its activity, notably in a transgenic context, using reporter assays as a segment of sequence sufficient to direct a spatiotemporal transcriptional activity resembling that of their original target gene (9–12). In developmental biology, enhancer boundaries are defined from a DNA sequence sufficient to recapitulate specific elements of the endogenous expression pattern of the corresponding gene. This definition has several limits. One limit, not addressed in this study, is that the biological context in which enhancer activity is assessed differs from the native genomic and transcriptional context. Another limit is that it focuses on the relative spatial distribution of the regulatory activity, the pattern, rather than on its quantitative aspects and is therefore likely to reveal only partial enhancer sequences and to miss pleiotropic effects. Moreover, fragments are often chosen either arbitrarily or based on sequence conservation or genomic marks to limit the risk of disrupting functional features. These fragments can pinpoint minimal enhancers but fail to determine whether the same sequences at their locus of origin are necessary and sufficient to recapitulate the transcriptional activity of their cognate target gene (13–15). Finally, the representation of enhancers as rectangular boxes or stretches of sequence eludes the actual distribution of regulatory information along the enhancer sequence with different segments contributing different inputs (activation, repression, permissivity) and different activity levels. In an attempt to overcome most of these limits, we examine here the molecular relationship that a new regulatory activity entertains with a nearby ancestral activity.

While the wings of *Drosophila* are uniformly shaded with light gray pigment, some species, including *Drosophila biarmipes*, have gained a pattern of dark pigmentation, a spot, at the wing tip (7). The expression of the gene *yellow* (*y*) in the wings during pupal life is necessary both to the wing blade shading and to the spot pattern (16). These two components of *yellow* wing expression result from two distinct regulatory regions, the ancestral *wing*

Significance

Form diversity is fueled by changes in the expression of genes that build organisms. New expression often results from the emergence of new DNA switches, known as transcriptional enhancers. Many enhancers are thought to appear through the recycling of older enhancers, a process called evolutionary co-option. Enhancer co-option is difficult to assess, and the molecular mechanisms explaining its prevalence are elusive. Using state-of-the-art quantification and analyses, we reveal that the sequences of an ancestral and a derived enhancer overlap extensively. They contain specific binding sites for regulators imparting spatial activities. We found that the two enhancers also share a site facilitating access to chromatin in a region where they overlap.

Author contributions: Y.X., Y.L.P., L.L., and N.G. designed research; Y.X., Y.L.P., L.L., M.M., B.M., R.J., and E.O. performed research; Y.X., Y.L.P., L.L., M.M., and N.G. analyzed data; and Y.L.P. and N.G. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

[1]Y.X. and Y.L.P. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: gompel@bio.lmu.de.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2004003117/-/DCSupplemental.

First published August 10, 2020.

blade enhancer (referred to as "*wing*" in other publications) and the recently evolved *spot* enhancer (6, 7, 17–21). In *D. biarmipes*, both activities map within 6 kb upstream of *y* transcription start site (6) (*y* 5′ region) (Fig. 1*A*). Two short adjacent regulatory fragments (~1.1 kb together) within this *y* 5′ region drive distinct spatial expression in the spot and uniformly in the wing blade, respectively (6, 16). It is, however, unclear to what extent sequences surrounding these fragments at their locus of origin also contribute to each transcriptional activity. It is equally unclear whether or not the contributing sequences of the two enhancers overlap. Because both activities are driven in the same tissue and developmental stage, it is technically and conceptually challenging to evaluate the distribution of regulatory information quantitatively and assess possible pleiotropic effects.

Testing the hypothesis of enhancer structural co-option in our system required us to link regulatory information distributed in DNA to activities measured with quantitative spatial reporter expression. Using classical reporter assays in transgenic *Drosophila*, we mapped regulatory information with two series of nested fragments, depleting sequence information from the 3′ end or the 5′ end. This approach reveals the contribution of DNA segments along the sequence, including sequences that cannot drive activity alone and whose activity depends on nearby sequences. A simple qualitative assessment of the reporter activity resulting from each construct is, however, insufficient to produce a precise regulatory map. Moreover, qualitative or semiquantitative approaches would not allow us to separately measure each regulatory activity because of the spatial and temporal overlap with the other activity. This prompted us to develop a generic quantification pipeline to comprehensively describe variation in reporter expression levels across the wing. Finally, with an appropriate analytical framework, we have mathematically separated the two activities, although they drive in the same tissue and developmental stage. Our results indicate that the regulatory information spans a much wider region than previously described and that, unexpectedly, the ancestral *wing blade* and the derived *spot* activities overlap extensively. Further, the molecular dissection of the overlapping region led us to uncover a site with pleiotropic effects in the core of the derived enhancer, which proved to regulate chromatin accessibility.

Results

To evaluate how the *wing blade* and the *spot* activities are distributed along *y* 5′ sequences of *D. biarmipes* and to test whether they are intertwined, we derived two series of reporter constructs from the *y* 5′ region (Fig. 1*B*) and tested them in *Drosophila melanogaster*. The first series (*D*) consists of distal (5′) truncations, while in the second series (*E*), we randomized increasingly longer segments of wild-type proximal (3′) sequence, keeping the total fragment size constant (identical to that of construct *D2*). In each series, the largest intact fragment is a reference for the complete regulatory information (*D0* in the 5′ dissection and *D2* in the 3′ dissection) (Fig. 1*B*). These two series allow us to measure how a segment modulates regulatory information, when the information in 3′ (*D* series) or in 5′ (*E* series) of this segment is preserved. We define as enhancer core any segment that, in its local genomic context (including the distance to the core promoter), is necessary and sufficient to drive significant levels of a given activity (see below).

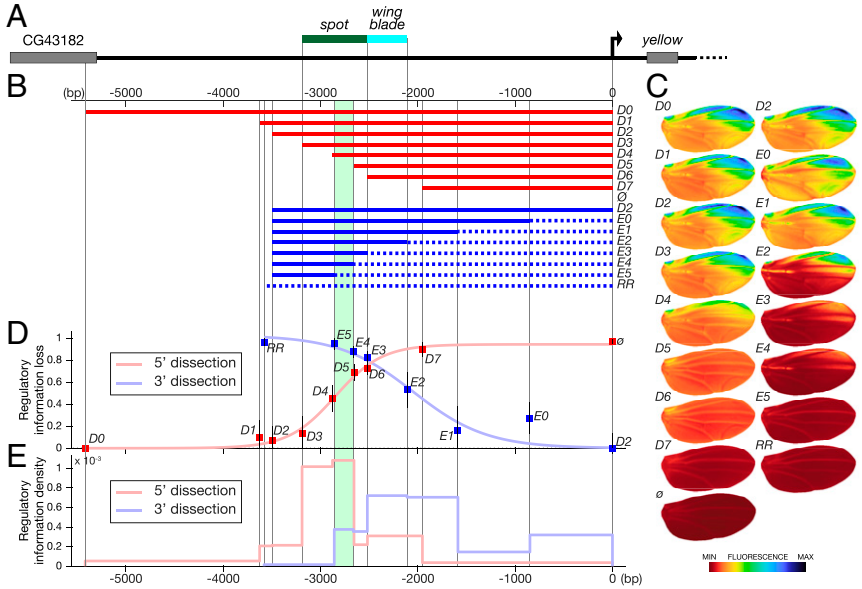

Fig. 1. Quantitative mapping of wing regulatory activities at the *yellow* locus. (*A*) The top line represents the 5′ region of the *yellow* locus from *D. biarmipes*; the green and blue bars indicate the respective locations of *spot* and *wing blade*, respectively, as originally mapped (6). (*B*) Two series of fragments derived from *y* 5′ region (*D* series, red; *E* series, blue) were tested in reporter constructs in *D. melanogaster*. The dotted lines in the *E* series represent randomized sequences (*Materials and Methods*); *ø* and *RR* stand for an empty reporter vector and a vector containing a completely randomized fragment, respectively. The area shaded in green in *B*, *D*, and *E* identifies a previously studied regulatory component (16), *spot*^{196}. (*C*) Images of average reporter expression of all individuals for each construct in the wing at emergence from the pupa according to the color map below. Note that *spot*^{196} appears strictly necessary to any activity in the spot region (compare *D4* with *D5* and compare *E4* with *E5*). (*D*) Overall loss of regulatory information (fluorescence levels) along the sequence (base pairs). The loss of phenotypic information measures how much truncating or randomizing a fragment affects the whole activity relatively. It is estimated by the ratio $\frac{d(P_x, P_{ref})}{d(P_\varnothing, P_{ref})}$, where P_x, P_{ref}, and $P_\varnothing$ are the phenotypes of construct *x*, construct *D0* or *D2* (the largest constructs of each series as a reference for that series), and the empty construct *ø* in the PCA space, respectively, plotted as a function of the distance to the starting point of the randomization (series *E*) or truncation (series *D*). Error bars represent the SD of the phenotype of each construct in PCA space normalized by the distance $d(P_\varnothing, P_{ref})$. (*E*) Density of regulatory information along the *y* 5′ region (fluorescence levels per base pair). It is technically the first derivative of the regulatory information loss shown in *D*. For each series, it represents the phenotypic distance (in PCA space) between two consecutive constructs divided by the number of base pairs that changed between those two constructs. It indicates the regulatory contribution per base pair of each DNA segment measured in each series.

We imaged 27 wings on average (minimum 22; maximum 39) for each construct and used them to precisely quantify spatial reporter expression (referred to as phenotype) driven by each construct in the wings of transgenic *D. melanogaster*, used here as an experimental recipient with site-specific transgenesis (22) (Fig. 1*C*). We summarized the variation in activity across the wing (both pattern and levels) from each series of constructs with principal component analysis (PCA), producing a comprehensive description of the phenotypic variation (*SI Appendix*, Fig. S1*A*). We define the overall loss of regulatory information for each construct as the amount of change in activity compared with the activity of a reference construct. To estimate this loss, we use the distance between the average phenotypes, as described in PCA space. This distance takes any change of activity into account. As this measure is more informative when represented relatively, we normalized the loss of regulatory information to the total amount of regulatory information brought by the enhancer, as estimated by the distance between the reference activity and the empty construct. The relative loss is therefore given by the following formula:

$$\frac{d(P_x, P_{ref})}{d(P_\emptyset, P_{ref})},$$

where P_x, P_{ref}, and $P_\emptyset$ are the average phenotypes of construct x, the reference construct (*D0* or *D2*, the largest constructs of each series), and the empty construct $\emptyset$, respectively, and $d(P_x,P_{ref})$ is the distance between these average phenotypes. Hence, this ratio estimates the loss of regulatory output of each construct compared with the largest construct of the series. In contrast to classical reporter assays testing the sole sufficiency of candidate regulatory fragments to produce a spatial pattern, the combined series reveal a surprisingly large stretch of the regulatory activities along *y* 5′ sequences (the regulatory activity of each construct is significantly different from that of the largest construct of the series) (*SI Appendix*, Table S1). Further, Fig. 1*E* establishes the contribution of each segment to these activity differences (intensity effect/base pair). Consistent with previous work (6), the 5′ series (*D*) shows that most of the regulatory activity maps within ~1.7 kb (−3.6 to −2 kb) (Fig. 1 *D* and *E*). The 3′ dissection, however, reveals additional regulatory information contributing to the activity, located proximally to this 1.7-kb segment and extending to *y* promoter region (Fig. 1 *D* and *E*). These results demonstrate that *y* regulatory activities in the wing extend over 3 kb (conservative) to 4 kb upstream of *y* promoter, a much broader region than previously assessed (6, 7).

To specifically address the question of regulatory co-option, we then examined the sequence relationship between *spot* and *wing blade* activities. It was first necessary, however, to mathematically separate the *wing blade* and the *spot* activities to then evaluate to what extent they map to distinct segments. In the PCA of all constructs, we found that both the *D* and the *E* series varied mostly along a combination of two additive directions in the phenotype space, explaining a large part (69%) of the phenotype variance resulting from the two dissection series. We noticed that these two directions correspond to a near-uniform increase in expression across the wing and an increased expression mostly at the anterior distal wing tip, respectively. These two directions map to overlapping sequence segments: −2,656 to 0 bp (ø to *D5*) and −3,496 to −2,519 bp (*RR* to *E2*, where RR is a segment of randomized sequence; see *Materials and Methods*), respectively (reference segments in Fig. 2 *B* and *C*). The segment driving a uniform pattern of activity fully includes the originally defined *wing blade* enhancer (6) but not the full original *spot* enhancer. Surprisingly, the segment driving a spotted pattern of activity includes both the originally defined *spot* and *wing blade* enhancers (6), despite its very low activity in the wing blade.

Hence, guided by the structure of the phenotypic space, we extracted representations of the actual patterns of activity driven by the *wing blade* and the *spot* enhancers, where *D5* and *E2* are representative segments of each direction, respectively (Fig. 2 *B* and *C* and *SI Appendix*, Fig. S1*A*). The segments defining the two activities (−3,496 to −2,519 bp for the *spot* activity and −2,656 to 0 bp for the *wing blade* activity) share regulatory information, indicating that our estimate of the structural co-option is conservative as it tends to minimize the measured sequence overlap between the two activities. It is important to note that the definition of those two directions (independently representing the *spot* and *wing blade* activities) (axes of Fig. 2*A*) is not linked to prior knowledge on these enhancers, neither from the phenotypic nor the sequence point of view. The fact that those data-driven directions correspond to uniform and spotted activities confirms that the two activities map mainly, when the two series are considered separately, to different segments. It also shows that the full 5′ region of *y* drives mainly two different activities, apparently relatively independently. Structural co-option implies that at least some segments of *y* 5′ contribute to the *wing blade* and *spot* activities simultaneously. Because the two activities overlap in space in the wing, they cannot be distinguished by simply measuring the separate reporter expression in their respective domains. To independently evaluate the uniform activity and the patterned, spotted activity, we projected the phenotype of each individual wing in the two-dimensional basis defined by these two phenotypic directions using a mathematical operation called change of basis (*Materials and Methods*, Fig. 2*A*, and *SI Appendix*, Fig. S1*A*). With the possibility to evaluate *wing blade* and *spot* activities independently, we quantified the contribution of each DNA segment to the respective activities.

We first tested whether, in the case of the *wing blade* and *spot* enhancers, the enhancer cores, as defined above, mapped to the same region. In our experimental system, the core of an enhancer is a segment sufficient to contribute a uniform or a spotted activity in the wing when either flanking 5′ or 3′ regions are missing. Because of the particular enhancer configuration in our system, each dissection series is simultaneously testing the sufficiency of a segment for one activity and its necessity for the other activity. This definition takes the preserved distance of regulatory information to the core promoter into account as well as the local genomic context at the *yellow* locus. We submit that this approach is more informative than testing the sole sufficiency of an isolated segment, as is classically done (21). These cores can logically be visualized in Fig. 2 *B* and *C* as the intersection between the 5′ and 3′ dissection curves. The core of the *spot* activity as revealed here coincides exactly with the $spot^{196}$ enhancer, defined in previous work (6, 16). For the *wing blade* enhancer, interestingly, there are two cores (from −2,111 to −1,953 bp and from −2,877 to −2,518 bp) flanking what was previously defined as the *wing blade* enhancer (6). Thus, there are two regions sufficient to drive a significant amount of *wing blade* activity when either 5′ or 3′ regulatory information is missing. Moreover, the overlap between the core of the *spot* enhancer and one of the cores of the *wing blade* enhancer reveals that a region inside the *spot* enhancer is sufficient to drive a substantial amount of expression in the wing blade.

Further investigating the interweaving of the two activities, we found, strikingly, that the sequences contributing to them largely overlap (Fig. 2 *B* and *C* and *SI Appendix*, Fig. S1 *C* and *D*). We asked whether sequences 3′ to the *spot* reference segment also contributed significant regulatory information to the *spot* activity. To this end, we compared *D2* (the largest fragment of the *E* series) with *E2*, in which these 3′ sequences are randomized (−2,111 to 0 bp) and found that this region contributes a substantial and unexpected amount of *spot* activity [22%, ANOVA: $F(1, 55) = 22.57$, $P = 1.4954\text{e-}05$] (horizontal double arrow in Fig. 2*A* and 3′ curve in Fig. 2*B*). Reciprocally, we asked whether

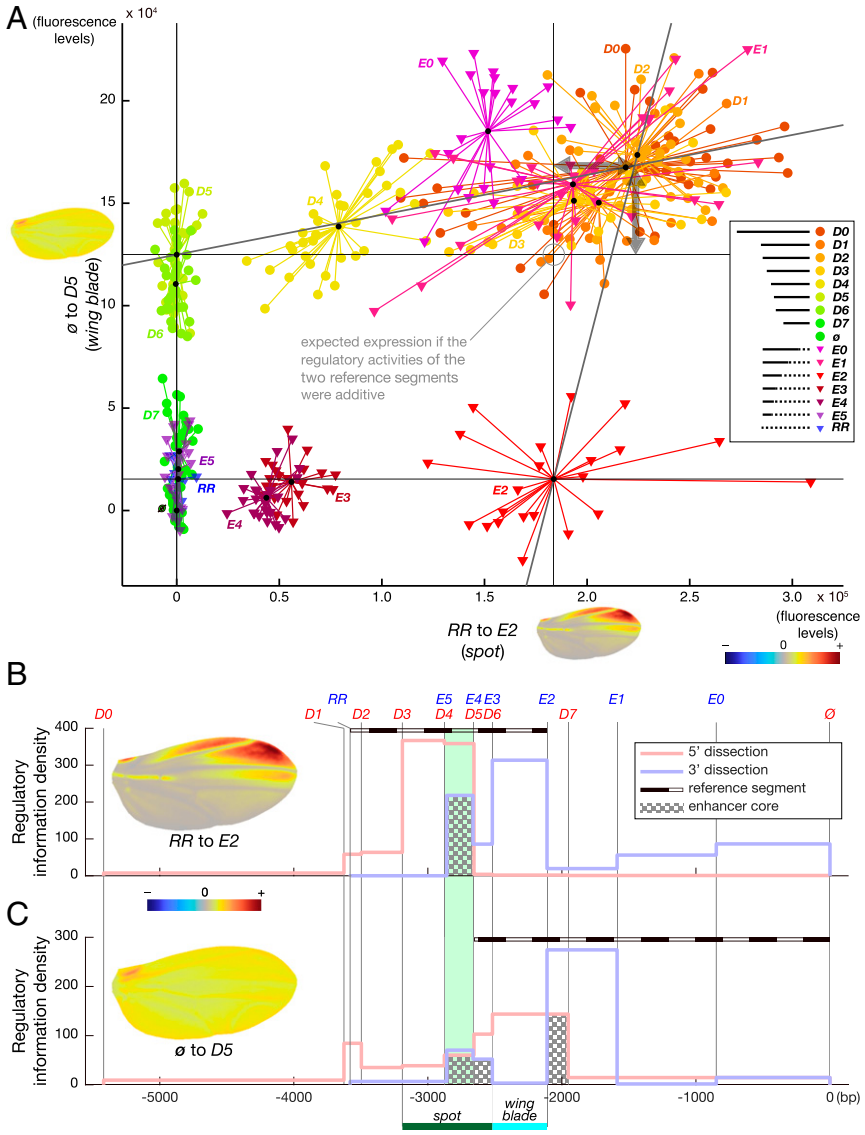

Fig. 2. *wing blade* and the *spot* activities map to overlapping sequences. (*A*) Representation of *wing blade* activity as a function of *spot* activity. Independent estimates were produced by projecting the PCA phenotypic space (PCA in *SI Appendix*, Fig. S1*A*) on a two-vector basis defined by two independent directions identified in *SI Appendix*, Fig. S1*A* (phenotypic directions with color map near each axis) and corresponding to *wing blade* (constructs *ø* to *D5*; dotted line in *B*) and *spot* (constructs *RR* to *E2*; dotted line in *C*) activities. The mathematical change from the PCA coordinate basis to this two-vector coordinate basis affords the separation and independent measurements of both activities, although they occur in the same tissue. This graph shows for each individual wing (dots and triangles) of each reporter line the contribution to the *wing blade* and *spot* activities. Small black dots mark the center of a cluster for each construct. Note that constructs driving both activities (*D0* to *D4*, *E0* to *E1*) produce more expression than expected if the activities were strictly additive (i.e., they lie above the point of strict additivity of the activities driven by the two reference segments of the *wing blade* and the *spot* activities; the resulting nonadditive effects are shown with double arrows). (*B* and *C*) Density of regulatory information along the *y* 5′ region (fluorescence levels per base pair) as measured specifically (*Materials and Methods*) for the *spot* activity (*B*) and the *wing blade* activity (*C*). Construct boundaries are delineated with vertical gray lines labeled with the construct name on top in *B* and *C*. The original *spot* and *wing blade* boundaries (6) are indicated by a green bar and a blue bar, respectively, for comparison. Color scheme is the same as in Fig. 1. Enhancer cores, defined in the results as the intersection between the 5′ and 3′ dissection curves, are highlighted with a checkerboard pattern in *B* and *C*.

sequences 5′ to the *wing blade* reference segment also contributed significant regulatory information to the *wing blade* activity. When comparing *D0* (the largest construct of the *D* series) with *D5*, in which these 5′ sequences are truncated, we observed an increase of *wing blade* activity of 34% [ANOVA: $F(1, 68) = 56.35$, $P = 1.7205\text{e-}10$] (vertical double arrow in Fig. 2*A* and 5′ curve in Fig. 2*C*). If activities driven by the truncated segment in *D5* (−5,419 to −2,656 bp) and the randomized segment in *E2* (−2,111 to 0 bp) were strictly additive, the phenotypes in Fig. 2*A* would form, conservatively, a perfect rectangle (indicated by four lines in the graph). Additivity would translate geometrically into the addition of the two vectors *ø* to *D5* and *RR* to *E2*, placing the maximum of each activity measured along each direction at the top right corner of this rectangle. Yet, this is not the case, indicating that the sequences contributing to the *spot* activity between −2.8 kb and the core promoter and those contributing to the *wing blade* activity between −5,419 and −2,656 bp are not sufficient to drive the maximum activity. Their effects require the presence of sequences in 5′ for the *spot* activity and sequence in 3′ for the *wing blade* activity, respectively. This is confirmed by the fact that those same sequences show very little to no effect in 5′ dissection for the *spot* activity and in the 3′ dissection for the

wing blade activity. We concluded from this analysis that, although their cores are partially distinct, the derived *spot* activity is largely intertwined in the DNA segment driving the ancestral *wing blade* activity. This strongly suggests that the *spot* enhancer evolved by co-opting the ancestral regulatory segment and raises the possibility that the two enhancer regions share pleiotropic inputs. The notion of enhancer pleiotropy is suggested or discussed as such by several other studies (23–26). In two cases, enhancer pleiotropy was shown to directly result from shared TFBSs in enhancers active in different tissues and at different times of development (3, 27). Although it is unclear whether the *wing blade* and *spot* activities share regulatory information that would result in enhancer pleiotropy, our observations prompted us to explore the modalities of these regulatory interactions further.

In principle, the *spot* and the *wing blade* enhancer, although intertwined, may be functionally independent, with separate sets of intermingled TFBSs. They may on the contrary share TFBSs. In our quantitative mapping (Fig. 1), we noticed that the overlap between the *spot* and *wing blade* activities encompasses a 196-bp fragment (the segment between *D4* and *D5*) (Fig. 1*B*) with interesting regulatory properties. It is indeed necessary for the overall *spot* activity (i.e., any construct missing this fragment displays no spot pattern) (Figs. 1 *B* and *C* and 2*B*, intersection between the 5′ and 3′ dissection curves). In addition, it contributes quantitative information both to the *spot* and the *wing*

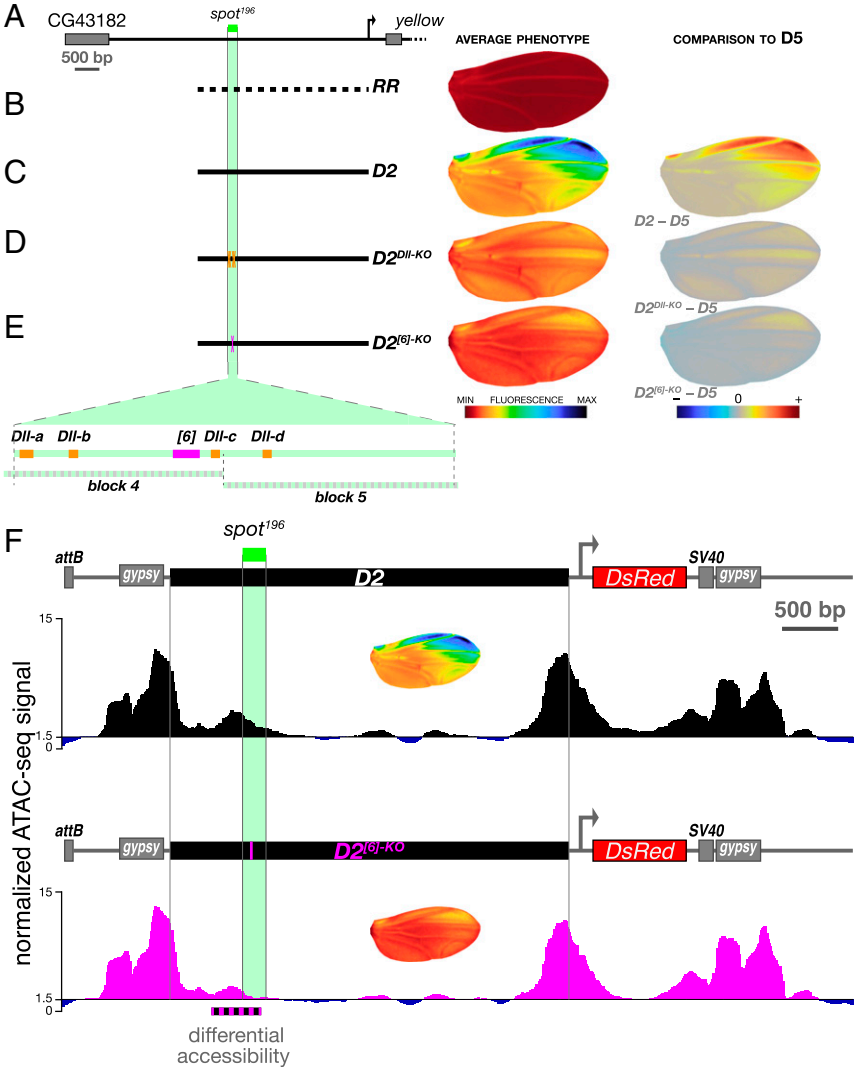

Fig. 3. Shared regulatory inputs of the *wing blade* and the *spot* activities. (*A*) A map of the *yellow* locus 5′ region highlighting the position of the $spot^{196}$ core. (*B–E*) The *wing blade* and the *spot* activities are strongly affected by discrete mutations in *D2*. (*Left*) Construct schematics. (*Center*) Average phenotype. (*Right*) Comparison with construct *D5* phenotype (difference). The positions of mutated sites as well as those of blocks 4 and 5 (*SI Appendix*, Fig. S2) are depicted on blown-up schematics of the $spot^{196}$ core in *E*. (*B*) *RR* is a negative control, the same randomized fragment as in Fig. 1 *B* and *C*. (*C*) *D2* is identical to Fig. 1 *B* and *C*. (*D*) Mutating all four characterized Dll binding sites (16) of $spot^{196}$ in the context of *D2* ($D2^{DII\text{-}KO}$) reduces the *spot* activity strongly and the *wing blade* activity moderately, as seen when comparing this mutant construct with *D5*. (*E*) Mutating a newly identified activator site (28) of $spot^{196}$ ($spot^{196\ [6]}$, 12 bp mutated) in the context of *D2* ($D2^{[6]\text{-}KO}$) reduces both *spot* and *wing blade* activities strongly, as seen when comparing this mutant construct with *D5*. (*F*) Chromatin accessibility measured with ATAC-seq at the *D2* and $D2^{[6]\text{-}KO}$ transgenes at the onset of *spot* activity (*SI Appendix*, Fig. S3) (66-h pupal wings) differs significantly in a 500-bp region overlapping $spot^{196\ [6]}$ (dotted black and magenta line). This is the only region in the entire locus identified as a differentiated site using diffBind (50, 51) and DEseq2 (52) analyses (*Materials and Methods*) (adjusted *P* value from the DESeq2 analysis: 7.21E-08). ATAC-seq traces represent the pooled signal of three replicates for each transgenic line (*SI Appendix*, Fig. S4). The discrepancy between the enhancer boundaries defined in Fig. 1 and the accessible region of *F* may stem in part from the different stages at which these properties were assessed. Average activity phenotypes of each construct also shown in *C* and *E* are indicated in *Insets* under each construct diagram as a reminder.

blade activities, as we have seen above (Figs. 1 and 2), and is a second enhancer core of the *wing blade* activity. We confirmed this core function of the *spot* activity when we randomized small blocks of sequence (100 bp) overlapping the 196-bp fragment in the context of *D2*. The randomization of the proximal half of this core element (*SI Appendix*, Figs. S1*B* and S2, $D2^{block5}$) reduces the *spot* activity by 61% [ANOVA *D2* vs. $D2^{block\ 5}$: $F(1, 44) = 516.84$, $P = 5.9730e\text{-}26$] without affecting the average levels of *wing blade* activity [ANOVA *D2* vs. $D2^{block\ 5}$: $F(1, 44) = 0.58$, $P = 0.452$]. By contrast, the randomization of the distal half of this core element (*SI Appendix*, Figs. S1*B* and S2, $D2^{block4}$) abolishes the *spot* activity completely and suppresses the nonadditive effects on *wing blade* activity described above [ANOVA $D2^{block\ 4}$ vs. *D5*, $F(1, 45) = 0.025$, $P = 0.876$] (*SI Appendix*, Fig. S2*D*). In previous studies (6, 16), we had analyzed these 196 bp (called $spot^{196}$) because they represented a minimal enhancer to understand the evolution of a spatial expression pattern (not the transcription levels). In particular, we found that this fragment was activated by the transcription factor (TF) Distal-less (Dll) through at least four TFBSs (16), three of which map to the region randomized in $D2^{block4}$ (Fig. 3). In a recent and independent dissection of $spot^{196}$, we identified a potential site for one or more unknown transcription factor(s), $spot^{196\ [6]}$, whose mutation (12 bp) nearly abolishes $spot^{196}$ activity completely (28). It is conceivable that these sites necessary for the *spot* activity also influence the *wing blade* activity, thereby producing pleiotropic effects. We mutated them in the context of *D2* to measure their relative contribution to the *spot* and the *wing blade* activities (Fig. 3). $D2^{Dll\text{-}KO}$ and $D2^{[6]\text{-}KO}$ resulted in strong effects on the *spot* (Fig. 3 *A*–*E* and *SI Appendix*, Fig. S1*B*), and both abolished the nonadditive *wing blade* activity, bringing it to the levels of *D5* (*SI Appendix*, Fig. S1*B*). Mutating the sole site $spot^{196\ [6]}$ in *D2*, along with abolishing 85% of the *spot* activity, also reduced the *wing blade* activity by 44% compared with *D2* (*SI Appendix*, Fig. S1*B*). As a comparison, $D2^{[6]\text{-}KO}$ has a stronger effect on *wing blade* than *D5*, from which the whole $spot^{196}$ segment was removed (Fig. 3*E* and *SI Appendix*, Fig. S1*B*). We were intrigued by these results, as the mutation $spot^{196\ [6]}$ had an effect on the *wing blade* activity only when the rest of the $spot^{196}$ was intact. This suggested that site $spot^{196\ [6]}$ could act indirectly on the *wing blade* activity by preventing, for example, the action of repressors regulating both activities. As the effect on the *wing blade* activity is not observed in $D2^{block4}$, which also randomizes site $spot^{196\ [6]}$, it is likely that sites for repressors acting on both activities are located within the 100 bp randomized in $D2^{block4}$. In our separate dissection of $spot^{196}$ (28), we reached a similar conclusion for the role of $spot^{196\ [6]}$. Even without knowing the molecular mechanism at work, our results suggest that $spot^{196\ [6]}$ could be the target site of a global, permissive activator of both activities in the context of segment $spot^{196}$. They demonstrate that *spot* and *wing blade* enhance transcription from shared, pleiotropic DNA sites. Because $spot^{196\ [6]}$ shows an effect on the *wing blade* activity not observed when mutating Dll TFBSs, we reasoned that the TFBSs for Dll and site $spot^{196\ [6]}$ may convey different information. We have previously shown that Dll primarily instructs the spatial pattern of the *spot* enhancer (16). The global spatial effect of site $spot^{196\ [6]}$, by contrast, suggests a permissive role such as the control of DNA accessibility in this regulatory region. To test this hypothesis, we compared the DNA accessibility of constructs *D2* and $D2^{[6]\text{-}KO}$ using ATAC-seq (Assay for Transposase-Accessible Chromatin with high-throughput sequencing) (29) in pupal wings at the onset of activation of the *wing blade* and the *spot* (Fig. 3*F* and *SI Appendix*, Fig. S3). While the genome-wide accessibility profiles of the two transgenic lines were similar, we observed a striking and specific disappearance of the accessibility peak overlapping the two activities in $D2^{[6]\text{-}KO}$ (Fig. 3*F*). These results suggest that the effect of site $spot^{196\ [6]}$ for the *wing blade* and the *spot* activities could stem from its effect on accessibility of a shared segment. We speculate that it could prime *yellow* regulatory activities in the wing by responding to a pioneer transcription factor (30–32), although its sequence does not resemble known motifs (33) of TFs expressed in pupal wings (16).

Discussion

Our results give a molecular snapshot of the evolutionary situation of two enhancers that today are entangled. In the 15 My since the emergence of the *spot* activity (7), the turnover of TFBSs in this region has likely been important, and there is no indication that the very inputs at work today are those involved in the original events of regulatory co-option. Our results, nevertheless, show that the sequences contributing the two activities largely overlap and that at least one site, $spot^{196\ [6]}$, influences both *wing blade* and *spot* activities in the wing. This is, therefore, a characterized case of enhancer pleiotropy. One molecular function associated with this site, as we have shown, is the regulation of chromatin accessibility. We envision the following sequence of events in this regulatory region during development. The regulatory region inaccessible to TFs at earlier developmental stages produces no activity in the wing (Fig. 4*A*). Site $spot^{196\ [6]}$ and probably several other sites, possibly through the interaction with a pioneer factor binding nucleosomal DNA, contribute to loosen local chromatin, resulting in enhancers poised for transcriptional activity (34). After the access to the enhancer sequences is granted, activator and repressor TFs bind to their cognate sites, and the respective enhancer activities start. This general developmental time line (silenced, poised, active enhancer) is supported by numerous recent publications (30, 35). In line with our results, the notion that enhancers control and fine tune their own accessibility is gaining rapid ground (30, 34). The pleiotropic effect of $spot^{196\ [6]}$ and its effect on chromatin opening suggest that, in contrast to the instructive role of Dll (this work and ref. 16) or Engrailed TFBS (6), it may be a site targeted by a pioneer transcription factor (32). As removing this site shows a pleiotropic effect only in the context of an intact $spot^{196}$, we suppose that its role on chromatin opening may give way to TFs preventing global repressors in the $spot^{196}$ acting pleiotropically on both activities.

The question of the evolutionary history of this pleiotropic site is still unclear, and to understand whether or not it is ancestral will require further work. The extensive interweaving that we observed between the *spot* and the *wing blade* enhancers, however, suggests that the evolution of the *spot* activity is tightly linked to the ancestral *wing blade* activity. TFBSs for spatial regulators of an enhancer emerge through random mutations. Mutations in an accessible region resulting in a TFBS for a spatial regulator, unlike mutations trapped in compacted chromatin, have the potential to contribute to a new spatial activity (Fig. 4*B*). In evolutionary terms, this means a shorter mutational path to gaining a regulatory activity (36) and therefore, an increased likelihood (37). Such shortcuts to the emergence of new regulatory activities may explain the apparent prevalence of enhancer co-option.

Materials and Methods

Fly Husbandry. Our *D. melanogaster* stocks were maintained on standard cornmeal medium at 25 °C with a 12:12 day:night light cycle.

Transgenesis. All reporter constructs were injected as in Arnoult et al. (16). We used ϕC31-mediated transgenesis (22) and integrated all constructs at the genomic *attP* site *VK00016* on chromosome 2 (38). The enhancer sequence of all transgenic stocks was genotyped before imaging.

Molecular Biology. Fragments of the *D* series were amplified by PCR from *D. biarmipes* [genome strain (39)] with Phusion polymerase (NEB) and cloned into our transformation vector pRedSA [a custom version of the transformation vector pRed H-Stinger (40) with a 284-bp *attB* site for ϕC31-mediated transgenesis (22) cloned at the AvrII site] digested with BamHI and EcoRI

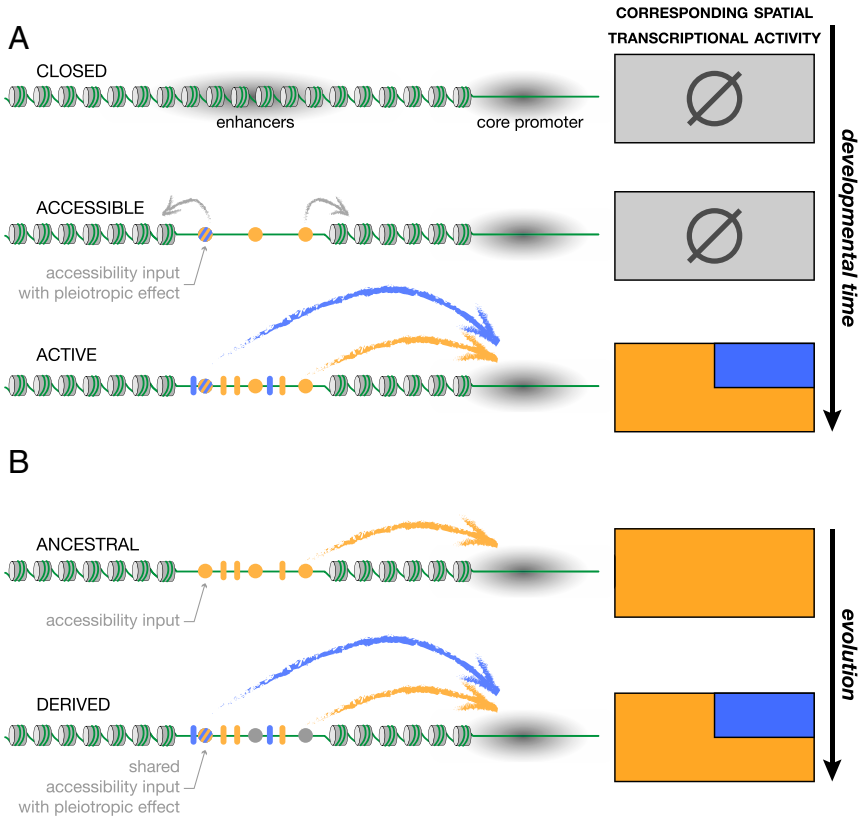

Fig. 4. Developmental enhancer pleiotropy and evolutionary enhancer co-option. (*A*) The developmental progression toward the activation of two interdependent enhancers inferred from our results. High nucleosome occupancy prevents access of transcription factors to the enhancers sequence (closed state; gray shading). Later during development, one or more specific sites in the regulatory sequence (pleiotropic accessibility input; colored circles) determine accessibility across a tissue [for instance, upon priming by a pioneer factor (34)], poising the region for transcriptional activity. Upon binding of specific regulators to their cognate TFBSs (blue and orange ovals), the enhancers become transcriptionally active, producing specific spatial activity patterns. (*B*) A speculative model of the emergence of a new enhancer by co-option. Some of the accessibility sites may be ancestral sites controlling the local accessibility of the regulatory region. During evolution, new TFBSs for spatial regulators, gained in the already accessible region, have the potential to promote the derived activity (blue), unlike TFBSs emerging from mutations in inaccessible regions. In this scenario, the derived activity co-opts an otherwise accessible ancestral activity, creating de facto pleiotropic regulatory information.

using In-Fusion HD Cloning Kits (Takara; catalog no. 121416). The fragment encompassing the four *Dll* sites in construct $D2^{Dll\text{-}KO}$ was synthetized in vitro by Integrated DNA Technologies. The mutations in construct $D2^{[06]\text{-}KO}$ were introduced by PCR through site-directed mutagenesis.

Constructs from the *E* series were produced similarly, but the fragments were made of two components stitched by PCR: a distal part amplified from *D. biarmipes* genome, as above, and a proximal part (dotted line in Fig. 1*A*) amplified from a unique randomized fragment (see below). Likewise, the randomized parts in constructs $D2^{block\ 4}$ and $D2^{block\ 5}$ were amplified from the same randomized fragment and stitched to *D. biarmipes* amplicons.

A randomized sequence was derived from the distal 4 kb of *D0* by randomizing 100-bp segments separately to preserve the local guanine–cytosine content and used for all constructs with randomized sequence. We generated it with an online DNA sequence randomizer (https://faculty.ucr.edu/~mmaduro/random.htm). The 4-kb fragment was synthetized in vitro by Integrated DNA Technologies and used as PCR template to amplify randomized spacers in *E*-series constructs as well as constructs $D2^{block\ 4}$, $D2^{block\ 5}$, and *RR*.

All primers are listed in *SI Appendix*, Table S2. The sequences of all fragments we tested are provided in *SI Appendix*, Table S3. Both *D* and *E* series keep the distance to the core promoter unaffected.

Imaging.

Sample preparation. All transgenic wings imaged in this study were heterozygous for the reporter construct. Males were selected minutes after emergence from pupa, a stage that we call "postemergence," when their wings are unfolded but still slightly curled. When flies were massively emerging from an amplified stock, we collected every 10 min and froze staged flies at −20 °C until we had reached a sufficient number of flies. Staged flies were processed after a maximum of 48 h at −20 °C. We dissected a single wing per male. Upon dissection, wings were immediately mounted onto a microscope slide coated with transparent glue (see below) and fixed for 30 min at room temperature in 4% paraformaldehyde diluted in phosphate buffer saline 1% Triton X-100 (PBST). Slides with mounted wings were then rinsed in PBST and kept in a PBST bath at 4 °C until the next day. Slides were then removed from PBST, and the wings were covered with Vectashield (Vector Laboratories). The samples were then covered with a coverslip. Preparations were stored for a maximum of 48 h at 4 °C until image acquisition.

The glue-coated slides were prepared immediately before wing mounting by dissolving adhesive tape (Tesa brand; tesafilm, reference 57912) in heptane (two rolls in 100 mL heptane) and spreading a thin layer of this solution onto a clean microscope slide. After the heptane had evaporated (under a fume hood), the slide was ready for wing mounting.

Microscopy. All wing images were acquired as 16-bit images on a Ti2-Eclipse Nikon microscope equipped with a 10× plan apochromatic lens (numerical aperture 0.45) and a 5.5-M scientific complementary metal oxide semiconductor camera (PCO). Each wing was imaged as a tile of several *z* stacks (*z* step = 4 μm) with 50% overlap between tiles. Each image comprises a fluorescent (TRITC-B filter cube) and a bright-field channel, the latter being used for later image alignment.

z Projection. Stitched three-dimensional stacks were projected to two-dimensional (2D) images for subsequent analysis. The local sharpness average of the bright-field channel was computed for each pixel position in each *z* slice, and an index of the slice with the maximum sharpness was recorded

and smoothed with a Gaussian kernel (sigma = 5 pixel). Both bright-field and fluorescent 2D images were reconstituted by taking the value of the sharpest slice for each pixel.

Image Quantification and Analysis.

Image alignment. Wing images were aligned using the veins as a reference. Fourteen landmarks placed on vein intersections and end points and 26 sliding landmarks equally spaced along the veins were placed on bright-field images using a semiautomatized pipeline. Landmark coordinates on the image were then used to warp bright-field and fluorescent images to match the landmarks of an arbitrarily chosen reference wing by the thin plate spline interpolation (41). All wings were then in the same coordinate system, defined by their venation.

Fluorescent signal description. A transgenic line with an empty reporter vector (*ø*) was used as a proxy to measure noise and tissue autofluorescence. The median raw fluorescent image was computed across all *ø* images and used to remove autofluorescence, subtracted from all raw images before the following steps. All variation of fluorescence below the median *ø* value was discarded. The DsRed (red fluorescent protein from *Discosoma*) reporter signal is mostly localized in the cell nuclei. We measured the local average fluorescent levels by smoothing fluorescence intensity through a Gaussian filter (sigma = 8 pixel) on the raw 2D fluorescent signal. The radius of the Gaussian filter, sigma, corresponded roughly to two times the distance between adjacent nuclei. To lower the memory requirement, images were then subsampled by a factor of two. We used the 89,735 pixels inside the wings as descriptors of the phenotype for all subsequence analyses.

Average phenotype images and differences, color maps, and normalization. Average reporter expression images were computed as the average smoothed fluorescence intensity at every pixel among all individuals in a given group (27 individuals per transgenic line on average). The difference between groups was computed as the difference between the average of the groups. Averages and difference images were represented using colors equally spaced in CIELAB perceptual color space (42). With these color maps, the perceived difference in colors corresponds to the actual difference in signal. Color maps were spread between the minimal and maximal signals across all averages for average phenotypes and between minus and plus the absolute value of all difference for the phenotype differences.

PCA. PCA was used to remove correlation between pixel intensities, to concentrate the variance on few variables, and therefore, to describe the variation in intensity and pattern of reporter gene expression in a comprehensive and unbiased way with few dimensions. PCA was calculated on the matrix of dimensions (n_individual $\times$ n pixels on the wing). The average phenotype of a construct was described as the average score in the PCA space among all wings of the construct, taking all components into account. Of note, in our calculations, working in the PCA space is equivalent to working directly in the image space. The variance of multidimensional phenotypes in PCA space was measured as the trace of the covariance matrix within each construct. SD was calculated as the square root of this variance.

Overall regulatory information loss. The overall amount of regulatory information lost or modified in successive fragments for each reporter construct series was approximated to the phenotypic distance to the respective largest fragment (*D0* for the *D* series, *D2* for the *E* series) in PCA space divided by the phenotypic distance between the largest construct of the series and the empty construct (*ø*) for normalization purpose. Consequently, while this phenotypic distance is zero for the largest construct, it increases as regulatory information is removed from the enhancer sequence as a result of truncation or randomization. The overall regulatory information loss reaches one when no regulatory information is left (i.e., when a construct has an average phenotype similar to that of the empty construct [*ø*]). A sigmoid curve of equation $\frac{Q}{1+e^{-\alpha*(t-t_{half})}}$, where t is the position along the enhancer sequence, was fitted to the measurements. The amount of regulatory information for each activity was calculated similarly but using *wing blade* and *spot* enhancer-independent measurement (see below) instead of the phenotypic distance described above.

Density of regulatory information per base. The amount of regulatory information brought by a segment of DNA was calculated as the absolute value of the difference between two consecutive fragments, of either the phenotypic distance to the full enhancer for the overall density or the *wing blade* and *spot* enhancer-independent measurements (see below) for the activity specific densities, divided by the differential fragment length. It represents the average amount of information (in terms of fluorescence intensity) per base pair, assuming that it is spread evenly across the modified sequence. To represent regulatory information, be it activating or repressing information, we used the absolute value of the change in the measure of activity, resulting in a similar representation of repression and activation.

Wing blade and spot enhancer-independent measurements. To measure independently the signal brought by the two enhancers, all individuals were projected from the PCA space onto a new two-vector basis, defined by the direction between *ø* and *D5* and the direction between *RR* and *E2*, both normalized to unit length. The coordinates in this two-vector basis represent directly reconstructed values for each activity as two independent measurements. These directions were chosen following the two independent directions of variations observed in the PCA space. Because *D5* and *E2* share 546 common nonmodified nucleotides, this is a conservative estimate of the independent effects in the context of measuring overlapping effect. The difference of expression of either activity between two groups was measured as the difference between the group average of the *wing blade* activity or *spot* activity coordinates described above.

Wing blade and spot regulatory information loss and density. The amount of regulatory information estimated specifically for each activity was calculated similarly to the overall regulatory information loss but using *wing blade* and *spot* enhancer-independent measurements (see above) instead of the phenotypic distance. The density of regulatory information specifically for the two activities was computed the same way as the overall regulatory information.

ATAC-Seq.

Buffers. Buffers for the purification of nuclei from pupal wings were prepared according to the omni–ATAC-seq protocol (43) with some modifications: 1× nuclei permeabilize buffer (NPB) buffer: 15 mM Tris·HCl, pH 7.5, 3 mM $MgCl_2$, 1× protease inhibitor mixture (Roche; cOmplete catalog no. 04693132001), ultrapure water (Invitrogen); 1× lysis buffer: NPB, 1% (vol/vol) Nonidet P-40 (Sigma), 1% (vol/vol) TWEEN 20 (Sigma), 0.1% (vol/vol) Digitonin (Promega), 1 mM dithiothreitol; and 1× wash buffer: NPB, 2% (vol/vol) Nonidet P-40, 10 mM NaCl.

Nuclei preparation. Male white pupa (0 to 1 h after puparium formation) were left to develop for 66 h at 25 °C. Twenty-four pupal wings were then dissected, rinsed twice in cold phosphate-buffered saline, and transferred into 100 μL cold 1× lysis buffer. The wings were cut coarsely into three to four pieces, transferred into a 2-mL Dounce homogenizer (Kimble), and further disrupted by 12 strokes using pestle A. The homogenate was let to rest on ice for 5 min and then further processed with 20 strokes using pestle B. After an additional 10 min of incubation on ice, 900 μL 1× wash buffer was added. A 20-mL syringe and a 20 1/2-gauge needle (Becton Dickinson) were employed to separate cells from the wing cuticle. The mixture was then filtered with a 40-μM strainer (Corning) and centrifuged at 4 °C at 1,000 × *g* for 10 min.

Tagmentation. Pelleted nuclei were gently resuspended in 45 μL ultrapure water and counted using a hemocytometer; 50,000 nuclei were then centrifuged at 4 °C at 1,000 × *g* for 10 min and resuspended in 8 μL 2× Tagment DNA (TD) buffer (Illumina; catalog no. 15027866). The tagmentation reaction followed the previous ATAC-seq protocol (29) with minor modifications: 10 μL 2× TD buffer with nuclei, 2 μL TD Enzyme (Illumina; catalog no. 15027865), 8 μL ultrapure water. The reaction was terminated by the addition of 5× volume PB buffer from the Qiagen MinElute kit, and the library was then purified following the kit's instruction. ATAC-seq libraries were amplified by NEBNext High-Fidelity 2× PCR Master Mix (NEB; catalog no. M0541S) for 9 to 11 PCR cycles and purified by Agencourt AMPure XP beads (Beckman Coulter) with double size selection (0.5× and 2.0×). Bioanalyzer with HS-DNA chip (Agilent) was used to determine the library quality and the final concentration for sequencing.

Sequencing and data processing. The sequencing was carried on an Illumina HiSeq1500 at LAFUGA (Laboratory for Functional Genome Analysis), Gene Center, Ludwig-Maximilians-Universität München, with pair-end settings. The reads for each library were around 50 to 70 million. The sequenced libraries were then demultiplexed, trimmed, and aligned to the reference genome UCSC (University of California, Santa Cruz) dm6 using Bowtie2 (44, 45) with following settings: –X 2000;–fr;–very-sensitive. The aligned reads were then filtered by Picard (46) with the following steps: clean sam, FixMate information, MarkDuplicate. The PCR duplicates were subsequently removed by SAMtools (47). Deeptools (48) was used to obtain the correlation among replicates. Peak calling was performed on three replicates together using MACS2 (49) with the following settings: –keep-dup all; –q 0.01;–nomodel;–shift −100;–extsize 200; −B –SPMR;–call-summits. The differentiated peak analysis was done with diffBind (50, 51) using DESeq2 (52) settings. Three replicates were used for each line. All counts were normalized with the setting bFullLinrarySize = TRUE. All raw and processed ATAC sequencing data have been submitted to the National Center for Biotechnology Information Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under the following accession numbers: pupal wing, D2_66 hAPF_rep1 (GSM4222134); pupal wing, D2_66hAPF_rep2 (GSM4222135); pupal

wing, D2_66hAPF_rep3 (GSM4222136); pupal wing, D206KO_66hAPF_rep1 (GSM4222137); pupal wing, D206KO_66hAPF_rep2 (GSM4222138); and pupal wing, D206KO_66hAPF_rep3 (GSM4222139).

Data Availability. ATAC-seq data have been deposited in GEO (accession nos. GSM4222134–GSM4222139).

ACKNOWLEDGMENTS. We thank Benjamin Prud'homme, Ilona Grunwald Kadow, Miltos Tsiantis, and Marta Bożek for insightful comments on the manuscript. We also thank S. Krebs and H. Blum (LAFUGA at Gene Center, Ludwig-Maximilians-Universität München) for support with sequencing. This work was supported by funds from the Ludwig Maximilians Universität München, The Graduate School of Quantitative Biosciences Munich, Human Frontiers Science Program Grant RGP0021/2018, and Deutsche Forschungsgemeinschaft Grants INST 86/1783-1 LAGG (to N.G.) and GO 2495/5-1 (to N.G.). Y.X. was supported by China Scholarship Council Fellowship 201506990003. M.M. was the recipient of a fellowship from the German Academic Exchange Service.

1. S. B. Carroll, Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell* **134**, 25–36 (2008).
2. M. Rebeiz, M. Tsiantis, Enhancer evolution and the origins of morphological novelty. *Curr. Opin. Genet. Dev.* **45**, 115–123 (2017).
3. W. J. Glassford *et al*., Co-option of an ancestral Hox-regulated network underlies a recently evolved morphological novelty. *Dev. Cell* **34**, 520–531 (2015).
4. J. Banerji, L. Olson, W. Schaffner, A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* **33**, 729–740 (1983).
5. G. Sabaris, I. Laiker, E. Preger-Ben Noon, N. Frankel, Actors with multiple roles: Pleiotropic enhancers and the paradigm of enhancer modularity. *Trends Genet.* **35**, 423–433 (2019).
6. N. Gompel, B. Prud'homme, P. J. Wittkopp, V. A. Kassner, S. B. Carroll, Chance caught on the wing: Cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* **433**, 481–487 (2005).
7. B. Prud'homme *et al*., Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature* **440**, 1050–1053 (2006).
8. A. Monteiro, O. Podlaha, Wings, horns, and butterfly eyespots: How do complex traits evolve? *PLoS Biol.* **7**, e37 (2009).
9. B. D. Pfeiffer *et al*., Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9715–9720 (2008).
10. E. Z. Kvon *et al*., Genome-scale functional characterization of *Drosophila* developmental enhancers in vivo. *Nature* **512**, 91–95 (2014).
11. A. Visel, S. Minovitsky, I. Dubchak, L. A. Pennacchio, VISTA Enhancer Browser–A database of tissue-specific human enhancers. *Nucleic Acids Res.* **35**, D88–D92 (2007).
12. B. P. Berman *et al*., Computational identification of developmental enhancers: Conservation and function of transcription factor binding-site clusters in *Drosophila melanogaster* and *Drosophila pseudoobscura*. *Genome Biol.* **5**, R61 (2004).
13. J. Crocker, D. L. Stern, Functional regulatory evolution outside of the minimal *even-skipped* stripe 2 enhancer. *Development* **144**, 3095–3101 (2017).
14. N. Frankel, Multiple layers of complexity in cis-regulatory regions of developmental genes. *Dev. Dyn.* **241**, 1857–1866 (2012).
15. M. Marinić, T. Aktas, S. Ruf, F. Spitz, An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. *Dev. Cell* **24**, 530–542 (2013).
16. L. Arnoult *et al*., Emergence and diversification of fly pigmentation through evolution of a gene regulatory module. *Science* **339**, 1423–1426 (2013).
17. P. K. Geyer, V. G. Corces, Separate regulatory elements are responsible for the complex pattern of tissue-specific and developmental transcription of the yellow locus in *Drosophila melanogaster*. *Genes Dev.* **1**, 996–1004 (1987).
18. M. F. Walter *et al*., Temporal and spatial expression of the *yellow* gene in correlation with cuticle formation and dopa decarboxylase activity in *Drosophila* development. *Dev. Biol.* **147**, 32–45 (1991).
19. P. J. Wittkopp, J. R. True, S. B. Carroll, Reciprocal functions of the *Drosophila* yellow and ebony proteins in the development and evolution of pigment patterns. *Development* **129**, 1849–1858 (2002).
20. P. J. Wittkopp, K. Vaccaro, S. B. Carroll, Evolution of *yellow* gene regulation and pigmentation in *Drosophila*. *Curr. Biol.* **12**, 1547–1556 (2002).
21. G. Kalay, J. Lachowiec, U. Rosas, M. R. Dome, P. Wittkopp, Redundant and cryptic enhancer activities of the *Drosophila yellow* gene. *Genetics* **212**, 343–360 (2019).
22. A. C. Groth, M. Fish, R. Nusse, M. P. Calos, Construction of transgenic *Drosophila* by using the site-specific integrase from phage *phiC31*. *Genetics* **166**, 1775–1782 (2004).
23. E. Preger-Ben Noon *et al*., Comprehensive analysis of a cis-regulatory region reveals pleiotropy in enhancer function. *Cell Rep.* **22**, 3021–3031 (2018).
24. R. Barrio, J. F. de Celis, S. Bolshakov, F. C. Kafatos, Identification of regulatory regions driving the expression of the *Drosophila spalt* complex at different developmental stages. *Dev. Biol.* **215**, 33–47 (1999).
25. R. B. Emmons, D. Duncan, I. Duncan, Regulation of the *Drosophila* distal antennal determinant spineless. *Dev. Biol.* **302**, 412–426 (2007).
26. J. T. Wagner-Bernholz, C. Wilson, G. Gibson, R. Schuh, W. J. Gehring, Identification of target genes of the homeotic gene *Antennapedia* by enhancer detection. *Genes Dev.* **5**, 2467–2480 (1991).
27. O. Nagy *et al*., Correlated evolution of two copulatory organs via a single cis-regulatory nucleotide change. *Curr. Biol.* **28**, 3450–3457.e13 (2018).
28. Y. Le Poul *et al*., Deciphering the regulatory logic of a *Drosophila* enhancer through systematic sequence mutagenesis and quantitative image analysis. https://doi.org/10.1101/2020.06.24.169748(25 June 2020).
29. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, W. J. Greenleaf, Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
30. J. Jacobs *et al*., The transcription factor *Grainy head* primes epithelial enhancers for spatiotemporal activation by displacing nucleosomes. *Nat. Genet.* **50**, 1011–1020 (2018).
31. Y. Sun *et al*., *Zelda* overcomes the high intrinsic nucleosome barrier at enhancers during *Drosophila* zygotic genome activation. *Genome Res.* **25**, 1703–1714 (2015).
32. K. S. Zaret, S. E. Mango, Pioneer transcription factors, chromatin dynamics, and cell fate control. *Curr. Opin. Genet. Dev.* **37**, 76–81 (2016).
33. T. L. Bailey *et al*., MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208 (2009).
34. M. Bozek, N. Gompel, Developmental transcriptional enhancers: A subtle interplay between accessibility and activity—considering quantitative accessibility changes between different regulatory states of an enhancer deconvolutes the complex relationship between accessibility and activity. *BioEssays* **42**, e1900188 (2020).
35. J. Crocker, A. Tsai, D. L. Stern, A fully synthetic transcriptional platform for a multicellular eukaryote. *Cell Rep.* **18**, 287–296 (2017).
36. N. Gompel, B. Prud'homme, The causes of repeated genetic evolution. *Dev. Biol.* **332**, 36–47 (2009).
37. I. Maeso, J. J. Tena, Favorable genomic environments for cis-regulatory evolution: A novel theoretical framework. *Semin. Cell Dev. Biol.* **57**, 2–10 (2016).
38. K. J. Venken, Y. He, R. A. Hoskins, H. J. Bellen, P[acman]: A BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science* **314**, 1747–1751 (2006).
39. Z. X. Chen *et al*., Comparative validation of the *D. melanogaster* modENCODE transcriptome annotation. *Genome Res.* **24**, 1209–1223 (2014).
40. S. Barolo, B. Castro, J. W. Posakony, New *Drosophila* transgenic reporters: Insulated P-element vectors expressing fast-maturing RFP. *Biotechniques* **36**, 436–440, 442 (2004).
41. M. F. Hutchinson, Interpolating mean rainfall using thin plate smoothing splines. *Int. J. Geogr. Inf. Syst.* **9**, 385–403 (1995).
42. CIE, *Colorimetry*, (CIE Central Bureau, Vienna, Austria, ed. 4, 2018).
43. M. R. Corces *et al*., An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).
44. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
45. B. Langmead, C. Wilks, V. Antonescu, R. Charles, Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* **35**, 421–432 (2019).
46. Broad Institute, Picard (2019). https://broadinstitute.github.io/picard/. Accessed 3 September 2019.
47. H. Li *et al*.; 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
48. F. Ramirez, F. Dundar, S. Diehl, B. A. Gruning, T. Manke, DeepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
49. Y. Zhang *et al*., Model-based analysis of ChIP-seq (MACS). *Genome Biol.* **9**, R137 (2008).
50. R. Stark, G. Brown, DiffBind: Differential binding analysis of ChIP-Seq peak data. http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf. Accessed 6 September 2019.
51. C. S. Ross-Innes *et al*., Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
52. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).