© 2018. Published by The Company of Biologists Ltd | Development (2018) 145, dev160663. doi:10.1242/dev.160663

TECHNIQUES AND RESOURCES

RESEARCH ARTICLE

Enhancer identification and activity evaluation in the red flour beetle, *Tribolium castaneum*

Yi-Ting Lai[1], Kevin D. Deem[1], Ferran Borràs-Castells[1], Nagraj Sambrani[1], Heike Rudolf[2], Kushal Suryamohan[3], Ezzat El-Sherif[2], Marc S. Halfon[3], Daniel J. McKay[4] and Yoshinori Tomoyasu[1,*]

ABSTRACT

Evolution of *cis*-regulatory elements (such as enhancers) plays an important role in the production of diverse morphology. However, a mechanistic understanding is often limited by the absence of methods for studying enhancers in species other than established model systems. Here, we sought to establish methods to identify and test enhancer activity in the red flour beetle, *Tribolium castaneum*. To identify possible enhancer regions, we first obtained genome-wide chromatin profiles from various tissues and stages of *Tribolium* using FAIRE (formaldehyde-assisted isolation of regulatory elements)-sequencing. Comparison of these profiles revealed a distinct set of open chromatin regions in each tissue and at each stage. In addition, comparison of the FAIRE data with sets of computationally predicted (i.e. supervised *cis*-regulatory module-predicted) enhancers revealed a very high overlap between the two datasets. Second, using *nubbin* in the wing and *hunchback* in the embryo as case studies, we established the first universal reporter assay system that works in various contexts in *Tribolium*, and in a cross-species context. Together, these advances will facilitate investigation of *cis*-evolution and morphological diversity in *Tribolium* and other insects.

KEY WORDS: Insects, Reporter assay, Chromatin profiling, FAIRE-seq

INTRODUCTION

Insects display some of the greatest diversity of morphology found among eukaryotic taxa, offering a variety of opportunities to investigate molecular and developmental mechanisms underlying morphological evolution. Decades of studies in evolutionary developmental biology (evo-devo) have revealed that changes in gene regulatory networks (GRNs) have been a major driving force in the production of the diverse morphology seen in insects, as well as in other taxa (Carroll, 2008; Carroll et al., 2005). In general, a GRN can be divided into two components: *trans* and *cis*. *trans* components are transcription factors (TFs) and their upstream regulators (such as signal transduction pathways) that provide instructive cues for patterning and differentiation to the tissues where they are expressed. In contrast, *cis* components are non-coding DNA elements (i.e. *cis*-regulatory elements, CREs) that gather and process the upstream *trans* information, and determine the spatial and temporal expression of the genes downstream in the genetic pathway. Changes in both *cis* and *trans* components have been implicated in morphological evolution (Carroll, 2008; Carroll et al., 2005; Halfon, 2017).

By using unparalleled genetic tools, biologists have analyzed both *cis* and *trans* components in great detail in the fruit fly *Drosophila melanogaster*. The accumulated knowledge obtained from *Drosophila* studies can be used as a reference (i.e. the *Drosophila* paradigm) when studying other insects and identifying the changes in GRNs responsible for morphological evolution. RNA interference (RNAi)-based gene knockdown techniques have allowed for an investigation of the *trans* components involved in development and their evolutionary conservation/diversification in various insects (Bellés, 2010). However, the lack of a reliable method for identifying *cis* components in non-*Drosophila* insects has made it difficult to study the evolution of *cis* regulation beyond the *Drosophila* species, even though this is important in order to gain a comprehensive view of the GRN changes that contribute to morphological evolution.

The major difficulty in identifying CREs, such as enhancers, stems from the labile nature of *cis* components. The genes that code for *trans* factors that are important for development are usually evolutionarily well-conserved; thus, it is relatively easy to identify these *trans* components in various insects based on their homologies (Carroll et al., 2005). In contrast, *cis* components appear to be more evolutionarily flexible in a variety of aspects. First, the order of TF-binding sites can vary widely within an enhancer region, and the location of enhancers relative to the target gene also appears to be variable. Second, there can be redundancy among multiple enhancers responsible for the same gene (i.e. shadow enhancers) (Hong et al., 2008), allowing the enhancers to evolve more rapidly. In addition, the function of each enhancer tends to exhibit low levels of pleiotropy (Carroll, 2008), resulting in the accumulation of more evolutionary changes in enhancers. These characteristics, along with the faster rate of genome evolution in insects compared with vertebrates (Zdobnov and Bork, 2007), make the identification of insect enhancers a challenging task.

Traditionally, enhancers have been identified using reporter assays, in which the transcriptional activation capability of genomic regions near the gene of interest is assessed via a reporter gene construct (see Suryamohan and Halfon, 2015 for a review of traditional and new methods for identifying enhancers). This is a time-consuming and arduous approach, as an enhancer can sometimes reside hundreds of thousands of base pairs away from the gene that it regulates (Shlyueva et al., 2014). Identification of evolutionarily conserved genomic regions outside of coding regions among several closely related species, such as the *Drosophila* species group, has been helpful in narrowing down regions to survey

[1]Department of Biology, Miami University, Oxford, OH 45056, USA. [2]Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen 91058, Germany. [3]Department of Biochemistry, State University of New York at Buffalo, Buffalo, NY 14214, USA. [4]Department of Biology, Department of Genetics, Integrative Program for Biological and Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.

*Author for correspondence (tomoyay@miamioh.edu)

F.B., 0000-0002-7585-8373; K.S., 0000-0002-6614-7182; Y.T., 0000-0001-9824-3454

Received 24 October 2017; Accepted 9 March 2018

Development (2018) 145, dev160663. doi:10.1242/dev.160663

for enhancer activity (phylogenetic footprinting) (Frazer et al., 2004; Mayor et al., 2000; Sosinsky et al., 2007; Stark et al., 2007). Enhancer predictions based on the TF-binding motifs have also been helpful in identifying potential enhancer regions, although the prediction appears to work more efficiently for embryonic enhancers because of the clustering tendency of TF-binding motifs within an enhancer that is active during the syncytial blastoderm stage, while enhancers for other stages might be more difficult to identify through current prediction methods (Li et al., 2007). Combinations of these approaches have allowed for successful identification of enhancers that are active in various developmental contexts in *Drosophila*. More recently, the reporter assay approach has been applied in a genome-wide fashion in *Drosophila* (as in the FlyLight project), identifying over 10,000 genomic regions capable of activating transcription (Jenett et al., 2012; Jory et al., 2012; Kvon et al., 2014; Pfeiffer et al., 2008). Unfortunately, many of these approaches are technically demanding and resource intensive, and thus are currently only possible in *Drosophila* (but also see Kazemian et al., 2014 for the successful application of enhancer prediction in non-*Drosophila* insects).

In parallel to the methods described above, several genomic approaches have been developed for the identification of possible enhancer regions in the *Drosophila* genome (reviewed by Shlyueva et al., 2014; Suryamohan and Halfon, 2015). One such method is formaldehyde-assisted isolation of regulatory elements (FAIRE) in combination with next-generation sequencing (FAIRE-seq), which identifies open chromatin regions across the genome (Simon et al., 2012). FAIRE-seq has been used in *Drosophila*, showing that open chromatin regions often correspond to enhancers and other CREs (McKay and Lieb, 2013; Pearson et al., 2016; Uyehara et al., 2017). In addition, FAIRE-seq requires less input material and does not rely on antibodies, thus making it less technically demanding than techniques such as chromatin immunoprecipitation-sequencing (ChIP-seq). These features make FAIRE a promising technique to apply to non-*Drosophila* insects. However, it is important to note that potential enhancers identified by FAIRE (or other genomic approaches) still require functional validations, such as with a reporter assay. This presents another significant hurdle when studying enhancers in non-*Drosophila* insects, as the availability of a modern genetic toolkit (such as a versatile reporter construct) is very limited for non-*Drosophila* species.

In this study, we aimed to establish an enhancer identification and evaluation method in the red flour beetle, *Tribolium castaneum*. A variety of genetic and genomic tools are available for *Tribolium*, making this insect a powerful model system for comparative developmental biology and evo-devo studies (Denell, 2008; Schmitt-Engel et al., 2015; Tribolium Genome Sequencing, 2008; Wang et al., 2007). The robust systemic RNAi response of *Tribolium* has allowed researchers to study *trans* components in detail (Brown et al., 1999; Bucher et al., 2002; Tomoyasu and Denell, 2004) and to identify changes in GRNs that are responsible for morphological evolution from the *trans* point of view (see Peel, 2008 for a review of the findings related to the evolution of insect segmentation; Tomoyasu et al., 2009 for insect wing evolution; and Angelini et al., 2012 and Smith et al., 2014 for the evolution of insect appendages). However, studies of *cis* components in *Tribolium* are currently limited because of the lack of reliable enhancer identification methods.

For the initial identification of possible enhancer regions, we first implemented FAIRE-seq and obtained genome-wide open chromatin profiles from various tissues and at different developmental stages of *Tribolium*. The comparison of chromatin profiles between different tissues and stages revealed a distinct set of open chromatin regions in each tissue and stage. In addition, the open chromatin regions detected by FAIRE matched very well with those identified in previous *Tribolium* enhancer studies (Cande et al., 2009; Eckert et al., 2004; Kazemian et al., 2014; Wolff et al., 1998), as well as with supervised *cis*-regulatory module (SCRMshaw)-predicted enhancers (Kantorovitz et al., 2009; Kazemian et al., 2011, 2014). Second, we chose the wing expression of *nubbin* (*nub*) (Fig. 1) as a case study, and established the first universal reporter assay system that works in *Tribolium* and also in a cross-species context. To our knowledge, the *T. castaneum-nub* (*Tc-nub*) wing enhancer identified here is the first post-embryonic enhancer that has been functionally evaluated in non-*Drosophila* insects. Furthermore, using *hunchback* (*hb*) as another example, we demonstrated that our reporter construct works in other developmental contexts in *Tribolium*. Together, these advances will facilitate investigation of enhancers in *Tribolium* and in other insects, which will provide a more comprehensive understanding of the molecular mechanisms underlying the production of the vast morphological diversity seen in insects.

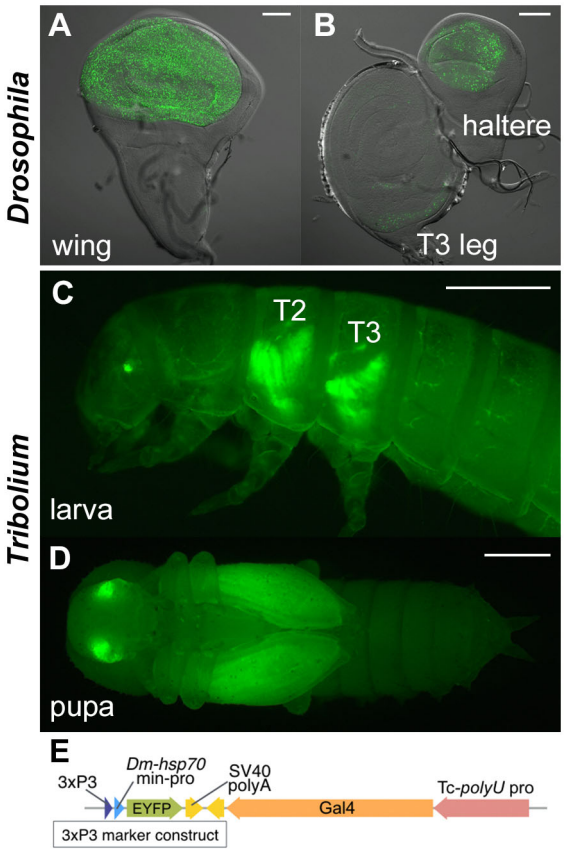

Fig. 1. *nub* enhancer trap expression in *Drosophila* and *Tribolium*. (A,B) The *nub* enhancer trap expression in the wing disc (A), and the haltere and T3 leg discs (B) in *Drosophila*. (C,D) Expression pattern of the *nub* enhancer trap line (*pu11*) at the larval (C) and pupal (D) stages in *Tribolium*. (E) The piggyBac construct inserted near the *nub* locus in the *pu11* beetles. The *Dm-hsp70* minimal promoter within the 3xP3 construct appears to have trapped the wing enhancer of *nub* in *pu11*, driving EYFP in the pattern identical to the endogenous *nub* wing expression. In contrast, the eye expression in *pu11* is due to the 3xP3 construct and is independent of *nub* enhancer activity. The *polyubiquitin* promoter-Gal4 construct appears to be non-functional in *pu11* (Y.T., unpublished observation). Scale bars: 0.5 mm.

Development (2018) 145, dev160663. doi:10.1242/dev.160663

RESULTS

FAIRE-seq revealed a spatially and temporally regulated chromatin profile in the *Tribolium* genome

To obtain chromatin profiles from diverse tissues and stages of *Tribolium*, we performed FAIRE-seq with the following six samples: three stages of embryos (0-24, 24-48 and 48-72 h), the second (T2) and third (T3) thoracic epidermal tissues of the last instar larvae that contain the forewing (elytron) and hindwing imaginal tissues, and the brain isolated from the last instar larvae. The sequence reads obtained from FAIRE-seq were then mapped to the *Tribolium* genome assembly (Tcas_3.0). Each sample displayed a unique set of open chromatin regions (referred to as 'peaks'; see Fig. 2A for an example), indicating that FAIRE-seq with *Tribolium* tissues was successful. The overall open chromatin characteristics were similar in *Tribolium* and *Drosophila*; however, we also noticed some features that were unique to the *Tribolium* chromatin profiles. We detected more than 40,000 open chromatin regions in the *Tribolium* genome across the samples (Table 1). To identify differences in open chromatin profiles between samples, we performed differential peak calling using DiffBind [false discovery rate (FDR)<0.05]. The number of differentially accessible peaks between pairs of samples varied widely. For example, there were over 26,000 differentially accessible peaks between 0-24 h embryos and T3 (Table 1, Fig. S1), reflecting the extensive differences in *cis*-regulatory control that likely exist between these two samples. By contrast, we found only four differentially accessible peaks between T2 and T3. The similarity in open chromatin profiles in T2 and T3 tissues is remarkable, given the dramatic differences in morphology between forewing and hindwing in *Tribolium*. However, similar findings were obtained in *Drosophila* (McKay and Lieb, 2013), suggesting that both species use shared sets of enhancers to shape their appendages. Intriguingly, although the level of nucleosome depletion in the FAIRE-isolated genomic regions is variable between stages and tissues, their positions appear to correlate highly with the guanine-cytosine (GC)-rich regions of the genome (Fig. S2A). Furthermore, these GC-rich and FAIRE-identified regions occur at regular intervals, producing a 'ruler-like' pattern of FAIRE peaks throughout the genome (Fig. S2B). This regular periodicity of the GC-rich and FAIRE-identified regions appears to be unique to the *Tribolium* lineage, as we did not detect a similar periodicity in other coleopteran genomes or in the genome of the lepidopteran *Bombyx mori* (see Fig. S2C for *Drosophila*; data not shown for other insects).

Comparison of the FAIRE data with previous enhancer studies in *Tribolium*

Several previous studies have investigated the activity of *Tribolium* enhancers. To our knowledge, the only study analyzing enhancer activity in the *Tribolium* native context is that of Eckert et al., 2004. This study identified enhancers that are important for the stripe expression of the *Tribolium hairy* gene. Some additional enhancers for *Tribolium* genes have also been identified, albeit in a cross-species context (i.e. *Drosophila*). These include enhancers for *hunchback* (Wolff et al., 1998), *single-minded*, *cactus* and *short gastrulation* (Cande et al., 2009), and *labial*, *Dichaete* and *wingless* (Kazemian et al., 2014). We analyzed the FAIRE profiles at these gene loci and found that our FAIRE peaks matched with many of the enhancer regions identified in these studies (Fig. S3).

More recently, Kazemian et al. applied their enhancer discovery approach, SCRMshaw, to the *Tribolium* genome and predicted 1214 genomic regions to be potential enhancers (Kantorovitz et al., 2009; Kazemian et al., 2011, 2014). Comparison of our FAIRE data with their SCRMshaw predictions revealed a striking degree of overlap between the two datasets: 78.8% (957/1214) of SCRMshaw predictions overlapped with at least one embryonic FAIRE peak, and 88.1% (1070/1214) of predictions overlapped with at least one larval FAIRE peak (Tables S1, S2; $P \approx 0$); overall, 1096 of the 1214 (90.3%) predicted *cis*-regulatory modules (CRMs) overlapped with at least one FAIRE peak. For certain sets of SCRMshaw predictions, the overlaps were even more extensive: e.g. 98% (97/99) of wing-specific predicted enhancers overlapped with a larval FAIRE peak (Table S1). Taken together, the high degree of overlap between the FAIRE peaks and previously identified enhancer regions, and between the FAIRE-peaks and SCRMshaw-predicted enhancers, verifies that FAIRE-seq is a powerful tool for identifying enhancers in *Tribolium*.

Identification of the *Tribolium nub* wing enhancer using a cross-species reporter assay

As mentioned in the Introduction, reporter assays are a time-consuming and laborious task, which makes them difficult to perform in non-*Drosophila* insects, including *Tribolium*. However, to fully exploit the benefit of the FAIRE profiling data, it is crucial to have a reliable method to evaluate the function of *Tribolium* enhancers. Previously, the activity of potential *Tribolium* enhancers has been successfully evaluated using reporter assays in *Drosophila* (Cande et al., 2009; Kazemian et al., 2014; Wolff et al., 1998; Zinzen et al., 2006). We reasoned that the enhancer of a gene that has a conserved expression pattern (both temporal and spatial) in *Drosophila* and *Tribolium* has the highest chance of being active, even in a cross-species context, and is thus ideal for a case study. The enhancer responsible for the wing expression of *nub* fits this criterion, as *nub* is expressed broadly in the tissues that give rise to the wings in both insects (Fig. 1) (Ng et al., 1995; Tomoyasu et al., 2009). In addition, an enhancer trap line for *nub* is available in *Tribolium* (*pu11*; Fig. 1C-E). We have previously determined that this enhancer trap line has a piggyBac construct inserted about 30 kb upstream of the *nub* transcription start site (Clark-Hachtel et al., 2013) (Fig. 2A), which can be used as a starting point to survey for the wing enhancer.

nub codes for an evolutionarily conserved TF that is important for the proliferation of wing cells (Ng et al., 1995). *Drosophila* has two *nub* paralogs (*nub* and *pdm2*), whereas *Tribolium* has one (*Tc-nub*). FAIRE analysis revealed a number of open chromatin regions located in and near the *Tc-nub* locus (Fig. 2A). Some of the open chromatin regions were shared across the six samples tested (such as the region corresponding to the promoter), but others were unique to specific tissues and stages. We tested the two open chromatin regions at or near the *pu11* insertion site (*Tc-nub3* and *Tc-nub2*) in *Drosophila* (Fig. 2A,B). In addition, we also tested another major open chromatin region located further upstream of the *pu11* insertion site (*Tc-nub1*). This region is open predominantly in the larval T2 and T3 epidermal tissues (containing the future wing tissues), but not in any of the embryonic samples, suggesting that this region contains enhancers that are specific to the post-embryonic stage (Fig. 2A,B).

The cross-species reporter assay showed that *Tc-nub2* and *Tc-nub3* do not have enhancer activity in the future wing-related tissues (wing and haltere imaginal discs) in *Drosophila* (Fig. 2C-F). *Tc-nub*3 showed activity in a small region near the hinge of the wing and haltere discs, but not in the region that gives rise to the wings (wing and haltere pouches) (Fig. 2C,D). *Tc-nub*2 drove reporter expression in the leg discs, but did not show any enhancer activity in the wing and haltere discs (Fig. 2E,F). In contrast, *Tc-nub1* showed significant enhancer activity in the pouch region of the wing disc

Development (2018) 145, dev160663. doi:10.1242/dev.160663

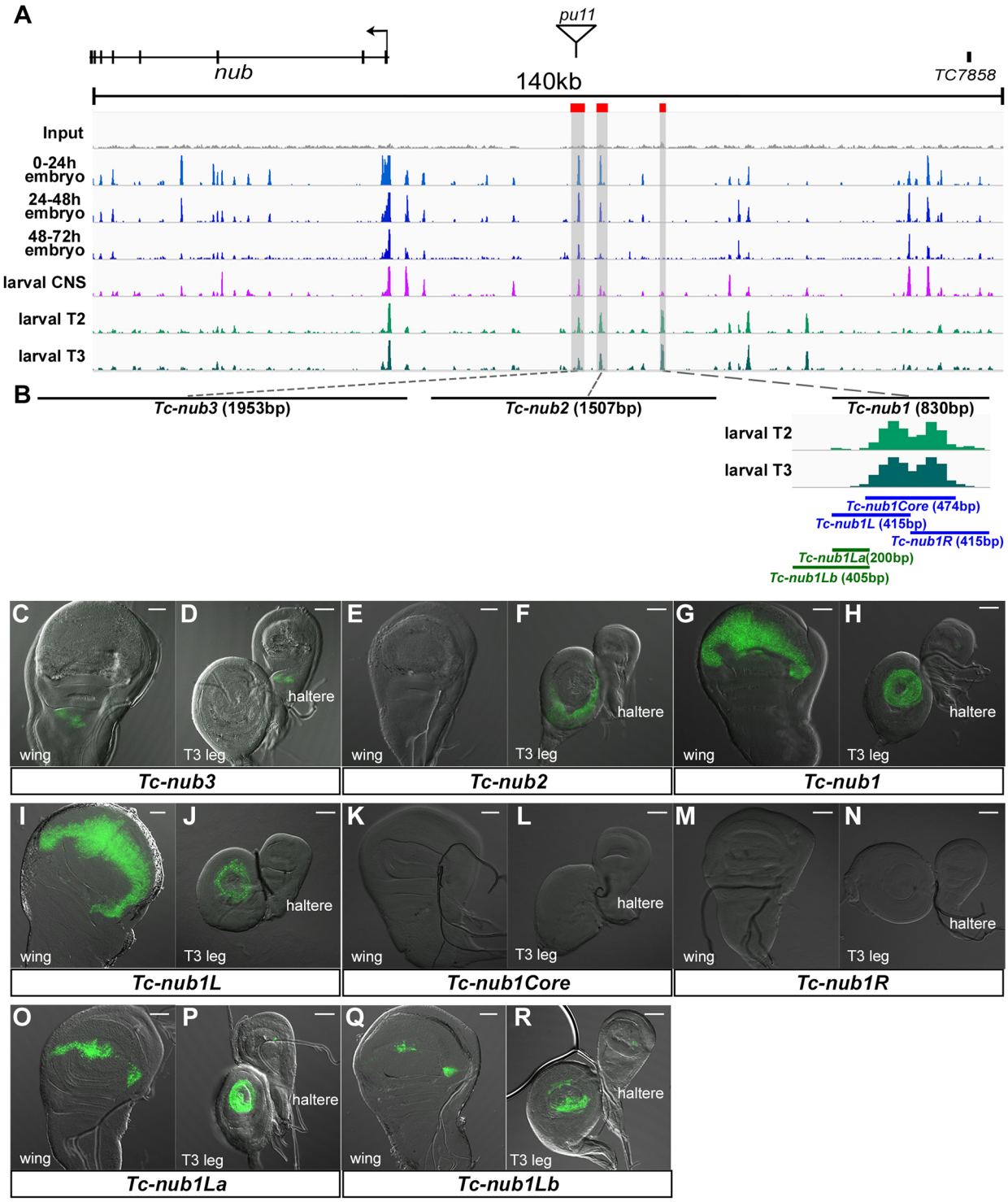

Fig. 2. Identification of the *Tribolium nub* wing enhancer using FAIRE and cross-species reporter assay. (A) FAIRE profiles at the *Tribolium nub* locus in six different tissues/stages. The *pu11* insertion site is indicated with a triangle. Three peaks near the *pu11* insertion site that were chosen for evaluating enhancer activity are marked with red boxes. (B) Summary of the regions that were tested by the reporter assay. The distance between *Tc-nub1*, *Tc-nub2* and *Tc-nub3* are not scaled. The magnified view of the FAIRE peak corresponding to *Tc-nub1L* is also presented. (C-R) Enhancer activity of each *Tribolium* genomic region tested in the *Drosophila* imaginal discs. Scale bars: 50 µm.

(Fig. 2G). *Tc-nub1* also drove reporter expression in the leg disc, but was not active in the haltere disc (Fig. 2H). Because *Tc-nub1* corresponds to the region that is uniquely open in the larval epidermis in *Tribolium*, the outcome of our cross-species reporter assay indicates that (1) the open chromatin profiling of various tissues and stages by FAIRE-seq in *Tribolium* can help predict tissue/stage-specific enhancers from the *Tribolium* genome, and (2) the cross-species reporter assay can be useful, at least for the enhancers responsible for the post-embryonic expression of *nub* in *Tribolium*.

Development (2018) 145, dev160663. doi:10.1242/dev.160663

Table 1. The number of differentially open peaks

	T3	T2	Brain	48-72 h	24-48 h
T2	1/**3**				
Brain	15427/**7262**	11634/**5258**			
48-72 h	6428/**8134**	5575/**7259**	1602/**9089**		
24-48 h	10380/**6729**	9572/**6031**	3162/**6689**	863/**40**	
0-24 h	17450/**8800**	14138/**6808**	9002/**8279**	7651/**1041**	2407/**586**

We next sought to minimize the *Tc-nub* wing enhancer by testing three shorter fragments within the *Tc-nub1* region (Fig. 2B). Interestingly, despite covering the main FAIRE peak region of *Tc-nub*1, *Tc-nub1Core* did not show any enhancer activity in the wing (Fig. 2K,L). Instead, *Tc-nub1L*, which corresponds to only a part of the major open chromatin region, drove reporter expression with a pattern and level almost identical to those driven by *Tc-nub1* (Fig. 2I,J). *Tc-nub1R* did not show any enhancer activity (Fig. 2M,N). These results suggest that the important elements for driving wing expression reside within the first 200 bp of *Tc-nub*1. We tested this idea by making a reporter construct using only the 200 bp region unique to *Tc-nub1L* (*Tc-nub1La*, Fig. 2B). This fragment drove reporter expression in the wing and leg discs, albeit with a more-restricted expression domain and/or a lower expression level compared with *Tc-nub1L* (Fig. 2O,P). We also tested a construct that contained the *Tc-nub1L* region along with an additional 200 bp sequence outside of *Tc-nub1* (*Tc-nub1Lb*, Fig. 2B), because the location of the functional *Tc-nub* wing enhancer may be slightly misaligned with respect to the open chromatin region predicted by FAIRE. However, *Tc-nub1Lb* showed even weaker enhancer activity (Fig. 2Q,R), suggesting that there might be a suppressor element next to the *Tc-nub1* region. The constructs we made also drove reporter expression outside the wing and leg imaginal discs. These results are summarized in Table S3.

Identification of the *Drosophila nub* wing enhancer using a combination of genomic resources, FAIRE profiling and the reporter assay approach in *Drosophila*

For comparison with the enhancer identified via a cross-species reporter assay described above, we sought to identify the *nub* wing enhancer that is native to the species used for the reporter assay (i.e. *Drosophila*). As mentioned earlier, there are two *nub* paralogs in *Drosophila* (*nub* and *pdm2*), both of which have similar expression in the wing pouch (Ng et al., 1995). We first took advantage of the FlyLight project and surveyed the *nub* and *pdm2* loci for a genomic region that has wing enhancer activity. Among the 33 constructs tested in FlyLight (Fig. 3A), one region (GMR11F02) has a record of enhancer activity in the wing and haltere pouches, along with additional expression in the leg disc (Fig. 3B,C). We then used the previously published FAIRE profile for *Drosophila* (McKay and Lieb, 2013), and identified three distinct regions within GMR11F02 that are open in the wing and haltere discs (Fig. 3A). We cloned these three regions [Fig. 3B; *D. melanogaster-nub* (*Dm-nub*)*1*, *Dm-nub2* and *Dm-nub3*] and tested their enhancer activity in *Drosophila*. Among the three regions, *Dm-nub2* displayed strong enhancer activity in the wing pouch region (Fig. 3G,H). *Dm-nub1* (Fig. 3E,F) and *Dm-nub3* (Fig. 3I,J) did not drive reporter expression in the wing and haltere discs. In addition, *Dm-nub3* was active in the leg disc, suggesting that the *Dm-nub* wing/haltere enhancer and leg enhancer are separable (Fig. 3J). To further minimize the *Dm-nub* wing enhancer, we tested three shorter fragments within *Dm-nub2* (*Dm-nub2a*, *Dm-nub2b* and *Dm-nub2c*; Fig. 3D). The wing-related expression is driven by *Dm-nub2a*, albeit at a weaker level (Fig. 3K,L). This suggests that, although *Dm-nub2a* contains sufficient components to drive wing expression, a broader genomic region is required for robust wing expression of *Dm-nub*. In contrast, *Dm-nub2b* and *Dm-nub2c* did not drive any expression (Fig. 3M-P). The expression patterns of these constructs in other tissues are summarized in Table S4. Taken together, the *Dm-nub2* region that we isolated (1.3 kb) is sufficient to drive a robust wing expression in *Drosophila*.

Establishing a reporter assay system and evaluating the *nub* wing enhancers in *Tribolium*

Although some *Tribolium* enhancers have been shown to be active in the cross-species context, these enhancers still need to be examined in their native species for functional validation. However, the lack of a reliable reporter construct has been a major obstacle in performing functional evaluation of enhancers in *Tribolium*. The Gateway system (Katzen, 2007) has been useful in quickly cloning genomic regions into a reporter construct and testing their enhancer activity in *Drosophila*. We sought to establish a Gateway-compatible reporter construct that is functional in *Tribolium*.

A key issue in establishing a reporter construct is the choice of promoters. Previous studies have raised concerns about using *Drosophila* promoters in *Tribolium* (Schinko et al., 2010). While establishing the Gal4/UAS system in *Tribolium*, Schinko et al. found that the core promoter isolated from a *Tribolium* endogenous gene, *Tc-hsp68*, worked more efficiently than the exogenous promoters that were tested (Schinko et al., 2010). We therefore made a Gateway-compatible piggyBac construct that contained the *Tc-hsp68* core promoter driving the *dsRed* gene [piggyBac Gateway Tc-hsp68 dsRed (piggyGHR), Fig. 4A]. In addition, we added the *gypsy* element, which is a *Drosophila* insulator, to either side of the the reporter assay construct to prevent position effects (Fig. 4A). We tested this piggyBac construct with the *Tribolium* and *Drosophila nub* wing enhancers (*Tc-nub1L* and *Dm-nub2*) in *Drosophila*. Both *Tc-nub1L* and *Dm-nub2* drove dsRed expression identical to the patterns obtained with the *Drosophila* reporter construct (compare Fig. 4B,C with Fig. 2I,J, and Fig. 4D,E with Fig. 3G,H), confirming that piggyGHR is functional. However, neither *Tc-nub1L* nor *Dm-nub2* in piggyGHR showed consistent enhancer activity in the wing tissues when transformed into *Tribolium* (Fig. 4F-M). None of the seven independent transgenic lines obtained for piggyGHR-*Tc-nub1L* had clear dsRed expression in the wing tissues (Fig. 4F-K). Instead, four lines had dsRed expression in non-wing tissues, with a distinct pattern in each line, likely because of trapping local enhancers (Fig. 4F-K). We obtained only two independent transgenic lines for piggyGHR-*Dm-nub*2, neither of which had dsRed expression in the wing tissue (Fig. 4L,M). These results indicate that our construct with the *Tc-hsp68* core promoter does not work well for reporter assays, at least in the wing-related tissues in *Tribolium*, although it does work in *Drosophila*. Alternatively, it is also possible that the *Drosophila gypsy* insulators that we added to the construct might not be functioning properly in *Tribolium*.

We next tested a synthetic promoter in *Tribolium*. Pfeiffer et al. modified the super core promoter 1 (SCP1) (Juven-Gershon et al., 2006) and constructed the *Drosophila* synthetic core promoter (DSCP), which was used for the FlyLight project as well as in other *Drosophila* reporter constructs, including pFUGG in this study (McKay and Lieb, 2013). We made a piggyBac construct with the DSCP driving mCherry [piggyBac Gateway universal promoter mCherry (piggyGUM); Fig. 5A]. We removed the *Drosophila gypsy* insulators from our construct to avoid possible cross-species issues. Similar to piggyGHR, piggyGUM with the *Drosophila* and

Development (2018) 145, dev160663. doi:10.1242/dev.160663

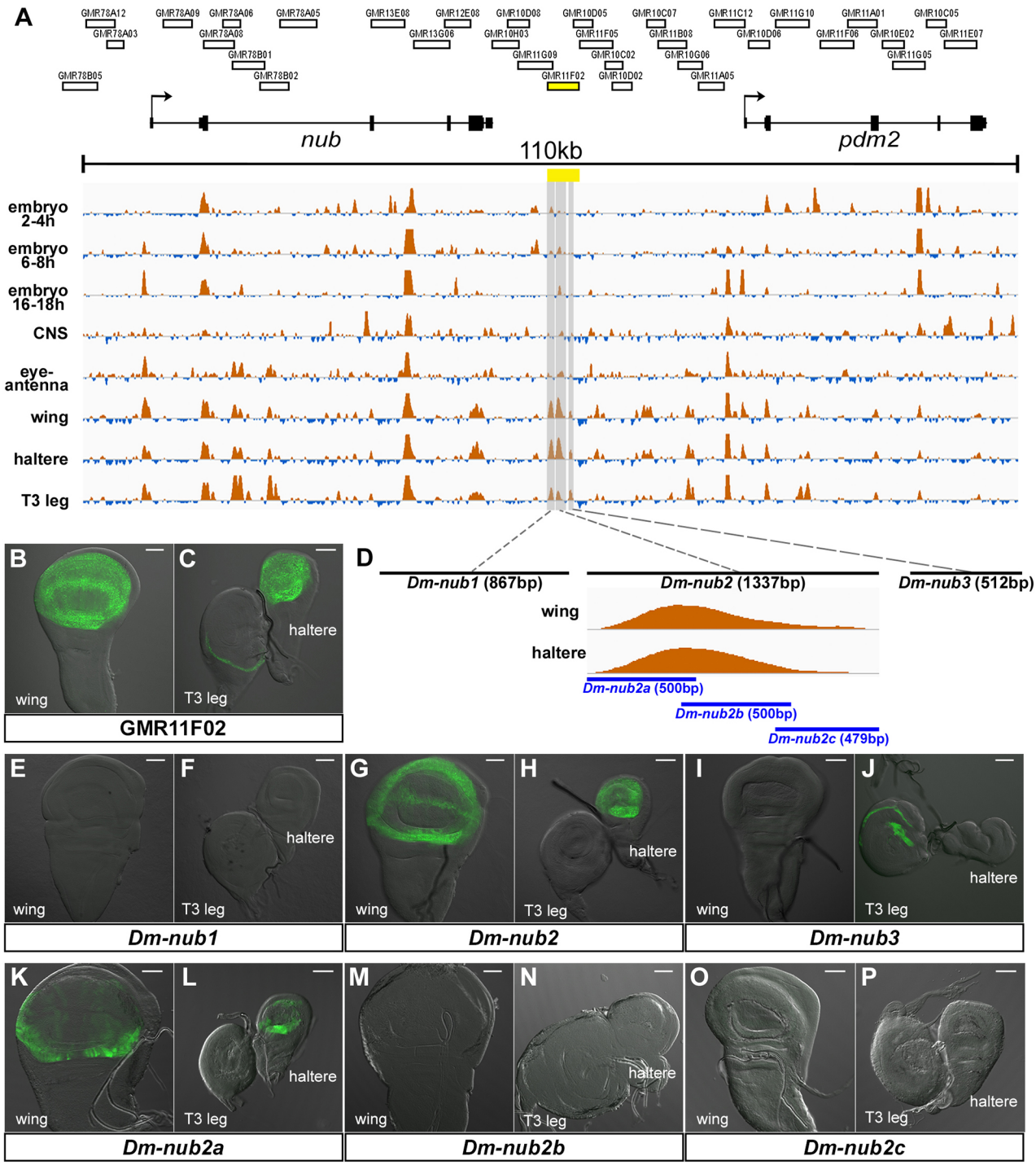

Fig. 3. Identification of the *Drosophila nub* wing enhancer. (A) FAIRE profiles from eight different tissues/stages at the *nub* and *pdm2* loci in *Drosophila*. The regions surveyed in the FlyLight project are also indicated. The region that shows wing enhancer activity is marked in yellow. (B,C) Expression driven by GMR11F02 in the *Drosophila* imaginal discs. (D) Summary of the regions within GMR11F02 that were tested by the reporter assay. The relative distance between *Dm-nub1, Dm-nub2* and *Dm-nub3* are not to scale. The magnified view of the *Dm-nub2* peak is also included. (E-P) Enhancer activity of each *Drosophila* genomic region tested in the *Drosophila* imaginal discs. Scale bars: 50 µm.

Tribolium nub wing enhancers drove reporter expression in the wing disc in *Drosophila* (Fig. 5B-E), confirming that piggyGUM is functional. In *Tribolium*, in contrast to the piggyGHR constructs, piggyGUM-*Tc-nub1L* successfully recaptured the expression pattern of the *nub* enhancer trap line (*pu11*) and drove reporter expression in the wing-related tissues (both in T2 and in T3) at both larval and pupal stages (Fig. 5F-I compared with Fig. 1C,D). piggyGUM-*Dm-nub2* also showed enhancer activity in the larval wing discs in *Tribolium* (Fig. 5J-L). The expression driven by *Dm-nub2* in *Tribolium* was mostly in the wing hinge and the margin regions, similar to the pattern observed for this enhancer in the *Drosophila* imaginal discs (Figs 3G,H and 5D,E). These results indicate that: (1) our Gateway-compatible DSCP piggyBac construct (piggyGUM) can be used for reporter assays both in *Tribolium* and in *Drosophila*;

Development (2018) 145, dev160663. doi:10.1242/dev.160663

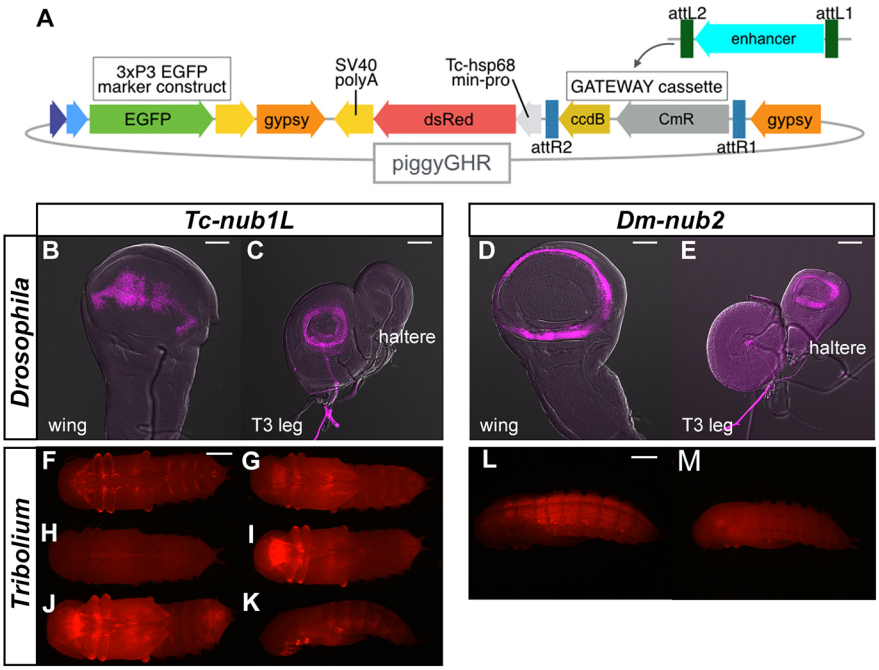

Fig. 4. Reporter assay with the *Tc-hsp68* promoter construct in *Drosophila* and *Tribolium*. (A) The piggyGHR construct. (B-E) Enhancer activity of *Tc-Nub1L* (B,C) and *Dm-nub2* (D,E) tested with the piggyGHR construct in *Drosophila*. (F-M) Enhancer activity of *Tc-nub1L* (F-K) and *Dm-nub2* (L,M) tested with piggyGHR at the pupal stage in *Tribolium*. Six independent lines for *Tc-nub1L* (F-K) and two for *Dm-nub2* (L,M) are shown. Scale bars: 50 µm (B-E); 0.5 mm (F-M); scale bars in F and L apply to F-K and L,M, respectively.

(2) the *Tribolium nub* wing enhancer that was identified through a cross-species reporter assay (*Tc-nub1L*) is indeed functional as a wing enhancer in *Tribolium*. In addition, some of the piggyGUM transgenic lines showed mCherry expression in tissues other than wings (data not shown). The expression patterns outside the wing-related tissues were not consistent among the transgenic lines, suggesting that the piggyGUM construct also occasionally traps endogenous enhancers.

We also tested whether the promoter that is endogenous to the enhancer works better for a reporter assay construct in *Tribolium*. We made a piggyBac construct with the 2 kb sequence upstream of the *Tc-nub* transcription start site [confirmed by 5′ rapid amplification of cDNA ends (RACE); Clark-Hachtel et al., 2013] as the promoter [piggyBac *nub* promoter dsRed (piggyNub-proR); Fig. 6A]. We also used the 2 kb sequence downstream of the *Tc-nub* stop codon (confirmed by 3′ RACE; Clark-Hachtel et al., 2013) as the 3′ untranslated region (UTR) and the polyA signal native to *Tc-nub* (Fig. 6A). We made a similar construct for *Tc-Act5c* (with the 1 kb sequence upstream of the transcription start site and the 1 kb sequence downstream of the stop codon as the native promoter and polyA signal, respectively) as a comparison (Fig. 6B). To our surprise, *Tc-nub1L* in piggyNub-proR did not drive any expression in *Tribolium* (Fig. 6C-F) or in *Drosophila* (Fig. 6G,H). Real-time qPCR analysis revealed that there is no transcription of dsRed in these transgenic lines in both species (data not shown), suggesting that the lack of reporter expression is not due to incompatibility of the reporter gene with the *Tc-nub* UTRs but, rather, to the *nub* wing enhancer failing to work with the endogenous promoter and/or polyA signal. In contrast to piggyNub-proR-*Tc-nub1L*, piggyAct5cR showed strong and ubiquitous dsRed expression in *Tribolium* (Fig. 6I), indicating that our strategy of incorporating the endogenous transcription and translation components is valid. Intriguingly, however, piggyAct5cR did not drive any expression in *Drosophila* (data not shown), implying a strict species-specific nature of the transcription and/or translation components (such as promoters), even for an evolutionarily highly conserved house-keeping gene that is uniformly expressed in various species, including *Drosophila* and *Tribolium* (Chung and Keller, 1990).

Testing the reporter construct in another context in *Tribolium*

We next tested whether our DSCP reporter system worked in a context other than wings in *Tribolium*. We chose *hb* as a case study, and tested the reporter activity during embryogenesis. *hb* expression in *Tribolium* starts as a broad posterior domain in the blastoderm, and subsequently clears from the posterior to form an anterior band of expression that covers the pre-gnathal and gnathal segments (Lynch et al., 2012; Marques-Souza et al., 2008). In the early germband stage, the band resolves into a stripe covering the labium (Fig. 7B) (Marques-Souza, 2007; Zhu et al., 2017). Wolff et al. previously identified a genomic region at the *Tribolium hb* locus that drives blastoderm expression when introduced in *Drosophila* (Fig. 7A, orange bar) (Wolff et al., 1998). This region corresponds to a SCRMshaw prediction (Fig. 7A, purple bars). Therefore, although the FAIRE signal at this region is weak (likely because of the wide time window of sampling during early embryogenesis), the outcomes of previous studies make this region an excellent candidate enhancer to test with our reporter system in *Tribolium*. We cloned a 1340 bp fragment containing this genomic region (*hb*-PE1, Fig. 7A, red bar) and tested its enhancer activity using the piggyGUM construct in *Tribolium*. *In situ* hybridization for the *mCherry* reporter gene revealed that the piggyGUM-*hb*-PE1 construct recapitulates the *hb* expression at the early germband stage in *Tribolium* (Fig. 7C). This result indicates that: (1) our DSCP reporter system works well even during embryogenesis in *Tribolium*; (2) *hb*-PE1 contains the *hb* early germband enhancer.

DISCUSSION

In this study, we demonstrated that FAIRE-based chromatin profiling is a powerful approach for identifying CREs, such as enhancers, in *Tribolium*. The *Tribolium nub* wing enhancer that we identified (*Tc-nub1L*) is over 40 kb away from the *nub* transcription start site and 10 kb away from the *pu11* insertion site, which would be very difficult to identify without the aid of open chromatin profiles. In addition, with the use of the DSCP, we were able to establish a functional reporter assay construct in *Tribolium*. A combination of FAIRE-based chromatin profiling with this reporter

Development (2018) 145, dev160663. doi:10.1242/dev.160663

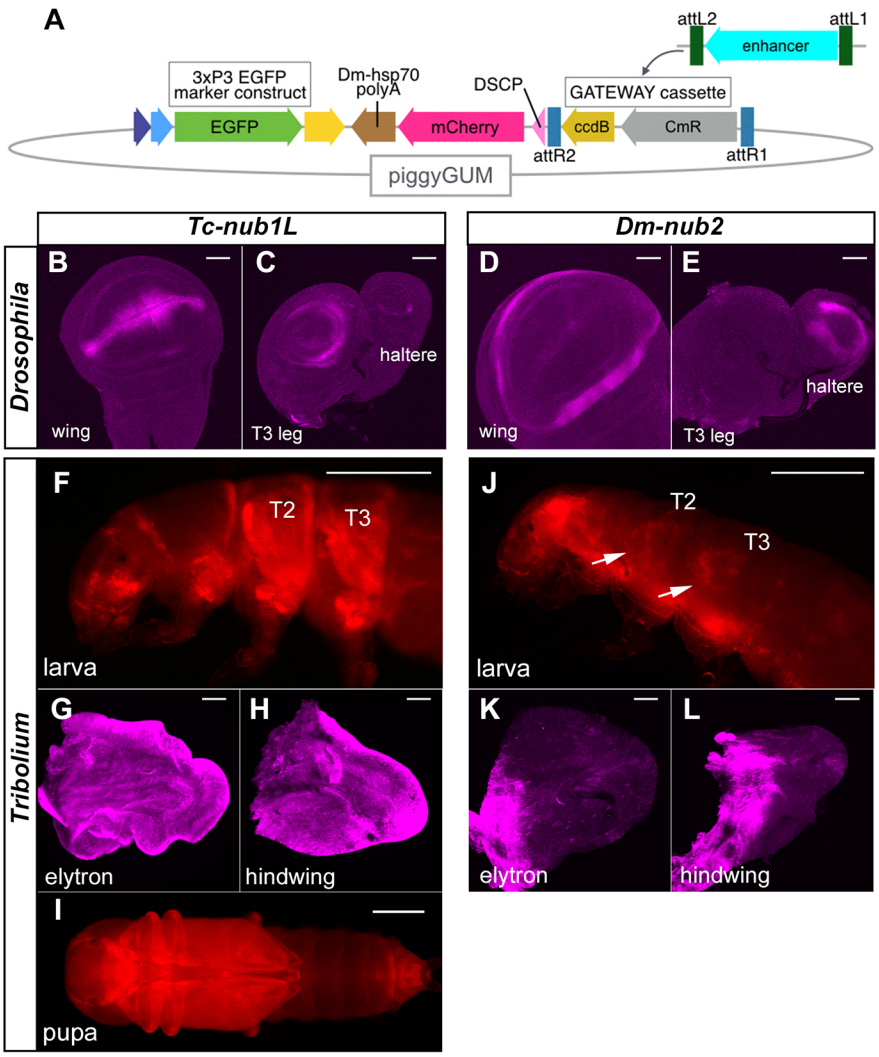

Fig. 5. Reporter assay with the DSCP construct in *Drosophila* and *Tribolium*. (A) The piggyGUM construct. (B-E) Enhancer activity of *Tc-Nub1L* (B,C) and *Dm-nub2* (D,E) tested with the piggyGUM construct in *Drosophila*. (F-L) Reporter expression of piggyGUM-*Tc-nub1L* (F-I) and piggyGUM-*Dm-nub2* (J-L) in *Tribolium*. Scale bars: 50 µm (B-E,G,H,K,L); 0.5 mm (F,I,J).

assay system will allow us to assess the function and evolution of enhancers in *Tribolium*.

FAIRE profiles in *Tribolium*

Genome-wide FAIRE profiling in *Tribolium* has identified a significant number of genomic regions whose chromatin status is regulated in a tissue- and stage-specific manner (Table 1, Fig. S1). These regions are promising candidates for future enhancer studies in *Tribolium*. In addition, our FAIRE analysis has revealed both evolutionarily conserved and diverged aspects of chromatin state regulation in *Drosophila* and *Tribolium*. For the conserved aspect, we saw similar chromatin profiles for the T2 and T3 epidermal samples, even though these two tissues differentiate into morphologically distinct structures (the elytron in T2 and hindwing in T3). This outcome echoes the message obtained from the *Drosophila* FAIRE study, i.e. that chromatin profiles are largely similar among the similar lineages of tissues (such as legs, wings and halteres), with the exception of a handful of ‘master control gene’ loci (McKay and Lieb, 2013). In fact, three of the four differentially open FAIRE peaks between T2 and T3 in our *Tribolium* FAIRE analysis were within the *Ultrabithorax* (the T3 selector gene) locus (Fig. S1) (for a review of the function of *Ultrabithorax* during wing development, see Tomoyasu, 2017). In contrast, the *Tribolium* FAIRE profiles during embryogenesis showed an interesting difference when compared with those in *Drosophila*. In *Drosophila*, the number of genomic regions that are open is fairly consistent throughout embryogenesis, with a distinct set of genomic regions being open in each stage (McKay and Lieb, 2013). In *Tribolium*, we noticed that a larger number of chromatin regions are open early in embryogenesis, and some of these regions are subsequently closed, resulting in a smaller number of open chromatin regions in later stages. This difference may be a reflection of the different modes of embryogenesis in the two insects (long versus short germband embryogenesis), although the significance of the difference in chromatin profiles has yet to be investigated.

We also noticed a strict overlap between the GC-rich regions and FAIRE-detected open chromatin regions. This raises interesting questions about the evolution of enhancers. Are these regions open because they are functionally important (such as enhancers)? Or have these regions become enhancers, because they were open owing to a bias in their nucleotide content and, thus, were accessible to TFs? There appears to be a similar correlation among the GC-rich regions, enhancers and FAIRE peaks in *Drosophila* (Li et al., 2007; McKay and Lieb, 2013). It will be interesting to investigate how GC-rich regions overlap with open chromatin regions in other insects. In addition, we found that the GC-rich and FAIRE-positive regions appear at regular intervals throughout the *Tribolium* genome. The molecular basis and functional implication of this

Development (2018) 145, dev160663. doi:10.1242/dev.160663

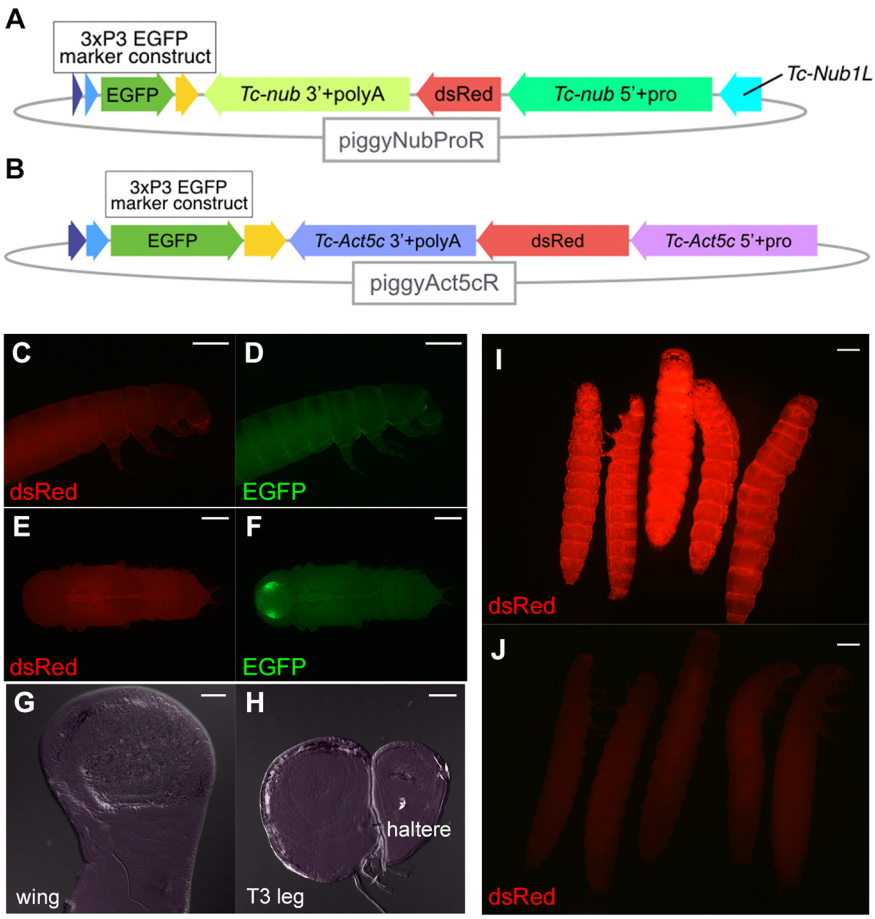

Fig. 6. Reporter assay with the *Tribolium* endogenous promoters in *Drosophila* and *Tribolium*. (A) The piggyNub-proR construct. (B) The piggyAct5cR construct. (C-F) Enhancer activity of *Tc-Nub1L* tested with the piggyNub-proR construct. dsRed reporter expression is completely absent (C,E), even though EGFP (D,F) confirms the presence of the transgenic construct. (G,H) The piggyNub-proR reporter expression in *Drosophila* imaginal discs. (I) dsRed expression of the piggyAct5cR at the larval stage in *Tribolium*. (J) dsRed expression of the piggyNub-proR with *Tc-Nub1L* at the larval stage in *Tribolium*, with the same exposure time as I. Scale bars: 0.5 mm (C-F,I,J); 50 µm (G,H).

periodicity is currently unknown; however, it is intriguing to speculate that a genome-wide event (such as transposon invasion) might have significantly influenced the chromatin state landscape in the *Tribolium* lineage.

Overlaps between FAIRE peaks and SCRMshaw enhancer predictions

The high degree of overlap observed between FAIRE peaks and enhancers predicted by the completely different, solely

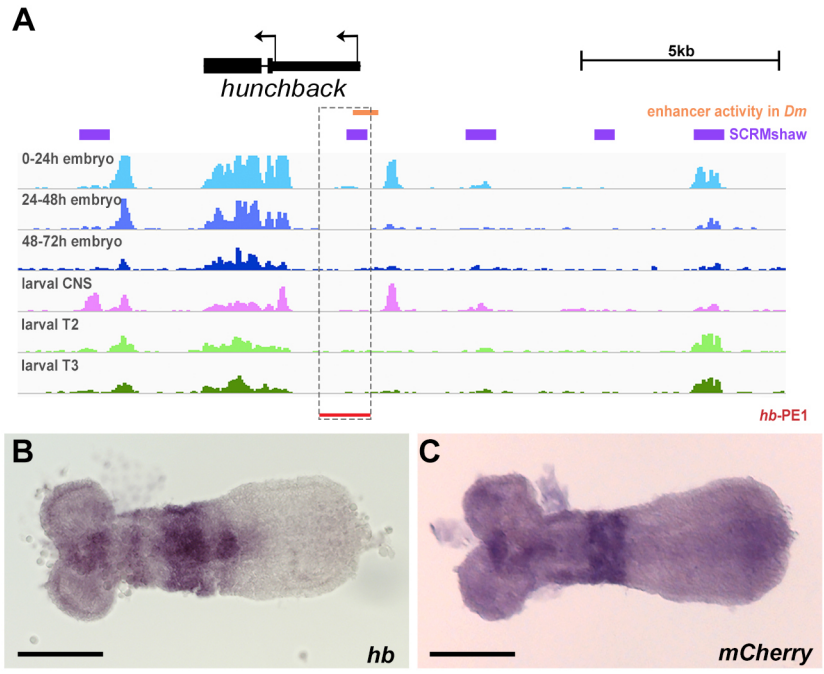

Fig. 7. *hb* enhancer analysis in *Tribolium*. (A) FAIRE profiles at the *hb* locus. Orange bar, blastoderm enhancer activity when introduced in *Drosophila*; purple, SCRMshaw predictions; red, the 1340 bp fragment tested in this study (*hb*-PE1). (B) *hb* expression at the early germband stage detected by *in situ* hybridization for *hb* transcript. (C) *mCherry* reporter gene expression of piggyGUM-*hb*-PE1 detected by *in situ* hybridization for *mCherry* transcript. Scale bars: 100 µm.

Development (2018) 145, dev160663. doi:10.1242/dev.160663

computational SCRMshaw method provides further confirmation that FAIRE is an effective means for enhancer discovery in *Tribolium*. Overall, the number of FAIRE peaks is well in excess of the number of SCRMshaw predictions. Several factors likely account for this result. First, the SCRMshaw predictions were performed at high stringency in order to minimize potential false-positive results (Kazemian et al., 2014); relaxing the prediction criteria would yield more predicted enhancers. Although this would potentially lead to more false positives, the >90% overlap seen for several specific datasets (Table S1) suggests that stringency could be relaxed in at least some cases. Second, SCRMshaw relies on training data from known *Drosophila* enhancers; therefore, enhancers with characteristics that deviate significantly from those of *Drosophila* enhancers will be found only by chromatin profiling, such as FAIRE. Finally, although FAIRE appears to be biased toward enhancers (Song et al., 2011), it also identifies other regions of open chromatin, such as promoters and insulator regions (Giresi et al., 2007), which are not predicted by the enhancer-specific SCRMshaw.

The twin issues of higher SCRMshaw false-positive rates at lower prediction stringencies and the lack of discrimination of FAIRE with respect to enhancers with specific spatial and temporal activity profiles suggest that considerable advantages could be obtained by using the methods in combination. Overlap with FAIRE peaks can be used to filter out false-positive SCRMshaw predictions, allowing predictions to be performed at lower stringency, and thus higher sensitivity. Conversely, SCRMshaw predictions can be used to focus on potentially more relevant FAIRE peaks – helping to avoid selecting enhancer sequences that are active in tissues other than the one of interest, enhancers for a neighboring housekeeping gene, insulators, and cryptic promoters or promoters for unannotated genes. This will be particularly useful in situations such as the one seen here for the larval samples, where cleanly separating the wing from body wall tissue was difficult: a common challenge when attempting to isolate tissues from small organisms such as insect embryos.

Enhancer activity in cross-species contexts and the limitation of non-native reporter assays

Our reporter assays in two insect species showed that both *Drosophila* and *Tribolium nub* wing enhancers were at least partially active in the cross-species context. We identified a 20 bp sequence that was shared between the two enhancers. This sequence contained binding sites of some wing-related TFs (such as Brinker and Mad) (Fig. S4), making it a promising candidate for an evolutionarily conserved enhancer motif. However, deletion of this sequence did not influence the activity of these enhancers in *Drosophila*, indicating that this sequence is dispensable for enhancer function (Fig. S4). We did not recognize any other significant sequence similarity or a conserved TF-binding site architecture between the two enhancers, suggesting that the regulatory landscape in the wings of the two species is evolutionarily maintained (as the *nub* enhancers can be functional in cross-species contexts), despite the lack of noticeable sequence conservation in the enhancer itself. A thorough examination of *trans* components that regulate the *nub* wing enhancers may give us insights into how enhancers evolve in a conserved regulatory landscape.

Although the *Tribolium* wing enhancer was active in *Drosophila*, we noticed that the activity of this enhancer was somewhat restricted, as it was active mainly at the dorsal-ventral (DV) compartmental boundary of the T2 wing, and in only a few cells in the haltere. This is in contrast with the expression in *Tribolium*, which showed a broader activity domain in the entire wing tissue in both the T2 and T3 segments. These differences in the activity domains suggest that some components that regulate the *Tribolium nub* wing enhancer are missing from the *Drosophila* T2 wing and are almost entirely absent in the haltere. This highlights the limitation of cross-species analyses and the importance of performing reporter assays in the native species. The reporter assay system we developed allows us to analyze enhancer activities in *Tribolium*. The successful demonstration of reporter analyses for *nub* in the wing and *hb* in the embryo suggest that our reporter construct works in various tissues; however, it is still crucial to evaluate the applicability of this system in diverse contexts.

Choice of core promoters in reporter constructs

Our study showed that the choice of promoters is crucial when assessing enhancer activity. *Tc-hsp68* was our first choice because it has successfully been used in the Gal4/UAS system in *Tribolium* (Schinko et al., 2010). However, although this promoter worked efficiently in *Drosophila*, in our reporter assay it failed to drive reporter expression even with a functional enhancer in *Tribolium* (at least in our hands). Interestingly, the transgenic beetles with the *Tc-hsp68* reporter construct showed a high occurrence of enhancer trap events (Fig. 4F-M), even though this promoter failed to work with the enhancer that we placed directly upstream of it. One explanation is that this promoter requires a certain distance for optimal interaction with enhancers in *Tribolium*. The situation might be less strict in *Drosophila* (for an unknown reason), allowing the *Tc-hsp68* promoter to overcome the distance requirement.

We also tried to assess the *nub* wing enhancer activity with the *nub* endogenous promoter but, to our surprise, this construct did not drive any expression. There are several possible explanations for this outcome. First, the region we selected might not contain the correct promoter for the *nub* transcript, although our 5′ RACE results (as well as the published *Tribolium* genome annotation, Tribolium Genome Sequencing et al., 2008) support our annotation of the *nub* transcription start site (Clark-Hachtel et al., 2013). Second, the 2 kb region we used as the promoter may contain a suppressor element, interfering with the enhancer to drive reporter expression. Third, the *nub* promoter might require a long distance to interact properly with the wing enhancer, as the wing enhancer that we identified was 40 kb away from the *nub* transcription start site. This characteristic might be similar to that of *Tc-hsp68*, which preferentially interacts with enhancers located at a certain distance. This may further support the idea that *Drosophila* are more permissive to changes in the enhancer/promoter distance. However, in the case of the *nub* endogenous promoter, there might be additional issues other than enhancer/promoter distance that prevent this reporter construct from working even in *Drosophila*.

The reporter construct with the DSCP (piggyGUM) worked efficiently both in *Drosophila* and in *Tribolium*. The DCSP is a synthetic core promoter, composed of several common core promoter motifs [i.e. TATA box, initiator element (Inr), motif ten element (MTE) and downstream promoter element (DPE)] isolated from the *Drosophila* genome. The DSCP has been shown to work efficiently with a diverse array of developmental enhancers in various contexts in *Drosophila* (Pfeiffer et al., 2008; Zabidi et al., 2015), suggesting that this promoter may also work well with other enhancers in *Tribolium*. However, it is worth mentioning that a synthetic promoter similar to the DCSP, SCP1 (composed of *Drosophila* and viral promoter motifs; Juven-Gershon et al., 2006), failed to work when tested in the Gal4/UAS system in *Tribolium* (Schinko et al., 2010). This again emphasizes the importance of choosing the correct promoter that fits the context of the study, which remains a crucial area for further exploration.

Development (2018) 145, dev160663. doi:10.1242/dev.160663

Enhancer studies in evo-devo

The study of enhancers and other CREs is crucial for understanding the molecular basis underlying morphological evolution, as changes in gene regulation, rather than the acquisition of new genes or the modification of protein structures, are often responsible for the evolution of diverse morphology (Carroll, 2008). For example, changes in enhancers can facilitate evolution of novel structures via co-opting pre-existing GRNs into a new context. Acquisition of enhancers *de novo* may also play a crucial role in morphological novelty. Therefore, studying both evolutionarily conserved and diverged enhancers will help further our understanding of morphological evolution (see Monteiro and Podlaha, 2009 for a comprehensive discussion of how *cis* studies can help elucidate the molecular basis for the evolution of novel traits). However, it has been a challenge to study enhancers in non-traditional model insects because of the lack of a reliable enhancer identification strategy. In this study, we showed that FAIRE-seq is readily applicable to non-traditional model species. Furthermore, the DSCP can be a useful promoter for establishing a reporter assay system and investigating the evolution of enhancers in non-*Drosophila* insects. Therefore, FAIRE-based chromatin profiling, along with reporter assay systems applicable to various insects and SCRMshaw enhancer prediction, will make the research on enhancers more accessible, which will provide us with more insights into the evolution of the regulatory mechanisms underlying morphological diversity.

MATERIALS AND METHODS

Fly stocks

The following two *Drosophila* strains used in this study were obtained from the Bloomington *Drosophila* Stock Center: *P{UAS-Dcr-2.D}*^{1}, *w*^{1118}; *P{GawB}nubbin-AC-62* and *y*^{1} *w**; *wg*$^{Sp-1}$/*CyO, P{Wee-P.ph0}Bacc*$^{Wee-P20}$; *P{20XUAS-6XGFP}attP2*.

Beetle cultures

The beetle cultures were reared on whole-wheat flour (+5% yeast) at 30°C in a temperature- and humidity-controlled incubator. The *nub* enhancer trap line *pu11*, which has enhanced yellow fluorescent protein (EYFP) expression in the hindwing and elytron discs (Clark-Hachtel et al., 2013; Lorenzen et al., 2003; Tomoyasu et al., 2005), was used to monitor *nub* expression in *Tribolium*.

Tissue preparation for FAIRE

For the *Tribolium* larval T2 and T3 wing tissues, the dorso-lateral region of the epidermal tissues that contain elytron (T2) and hindwing (T3) discs were dissected from the last instar larvae. Although these samples largely consisted of tissues that give rise to wing structures, they also contained body wall tissues as well as larval muscles because of the difficulty of precisely dissecting the wing tissues from larvae. About 50 larvae (100 dissected tissues) were used for each biological replicate, and three replicates were prepared for each wing sample. The brains were dissected from the head of the last instar larvae. About 40 brains were used for each biological replicate, and two replicates were prepared. Embryos were collected in whole-wheat flour (+5% yeast) for 24 h at 30°C. The collected embryos were cultured for 1 and 2 days at 30°C for the 24-48 h and 48-72 h samples, respectively; 0.1 g of embryos was used for each biological replicate, and three replicates were prepared for each sample. These tissues and embryos were crosslinked with 4% formaldehyde for 30 min (larval tissues) or 8% formaldehyde for 30 min (embryos).

FAIRE-seq analysis

FAIRE was performed as previously described (McKay and Lieb, 2013). FAIRE-seq libraries were sequenced on an Illumina HiSeq 2000 at the University of North Carolina High-Throughput Sequencing Facility. 50 bp single-end Illumina reads were obtained for FAIRE-treated samples and two non-FAIRE-treated input samples. Reads were trimmed to remove the index sequence and mapped to the *Tribolium* reference genome (version 3.0) with bowtie2 (Langmead and Salzberg, 2012). Read alignments were quality filtered (Q<10 dropped), and duplicate reads were removed using SAMtools (http://samtools.sourceforge.net/). For visualization of FAIRE signal, bigwig files were produced by merging tissue/stage-specific replicate bam files with SAMtools and normalizing reads to sequencing depth using deepTools (https://deeptools.readthedocs.io/en/develop/). These files were then visualized with the IGV genome viewer (Robinson et al., 2011; Thorvaldsdóttir et al., 2013). Peaks were called on individual replicates using MACS2 (https://github.com/taoliu/MACS), with the merged input sample bam files as the control. The *Drosophila* FAIRE profiles used in this study have been previously published (McKay and Lieb, 2013). For differentially open peak analysis, mapped reads (.bam files) for each replicate and the merged input, along with the MACS2 peaks (.narrowPeak files) called for each replicate, were provided as input for DiffBind (http://bioconductor.org/packages/release/bioc/html/DiffBind.html). DiffBind creates a consensus peakset for all replicates provided, requiring a consensus peak to be present in at least two replicates of a sample. An experiment-wide consensus peakset was produced using all samples. Pairwise analysis of differentially open peaks between samples was performed within DiffBind with the DESeq2 method for all consensus peaksets, and plotted using the dba.plotMA() function. The differentially open peaks are listed in Table S6.

Genome-wide GC-contents analysis

Using the experiment-wide consensus peakset described above, 1 kb of sequence upstream and downstream of each peak center was extracted from the genome using BEDTools (Quinlan and Hall, 2010) and custom Python (https://www.python.org/) scripts. For these 2 kb fragments, those free of 'N's were subjected to GC analysis. Changes in local GC content (250 bp sliding window, 10 bp step) were plotted against the whole-fragment average of GC content for all fragments. For the GC-rich region distance analysis, first, bedGraphs of GC content fluctuations above and below the genome wide average were computed at 70 and 60 bp resolution for the *Tribolium* and *Drosophila* genomes, respectively. The genome of *Bombyx mori* as well as those of several coleopteran insects (*Agrilus planipennis*, *Dendroctonus ponderosae*, *Anoplophora glabripennis*, *Leptinotarsa decemlineata*, *Nicrophorus vespilloides* and *Onthophagus taurus*) were analyzed at 70 bp resolution. Peaks were then called using the bdgcallpeak command in MACS2. The distance between the edges of adjacent peaks was categorized into 100 bp bins, and the natural logarithm of the number of occurrences was plotted. For the FAIRE peak distance analysis, distances between FAIRE peaks were collected and plotted in the same manner as the GC peaks. A consensus *Drosophila* FAIRE peakset was obtained from DiffBind with the same setting as those for the *Tribolium* data applied to the previously published data (McKay and Lieb, 2013).

Comparison between FAIRE and SCRMshaw

Enhancers predicted by SCRMshaw were taken from Kazemian et al. (2014) and converted into BED format. BEDTools (http://bedtools.readthedocs.io/en/latest/) *merge* was used to combine overlapping and/or redundant (i.e. from more than one SCRMshaw scoring method) predictions, reducing the total number of predicted enhancers to 1214. BEDTools *intersect* was then used to determine all predicted enhancers with at least 50 bp overlap with a FAIRE peak (-f 0.10). FAIRE peaks that were not assigned to a *Tribolium* chromosome (i.e. not starting with 'ChLG') were omitted. Significance of overlaps was determined using BEDTools *fisher*; all overlaps were highly significant with $-\log(P) \geq 19$. Because this method provides only an approximation, a selection of datasets was tested via randomization. BEDTools *shuffle* was used to generate 1000 random intervals, and the intersections were determined as above. The mean and standard deviation of the randomized intersections were calculated and used with the observed (SCRMshaw) intersection value to determine a *z* score. *P* values from all randomization tests were highly significant.

Drosophila reporter assay constructs

pFUGG, a *Drosophila* Gateway-compatible phiC31 transformation plasmid, was used for reporter assays in *Drosophila* (McKay and Lieb, 2013). The phiC31 system allows site-specific integration (Bischof et al.,

Development (2018) 145, dev160663. doi:10.1242/dev.160663

2007), thus preventing position effects due to different insertion sites. An enhancer cloned into pFUGG drives Gal4 as the reporter, whose expression domains can then be visualized by crossing to UAS-EGFP flies.

Gateway-compatible piggyBac reporter constructs

The piggyBac plasmid with the 3×P3-EGFP marker construct and the *Fse*I/*Asc*I cloning site (Horn and Wimmer, 2000) was used to make all piggyBac constructs used in this study. For piggyGHR (piggyBac Gateway Tc-hsp68 dsRed), the *gypsy* element, the *Tc-hsp68* core promoter, *dsRed* and the SV40 polyA signal were amplified by PCR, assembled through ligation and inserted into the *Fse*I/*Asc*I site of the piggyBac plasmid. The assembled plasmid was then converted to a Gateway-compatible plasmid by Gateway Vector Conversion System (ThermoFisher Science). For piggyGUM (piggyBac Gateway Universal promoter mCherry), the reporter construct including the Gateway cassette was amplified from a *Drosophila* Gateway-compatible phiC31 transformation vector and inserted into the *Fse*I/*Asc*I site of the piggyBac plasmid. The primers used to make piggyGUM are listed in Table S5. The annotated sequence of the DSCP used in piggyGUM is shown in Fig. S5. The reporter constructs in piggyNub-proR (piggyBac *nub* promoter dsRed) and piggyAct5cR (*piggy*Bac *Act5c* promoter dsRed) were *de novo* synthesized and inserted into the *Fse*I/*Asc*I site of the piggyBac plasmid.

Enhancer cloning

Genomic fragments corresponding to possible enhancer regions were PCR amplified and cloned into pENTR using pENTR-D Directional TOPO Cloning kit (Thermo-Fisher Scientific, K240020). The primers used to clone the enhancer regions from the *Drosophila* and *Tribolium* genome are listed in Table S5. Cloned genomic fragments were then inserted into reporter constructs via Gateway Clonase reaction (Thermo-Fisher Scientific, 11791-019).

Drosophila and *Tribolium* transgenesis

For *Drosophila* transgenesis, pFUGG constructs were transformed into the attP2 site (68A4) through phiC31 integrase-mediated transgenesis system, and piggyBac constructs were transformed into w^{1118} with EGFP as a visible marker (BestGene *Drosophila* transgenic service). For *Tribolium* transgenesis, piggyBac constructs were transformed into *vermilion*^{white} with EGFP as a visible marker (TriGenES *Tribolium* Genome Editing Service for the *nub* and *Act5c* constructs, Friedrich-Alexander-Universität Erlangen-Nürnberg for the *hb* construct).

Immunohistochemistry and *in situ* hybridization

Drosophila imaginal discs were dissected from the third instar larvae and fixed with 4% formaldehyde for 25 min. *Tribolium* elytron and hindwing discs were dissected from the last instar larvae and fixed with 4% formaldehyde for 25 min. Dissected tissues were then washed and blocked with 10% BSA, and incubated with rabbit anti-mCherry antibody (1:500; Abcam, ab167453) at 4°C overnight. After washing for 1 h, the tissues were incubated with the Alexa 555-conjugated goat anti-rabbit antibody (1:500) for 2 h at room temperature. All the discs were mounted on glass slides with ProLong Gold antifade reagent (Life Technologies) for documentation. *in situ* hybridization was performed as previously described (Shippy et al., 2009), with digoxigenin (DIG)-labeled riboprobes and alkaline phosphatase-conjugated anti-DIG antibody (Sigma-Aldrich 11093274910). The signal was developed using BM-Purple (Sigma-Aldrich 11442074001). The primers used to amplify the *mCherry* fragment for riboprobe synthesis are included in Table S5. The *hb* riboprobe used in this study has been previously described (Wolff et al., 1998).

Image processing and documentation

The images were captured with a Zeiss 710 confocal microscope (mounted discs) and Zeiss AxioCam MRc5 with Zeiss Discovery V12 (*Tribolium* larvae and pupae). A filter set specific to mCherry (575/50×, 640/50 m) was used to visualize the mCherry expression driven by piggyGUM constructs. *Tribolium* germband embryos were imaged with a ProgRes CFcool camera on a Zeiss Axio Scope.A1 microscope using ProgRes CapturePro image acquisition software. Some pictures were enhanced only for brightness and contrast with Adobe Photoshop.

Acknowledgements

We thank the Bloomington Stock Center for fly stocks, Johannes Schinko for discussions on *Tribolium* transgenesis, and the Center for Bioinformatics and Functional Genomics (CBFG) and Center for Advanced Microscopy and Imaging (CAMI) at Miami University for technical support. We also thank Shuxia Yi for technical support, and Courtney Clark-Hachtel, David Linz and other members of Tomoyasu lab for helpful discussions.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: D.J.M., Y.T.; Methodology: Y.-T.L., K.D.D., F.B.-C., N.S., M.S.H., D.J.M., Y.T.; Software: K.D.D., K.S., M.S.H., D.J.M., Y.T.; Validation: K.D.D., F.B.-C., N.S., M.S.H., D.J.M., Y.T.; Formal analysis: Y.-T.L., K.D.D., F.B.-C., M.S.H., D.J.M., Y.T.; Investigation: Y.-T.L., K.D.D., F.B.-C., N.S., H.R., K.S., E.E.-S., M.S.H., D.J.M., Y.T.; Resources: D.J.M., Y.T.; Data curation: Y.-T.L., K.D.D., M.S.H., D.J.M., Y.T.; Writing - original draft: Y.-T.L., M.S.H., D.J.M., Y.T.; Writing - review & editing: Y.-T.L., K.D.D., E.E.-S., M.S.H., D.J.M., Y.T.; Visualization: Y.-T.L., K.D.D., H.R., E.E.-S., M.S.H., D.J.M., Y.T.; Supervision: M.S.H., D.J.M., Y.T.; Project administration: M.S.H., D.J.M., Y.T.; Funding acquisition: M.S.H., Y.T.

Funding

This project was supported by a National Science Foundation (NSF) grant (IOS0950964 and IOS1557936 to Y.T.), a U.S. Department of Agriculture (USDA) grant (2012-67013-19361 to M.S.H.) and The University of North Carolina at Chapel Hill (UNC-CH) start-up funds to D.J.M.

Data availability

FAIRE-seq data have been deposited in Gene Expression Omnibus (GEO) under accession number GSE104495. Sequences of the piggyBac plasmids established in this study are available at Figshare (https://figshare.com/articles/Insect_Transformation_Plasmids/6050486).

Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.160663.supplemental

References

Angelini, D. R., Smith, F. W. and Jockusch, E. L. (2012). Extent with modification: leg patterning in the beetle Tribolium castaneum and the evolution of serial homologs. *G3* **2**, 235-248.

Bellés, X. (2010). Beyond Drosophila: RNAi in vivo and functional genomics in insects. *Annu. Rev. Entomol.* **55**, 111-128.

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

Brown, S. J., Mahaffey, J. P., Lorenzen, M. D., Denell, R. E. and Mahaffey, J. W. (1999). Using RNAi to investigate orthologous homeotic gene function during development of distantly related insects. *Evol. Dev.* **1**, 11-15.

Bucher, G., Scholten, J. and Klingler, M. (2002). Parental RNAi in tribolium (Coleoptera). *Curr. Biol.* **12**, R85-R86.

Cande, J., Goltsev, Y. and Levine, M. S. (2009). Conservation of enhancer location in divergent insects. *Proc. Natl. Acad. Sci. USA* **106**, 14414-14419.

Carroll, S. B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25-36.

Carroll, S., Grenier, J. K. and Weatherbee, S. D. (2005). *From DNA to Diversity*, 2nd edn. Malden, MA: Blackwell Publishing.

Chung, Y. T. and Keller, E. B. (1990). Positive and negative regulatory elements mediating transcription from the Drosophila melanogaster actin 5C distal promoter. *Mol. Cell. Biol.* **10**, 6172-6180.

Clark-Hachtel, C. M., Linz, D. M. and Tomoyasu, Y. (2013). Insights into insect wing origin provided by functional analysis of vestigial in the red flour beetle, Tribolium castaneum. *Proc. Natl. Acad. Sci. USA* **110**, 16951-16956.

Denell, R. (2008). Establishment of tribolium as a genetic model system and its early contributions to evo-devo. *Genetics* **180**, 1779-1786.

Eckert, C., Aranda, M., Wolff, C. and Tautz, D. (2004). Separable stripe enhancer elements for the pair-rule gene hairy in the beetle Tribolium. *EMBO Rep.* **5**, 638-642.

Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M. and Dubchak, I. (2004). VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* **32**, W273-W279.

Giresi, P. G., Kim, J., McDaniell, R. M., Iyer, V. R. and Lieb, J. D. (2007). FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res.* **17**, 877-885.
Halfon, M. S. (2017). Perspectives on gene regulatory network evolution. *Trends Genet.* **33**, 436-447.
Hong, J.-W., Hendrix, D. A. and Levine, M. S. (2008). Shadow enhancers as a source of evolutionary novelty. *Science* **321**, 1314.
Horn, C. and Wimmer, E. A. (2000). A versatile vector set for animal transgenesis. *Dev. Genes Evol.* **210**, 630-637.
Jenett, A., Rubin, G. M., Ngo, T.-T. B., Shepherd, D., Murphy, C., Dionne, H., Pfeiffer, B. D., Cavallaro, A., Hall, D., Jeter, J. et al. (2012). A GAL4-driver line resource for Drosophila neurobiology. *Cell Rep.* **2**, 991-1001.
Jory, A., Estella, C., Giorgianni, M. W., Slattery, M., Laverty, T. R., Rubin, G. M. and Mann, R. S. (2012). A survey of 6,300 genomic fragments for cis-regulatory activity in the imaginal discs of Drosophila melanogaster. *Cell Rep.* **2**, 1014-1024.
Juven-Gershon, T., Cheng, S. and Kadonaga, J. T. (2006). Rational design of a super core promoter that enhances gene expression. *Nat. Methods* **3**, 917-922.
Kantorovitz, M. R., Kazemian, M., Kinston, S., Miranda-Saavedra, D., Zhu, Q., Robinson, G. E., Göttgens, B., Halfon, M. S. and Sinha, S. (2009). Motif-blind, genome-wide discovery of cis-regulatory modules in Drosophila and mouse. *Dev. Cell* **17**, 568-579.
Katzen, F. (2007). Gateway® recombinational cloning: a biological operating system. *Expert Opin. Drug Discov.* **2**, 571-589.
Kazemian, M., Zhu, Q., Halfon, M. S. and Sinha, S. (2011). Improved accuracy of supervised CRM discovery with interpolated Markov models and cross-species comparison. *Nucleic Acids Res.* **39**, 9463-9472.
Kazemian, M., Suryamohan, K., Chen, J.-Y., Zhang, Y., Samee, M. A. H., Halfon, M. S. and Sinha, S. (2014). Evidence for deep regulatory similarities in early developmental programs across highly diverged insects. *Genome Biol. Evol.* **6**, 2301-2320.
Kvon, E. Z., Kazmar, T., Stampfel, G., Yáñez-Cuna, J. O., Pagani, M., Schernhuber, K., Dickson, B. J. and Stark, A. (2014). Genome-scale functional characterization of Drosophila developmental enhancers in vivo. *Nature* **512**, 91-95.
Langmead, B. and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 367-359.
Li, L., Zhu, Q., He, X., Sinha, S. and Halfon, M. S. (2007). Large-scale analysis of transcriptional cis-regulatory modules reveals both common features and distinct subclasses. *Genome Biol.* **8**, R101.
Lorenzen, M. D., Berghammer, A. J., Brown, S. J., Denell, R. E., Klingler, M. and Beeman, R. W. (2003). piggyBac-mediated germline transformation in the beetle Tribolium castaneum. *Insect Mol. Biol.* **12**, 433-440.
Lynch, J. A., El-Sherif, E. and Brown, S. J. (2012). Comparisons of the embryonic development of *Drosophila*, *Nasonia*, and *Tribolium*. *Wiley Interdiscip. Rev. Dev. Biol.* **1**, 16-39.
Marques-Souza, H. (2007). Evolution of the gene regulatory network controlling trunk segmentation in insects.
Marques-Souza, H., Aranda, M. and Tautz, D. (2008). Delimiting the conserved features of hunchback function for the trunk organization of insects. *Development* **135**, 881-888.
Mayor, C., Brudno, M., Schwartz, J. R., Poliakov, A., Rubin, E. M., Frazer, K. A., Pachter, L. S. and Dubchak, I. (2000). VISTA : visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* **16**, 1046-1047.
McKay, D. J. and Lieb, J. D. (2013). A common set of DNA regulatory elements shapes Drosophila appendages. *Dev. Cell* **27**, 306-318.
Monteiro, A. and Podlaha, O. (2009). Wings, horns, and butterfly eyespots: how do complex traits evolve? *PLoS Biol.* **7**, e1000037.
Ng, M., Diaz-Benjumea, F. J. and Cohen, S. M. (1995). Nubbin encodes a POU-domain protein required for proximal-distal patterning in the Drosophila wing. *Development* **121**, 589-599.
Pearson, J. C., McKay, D. J., Lieb, J. D. and Crews, S. T. (2016). Chromatin profiling of Drosophila CNS subpopulations identifies active transcriptional enhancers. *Development* **143**, 3723-3732.
Peel, A. D. (2008). The evolution of developmental gene networks: lessons from comparative studies on holometabolous insects. *Philos. Trans. R. Soc. B Biol. Sci.* **363**, 1539-1547.
Pfeiffer, B. D., Jenett, A., Hammonds, A. S., Ngo, T.-T. B., Misra, S., Murphy, C., Scully, A., Carlson, J. W., Wan, K. H., Laverty, T. R. et al. (2008). Tools for neuroanatomy and neurogenetics in Drosophila. *Proc. Natl. Acad. Sci. USA* **105**, 9715-9720.
Quinlan, A. R. and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842.
Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. and Mesirov, J. P. (2011). Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24-26.
Schinko, J. B., Weber, M., Viktorinova, I., Kiupakis, A., Averof, M., Klingler, M., Wimmer, E. A. and Bucher, G. (2010). Functionality of the GAL4/UAS system in Tribolium requires the use of endogenous core promoters. *BMC Dev. Biol.* **10**, 53.
Schmitt-Engel, C., Schultheis, D., Schwirz, J., Ströhlein, N., Troelenberg, N., Majumdar, U., Dao, V. A., Grossmann, D., Richter, T., Tech, M. et al. (2015). The iBeetle large-scale RNAi screen reveals gene functions for insect development and physiology. *Nat. Commun.* **6**, 7822.
Shippy, T. D., Coleman, C. M., Tomoyasu, Y. and Brown, S. J. (2009). Concurrent in situ hybridization and antibody staining in red flour beetle (Tribolium) embryos. *Cold Spring Harb. Protoc.* **2009**, pdb prot5257.
Shlyueva, D., Stampfel, G. and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* **15**, 272-286.
Simon, J. M., Giresi, P. G., Davis, I. J. and Lieb, J. D. (2012). Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. *Nat. Protoc.* **7**, 256-267.
Smith, F. W., Angelini, D. R., Gaudio, M. S. and Jockusch, E. L. (2014). Metamorphic labral axis patterning in the beetle *Tribolium castaneum* requires multiple upstream, but few downstream, genes in the appendage patterning network. *Evol. Dev.* **16**, 78-91.
Song, L., Zhang, Z., Grasfeder, L. L., Boyle, A. P., Giresi, P. G., Lee, B.-K., Sheffield, N. C., Gräf, S., Huss, M., Keefe, D. et al. (2011). Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res.* **21**, 1757-1767.
Sosinsky, A., Honig, B., Mann, R. S. and Califano, A. (2007). Discovering transcriptional regulatory regions in Drosophila by a nonalignment method for phylogenetic footprinting. *Proc. Natl. Acad. Sci. USA* **104**, 6305-6310.
Stark, A., Lin, M. F., Kheradpour, P., Pedersen, J. S., Parts, L., Carlson, J. W., Crosby, M. A., Rasmussen, M. D., Roy, S., Deoras, A. N. et al. (2007). Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. *Nature* **450**, 219-232.
Suryamohan, K. and Halfon, M. S. (2015). Identifying transcriptional *cis* -regulatory modules in animal genomes. *Wiley Interdiscip. Rev. Dev. Biol.* **4**, 59-84.
Thorvaldsdóttir, H., Robinson, J. T. and Mesirov, J. P. (2013). Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178-192.
Tomoyasu, Y. (2017). Ultrabithorax and the evolution of insect forewing/hindwing differentiation. *Curr. Opin. Insect Sci.* **19**, 8-15.
Tomoyasu, Y. and Denell, R. E. (2004). Larval RNAi in Tribolium (Coleoptera) for analyzing adult development. *Dev. Genes Evol.* **214**, 575-578.
Tomoyasu, Y., Wheeler, S. R. and Denell, R. E. (2005). Ultrabithorax is required for membranous wing identity in the beetle Tribolium castaneum. *Nature* **433**, 643-647.
Tomoyasu, Y., Arakane, Y., Kramer, K. J. and Denell, R. E. (2009). Repeated co-options of exoskeleton formation during wing-to-elytron evolution in beetles. *Curr. Biol.* **19**, 2057-2065.
Tribolium Genome Sequencing, C., Richards, S., Gibbs, R. A., Weinstock, G. M., Brown, S. J., Denell, R., Beeman, R. W., Gibbs, R., Beeman, R. W., Brown, S. J. et al. (2008). The genome of the model beetle and pest Tribolium castaneum. *Nature* **452**, 949-955.
Uyehara, C. M., Nystrom, S. L., Niederhuber, M. J., Leatham-Jensen, M., Ma, Y., Buttitta, L. A. and McKay, D. J. (2017). Hormone-dependent control of developmental timing through regulation of chromatin accessibility. *Genes Dev.* **31**, 862-875.
Wang, L., Wang, S., Li, Y., Paradesi, M. S. R. and Brown, S. J. (2007). BeetleBase: the model organism database for Tribolium castaneum. *Nucleic Acids Res.* **35**, D476-D479.
Wolff, C., Schröder, R., Schulz, C., Tautz, D. and Klingler, M. (1998). Regulation of the Tribolium homologues of caudal and hunchback in Drosophila: evidence for maternal gradient systems in a short germ embryo. *Development* **125**, 3645-3654.
Zabidi, M. A., Arnold, C. D., Schernhuber, K., Pagani, M., Rath, M., Frank, O. and Stark, A. (2015). Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518**, 556-559.
Zdobnov, E. M. and Bork, P. (2007). Quantification of insect genome divergence. *Trends Genet.* **23**, 16-20.
Zhu, X., Rudolf, H., Healey, L., François, P., Brown, S. J., Klingler, M. and El-Sherif, E. (2017). Speed regulation of genetic cascades allows for evolvability in the body plan specification of insects. *Proc. Natl. Acad. Sci. USA* **114**, E8646-E8655.
Zinzen, R. P., Cande, J., Ronshaugen, M., Papatsenko, D. and Levine, M. (2006). Evolution of the Ventral Midline in Insect Embryos. *Dev. Cell* **11**, 895-902.