© 2022. Published by The Company of Biologists Ltd | Development (2022) 149, dev201016. doi:10.1242/dev.201016

RESEARCH REPORT

Shared *cis*-regulatory modules control expression of the tandem paralogs *midline* and *H15* in the follicular epithelium

Cody A. Stevens[1], Helen L. Stott[1], Shreya V. Desai[2] and Nir Yakoby[1,2,*]

ABSTRACT

The posterior end of the follicular epithelium is patterned by *midline* (MID) and its paralog *H15*, the *Drosophila* homologs of the mammalian Tbx20 transcription factor. We have previously identified two *cis*-regulatory modules (CRMs) that recapitulate the endogenous pattern of *mid* in the follicular epithelium. Here, using CRISPR/Cas9 genome editing, we demonstrate redundant activity of these *mid* CRMs. Although the deletion of either CRM alone generated marginal change in *mid* expression, the deletion of both CRMs reduced expression by 60%. Unexpectedly, the deletion of the 5′ proximal CRM of *mid* eliminated *H15* expression. Interestingly, expression of these paralogs in other tissues remained unaffected in the CRM deletion backgrounds. These results suggest that the paralogs are regulated by a shared CRM that coordinates gene expression during posterior fate determination. The consistent overlapping expression of *mid* and *H15* in various tissues may indicate that the paralogs could also be under shared regulation by other CRMs in these tissues.

KEY WORDS: Developmental genes, Anterior-posterior axis coordination, Genes tagging, Gene regulatory network, Shared enhancers, *Cis*-regulation

INTRODUCTION

The follicular epithelium of the *Drosophila* egg chamber is an established model system for studying the genetic coordination of tissue patterning and axis formation (Bastock and St Johnston, 2008; Berg, 2005; Horne-Badovinac and Bilder, 2005; Revaitis et al., 2017; Ward and Berg, 2005; Yakoby et al., 2008b). In particular, the posterior end of the adult fly is set during early oogenesis by the secretion of Gurken, a transforming growth factor α-like ligand, from around the oocyte nucleus, and the subsequent activation of the epidermal growth factor receptor (EGFR) in the overlying follicle cells (Gonzalez-Reyes and St Johnston, 1998; Neuman-Silberberg and Schupbach, 1993; Ray and Schupbach, 1996; Sapir et al., 1998). This activation of EGFR establishes posterior fate by driving the expression of the ETS-transcription factor gene *pointed*, which regulates the *Drosophila* Tbx20 homolog *midline* (*mid*) (Fregoso Lomas et al., 2016, 2013; Stevens et al., 2020). As oogenesis proceeds, these ‘pre-patterned’ follicle cells establish the boundary between the future dorsal anterior and posterior domains by inhibiting *broad* (*br*): a marker of the future dorsal appendage domain of the *Drosophila* eggshell (Cheung et al., 2013; Deng and Bownes, 1997; Fregoso Lomas et al., 2013; Fuchs et al., 2012; Pyrowolakis et al., 2017; Tzolovsky et al., 1999). The *mid* paralog *H15* is expressed in a partially overlapping posterior domain (Fregoso Lomas et al., 2013), which may indicate a common regulatory mechanism. In addition, a basic alignment of *mid* and *H15* nucleotide-coding sequences reveals a 63% alignment, of which 76% was identical, further supporting the origin of the paralogs by gene duplication (Altschul et al., 1997; Gramates et al., 2022; Sebe-Pedros et al., 2013).

Recently published findings show that paralogous genes can maintain co-expression via shared CRMs in different developing tissues, including the embryo, leg and wing imaginal discs, indicating this mechanism is fundamental for proper animal development (Baudouin-Gonzalez et al., 2017; Bourbon et al., 2022; Lan and Pritchard, 2016; Levo et al., 2022; Loker and Mann, 2022). Here, we have analyzed two *cis*-regulatory modules (CRMs) that recapitulate the posterior pattern of *mid* in the follicle cells (Revaitis et al., 2017). We show using CRISPR/Cas9 genome editing that these CRMs are redundant. The deletion of both regions resulted in a 60% reduction of *mid* expression. Remarkably, the proximal upstream CRM is functionally shared between the two paralogs; its deletion eliminates *H15* expression in the follicle cells. The shared function of this CRM is limited to the regulation of *H15* and *mid* in the follicular epithelium, as its deletion did not affect the pattern of these genes in other developing tissues. Here, we have identified a coordinating-CRM that regulates tandem paralog genes during posterior fate determination in the female ovaries.

RESULTS AND DISCUSSION

Two CRMs recapitulate the endogenous pattern of MID/H15

The tandem paralogs *mid* and *H15* reside on the second chromosome about 55 kb apart (Fig. S1A). In a previous screen for oogenesis CRMs using the FlyLight collection, two CRMs recapitulated the endogenous posterior pattern of *mid* in the follicular epithelium (Fig. S1) (Jenett et al., 2012; Pfeiffer et al., 2008, 2010; Revaitis et al., 2017). One CRM, GMR86G04 (named here G04), is located upstream proximal to the *mid* promoter, and the other, GMR86F11 (named here F11), is located in the first intron of *mid* (Fig. S1B) (Gramates et al., 2022). We also tested whether the Ventral Leg Enhancer (VLE) is activated during oogenesis; however, no expression was detected (Fig. S1J-L′) (Svendsen et al., 2015, 2019). During stages 8/9 (S8/9) of oogenesis, both G04 and F11 drive a pattern of GFP expression similar to that of endogenous MID (Fig. S1D′,F′,H′). At S10, both CRMs drive a posterior pattern that remains outside the dorsolateral patches of Broad (BR) (Fig. S1E′, G′, I′), as was previously described for MID (Fregoso Lomas et al., 2013, 2016; Stevens et al., 2020). We note that the expression of GFP for both G04 and F11 is elevated in some

[1]Center for Computational and Integrative Biology, Rutgers, The State University of New Jersey, Camden, NJ 08103, USA. [2]Department of Biology, Rutgers, The State University of New Jersey, Camden, NJ 08103, USA.

*Author for correspondence (yakoby@camden.rutgers.edu)

N.Y., 0000-0002-7959-7588

Handling Editor: Cassandra Extavour
Received 9 June 2022; Accepted 11 October 2022

Development (2022) 149, dev201016. doi:10.1242/dev.201016

cells compared with others. This is likely influenced by the insertion of the GMR drivers outside the genome context of the *mid* locus, thus encompassing other elements non-native to the gene. In comparison with the native genome context, multiple CRMs reinforce the transcription of the endogenous gene, resulting in a more-robust expression pattern. In addition, the use of UAS/Gal4 system may lead to slight delays in reporter expression, as the system depends on the transcription and translation of the GAL4 transcription factor before driving the expression of the reporter gene.

MID and H15 exhibit overlapping patterns in multiple tissues

The patterns of *mid* and *H15* partially overlap during early developmental stages at the posterior follicular epithelium, suggesting a potential coordinating transcriptional mechanism (Fregoso Lomas et al., 2013). In the absence of a reliable antibody against H15, and with the consideration that high homology between *mid* and *H15* could allow the MID antibody to detect both proteins, we decided to differentially tag both proteins. We used CRISPR/Cas9 genome editing to tag the genes. The *H15* C terminus was HA-tagged (Fig. 1A) and the *mid* C terminus was tagged with a super-folder GFP (sfGFP) (Fig. 1B). A line containing both tagged genes was generated via co-injection of both donor vector plasmids (see Material and Methods, not shown in the schematic diagram – *H15*-HA, *mid*-sfGFP). All transgenic lines are homozygous fertile and viable healthy stocks that exhibited no apparent abnormalities in eggshell development (Fig. 1C) or in the adult flies when compared with the wild type.

During S7/8 of oogenesis, H15-HA and MID-sfGFP are detected in the posterior follicle cells (Fig. 1D,D′). In agreement with the previously reported patterns (Fregoso Lomas et al., 2013), H15-HA is more posteriorly restricted than the MID-sfGFP pattern. Beyond S9, H15-HA is no longer detected in the posterior (Fig. 1E); however, the MID-sfGFP is still expressed in the posterior end (Fig. 1E′), as was previously reported (Fregoso Lomas et al., 2013, 2016). The boundary between the posterior and dorsal anterior domains is established through the inhibition of the late expression of *br* by MID (Cheung et al., 2013; Deng and Bownes, 1997; Fregoso Lomas et al., 2013; Fuchs et al., 2012; Pyrowolakis et al., 2017; Tzolovsky et al., 1999). In agreement with this mechanism, MID-sfGFP does not overlap the dorsolateral patches of BR (Fig. 1F,F′). We detected a greater invasion of MID into the dorsal midline (compare Fig. S1E,E′ with Fig. 1F,F′, see also Fig. 2C′,C″), which may indicate that detection of MID-sfGFP is more sensitive than detection of MID using anti-MID antibodies.

To increase confidence in our tagged constructs, we looked for the patterns of MID and H15 in other developing tissues. Both H15-HA and MID-sfGFP express in the pre-cardioblast cells

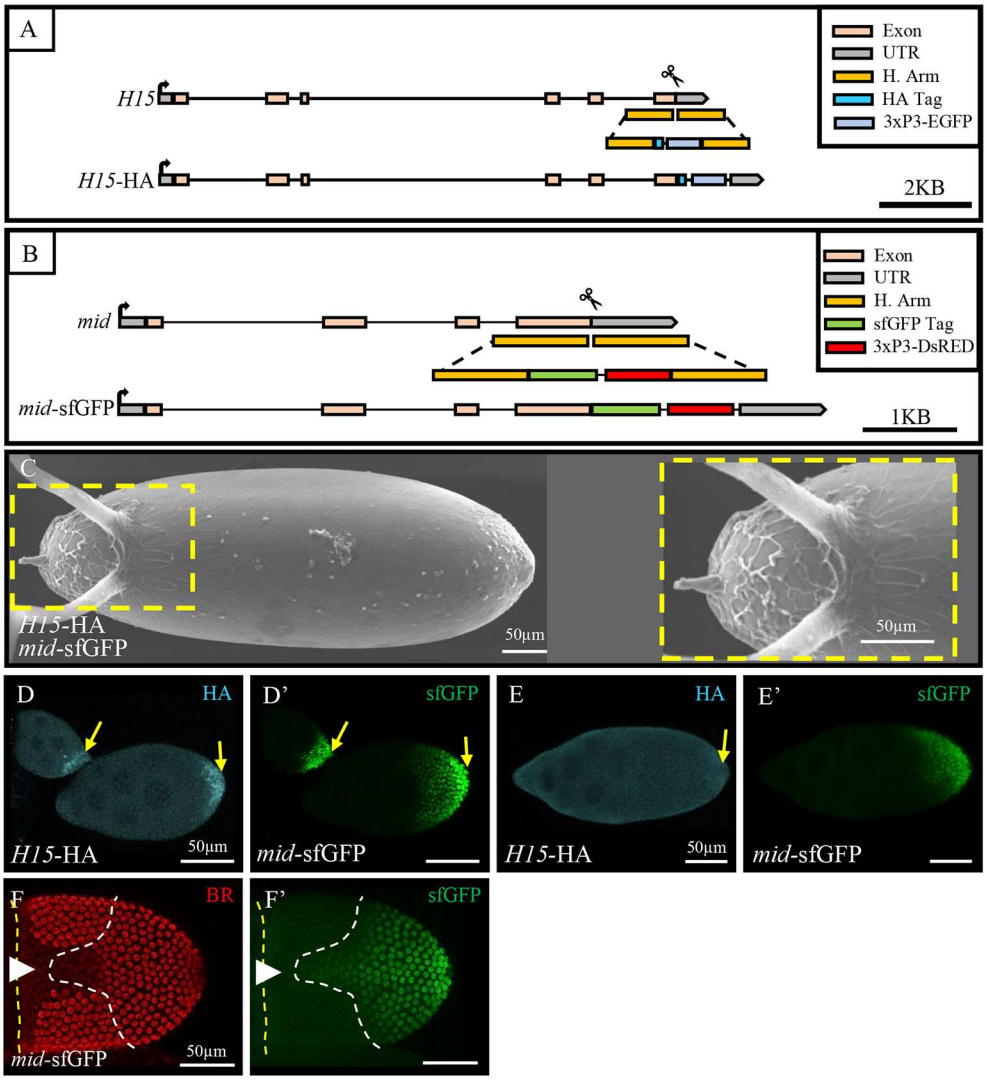

Fig. 1. CRISPR/Cas9 endogenous tagging of *H15* and *mid*. (A,B) Schematic diagram of CRISPR/Cas9 genome editing of *H15* and *mid* loci. H. Arm, homology arm. Black arrows indicate the direction of transcription; scissors indicate the location of a guide RNA double-stranded break. (C) *H15*-HA and *mid*-sfGFP eggshell (*n*=12) exhibiting wild-type eggshell morphology (compare with Fig. 2J). Dorsal view. Dashed yellow line indicates the area shown on the right. (D-E′) The patterns of H15-HA and MID-sfGFP at stages 7 and 8 (D,D′) (*n*=15) and stage 9 (E,E′) (*n*=8), yellow arrows mark the posterior end of the egg chamber. (F,F′) The pattern of MID-sfGFP at stage 10B. White dashed line indicates the boundary of MID-sfGFP pattern. Broad (BR, red) (*n*=8). Yellow dashed lines indicate the anterior boundary; white arrowheads indicate the dorsal midline. All images are shown with anterior towards the left; *n*, number of images with same results. Scale bars: 50 µm.

Development (2022) 149, dev201016. doi:10.1242/dev.201016

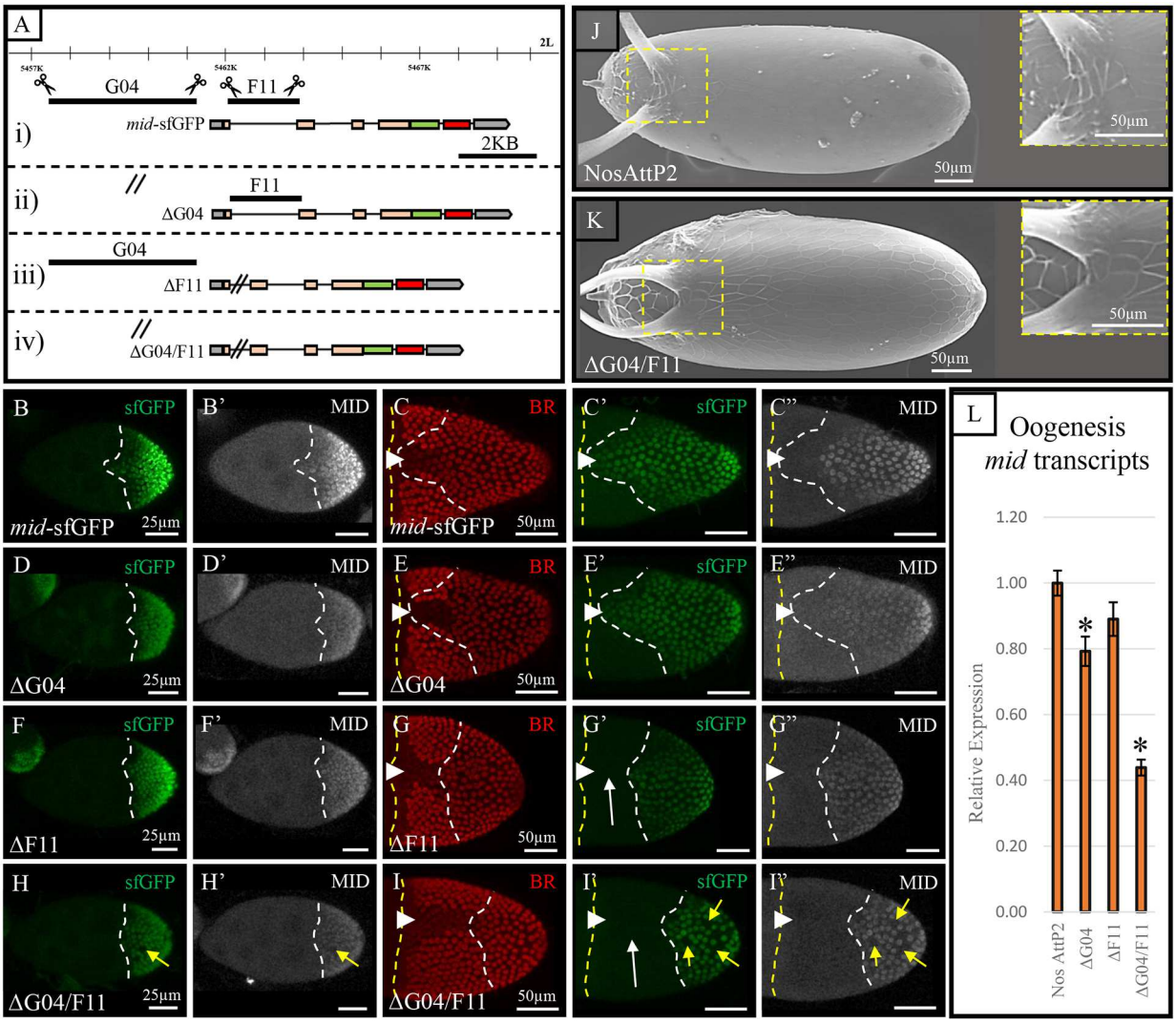

Fig. 2. Testing the function of G04 and F11 CRMs in the regulation of *mid* expression. (A) Schematic diagram of CRISPR/Cas9 genome editing of CRM deletions in the *mid* locus; horizontal line with small vertical lines marks the scaffold location on left arm 2nd chromosome. Scissors indicate the location of guide RNAs. Double diagonal lines indicate CRM deletion (Δ) regions. (B-C″) MID-sfGFP within the control at stage 8 (B,B′) (*n*=8) and stage 10B (C-C″) (*n*=16). (D-E″) MID-sfGFP within the G04 CRM deletion (ΔG04) at stage 8 (D,D′) (*n*=18) and stage 10B (E-E″) (*n*=10). (F-G″) MID-sfGFP within the F11 CRM deletion (ΔF11) at stage 8 (F,F′) (*n*=11) and stage 10B (G-G″) (*n*=6). White arrows in G′,I′ indicate loss of MID-sfGFP in dorsal midline domain. MID-sfGFP within G04 and F11 CRM deletion (ΔG04/F11) at stage 8 (*n*=6) and stage 10B (G,G′,I,I′) (*n*=5). Yellow arrows indicate loss of observable MID-sfGFP and MID. (C,E,G,I) Broad (BR, red). Yellow dashed line marks the anterior boundary; white arrowheads mark the dorsal midline; white dashed lines mark the boundary of the MID-sfGFP pattern. (B-I″) Using the antiGFP antibody together with the anti-MID increases the detection sensitivity of the anti-GFP antibody. (J) Eggshell of control (*n*=23) and (K) ΔG04/F11 line (*n*=70). The areas outlined by the dashed yellow boxes at the base of the two dorsal appendages are shown on the right. (L) qPCR for relative expression of *mid* transcripts compared with control NosAttP2 for each respective background. *mid* is reduced compared with control NosAttP2 expression (*$P \leq 0.05$; unpaired Student's *t*-test). All images are shown with anterior towards the left; *n*, number of images with same results. Scale bars: 25 µm (B,B′,D,D′,F,F′,H,H′); 50 µm (C-C″,E-E″,G-G″,I-K).

(Fig. S2A-A″), in agreement with their embryonic patterns (Buescher et al., 2004; Miskolczi-McCallum et al., 2005; Qian et al., 2005). We also later show (Fig. 3) that these patterns are conserved in the larvae brain and leg imaginal disc. We conclude that the tagged proteins are expressed and function correctly during development.

mid CRMs showcase redundancy in oogenesis

Using CRISPR/Cas9, we generated single and double deletions of G04 and F11 within the background of the *mid*-sfGFP and *H15*-HA lines (Fig. 2A$^{i\text{-}iv}$, *H15*-HA is not represented in the schematic diagram). The F11 CRM overlaps the first and second exons of *mid*; hence, we targeted the intron while ensuring the intron splice sites remained intact (Fig. 2A^{iii}). This removed ∼70% of the intron. Adult homozygous flies containing a single and double deletion (ΔG04, ΔF11 and ΔG04/F11) are viable, fertile and phenotypically wild type.

At S8, all deletion backgrounds express MID-sfGFP in a posterior pattern (Fig. 2B,D,F,H). However, sporadic cell-autonomous expression loss of MID-sfGFP was observed in ΔG04/F11 background (Fig. 2H). At S10, the MID-sfGFP pattern in ΔG04 is comparable with the control (Fig. 2C′,E′); however, ΔF11 shows loss of expression in the dorsal midline (compare Fig. 2C′,E′ with G′) and reflects by the MID pattern (Fig. 2G″). In the ΔG04/F11 background, MID-sfGFP was lost in the dorsal midline (as in the ΔF11, Fig. 2G′,I′) and in some posterior cells

Development (2022) 149, dev201016. doi:10.1242/dev.201016

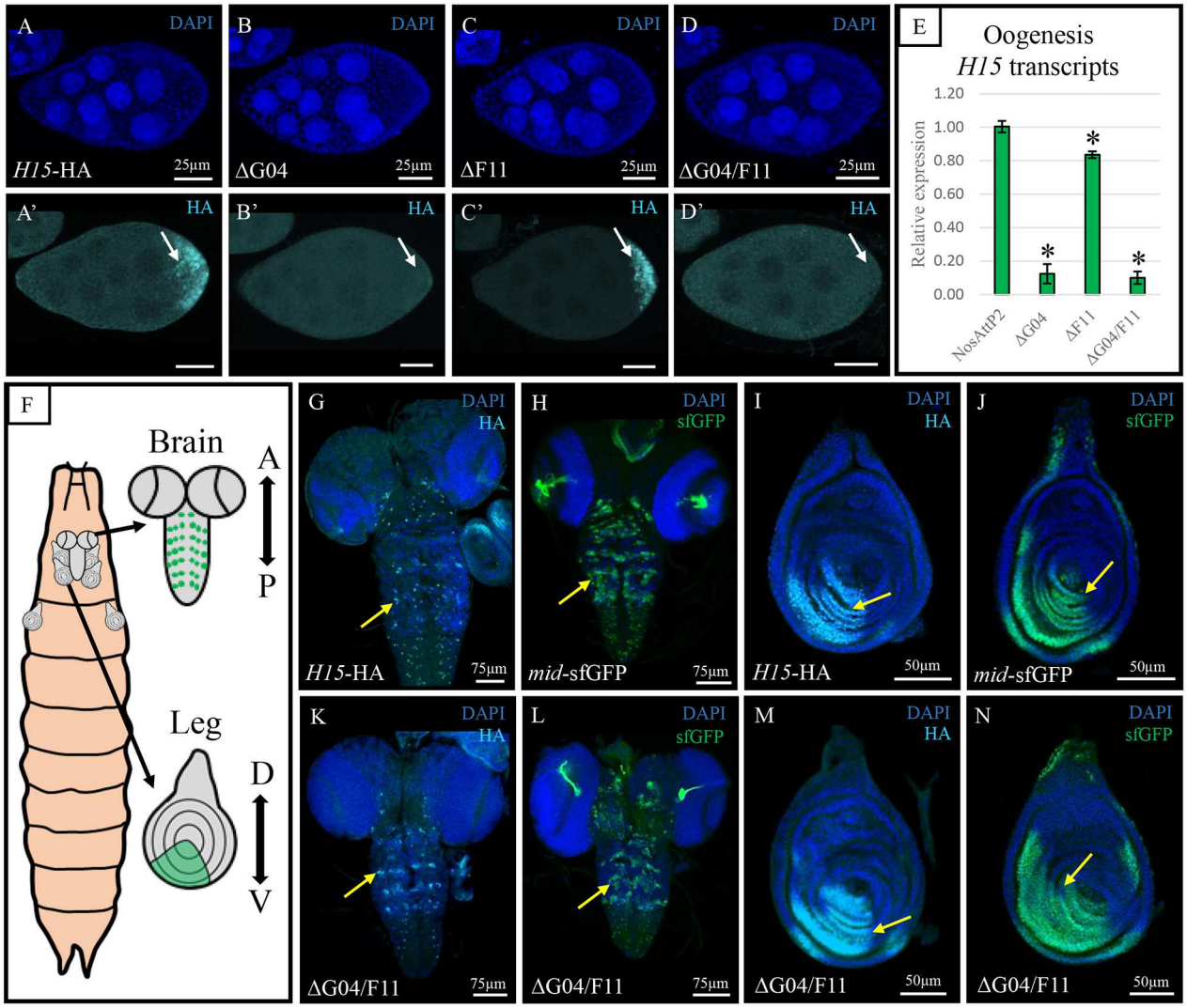

Fig. 3. The G04 CRM co-regulates *mid* and *H15* expression. The pattern of H15-HA in the control (A,A′) (*n*=21), (B,B′) ΔG04 (*n*=11), (C,C′) ΔF11 (*n*=12), (D,D′) ΔG04/F11 (*n*=10). (A,B,C,D) Corresponding DAPI staining of nuclei (blue). (E) qPCR for relative expression of *H15* transcripts compared with control NosAttP2 during oogenesis in each background. *H15* transcripts are significantly reduced compared with the control NosAttP2 (*$P \leq 0.05$; unpaired Student's *t*-test). (F) Schematic diagram of 3rd instar larval brain and leg imaginal discs. Green indicates known *mid*/*H15* expression. A, anterior, P, posterior, D, dorsal, V, ventral. (G-J) Control line exhibiting the patterns of H15-HA and MID-sfGFP in larval brain ventral nerve cord and leg imaginal disc (yellow arrows) (*n*=7, 5, 10 and 8, respectively). (K-N) ΔG04/F11 line exhibiting the pattern of H15-HA and MID-sfGFP as in G-J (*n*=3, 4, 14 and 8, respectively). For egg chambers, anterior is towards the left. *n*, number of images with same results. Yellow arrow indicates the protein pattern. Scale bars: 25 µm in A-D′; 50 µm in I,J,M,N; 75 µm in G,H,K,L.

(Fig. 2I,I′). In agreement, loss of MID was observed with anti-MID antibodies (Fig. 2H′,I″). Of note, our results show the greater sensitivity of the tagged MID-sfGFP in comparison with the use of antibodies (Fig. 2C′,E′ compared with C″,E″, respectively). Eggshells from the ΔG04/F11 backgrounds also exhibited dorsal appendages with a slightly connected base (56% connected, 44% wild type, *n*=70) (Fig. 2J,K), in agreement with the loss of MID in this domain (Fregoso Lomas et al., 2013).

We used a quantitative PCR (qPCR) assay of whole ovaries to determine the reduction in *mid* in these deletion backgrounds. We found a mild but significant ($P<0.05$) 21% reduction of *mid* in ΔG04 compared with the NosAttP2 control (Fig. 2L). No significant reduction was observed in ΔF11 ($P>0.05$) compared with the NosAttP2 control. We noted that the reduction of the MID pattern seen in the dorsal midline domain of stage 10 egg chambers in the ΔF11 background (Fig. 2G) might be compensated for by the elevated early expression of MID in this background (Fig. 2F). The difference between ΔG04 and ΔF11 was not significant (Fig. 2L). In agreement with the morphological changes observed in the eggshells (Fig. 2K), ΔG04/F11 shows a 56% reduction ($P<0.05$) in *mid* transcripts compared with the NosAttP2 control, with 44% reduction compared with ΔG04 and 50% reduction compared with ΔF11 (Fig. 2L). Although *mid* is significantly, but slightly, reduced in the ΔG04 background, this is not sufficient to change the eggshell structure. Hence, we propose that G04 and F11 are redundant CRMs in oogenesis; although each may be sufficient to drive biological functions in wild-type flies, the deletion of both reduces MID to a level where eggshell morphological defects are detected. The retention of *mid* transcription within the ΔG04/F11 background means that we cannot rule out the possibility of a third, or more, unknown CRMs. However, none were found when the surrounding locus was screened (Fig. S1B) (Revaitis et al., 2017). Of note, as is the case with shadow enhancers, these transgenic flies are housed under laboratory conditions. However, in the natural

Development (2022) 149, dev201016. doi:10.1242/dev.201016

environment, under stress, the severity of losing both CRMs could dramatically reduce *mid* transcription, as described for other genes (Frankel et al., 2010). Such experiments are beyond the scope of this report.

The G04 CRM is a coordinating *cis*-regulatory module in the posterior follicular epithelium

It was recently demonstrated that tandem paralogs can be regulated by shared CRMs (Bourbon et al., 2022; Levo et al., 2022; Loker and Mann, 2022). MID and H15 have a redundant role in *Drosophila* development (Miskolczi-McCallum et al., 2005; Qian et al., 2005; Svendsen et al., 2009, 2015, 2019). The patterns of MID and H15 overlap in the posterior follicular epithelium, which suggests a coordinating regulatory mechanism (Fig. 1D,D′) (Fregoso Lomas et al., 2013, 2016; Yakoby et al., 2008a). Additionally, we could not find any CRMs near *H15* that drive GFP expression during oogenesis (Fig. S1) (Revaitis et al., 2017). Accordingly, we aimed to determine whether the G04 and/or F11 functionally evolved to also regulate *H15*. Remarkably, the pattern of H15-HA was eliminated in the ΔG04 and ΔG04/ΔF11 backgrounds and reduced in the ΔF11background (Fig. 3A-D′).

Using qPCR, we quantified the levels of *H15* mRNA in the deletion backgrounds. As expected, both ΔG04 and ΔG04/F11 exhibited a dramatic reduction in *H15* expression ($P<0.05$, 88% and 90%, respectively) when compared with the control (Fig. 3E). Whereas, in ΔF11, we found a 16% reduction in *H15* expression ($P=0.04$). Deletion of F11 primarily drove the posterior retraction of MID-sfGFP (Fig. 2G′, white arrow), suggesting that F11 is required for MID expression in the posterior region of the dorsal midline. If indeed F11 reinforces the boundary between MID and BR, this may explain why H15, which is more posteriorly restricted than MID, is less affected by the deletion of F11 than the deletion of G04 (Fig. 3E). We conclude that, although both CRMs contribute to *H15* expression, G04 appears to be the primary driver of *mid* and *H15* expression during posterior fate determination.

The expression of *mid* and *H15* is regulated by other CRMs in other tissues

Developmental genes, like *mid* and *H15*, are often regulated by multiple CRMs (Andersson and Sandelin, 2020; Field and Adelman, 2020; Furlong and Levine, 2018; Long et al., 2016; Schoenfelder and Fraser, 2019; Shlyueva et al., 2014; Visel et al., 2009). These genes are also expressed in different tissues during animal development, which raises a question about the underlying mechanism controlling their expression. Are the identified CRMs for these genes tissue-specific or used across multiple tissues? Both *mid* and *H15* are expressed in numerous tissues, such as the embryo pre-cardioblast, the ventral nerve cord and leg imaginal discs of the developing larvae (Fig. 3F and Fig. S2) (Buescher et al., 2004; Svendsen et al., 2009). Our tagged proteins are detected in the control background patterning the ventral nerve cord neuroblasts and the leg imaginal disc (Fig. 3G-J). Interestingly, both proteins are also detected in the ΔG04/F11 background in these tissues (Fig. 3K-N). This indicates that other CRMs control *mid* and *H15* in these tissues, highlighting the necessity of multiple regulatory safeguards for these genes during development (Buescher et al., 2004). These results also explain how deletion of both CRMs still produced viable flies. The resultant flies produce fertile offspring, with changes at the base of the dorsal appendages (Fig. 2K). We conclude that the G04 and F11 CRMs are necessary for correct patterning in oogenesis, but *mid* and *H15* are regulated by different CRMs in other tissues. Support for this hypothesis was recently published; the VLE is predicted to be a shared CRM for the regulation of *mid* and *H15* in the leg disc (Levo et al., 2022).

Using a shared CRM to maintain the spatiotemporal overlapping activity of tandem paralogs is a simple mechanism reminiscent of the way a bacterial operon coordinates the expression of functionally linked genes (Kuhlman et al., 2007; Levo et al., 2022; Mayo et al., 2006; Setty et al., 2003). In this manner, MID and H15 reinforce central functions, including defining the posterior end, controlling heart development and setting the ventral side of the leg. Here, we show that although the contribution of G04 to the expression of the proximal *mid* gene is buffered by redundancy with F11, G04 is necessary for the expression of *H15* (residing 55 kb away) during posterior fate determination (Frankel et al., 2010; Perry et al., 2010, 2011).

Multiple CRMs driving overlapping patterns have been shown to buffer crucial gene expression under unstable environmental conditions (Frankel et al., 2010; Hong et al., 2008). Additionally, these redundant CRMs can provide ‘safeguards’ in the event of the loss of function of one CRM due to mutation. This phenomenon is supported in work on the transcription factor *dorsal*, for which it is estimated that 50% of its targets contain multiple CRMs (Perry et al., 2009). In contrast to the duplication of regulatory elements, gene duplication may double the amount of transcript, and in some cases could be deleterious (Lan and Pritchard, 2016). Given that *mid* and *H15* are expressed in multiple tissues during development, shared regulation of the paralogs may initially act as a protective mechanism that prevents overexpression of similar proteins before divergence.

A few compelling examples of the regulation of tandem paralogs by shared CRMs have recently been published (Bourbon et al., 2022; Levo et al., 2022; Loker and Mann, 2022). Here, we provide the first example of *mid* and *H15* regulation by a coordinating CRM in oogenesis. Recent work has demonstrated the importance of chromatin domain level effects on enhancer activity (Bolt et al., 2022). It is currently unknown whether the dramatic loss of *H15* seen with the deletion of the G04 CRM is due to a regulatory composition or to a geographical position. Future investigation into the DNA topology around the *mid* and *H15* will elucidate the mechanism underlying the remote regulation of tandem paralog genes by shared enhancers.

MATERIALS AND METHODS

Flies and reagents

Flies were raised on standard cornmeal agar and kept at room temperature. Wild-type *D. melanogaster* (25211), GMR^{86G04}-Gal4 (40467), GMR^{86F11}-Gal4 (40464), GMR^{86G06}-Gal4 (93199), GMR^{86G02}-Gal4 (40465), y[1]w[*]ubi-Cre;sna[Sco]/CyO;Dr/TM3,sb (34516) and nos-Cas9-AttP40 (78782) were obtained from the Bloomington *Drosophila* Stock Center. y[1]w[*]; if/SM6a;D/TM3,sb was a gift from Miki Fujioka (Thomas Jefferson University, Philadelphia, PA, USA). Flies generated in this study were: MID-VLE-LacZ, ΔG04, ΔF11, ΔG04/F11, *mid*-sfGFP/*H15*-HA, ΔG04-*mid*-sfGFP, ΔG04-*H15*-HA, ΔF11-*mid*-sfGFP, ΔF11-*H15*-HA, ΔG04/F11-*mid*-sfGFP and ΔG04/F11-*H15*-HA (all Rainbow Transgenics).

Immunohistochemistry

Immunohistochemistry analysis was carried out on 2- to 7-day-old flies raised on active yeast for 24 h at room temperature (23°C) before dissection. Ovaries were dissected in 1 ml Schneider’s media and fixed in a 4% paraformaldehyde/heptane/0.2% Triton X-100 in PBS (PBST) solution for 20 min. Samples were rinsed three times, for 5 min each time, in 0.2% PBST solution, then permeabilized in 1% PBST solution for 1 h. Samples were rinsed once in 0.2% PBST then blocked in 0.2% PBST with 1% bovine serum albumin (BSA) solution for 1 h. Samples were incubated overnight at 4°C in primary antibody cocktail with 0.2% PBST and 1% BSA. After

Development (2022) 149, dev201016. doi:10.1242/dev.201016

incubation, samples were washed three times for 20 min each in 0.2% PBST, then secondary antibody cocktail was added with 0.2% PBST and 1% BSA, and incubated for 1 h protected from light at room temperature. Samples were then washed three times for 20 min each time in 0.2% PBST and mounted in Fluoromount-G mounting media. Imaginal discs and nervous system were dissected from 3rd instar larva and immunohistochemistry performed using the above outlined procedure. For the embryo dissection procedure, approximately 3- to 7-day-old females were placed on grape juice agar plates and left overnight. Embryos were collected 24 h later and dechorionated in 4.125% sodium hypochlorite solution, subsequently rinsed with deionized water. Embryos were then placed in fix solution containing 4% paraformaldehyde. After fixation, embryos were washed in methanol to remove the vitelline membrane, then rehydrated in successive dilutions of methanol and 0.2% PBST. After rehydration, standard protocol for egg chambers was followed. Primary antibodies were mouse anti-Broad (1:250; DSHB, 25E9.D7), sheep anti-GFP (1:1000, Bio-Rad, 4745-1051), rabbit anti-β-galactosidase (1:1000; Invitrogen, A-11132), guinea pig anti-MID (1:1000; a gift from L. Nilson, McGill University, Montreal Quebec, Canada), and rat-anti HA (1:75; Roche, 11867423001). Secondary antibodies were Alex Fluor 488 donkey anti-mouse, Alexa Fluor 488 donkey anti-sheep, Alexa Fluor 568 donkey anti-rabbit, Alexa Fluor 568 donkey anti-mouse, Alexa Fluor 633 goat anti-guinea pig, and Alexa Fluor 568 donkey anti-rat (all at 1:1250; Invitrogen, A-21202, A-11015, A-10042, A-10037, A-21105, A-78946, respectively). Samples were imaged on Lecia SP8 confocal microscope with a 20× objective (Rutgers-Camden Imaging Core Facility). Images were processed using FIJI software (Schindelin et al., 2012). Eggshells were mounted on double-sided carbon tape, sputter coated with gold (JEOL Smart Coater) and imaged using a JCM-6000 JEOL NeoScope Scanning Electron Microscope (Rutgers-Camden Imaging Core Facility).

Generation of MID-VLE Reporter

All primers used for molecular manipulations are in Table S1. The *mid* ventral leg enhancer (VLE) (Svendsen et al., 2019, 2015) was amplified from OreR genomic-DNA. The fragment was Gibson assembled (NEB) into Xho1- and Not1-digested pattBGWhZn vector, upstream of LacZ alpha (Marmion et al., 2013). Reporter construct was injected into an attP40 line (Stock R8621, Rainbow Transgenics), integrating via PhiC31/attB-mediated integration (Groth et al., 2004). The generated fly was labeled MID-VLE-LacZ.

Generation of H15-HA and MID-sfGFP tagged transgenic flies

The H15-HA and MID-sfGFP transgenic flies were generated using one guide RNA (Gratz et al., 2014) to create the double-stranded break between the CDS and 3′UTR of each respective gene, and cloned into separate pU6-BbsI-chiRNA vectors (Gratz et al., 2013). Two 1 kb homology arms flanking the gRNA cut site were cloned into a modified pHd-HA-scarless vector for H15, and a phd-sfGFP-scarless vector for MID (Addgene, 80811). Flies were co-injected with both respective plasmids. Emerging progeny were crossed to y[1]w[*]; if/SM6a;D/TM3,sb and screened for an EGFP marker for H15-HA or a DsRED marker for MID-sfGFP, driven by a 3xP3 promoter in the adult eye. Transgenics expressed either EGFP or DsRED for a single integration, or both EGFP and DsRED for dual integration in the same injection background. Positive transgenics were PCR validated by comparing unadulterated Nos-AttP2 control and mutant gDNA. All amplicons were sequence validated (GeneWiz). Generated flies were labeled *mid*-sfGFP/*H15*-HA.

Generation of CRM deletions

To generate the ΔG04 and ΔF11 transgenic flies, two guide RNAs were cloned into pU6-BbsI chiRNA vectors (Gratz et al., 2014). Two 1 kb homology arms flanking the gRNA cut sites were cloned into a pHD-dsRED-attP vector (Gratz et al., 2014). Flies were injected with both guide and donor vector plasmids. Emerging progeny were crossed to y[1]w[*]; if/SM6a;D/TM3,sb and screened for the DsRED marker driven by a 3xP3 promoter. Once validated, the DsRED cassette was removed by crossing homozygous transgenic flies to y[1]w[*]ubi-Cre;sna[Sco]/CyO;Dr/TM3,sb, and screened for loss of DsRED. Positive transgenics were PCR validated by comparing Nos-AttP2 control and mutant gDNA. All amplicons were sequence validated (GeneWiz). Generated flies were labeled ΔG04, ΔF11.

To generate the ΔG04/F11 transgenic fly, the ΔF11 transgenic fly with the DsRED selection cassette removed was used as an injection background. The ΔG04 CRM was then targeted for deletion as previously described. Flies were injected with both guide and donor vector, as well as 400 ng/µl TrueCut Cas9 protein (Thermo Fisher; A36496). Positive transgenics were PCR validated by comparing Nos-AttP2 control and mutant gDNA. All amplicons were sequence validated (GeneWiz). Generated flies were labeled ΔG04/F11.

RT-qPCR

Quantitative PCR (qPCR) analysis on whole-ovary cDNA was conducted using an Applied Biosystems Quant-studio 6 Flex qPCR machine. 1 µg total RNA (Zymo RNA-easy Kit, standard protocol) was used to generate cDNA libraries (Protoscript II cDNA synthesis kit, standard protocol). SYBR-green PowerUp Master Mix (Thermo Scientific) was used for qPCR amplification of transcripts. Mitochondrial gene *RPL32* was used as housekeeping gene for fold-change calculations. For *H15* and *midline* genes, two primer sets were used and cT values averaged together for calculation of fold change. Three independent qPCR analyses, each comprising three biological replicates for each representative group, were averaged to represent the fold change. An unpaired Student's *t*-test was used for calculation of significant differences between the perturbation and the control.

Acknowledgements

We thank Laura Nilson for the *midline* antibody. We acknowledge Rainbow Transgenics for transgenic injections. We thank members of the Yakoby lab for fruitful discussions. SEM and confocal imaging were carried out at the Imaging Core Facility, Rutgers University, Camden, NJ, USA.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: C.A.S., N.Y.; Validation: C.A.S., H.L.S., S.V.D.; Formal analysis: C.A.S., H.L.S.; Investigation: C.A.S., H.L.S.; Resources: C.A.S., H.L.S.; Writing - original draft: C.A.S.; Writing - review & editing: C.A.S., H.L.S., N.Y.; Supervision: N.Y.; Project administration: N.Y.; Funding acquisition: N.Y.

Funding

C.A.S and H.L.S were partially supported by the Center for Computational and Integrative Biology at Rutgers, The State University of New Jersey, Camden. This research was supported by the National Institute of General Medical Sciences of the National Institutes of Health (2R15GM101597-02 to N.Y.) and by the National Science Foundation (IOS-1926802 to N.Y.). Deposited in PMC for release after 12 months.

Peer review history

The peer review history is available online at https://journals.biologists.com/dev/lookup/doi/10.1242/dev.201016.reviewer-comments.pdf.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402. doi:10.1093/nar/25.17.3389

Andersson, R. and Sandelin, A. (2020). Determinants of enhancer and promoter activities of regulatory elements. *Nat. Rev. Genet.* **21**, 71-87. doi:10.1038/s41576-019-0173-8

Bastock, R. and St Johnston, D. (2008). Drosophila oogenesis. *Curr. Biol.* **18**, R1082-R1087. doi:10.1016/j.cub.2008.09.011

Baudouin-Gonzalez, L., Santos, M. A., Tempesta, C., Sucena, E., Roch, F. and Tanaka, K. (2017). Diverse Cis-Regulatory mechanisms contribute to expression evolution of tandem gene duplicates. *Mol. Biol. Evol.* **34**, 3132-3147. doi:10.1093/molbev/msx237

Berg, C. A. (2005). The Drosophila shell game: patterning genes and morphological change. *Trends Genet.* **21**, 346-355. doi:10.1016/j.tig.2005.04.010

Bolt, C. C., Lopez-Delisle, L., Hintermann, A., Mascrez, B., Rauseo, A., Andrey, G. and Duboule, D. (2022). Context-dependent enhancer function

revealed by targeted inter-TAD relocation. *Nat. Commun.* **13**, 3488. doi:10.1038/s41467-022-31241-3

Bourbon, H. G., Benetah, M. H., Guillou, E., Mojica-Vazquez, L. H., Baanannou, A., Bernat-Fabre, S., Loubiere, V., Bantignies, F., Cavalli, G. and Boube, M. (2022). A shared ancient enhancer element differentially regulates the bric-a-brac tandem gene duplicates in the developing Drosophila leg. *PLoS Genet.* **18**, e1010083. doi:10.1371/journal.pgen.1010083

Buescher, M., Svendsen, P. C., Tio, M., Miskolczi-Mccallum, C., Tear, G., Brook, W. J. and Chia, W. (2004). Drosophila T box proteins break the symmetry of hedgehog-dependent activation of wingless. *Curr. Biol.* **14**, 1694-1702. doi:10.1016/j.cub.2004.09.048

Cheung, L. S., Simakov, D. S., Fuchs, A., Pyrowolakis, G. and Shvartsman, S. Y. (2013). Dynamic model for the coordination of two enhancers of broad by EGFR signaling. *Proc. Natl. Acad. Sci. USA* **110**, 17939-17944. doi:10.1073/pnas.1304753110

Deng, W. M. and Bownes, M. (1997). Two signalling pathways specify localised expression of the Broad-Complex in Drosophila eggshell patterning and morphogenesis. *Development* **124**, 4639-4647. doi:10.1242/dev.124.22.4639

Field, A. and Adelman, K. (2020). Evaluating enhancer function and transcription. *Annu. Rev. Biochem.* **89**, 213-234. doi:10.1146/annurev-biochem-011420-095916

Frankel, N., Davis, G. K., Vargas, D., Wang, S., Payre, F. and Stern, D. L. (2010). Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466**, 490-493. doi:10.1038/nature09158

Fregoso Lomas, M., Hails, F., Lachance, J. F. and Nilson, L. A. (2013). Response to the dorsal anterior gradient of EGFR signaling in Drosophila oogenesis is prepatterned by earlier posterior EGFR activation. *Cell Rep* **4**, 791-802. doi:10.1016/j.celrep.2013.07.038

Fregoso Lomas, M., De Vito, S., Boisclair Lachance, J. F., Houde, J. and Nilson, L. A. (2016). Determination of EGFR signaling output by opposing gradients of BMP and JAK/STAT activity. *Curr. Biol.* **26**, 2572-2582. doi:10.1016/j.cub.2016.07.073

Fuchs, A., Cheung, L. S., Charbonnier, E., Shvartsman, S. Y. and Pyrowolakis, G. (2012). Transcriptional interpretation of the EGF receptor signaling gradient. *Proc. Natl. Acad. Sci. USA* **109**, 1572-1577. doi:10.1073/pnas.1115190109

Furlong, E. E. M. and Levine, M. (2018). Developmental enhancers and chromosome topology. *Science* **361**, 1341-1345. doi:10.1126/science.aau0320

Gonzalez-Reyes, A. and St Johnston, D. (1998). The Drosophila AP axis is polarised by the cadherin-mediated positioning of the oocyte. *Development* **125**, 3635-3644. doi:10.1242/dev.125.18.3635

Gramates, L. S., Agapite, J., Attrill, H., Calvi, B. R., Crosby, M. A., Dos Santos, G., Goodman, J. L., Goutte-Gattat, D., Jenkins, V. K., Kaufman, T. et al. (2022). FlyBase: a guided tour of highlighted features. *Genetics* **220**, iyac035. doi:10.1093/genetics/iyac035

Gratz, S. J., Cummings, A. M., Nguyen, J. N., Hamm, D. C., Donohue, L. K., Harrison, M. M., Wildonger, J. and O'connor-Giles, K. M. (2013). Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **194**, 1029-1035. doi:10.1534/genetics.113.152710

Gratz, S. J., Ukken, F. P., Rubinstein, C. D., Thiede, G., Donohue, L. K., Cummings, A. M. and O'connor-Giles, K. M. (2014). Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. *Genetics* **196**, 961-971. doi:10.1534/genetics.113.160713

Groth, A. C., Fish, M., Nusse, R. and Calos, M. P. (2004). Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. *Genetics* **166**, 1775-1782. doi:10.1093/genetics/166.4.1775

Hong, J.-W., Hendrix, D. A. and Levine, M. S. (2008). Shadow enhancers as a source of evolutionary novelty. *Science* **321**, 1314. doi:10.1126/science.1160631

Horne-Badovinac, S. and Bilder, D. (2005). Mass transit: epithelial morphogenesis in the Drosophila egg chamber. *Dev. Dyn.* **232**, 559-574. doi:10.1002/dvdy.20286

Jenett, A., Rubin, G. M., Ngo, T. T., 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. doi:10.1016/j.celrep.2012.09.011

Kuhlman, T., Zhang, Z., Saier, M. H., Jr and Hwa, T. (2007). Combinatorial transcriptional control of the lactose operon of Escherichia coli. *Proc. Natl. Acad. Sci. USA* **104**, 6043-6048. doi:10.1073/pnas.0606717104

Lan, X. and Pritchard, J. K. (2016). Coregulation of tandem duplicate genes slows evolution of subfunctionalization in mammals. *Science* **352**, 1009-1013. doi:10.1126/science.aad8411

Levo, M., Raimundo, J., Bing, X. Y., Sisco, Z., Batut, P. J., Ryabichko, S., Gregor, T. and Levine, M. S. (2022). Transcriptional coupling of distant regulatory genes in living embryos. *Nature* **605**, 754-760. doi:10.1038/s41586-022-04680-7

Loker, R. and Mann, R. S. (2022). Divergent expression of paralogous genes by modification of shared enhancer activity through a promoter-proximal silencer. *Curr. Biol.* **32**, 3545-3555.e4. doi:10.1016/j.cub.2022.06.069

Long, H. K., Prescott, S. L. and Wysocka, J. (2016). Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* **167**, 1170-1187. doi:10.1016/j.cell.2016.09.018

Marmion, R. A., Jevtic, M., Springhorn, A., Pyrowolakis, G. and Yakoby, N. (2013). The Drosophila BMPRII, wishful thinking, is required for eggshell patterning. *Dev. Biol.* **375**, 45-53. doi:10.1016/j.ydbio.2012.12.011

Mayo, A. E., Setty, Y., Shavit, S., Zaslaver, A. and Alon, U. (2006). Plasticity of the cis-regulatory input function of a gene. *PLoS Biol.* **4**, e45. doi:10.1371/journal.pbio.0040045

Miskolczi-Mccallum, C. M., Scavetta, R. J., Svendsen, P. C., Soanes, K. H. and Brook, W. J. (2005). The Drosophila melanogaster T-box genes midline and H15 are conserved regulators of heart development. *Dev. Biol.* **278**, 459-472. doi:10.1016/j.ydbio.2004.11.026

Neuman-Silberberg, F. S. and Schupbach, T. (1993). The Drosophila dorsoventral patterning gene gurken produces a dorsally localized RNA and encodes a TGF alpha-like protein. *Cell* **75**, 165-174. doi:10.1016/S0092-8674(05)80093-5

Perry, M. W., Cande, J. D., Boettiger, A. N. and Levine, M. (2009). Evolution of insect dorsoventral patterning mechanisms. *Cold Spring Harb. Symp. Quant. Biol.* **74**, 275-279. doi:10.1101/sqb.2009.74.021

Perry, M. W., Boettiger, A. N., Bothma, J. P. and Levine, M. (2010). Shadow enhancers foster robustness of Drosophila gastrulation. *Curr. Biol.* **20**, 1562-1567. doi:10.1016/j.cub.2010.07.043

Perry, M. W., Boettiger, A. N. and Levine, M. (2011). Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. *Proc. Natl. Acad. Sci. USA* **108**, 13570-13575. doi:10.1073/pnas.1109873108

Pfeiffer, B. D., Jenett, A., Hammonds, A. S., Ngo, T. T., 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. doi:10.1073/pnas.0803697105

Pfeiffer, B. D., Ngo, T. T., Hibbard, K. L., Murphy, C., Jenett, A., Truman, J. W. and Rubin, G. M. (2010). Refinement of tools for targeted gene expression in Drosophila. *Genetics* **186**, 735-755. doi:10.1534/genetics.110.119917

Pyrowolakis, G., Veikkolainen, V., Yakoby, N. and Shvartsman, S. Y. (2017). Gene regulation during Drosophila eggshell patterning. *Proc. Natl. Acad. Sci. USA* **114**, 5808-5813. doi:10.1073/pnas.1610619114

Qian, L., Liu, J. and Bodmer, R. (2005). Neuromancer Tbx20-related genes (H15/midline) promote cell fate specification and morphogenesis of the Drosophila heart. *Dev. Biol.* **279**, 509-524. doi:10.1016/j.ydbio.2005.01.013

Ray, R. P. and Schupbach, T. (1996). Intercellular signaling and the polarization of body axes during Drosophila oogenesis. *Genes Dev.* **10**, 1711-1723. doi:10.1101/gad.10.14.1711

Revaitis, N. T., Marmion, R. A., Farhat, M., Ekiz, V., Wang, W. and Yakoby, N. (2017). Simple expression domains are regulated by discrete CRMs during drosophila oogenesis. *G3 (Bethesda)* **7**, 2705-2718. doi:10.1534/g3.117.043810

Sapir, A., Schweitzer, R. and Shilo, B. Z. (1998). Sequential activation of the EGF receptor pathway during Drosophila oogenesis establishes the dorsoventral axis. *Development* **125**, 191-200. doi:10.1242/dev.125.2.191

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B. et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676-682. doi:10.1038/nmeth.2019

Schoenfelder, S. and Fraser, P. (2019). Long-range enhancer-promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437-455. doi:10.1038/s41576-019-0128-0

Sebe-Pedros, A., Ariza-Cosano, A., Weirauch, M. T., Leininger, S., Yang, A., Torruella, G., Adamski, M., Adamska, M., Hughes, T. R., Gomez-Skarmeta, J. L. et al. (2013). Early evolution of the T-box transcription factor family. *Proc. Natl. Acad. Sci. USA* **110**, 16050-16055. doi:10.1073/pnas.1309748110

Setty, Y., Mayo, A. E., Surette, M. G. and Alon, U. (2003). Detailed map of a cis-regulatory input function. *Proc. Natl. Acad. Sci. USA* **100**, 7702-7707. doi:10.1073/pnas.1230759100

Shlyueva, D., Stampfel, G. and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* **15**, 272-286. doi:10.1038/nrg3682

Stevens, C. A., Revaitis, N. T., Caur, R. and Yakoby, N. (2020). The ETS-transcription factor Pointed is sufficient to regulate the posterior fate of the follicular epithelium. *Development* **147**, dev189787. doi:10.1242/dev.189787

Svendsen, P. C., Formaz-Preston, A., Leal, S. M. and Brook, W. J. (2009). The Tbx20 homologs midline and H15 specify ventral fate in the Drosophila melanogaster leg. *Development* **136**, 2689-2693. doi:10.1242/dev.037911

Svendsen, P. C., Phillips, L. A., Deshwar, A. R., Ryu, J. R., Najand, N. and Brook, W. J. (2019). The selector genes midline and H15 control ventral leg pattern by both inhibiting Dpp signaling and specifying ventral fate. *Dev. Biol.* **455**, 19-31. doi:10.1016/j.ydbio.2019.05.012

Svendsen, P. C., Ryu, J. R. and Brook, W. J. (2015). The expression of the T-box selector gene midline in the leg imaginal disc is controlled by both transcriptional regulation and cell lineage. *Biol. Open* **4**, 1707-1714. doi:10.1242/bio.013565

Tzolovsky, G., Deng, W. M., Schlitt, T. and Bownes, M. (1999). The function of the broad-complex during Drosophila melanogaster oogenesis. *Genetics* **153**, 1371-1383. doi:10.1093/genetics/153.3.1371

Visel, A., Rubin, E. M. and Pennacchio, L. A. (2009). Genomic views of distant-acting enhancers. *Nature* **461**, 199-205. doi:10.1038/nature08451

Ward, E. J. and Berg, C. A. (2005). Juxtaposition between two cell types is necessary for dorsal appendage tube formation. *Mech. Dev.* **122**, 241-255. doi:10.1016/j.mod.2004.10.006

Yakoby, N., Bristow, C. A., Gong, D., Schafer, X., Lembong, J., Zartman, J. J., Halfon, M. S., Schupbach, T. and Shvartsman, S. Y. (2008a). A combinatorial code for pattern formation in Drosophila oogenesis. *Dev. Cell* **15**, 725-737. doi:10.1016/j.devcel.2008.09.008

Yakoby, N., Lembong, J., Schupbach, T. and Shvartsman, S. Y. (2008b). Drosophila eggshell is patterned by sequential action of feedforward and feedback loops. *Development* **135**, 343-351. doi:10.1242/dev.008920