1 **Title:**

2

3 **The role of insulators in transgene transvection in *Drosophila***

4

5 Pawel Piwko^{1,2}, Ilektra Vitsaki^{1,2}, Ioannis Livadaras¹ and Christos Delidakis^{1,2}

6

7 1: Institute of Molecular Biology & Biotechnology, Foundation for Research &

8 Technology Hellas. Heraklion, Crete, Greece 70013.

9 2: Department of Biology, University of Crete. Heraklion, Crete, Greece 70013.

10

Second Page

short running title : Role of insulators in transvection

five key words or phrases: *Drosophila*, insulators, transvection, *E(spl)* locus, homolog pairing

the corresponding author's name, office mailing address including street name and number, phone number, and email address:

Christos Delidakis

IMBB, FORTH

100 N. Plastira St

Heraklion 70013

Crete, GREECE

tel: +30-2810-391112

email: delidaki@imbb.forth.gr

Abstract

Transvection is the phenomenon where a transcriptional enhancer activates a promoter located on the homologous chromosome. It has been amply documented in *Drosophila* where homologues are closely paired in most, if not all, somatic nuclei, but it has been known to rarely occur in mammals as well. We have taken advantage of site-directed transgenesis to insert reporter constructs into the same genetic locus in *Drosophila* and have evaluated their ability to engage in transvection by testing many heterozygous combinations. We find that transvection requires the presence of an insulator element on both homologues. Homotypic *trans*-interactions between four different insulators can support transvection: the *gypsy insulator* (*GI*), *Wari*, *Fab-8* and *1A2*; *GI* and *Fab-8* are more effective than *Wari* or *1A2*. We show that in the presence of insulators, transvection displays the characteristics that have been previously described: it requires homologue pairing, but can happen at any of several loci in the genome; a solitary enhancer confronted with an enhancerless reporter is sufficient to drive transcription; it is weaker than the action of the same enhancer-promoter pair in *cis* and it is further suppressed by *cis*-promoter competition. Though necessary, the presence of homotypic insulators is not sufficient for transvection; their position, number and orientation matters. A single *GI* adjacent to both enhancer and promoter is the optimal configuration. The addition of a heterologous insulator in one homolog can positively or negatively influence transvection strength. The local landscape of enhancers and promoters is also important, indicative of complex insulator-enhancer-promoter interactions.

52 Introduction

53 Transgenesis and reporter genes are crucial tools for studying gene regulation.
 54 However, a transgene is often susceptible to interactions with surrounding
 55 chromatin leading to significant variations in expression patterns and levels,
 56 depending on the site of insertion in the genome (Levis *et al.* 1985). In order to
 57 factor-out these "position effects" multiple lines of transgenic animals must be
 58 studied. More recently, this problem has been tackled by two strategies now widely
 59 used in *Drosophila*. One relies on the ability to guide integration of a transgene to a
 60 specific location in the genome via the use of the Φ C31-mediated integration system
 61 (Thorpe and Smith 1998; Groth *et al.* 2004; Pfeiffer *et al.* 2008; Kvon 2015; Markstein
 62 *et al.* 2008). By targeting all reporters to the same "landing site", they are directly
 63 comparable and only a single transgenic line per construct needs to be analysed. The
 64 second way to minimize genomic position effects utilizes insulator sequences in
 65 transgenesis vectors. Insulator DNA elements were first identified by their potential
 66 to block gene function when interposed between enhancer and promoter (Kellum
 67 and Schedl 1992; Cai and Levine 1995; Kuhn *et al.* 2003) and to insulate transgenes
 68 from effects of surrounding chromatin (Kellum and Schedl 1991; Roseman *et al.*
 69 1993, 1995). These faculties of insulators are linked to their strong propensity for
 70 interaction with each other (Kuhn *et al.* 2003; Chetverina *et al.* 2008; Kyrchanova *et*
 71 *al.* 2008a, 2011). Such interactions are proposed to form chromatin loops bridging
 72 even distant loci at the base of the loop and isolating interactions inside the loop
 73 from interactions outside the loop (Byrd and Corces 2003; Blanton *et al.* 2003; Doyle
 74 *et al.* 2014; Cubeñas-Potts and Corces 2015).

75 We started using both of these approaches to characterize enhancer modules of two
76 neighboring *Drosophila* genes *E(spl)m7* and *E(spl)m8*. To that end we generated a
77 series of reporter constructs flanked by two copies of the insulator sequence from
78 the *gypsy* transposon (*gypsy* insulator, GI) and integrated each construct into the
79 same *attP* locus. When we tested two different reporters in a heterozygous
80 configuration, we were surprised to find they were able to affect each other's
81 expression in *trans*. The ability of enhancers to activate promoters in the
82 homologous locus is called transvection and was first reported in *Drosophila*, as a
83 phenomenon of pairing-dependent intragenic (unexpected) complementation in loci
84 like *Ultrabithorax (Ubx)* (Lewis 1954; Martínez-Laborda *et al.* 1992), *decapentaplegic*
85 (*dpp*) (Gelbart 1982), *yellow* (Geyer *et al.* 1990) and *white* (Babu and Bhat 1980).
86 Subsequent studies employing randomly integrated P-element transgenes showed
87 that the *Drosophila* genome is generally permissive to enhancer action in *trans* (Chen
88 *et al.* 2002; Kravchenko *et al.* 2005) offering a first glimpse of the molecular basis of
89 this phenomenon and reconfirming the need for somatic homologue pairing (or
90 synapsis). Homologue synapsis is not limited to germline meiotic cells in *Drosophila*
91 (and dipterans in general), but is very common in somatic tissues. This is in contrast
92 to mammals, where somatic homologue synapsis seems to happen only in special
93 occasions and accordingly only sporadic cases of transvection have been reported
94 (McKee 2004; Heride *et al.* 2010; Apte and Meller 2012; Stratigi *et al.* 2015; Joyce *et*
95 *al.* 2016). The development of site-specific integration methods in *Drosophila* (such
96 as the Φ C31-based recombination method) has recently rekindled the interest in
97 transvection (Lee and Wu 2006; Bateman *et al.* 2012a; Mellert and Truman 2012;

Fujioka *et al.* 2016). However, the mechanism of transvection is still poorly understood.

Using our series of reporter transgenes, we decided to search for sequence determinants of transvection. We found that this interaction is dependent on the presence of homotypic insulator DNA elements on both homologues. Our transgenesis vectors contained two such insulators: the Gypsy Insulator (GI) commonly utilized to protect transgenes from genomic position effects, and the Wari Insulator (WI) carried in the 3' part of the mini-*white* marker gene. Two other insulators were also found to support transvection, while insulator removal from either of the *trans*-interacting transgenes abrogated transvection at five discrete genomic loci. While necessary, the presence of insulators was not sufficient to produce a robust transvection outcome. Parameters like the number, position and orientation of GIs relative to the *trans*-interacting enhancers and promoters proved to be of paramount importance. The implications of these results on the design of transgenes and on the broader role of insulators in transcription will be discussed.

Materials and Methods

Plasmid constructs

See attached **Supplementary Materials and Methods** file.

Fly maintenance and stocks

Flies were maintained under standard conditions at 25°C. Stocks containing attP docking sites used for the integration of *attB* plasmids: *attP40* (RRID:BDSC_25709),

120 *attP2* (RRID:BDSC_25710) (Groth *et al.* 2004; Markstein *et al.* 2008), VK2
 121 (RRID:BDSC_9723), VK13 (RRID:BDSC_24864), VK37 (RRID:BDSC_24872), VK40
 122 (RRID:BDSC_35568) (Venken *et al.* 2006); each carrying, and if not - crossed to, a
 123 chromosome expressing $\Phi C31$ integrase under the control of *nanos* derived from
 124 *attP40* stock (RRID:BDSC_25709) (Bischof *et al.* 2007). The *su(Hw)* mutant effects
 125 were assayed in the animals transheterozygous for *su(Hw)*^{e04061} null allele
 126 (RRID:BDSC_18224) (Thibault *et al.* 2004) and *su(Hw)*² strong hypomorphic allele
 127 resulting in the ten times decreased *su(Hw)* expression (RRID:BDSC_983) (Parkhurst
 128 *et al.* 1988; Georgiev *et al.* 1997).

129 ***Integration of attB plasmids into attP fly lines***

130 All plasmids in this study were integrated into a unique attP landing site specified in
 131 the text and figure legends for each transgene. Microinjection was performed as
 132 previously described (Ringrose 2009). A solution of 500 ng/μl plasmid DNA was
 133 microinjected into *nanos-ΦC31; attP* fly embryos. Flies that grew to adulthood were
 134 crossed with *y w* flies. Depending on the injected DNA construct, the mini-*white* or
 135 *3xP3-dsRed* marker was used for subsequent screening and tracking the transgene.

136 ***Immunostaining and microscopy***

137 Fixation and immunohistochemistry of embryos and larval tissues was performed
 138 according to standard protocols. CNSs and imaginal disks were dissected from late
 139 third instar larvae, fixed with 4% paraformaldehyde and labeled with rabbit
 140 polyclonal anti-GFP (Minotech Biotechnology) and mouse anti-β-galactosidase
 141 (Promega Cat# Z3781, RRID:AB_430877) primary antibodies. Goat anti-rabbit IgG
 142 secondary antibody, Alexa488-conjugated (Thermo Fisher Scientific Cat# A-11034,

143 RRID:[AB 2576217](#)) was used for GFP detection. Goat anti-mouse, Alexa633-
144 conjugated (Thermo Fisher Scientific Cat# A-21052, RRID:[AB 2535719](#)) or donkey
145 anti-mouse, Alexa647-conjugated (Thermo Fisher Scientific Cat# A-31571,
146 RRID:[AB 162542](#)) secondary antibodies were used for β -galactosidase detection.
147 Samples were imaged on a Leica SP8 confocal platform using a 20x Oil Immersion
148 Objective with fixed zoom levels for each tissue type (CNS, wing and eye disks). The
149 images were pseudo-colored in green (GFP), red (LacZ) and blue (DsRed). All samples
150 within each figure were fixed and immunostained at the same time. Scanning of all
151 figure samples was performed using identical microscope and software settings and,
152 when possible, completed within one imaging session to enable semi-quantitative
153 comparison. Where scanning of all figure samples within one session was not
154 possible, replicates of the samples from two chosen genotypes were re-scanned
155 together with the remaining samples in the next scanning session ensuring that the
156 replicated samples are comparable. Images were manipulated using ImageJ
157 (pseudocoloring, rotation and maximum intensity projection z-stacks) and arranged
158 into data sets using Adobe Photoshop CC 2017 and Microsoft PowerPoint. Note that
159 we used two different z-projections for some wing disk images. For the top part,
160 containing the wing pouch, a full z-projection of the sample was done, while the
161 bottom part, containing the notum and hinge, encompassed only sections containing
162 the adult muscle precursors (AMPs). This was done to avoid confusing AMP
163 expression with expression in the overlying tegula, a sensory organ primordium.
164 Whereas AMPs and tegula can be easily distinguished in the 3D confocal stacks, they
165 merge to one cluster upon z-projection. Since enhancer e7 activity is specific for the

AMPs and not the tegula, we excluded the tegula sections from the z-projections shown.

Luciferase assays

Luciferase activity was measured using the Promega Luciferase Assay System Kit (Cat# E153A). CNSs and imaginal disks from ten late third instar larvae were collected in 200 µl of 1X lysis reagent CCLR for each sample. Samples were collected over a series of days and stored at -80°C until five independent samples were collected for each genotype. Samples were defrosted, put on ice and homogenized using Kontes pestles. Homogenized samples were incubated at room temperature for 10 minutes and then centrifuged for 5 minutes to pellet tissue remains. Obtained homogenates were subsequently measured for luciferase activity and total protein content. 20 µl of each homogenized sample was mixed with 50 µl of Promega Luciferase Assay Reagent and promptly measured on single tube luminometer (Turner Designs TD-20/20). Total protein was measured using the Pierce BCA Protein Assay Kit (Cat# 23225). 10 µl of each homogenized sample was mixed with 200 µl BCA Working Reagent on clear-bottomed 96-well plates (Costar) and incubated at 37°C for 1 hr. The plates were transferred at room temperature for 10 min to allow the reactions to stabilize. Absorbance was measured on the Awareness Technology ChroMate Microplate Reader at 562 nm. Three replica plates were averaged for each sample. A standard curve was produced with BSA dilutions in Promega 1X lysis reagent CCLR. Luciferase activity from each sample was normalized to its total protein content.

Data Availability

All plasmids and fly strains are available upon request. The authors affirm that all data necessary for confirming the conclusions of the article are present within the article, figures, and tables.

Results

Enhancer element analysis of $E(spl)m7$ and $E(spl)m8$ reveals a transvection phenomenon

This work began with the aim to dissect the transcriptional regulation of two adjacent genes in the $E(spl)$ complex, $E(spl)m7$ and $E(spl)m8$, both encoding transcription factors involved in many aspects of the response to Notch signaling. These are small, intronless genes and are thought to be regulated by their proximal flanking sequences (Delidakis *et al.* 2014). We cloned the 7 kb genomic fragment encompassing these two genes and tagged them with EGFP independently in two otherwise identical genomic constructs: $GFPm7-m8$ (EGFP fused to the open reading frame (ORF) of $E(spl)m7$, **Figure 1 A**) and $m7-GFPm8$ (EGFP fused to the $E(spl)m8$ ORF, **Figure 1 B**) (see Supplementary Materials for further features of these transgenes). Consistent with the known *in situ* hybridization expression patterns (de Celis *et al.*, 1996), $GFP-E(spl)m7$ and $GFP-E(spl)m8$ displayed the same pattern in wing imaginal disks from third instar larvae: (1) in the region of wing margin (WM), and (2) in the adult muscle precursors (AMPs, or adepithelial cells) of the thorax, among other cells (**Figure 1 A, B**). Likewise, both constructs expressed GFP similarly in eye-antennal imaginal disks, whereas their central nervous system (CNS) patterns were different, especially apparent in the Ventral Nerve Cord (VNC) where $GFPm7$

was expressed strongly in the midline, while *GFPm8* was mainly expressed in the neuroblasts.

We went on to characterize the patterns produced from individual enhancers located immediately upstream of *E(spl)m7* and *E(spl)m8* by generating shorter genomic constructs, *GFPm7* and *GFPm8* (**Figure 1 C** and **D**, respectively). *GFPm7* contains *ep7* – the 2.1 kb sequence upstream of *E(spl)m7* containing its putative enhancer, *e7*, and promoter, *p7*. *GFPm8* contains *ep8* – the 1.3 kb 5' flanking *E(spl)m8* sequence, containing its putative enhancer, *e8*, and promoter, *p8*. *GFPm7* and *GFPm8* recapitulated the expression patterns seen for these genes in the longer *m7-m8* transgenes, with two notable exceptions in the wing disk: *GFPm7* lacked the wing margin (WM) (**Figure 1 C**) and *GFPm8* lacked the muscle precursors (AMPs) (**Figure 1 D**). Additional less conspicuous differences in the bristle proneural clusters and the antennal primordium were noted, but we did not consider these any further.

We conclude that these *E(spl)* genes contain two upstream enhancers, *e7* and *e8*, which drive distinct expression patterns in the CNS and in the wing disk and similar patterns in the eye disk. In the context of a large genomic fragment encompassing both genes, *e7* and *e8* are shared between promoters of both genes in the WM and AMPs (both genes expressed), but they act exclusively on their downstream gene in the VNC midline (only *m7*) and the neuroblasts (only *m8*).

We also created an *ep7-lacZ* construct to permit simultaneous detection of *ep7*-driven expression of β -galactosidase with *ep8*-driven GFP from *GFPm8* genomic constructs. Both constructs were inserted into the same chromosomal locus via the Φ C31 integration method (Markstein *et al.* 2008; Pfeiffer *et al.* 2010). *ep7-lacZ* faithfully recapitulated the expression pattern of *GFPm7*. When we made

heterozygous animals containing both transgenes (*ep7-lacZ* and *ep8-GFPm8*), we were surprised to find that *ep7-lacZ* displayed novel expression in the wing margin (characteristic of *e8*), while *ep8-GFP* was expressed in the AMPs (characteristic of *e7*) (**Figure 1 G**, compare to **Figure 1 E** and **F**). This effect was observed when both transgenes were inserted into the same attP landing site, either *attP40* (chromosome 2) or *attP2* (chromosome 3) (**Figure 1 G** and **H**). No such inter-transgene interaction was observed when one transgene was inserted in *attP40* and the other in *attP2* (**Figure 1 I** and **J**), suggesting that homologue pairing is required for this interaction; in other words, we are observing a transvection phenomenon.

Transvection is mediated by homotypic interactions between GIs or Wari insulators

Having encountered robust bi-directional gene activation in *trans* (*ep8* activated by *e7* in the AMPs and *ep7* activated by *e8* at the WM), we decided to dissect the sequences that mediate this phenomenon. Firstly, we asked whether the observed transvection is a general phenomenon requiring only an enhancer paired to a promoter, or whether it depends on the inserted transgene sequences. The constructs in **Figure 1** were based on the *pPelican* vector (Barolo *et al.* 2000), which contains a *pUC8* backbone, the mini-*white* gene, two *gypsy* insulator (GI) sequences flanking a multiple cloning site (MCS), as well as a $\Phi C31$ attB site inserted by us. In order to systematically address the role of vector sequences we re-cloned the *ep7-lacZ* reporter construct into a minimal *pBluescript* backbone to which we added a $\Phi C31$ attB site to enable fly transgenesis and either the mini-*white* gene or a *3xP3-dsRed* marker which expresses DsRed in the adult eye in response to an artificial

258 Pax6 (Toy/Eyeless) responsive enhancer (Berghammer *et al.* 1999; Horn *et al.* 2000).

259 All constructs for this analysis were introduced into the *attP40* locus.

260 Both *pBluescript*-based *ep7-lacZ* constructs (*mini-white* and *3xP3-dsRed*) expressed

261 LacZ in the same pattern as the original *pPelican*-based reporter. However, when

262 they were tested in *trans* to the original *ep8-GFPm8*, no transvection was observed,

263 namely we detected neither LacZ in the WM nor GFP in the AMPs (**Figure 2 A-C**). We

264 concluded that transvection at *attP40* does not happen whenever an enhancer-

265 promoter pair is placed in *trans*, but needs additional sequences from the vectors.

266 Upon flanking the *ep7-lacZ-3xP3-dsRed* reporter with two GIs in the same

267 orientation as in *pPelican*, bidirectional transvection (*e8→p7*, blue dotted rectangle

268 and *e7→p8*, red dotted rectangle in **Figure 2 D**) was restored implicating the paired

269 GIs as mediators of the effect. Because the GI is known to be bound by the Su(Hw)

270 zinc-finger protein, which mediates its insulator activity (Parkhurst *et al.* 1988; Spana

271 *et al.* 1988; Spana and Corces 1990; Holdridge and Dorsett 1991; Geyer and Corces

272 1992; Gerasimova *et al.* 1995; Ramos *et al.* 2006), we tested our transgene pairs in a

273 *su(Hw)* mutant background. Indeed, *e7→p8* (AMP) transvection was lost in this

274 background, but, surprisingly, *e8→p7* (WM) transvection was now apparent in all

275 transgene combinations that had a *mini-white* marker in both homologues (**Figure 2**

276 **A', B' and E'**), even the one that had not displayed this effect in wild type (wt)

277 genetic background (**Figure 2 B vs B'**). We then modified the *pPelican*-based *ep8-*

278 *GFPm8* construct by removing its GIs. This construct was capable of supporting

279 *e8→p7* transvection with the *mini-white* but not the *3xP3-dsRed* version of the

280 *pBluescript-ep7-lacZ* reporter in both wt and *su(Hw)* genetic backgrounds (**Figure 2**

E/E' - F/F'). The simplest conclusion from these results is that *e8*→*p7* transvection in the WM occurs when mini-*white* is present on both homologues, but this effect is annulled when GIs are placed nearby. Only upon removal of GIs from both homologues, or their inactivation by *su(Hw)* loss, is this effect observed. On the other hand, the presence of GIs in both homologues can sustain transvection in both directions, not only *e8*→*p7* but also *e7*→*p8*.

We mapped the transvection-mediating element of the *white* locus within its 3' part: out of 3 subfragments derived from mini-*white* in **Figure S1 A-C**, only its 3'most 0.9 kb recapitulated *white-white*-mediated transvection (**Figure S1 C**). It was previously reported that this sequence contains the 3'-UTR as well as an insulator element, dubbed Wari (hereafter referred to WI) (Chetverina *et al.* 2008). Therefore either of two different insulators, GI and WI, can promote transvection when placed in a paired configuration (in both homologues) near an enhancer-promoter pair. WI has been shown to have *su(Hw)*-independent insulator activity, but also to interact with GI in *cis* (Chetverina *et al.* 2008). When we confronted a WI-containing *ep7-lacZ* with a WI-containing (mini-*white*) *ep8-GFPm8* flanked by GIs (**Figure S1 D**), transvection was abolished, consistent with what we had observed earlier with the entire mini-*white* (**Figure 2B**). We hypothesized that this inhibition could result from presumptive insulation imposed by the GI located between *e8* and WI in *ep8-GFPm8*, thereby restricting access of WI to *e8*. This was not the case, as the inhibition of WI-mediated transvection was sustained even when we deleted the 3' GI, leaving only the 5' GI intact (**Figure S1 E**). Thus, a heterotypic GI-WI interaction in *cis* can disable the homotypic WI-WI-mediated transvection but not the GI-GI-mediated

transvection. However, we should emphasize that this inhibition was context-dependent and quite an opposite action of GI, namely enhancement of WI-WI-mediated transvection, was also possible in a different context (described later in this manuscript – see **Figure 8**).

GI-mediated transvection is promiscuous, whereas WI-mediated transvection is highly selective

Because *ep7* and *ep8* interact in *cis* in their native context (**Figure 1A, B**), it raises the possibility that the transvection we observed is tied to the specificity of these two regulatory modules for each other. Thus, we sought to determine whether GIs and WIs could mediate transvection of *ep7* and *ep8* to an unrelated, heterologous promoter. To address this question, we generated an enhancerless construct containing a basal promoter commonly used in assaying enhancer activity, derived from the *Drosophila hsp70* gene (hereafter referred to as *pH*). As expected, a transgene carrying this minimal *pH* promoter fused to GFP, flanked by GIs and marked with mini-*white* displayed, on its own, no expression in wing disks or CNS (**Figure 3 A**). When crossed to an *ep7-lacZ* construct, congruent in two flanking GIs and mini-*white*, GFP was detected with the *e7*-specific pattern (**Figure 3 B** compare GFP (*trans*) to LacZ (*cis*)). This activation of *pH* in *trans* relies on homotypic interaction between GIs, as disparity between transgenesis markers (but retention of congruent GIs) did not affect transvection of *e7* (**Figure 3 C**), while removal of GIs from one of the transgenes (with congruent mini-*white* markers) abolished the effect (**Figure 3 D**). Unlike the *ep8*→*ep7* transvection, removing GIs entirely, but keeping

congruent mini-*white* markers did not support transvection (**Figure 3 H**). Therefore, paired WIs cannot support transvection between *e7* and *pH*.

Interestingly, in this set of experiments, GIs mediated *trans*-activation of the *pH* promoter not only by the *e7* enhancer located in between the two GIs, but also by an enhancer exterior to the two GIs: the *3xP3*, which displays strong expression in a subset of glia in the CNS (**Figure 3C**). Depletion of Su(Hw) suppressed transvection of both enhancers, *e7* (**Figures 3 E, F**) and *3xP3* (CNS in **Figure 3 F**), and allowed *pH* to trap a tracheal enhancer (in *cis*) in the vicinity of *attP40* (white arrows in **Figures 3 E, F**). The fact that two unrelated enhancers, *e7* and *3xP3*, can transvect to a heterologous promoter, *pH*, suggested that GI-mediated transvection is unselective for enhancer-promoter pairs. This prompted us to use the GIs-containing transgene system to screen for putative enhancer elements across the 50 kb long *E(spl)* locus. A collection of 18 fragments of this locus inserted in a transgene between two GIs activated specific patterns of expression in *trans* from the GI-flanked *pH-gfp* transgene (**Figure S2**) – a full 15 out of these 18 fragments displayed robust enhancer activity in third instar larval disk/CNS tissue. Additionally, using this system, we were able to recapitulate in *trans* the *cis*-pattern of another enhancer unrelated to the *E(spl)* locus, the *vestigial* quadrant enhancer (last column, **Figure S2**).

Unlike GI-, WI-mediated transvection was specific for the *ep8/ep7* combination and mediated *trans*-activation of *p7* by *e8* unidirectionally. To test the possibility that WI-mediated transvection is specific for the *E(spl)m7* promoter (*p7*), we tested various minimal GFP reporters driven by different basal promoters, *pH*, *p7* and *E(spl)m8* promoter (*p8*). We confronted these enhancerless non-expressing reporters with an

350 *ep8-lacZ* transgene. We made sure that all combinations were congruent for both
351 GIs and mini-*white*, which enabled us to simultaneously test GI-mediated
352 transvection in a wt background and WI-mediated transvection in the absence of
353 Su(Hw) (**Figure S3**). All three basal reporters responded to *ep8* enhancer in a
354 *su(Hw)*^{+/−} background (GI-dependent transvection); the *p7* was by far the strongest
355 responder, with *pH* following and *p8* showing a very weak activation (**Figure S3 A-C**).
356 However, in the *su(Hw)*^{−/−} background (WI-dependent transvection), the *e8* enhancer
357 did not transfect to any of the three promoters (**Figure S3 M-O**, compare to **Figure**
358 **S3 G-I**). Therefore, the unidirectional *ep8*→*ep7* transvection supported by WIs was
359 not due to a selectivity of WI for *p7*. When the three basal promoters were fused to
360 the *e7* enhancer (**Figure S3 D-F**), these reporters started expressing (as hemizygotes)
361 in the AMPs and some proneural clusters, but not in the WM, consistent with the
362 activity of the *e7* enhancer (**Figure S3 D-F**). Once confronted in *trans* with *ep8-lacZ*,
363 these *e7*-bearing reporters were able to robustly express GFP in the WM (**Figure S3 J-**
364 **L**) and this expression was retained in the *su(Hw)* mutant genetic background
365 (**Figures S3 P-R**). Thus, the transvection mediated by the interaction between WIs
366 was not promoter context specific, but rather enhancer context specific, with the
367 responding gene requiring the presence of the *e7* enhancer in *cis* in order to sustain
368 WI-mediated transvection. In conclusion, whereas *trans*-paired GIs mediated
369 transvection between any enhancer-promoter pair tested, *trans*-paired WIs were
370 more selective and mediated only *e8*→*e7* transvection, an effect that was
371 dominantly suppressed by the presence of GI elements. This *e8*→*e7* transvection
372 may reflect some intrinsic affinity of these two enhancers for each other, but it still
373 requires the presence of GIs or WIs in order to materialize.

374

375 ***Transvection is weaker than cis enhancer-promoter activity and is suppressed by***
 376 ***promoter cis-preference***

377 To gauge the relative strength of transvection compared to *cis* enhancer-promoter
 378 (e-p) interaction we generated e-p pairs driving GFP expression in *cis* and compared
 379 them to the same e-p pairs driven in *trans*. All constructs designed for this purpose
 380 were based on the backbone of *pPelican/pStinger* vectors (Barolo *et al.* 2000) which
 381 contain two GIs to enable GI-mediated transvection, and subsequently inserted into
 382 the *attP40* locus. These flanking GIs also provided efficient insulation: all
 383 enhancerless promoter-reporter constructs had undetectable levels of expression in
 384 all three larval tissues tested (wing disk, eye disk and CNS; data not shown). In all
 385 cases we observed lower GFP levels from transvection than from the *cis*-combination
 386 (**Figure 4**). We tested three promoters, *p7*, *p8* and *pH*. Regardless of the enhancer
 387 assayed (*e7* or *e8*) the strongest expression levels, both in *cis* and in *trans*, were
 388 produced by *p7*, whereas *p8* was the weakest out of the three promoters. This
 389 suggests that a promoter's strength for driving transcription is its intrinsic property
 390 and does not depend on the enhancer activating it, at least for the two enhancers
 391 tested.

392 We also studied the embryonic *cis* versus *trans* expression of *e7* using the *pH*
 393 promoter constructs. In *cis*, *e7* displayed very dynamic embryonic expression,
 394 starting at the mesectoderm in stage 7, then in a neuroectodermal cluster pattern up
 395 to stage 10 and later in the VNC midline and the epidermis in a complex segmentally
 396 repeated pattern. Interestingly, the earlier patterns could not be transvected or

were transvected only in a random subset of the cells where the enhancer was active. From stage 11 onward, this sporadic transvection gave way to a more complete one, where the *e7 trans* pattern recapitulated the *cis* pattern (**Figure S4**). This correlates nicely with what is known about somatic homologue pairing: whereas in early embryonic stages paternal and maternal homologues start out unpaired, they gradually increase their pairing and reach maximum levels by about stage 11 (Hiraoka *et al.* 1993; Fung *et al.* 1998; Gemkow *et al.* 1998). Therefore, this observation, corroborates the need for homologue pairing in order for transvection to take place.

To obtain a more quantitative measure of the *cis* versus *trans* activity of an enhancer, we used a luciferase (*luc*) reporter (instead of GFP) and measured its activity in extracts of larval disk-brain complexes. The GI-flanked *e7-pH-luc* reporter showed 5 times higher activity in *cis* than the *e7* driving an enhancerless *pH-luc* reporter in *trans* (both GI-flanked, **Figure 5 A**). This *trans* activity was still much higher (~26x) than the basal levels of the *pH-luc* reporter. Interestingly, in this assay, *pH-luc* basal levels were low, yet detectable, even though the GFP counterpart had undetectable levels of GFP in the same tissues – this probably reflects the higher sensitivity of the *luc* versus the GFP reporter (Arnone *et al.* 2004). Upon removal of the GIs, *luc* reporter activity dropped to almost undetectable levels, which is consistent with previous observations that GIs can stimulate basal transcription from some promoters (Wei and Brennan 2001; Golovnin *et al.* 2005; Markstein *et al.* 2008; Soshnev *et al.* 2008).

419 In the above experiments we noticed that confronting an enhancerless reporter in
 420 *trans* to a solitary enhancer gave more robust expression compared to all our
 421 previous experiments, where the transfecting enhancer was linked in *cis* to a
 422 promoter (**Figure 5**, compare **B1** to **B2** for *e8*, and, **B3** and **B4** to **B5** for *e7*). This came
 423 as no surprise, since numerous earlier studies on transvection have indicated that an
 424 enhancer's action in *trans* is suppressed by the presence of a promoter in *cis* (Geyer
 425 *et al.* 1990; Martínez-Laborda *et al.* 1992; Hendrickson and Sakonju 1995; Casares *et*
 426 *al.* 1997; Sipos *et al.* 1998; Morris *et al.* 1999a; b, 2004; Bateman *et al.* 2012a;
 427 Kravchuk *et al.* 2016). Different promoters of varying core element composition have
 428 been reported to display *cis*-preference (i.e. to attenuate transvection). As a general
 429 rule, mutations compromising transcriptional strength of a *cis*-promoter usually
 430 release the enhancer towards *trans* action (Morris *et al.* 1999b, 2004; Lee and Wu
 431 2006). We made two mutations on the *p7* promoter in an attempt to compromise its
 432 strength without completely inactivating it. *p7* is a multi-element promoter,
 433 containing a TATA box, an initiator (Inr) and a downstream promoter element (DPE)
 434 (Klämbt *et al.* 1989; Kutach and Kadonaga 2000). We introduced two deletions into
 435 the *ep7-lacZ* construct aiming to disrupt each of these activities; one, *ep7-ΔTATA-*
 436 *lacZ*, removed the TATA box [deletion of -41 to -22 bp relative to the transcription
 437 start site (TSS)] and another, *ep7-ΔDPE-lacZ*, removed the Inr and DPE elements
 438 (deletion of -16 to +67 bp relative to TSS). Both of these promoter mutations
 439 retained weak yet detectable transcriptional activity (**Figure 5, B5-B7, cis** column).
 440 Even though the reduction in *cis* promoter activity was comparable between Δ TATA
 441 and Δ DPE, the two had dramatically different effects on transvection of the linked *e7*
 442 enhancer (**Figure 5, B5-B7, trans** column). *p7-ΔTATA* partially relieved *e7* from *cis*-

preference inhibition, leading to much higher *trans*-activation of *pH-gfp* than that observed with *p7-ΔDPE*. These effects of the mutant *cis-p7* promoters on *e7* transvection were independent of the identity of the *trans* promoter, as *p7* and *p8* – based enhancerless reporters responded with a similar trend (**Figure S5**). When the various *ep7-lacZ* versions were confronted with the *pH-luc* reporter, we confirmed that the DPE deletion was comparable to the unmutated promoter in strongly suppressing transvection (11-15x weaker than a promoterless *e7-lacZ*), whereas the TATA deletion released *e7* from *cis*-preference giving 6-7x stronger *trans* reporter expression (compared to the unmutated *ep7* or the *ep7ΔDPE*) (**Figure 5 A**). Interestingly, in this assay the intact *ep7* (and *ep7-ΔDPE*) produced very low *trans*-activation of the *pH-luc* reporter, only 1.4-2x higher than its basal levels attained in the absence of a transfecting enhancer. We speculate that this reflects the ability of the transfecting *e7* enhancer to activate *pH-luc* in a number of cells (as visualized by the *pH-GFP* reporter) but at the same time to *repress* the basal *pH-luc* activity in the remaining cells. Unfortunately, we can only measure the resultant luciferase activity in the whole brain-disk extract with no cell-to-cell resolution, so we cannot test this scenario. Regardless, these luciferase constructs enabled us to obtain a quantitative measure of transvection strength, which ranged from 5x to almost 100x lower than the *cis* output of the same promoter-enhancer pair, depending on the presence of a *cis*-linked promoter and in particular the integrity of its TATA box.

Relative position, number and orientation of GIs determine transvection effects

In all the previous experiments, all our GI-based transgenes contained two GIs each in forward orientation (GIs^{FOR}) as in the *pPelican* and *pStinger* series vectors: the '5' GI^{FOR}, at the 5' end of each construct (following only the ϕ C31 attB integration site), and the '3' GI^{FOR}, 3' to the *lacZ* or *gfp* reporters (preceding the *3xP3* or *mini-white* marker genes, **Figure S6**). The "forward" GI orientation is the same as the one found in the original *gypsy* transposon, where GI is located shortly downstream of the 5' LTR (ref). The two transgenes (*ep7-lacZ* and *pH-gfp*) in this starting configuration of GIs (hereafter referred to as the 'dual-GIs^{FOR}' configuration), when presented to each other in *trans*, result in expression of GFP in two distinct patterns: *e7*-specific in all tested tissues (wing disk, eye disk and CNS) and *3xP3*-specific in the CNS (**Figure 3** and **Figure S6**). Thus, the *pH* located in between the two GIs^{FOR} receives input from two enhancers in *trans*: *e7* – located downstream of the 5' GI^{FOR}, and *3xP3* located downstream of the 3' GI^{FOR}. We have introduced a series of modifications in the configuration of the GIs within these two constructs in order to understand how the relative number, position and orientation of the GIs influence transvection. All resultant constructs for this analysis were introduced into the *attP40* locus.

Deletion of the 3' GI^{FOR} in the transvection-receiving construct, *pH-gfp*, while preserving the dual-GIs^{FOR} configuration in the 'sending' *ep7-lacZ* construct, caused a reduction in transvection of *e7*, with concomitant increase of *3xP3* transvection (compare **Figure 6 A** to **B**). When the 3' GI^{FOR} was deleted in the 'sending' construct, and the 'receiving' construct was kept in its initial dual GIs configuration, the transvection of *e7* seemed unaffected, while transvection of *3xP3* was nearly lost

(compare **Figure 6 C to A**). Finally, deletion of the 3' Gl^{FOR} in both constructs led to augmented GFP expression with an *e7* pattern and an almost undetectable *3xP3* pattern (compare **Figure 6 D to A**). These data demonstrate that a robust *trans*-activation of the *pH* promoter by the *e7* enhancer is mediated via an interaction between the 5' Gl^{FOR} of the two transgenes. This interaction seems to be weakened by the presence of a 3' Gl^{FOR} in either or both of the interacting transgenes. However, the 3' Gl^{FOR} in the sending construct is required for effective transvection of the *3xP3*, suggesting that the *3xP3* enhancer, like *e7*, needs an adjacent Gl in order to robustly act in *trans*.

Consistently, when the dual-enhancer construct (*ep7-lacZ-3xP3-dsRed*) contained a single Gl^{FOR}, only its downstream enhancer was transacted to a Gl^{FOR}-preceded *pH-gfp: e7* (with weak sporadic activity of *3xP3*) (**Figure 6 D,E,I**) or *3xP3* (**Figure 6 G**). Moreover, the presence of the Gl^{FOR} upstream of the *pH* promoter was essential for its activation by a *trans*-enhancer, as its deletion abolished transvection altogether, even when another Gl^{FOR} was present at the 3' position (**Figures 6 F and H**). It therefore seems that the *trans*-activity of both enhancers (*e7* and *3xP3*) obtained from the dual-Gls^{FOR} *ep7-lacZ* transgene (**Figures 6 A and B**) resulted from the interaction between the 5' Gl^{FOR} preceding *pH* with the two Gl^{FOR} of *ep7-lacZ* in *trans*, each upstream of each enhancer.

We next addressed the significance of the orientation of Gl^s in mediating transvection. Surprisingly, the presence of a reversed Gl in the 5' position (5' Gl^{REV}) in either of the two transgenes strongly reduced transvection effects, even when both transgenes contained 5' Gl^{REV} (**Figure 6 J-L** compared to **I**). Therefore, *trans*

interaction between enhancers and promoters is favored when both are located on the 3' side of GI. Weak transvection, on the other hand, can be sustained regardless of 5' GI orientation. In fact, even incongruent combinations of 5' GIs (one FOR and the other REV, **Figure 6 J, K**) displayed weak transvection, suggesting that it is not so much the congruence of the *trans*-insulator pair, but rather the absolute orientation of both GIs that is needed for robust transvection.

The fact that the preferred position of the *trans*-interacting enhancer-promoter is on the 3' side of the two GIs made us consider the possibility that placing the 5' and 3' GIs in a convergent orientation (i.e., 5' GI^{FOR} / 3' GI^{REV}) in both constructs might strengthen *trans*-interaction. However, this was not the case, as such transgenes produced equal levels of transvection to those with 5' and 3' GIs in the forward orientation (**Figure 6 M** compared to **A**) and less than the combination where 3' GIs are absent altogether (**Figure 6 D**). These results suggest that the interaction between 5' GI^{FOR}-preceded enhancer and *trans*-promoter is weakened by the presence of a second (3') GI in *cis*, irrespective of its orientation. Moreover, transgenes with a divergent configuration of GIs (5' GI^{REV} / 3' GI^{FOR}) did not improve the weak transvection observed between transgenes with a single 5' GI^{REV} (compare **Figure 6 P** to **L**), nor were incongruent 5' GI configurations improved by a 3' GI (**Figure 6 N, O**) (note, however, that *3xP3* was efficiently transvected to *pH* in **Figure 6 O** as both *3xP3* and *pH* are preceded by GI^{FOR}).

Is the ability of a single 5' GI^{FOR} to support transvection a peculiarity of the *attp40* locus or can it happen in more genomic loci? When we tested constructs with a single 5' GI^{FOR} in four more *attP* loci and we got robust transvection in all cases

(**Figure S7**). Importantly, removing the 5' GIs^{FOR} from the transgenes abolished transvection in all loci, reconfirming the need for paired homotypic insulators in both homologues. As a corollary, we conclude that, at least in the five genomic loci we tested, nearby endogenous genomic insulators were not capable of mediating transvection from the GI-less constructs, suggesting that "insulator trapping" is probably not a common phenomenon in the *Drosophila* genome. On the contrary, enhancer trapping is very common; we were able to detect some non-*e7* dependent patterns of expression of our transgenes in four out of the five loci tested (all except *VK40*; see **Figure S7**).

In summary, transvection needs homotypic insulators in both homologues, but having homotypic insulators is not sufficient. The outcome is also influenced by the *position, orientation* and *number* of these insulators. In the context of the *ep7-lacZ-3xP3* → *pH-gfp* transvection, both GIs have to be 5' of the *pH* promoter and directly adjacent to the transfecting enhancer, *e7* and/or *3xP3*. The FOR orientation is greatly favoured for both homologues; the REV orientation produces a much weaker effect. Finally adding another GI in one or both transgenes weakens the 5'GI^{FOR} mediated transvection. It should be kept in mind, however, that the transcriptional outcome of a *trans*-interacting insulator pair, is also greatly dependent on the enhancers and promoters located in the vicinity of these GIs: when the same *ep7-lacZ* series of transgenes was tested in *trans* to an *ep8-m8GFP* series (instead of the *pH-GFP*), the *ep7*→*ep8* transvection largely obeyed the above rules, but the *ep8*→*ep7* transvection was detectable even with a single 3' GI, regardless of orientation (**Figure S8**). Still, the single 5' GI^{FOR} configuration gave the strongest

556 *trans*-effect even with this transgene combination. The removal of GI from either
 557 homologue completely abolished the effect, as expected.

558 Why was the addition of a 3' GI^{FOR} detrimental to *ep7*→*pH* transvection (**Figure 6 A-**
 559 **D, M-P**)? Could it potentially engage in homotypic interactions with either the *cis* or
 560 the *trans* 5' GI that might compete with the ability of these 5' GIs to support
 561 transvection? To gain insight on the activity of this 3' GI^{FOR} we appended a "tester"
 562 module to our GI^{FOR}-*ep7-lacZ*-GI^{FOR}-3xP3-*dsRed* transgene. Two different modules
 563 were cloned immediately 3' of the 3' GI^{FOR}: (1) a 'receiving' *pH-gfp* module or (2) a
 564 'sending', *ep8-m8* module (*ep8* driving an untagged *E(spl)m8* CDS). **Figure 7** presents
 565 the results obtained in wing disks, which are consistent with those obtained in the
 566 CNS and eye-antennal disk (data not shown). The tester modules did not influence
 567 *ep7-lacZ* expression: both transgenes, on their own, expressed LacZ in the AMPs, as
 568 expected (**Figure 7 A/A'**); an *e8*-specific wing margin LacZ pattern was not detected,
 569 consistent with insulation of *ep7-lacZ* from *e8* (**Figure 7 A**); similarly, the transgene
 570 containing the insulated enhancerless *pH-gfp* module showed no GFP expression
 571 (**Figure 7 A'**), suggesting that the 3' GI in this construct insulates *ep7* from *pH*,
 572 instead of enabling their interaction. It therefore seems that the two GIs^{FOR} in these
 573 tester constructs do not productively interact in *cis* (see also Cai and Shen 2001;
 574 Kyrchanova *et al.* 2008a). When tested against each other, we observed a robust
 575 activity of the *pH-gfp* module in the WM, indicative of a *trans* interaction between
 576 the two 3' GIs^{FOR} resulting in *ep8-m8*→*pH-gfp* transvection (**Figure 7 B**). Therefore, 3'
 577 GIs prefer to homotypically interact in *trans*. The presence of only few GFP positive
 578 AMPs and no apparent WM LacZ expression in this combination (**Figure 7 B**)

579 demonstrate that the "diagonal" interactions between Gl^{FOR} (i.e., 5' Gl^{FOR}-3' Gl^{FOR})
580 are less favored than the "vertical" ones (5'-5' and 3'-3').

581 Placing the 5' Gl^{FOR} *pH-gfp* transgene in *trans* to the testers gave strong transvection
582 of *e8* (via the 3' Gl^{FOR}; **Figure 7 C**) but only weak or no transvection of *e7* (via the 5'
583 Gl^{FOR}; **Figure 7 C, C'**). Interestingly, *e7*→*pH* transvection with both testers was
584 augmented when a 3' Gl^{FOR} was added to the responding *pH-GFP* transgene, as
585 evidenced by broadly expressed GFP in the AMPs (**Figure 7 D, D'**). Therefore, when
586 confronted with dual Gl^{FOR}, a single Gl^{FOR} preferentially interacts with the *trans* 3'
587 Gl^{FOR}, but this shifts to the 5' Gl^{FOR} when a second Gl^{FOR} is added resulting in a
588 dual/dual configuration. This conclusion was supported by confronting the tester
589 transgenes with a single vs dual Gl^{FOR} configured *ep7-lacZ* reporter. Only the single 5'
590 Gl^{FOR} was able to interact in *trans* with the two tester modules, as evidenced by *ep8*-
591 *m8*→*ep7-lacZ* and *ep7-lacZ*→*pH-gfp* transvection (**Figure 7 E, E'**), whereas both of
592 these transvection effects were lost in the dual/dual configuration (**Figure 7 F, F'**).

593 Surprisingly, the *ep7-lacZ*→*pH-gfp* transvection was regained when the *p7* promoter
594 was compromised or (better) deleted in the dually Gl^{FOR} flanked *ep7-lacZ* transgene
595 (**Figure 7 G-I**). We therefore conclude that Gl^{FOR} sample different homotypic
596 interactions in *trans*, both vertical and diagonal. The latter are disfavored, but can
597 still occur and their ability to support transvection is influenced by the enhancer-
598 promoter interactions in their vicinity, e.g. the *ep7-pH* is not sufficient to produce
599 diagonal transvection (**Figure 7 F'**), unless relieved from *cis* promoter preference
600 (**Figure 7 H, I**). As another example, a similar diagonal Gl-Gl interaction can produce
601 transvection of *e8* (from the *ep8-m8* tester) to *pH-gfp* (**Figure 7 D**), but not to *ep7*-

lacZ (**Figure 7 F**). Such alternative *trans*-interactions in dual-GI combinations probably compete with 5'GI - 5'GI interactions and could underlie the suppression of transvection produced by the addition of 3' GIs in **Figure 6 A-D**.

Finally, we note that the *3xP3*→*pH* transvection in **Figure 6 A** is also mediated by a 5' – 3' (diagonal, less favored) GI^{FOR} interaction. We tested the same configuration of enhancers and insulators and changed only the responding promoter on the dual GIs^{FOR} *p-gfp* construct. Only *pH* was able to strongly respond to *3xP3*, whereas two other promoters, *p7* and *p8*, showed no or very weak response (**Figure S9**), consistent with an important role for enhancer-promoter combination to bias the transvection outcome of a less favorable GI-GI homotypic interaction.

Other insulators also mediate transvection

To determine whether other *Drosophila* insulators also mediate transvection, we generated two series of constructs based on the *pLacZattB* vector (Bischof *et al.* 2007): 'sender' constructs containing an insulator cloned in between the *e7* and *e8* enhancers, and 'responder' constructs containing an insulator between divergently oriented *pH-lacZ* and *pD-gfp*, a reporter driven by the *Drosophila Synthetic Core Promoter* (DSCP, abbreviated as *pD*) (Pfeiffer *et al.* 2008). Besides GI, we tested two new insulators: (1) the Fab8 insulator isolated from the *Abdominal-B* region of the *bithorax* complex; Fab8 insulator activity depends on dCTCF, the ortholog of the vertebrate CTCF protein (Barges *et al.* 2000; Moon *et al.* 2005; Kyrchanova *et al.* 2008b, 2016), and (2) 1A2, an insulator located downstream of the *yellow* gene, containing two Su(Hw) binding sites (Golovnin *et al.* 2003; Parnell *et al.* 2003). Fab8

and 1A2 exemplify two major classes of endogenous insulators [defined by binding of dCTCF and Su(Hw), respectively] which are abundantly represented in the *Drosophila* genome (Parnell *et al.* 2006; Adryan *et al.* 2007; Negre *et al.* 2010; Baxley *et al.* 2017). In addition, the resultant constructs contain WI carried in the mini-*white* marker gene. All transgenes were inserted into the *attP40* locus and we present the results from wing disks, which are consistent with the results obtained from the CNS and eye-antennal disks (not shown).

The responder transgene without an insulator between *pH* and *pD* promoters ("blank" responder), on its own, showed trachea-specific activity of both promoters (both LacZ and GFP, **Figure 8 A1**), similar to the expression of the uninsulated *pH-gfp* reporter at the *attP40* locus (**Figure 3 G**). Inserting GI between the two promoters insulated *pD-gfp* from the tracheal enhancer and resulted in the trapping of (an)other enhancer(s) at the *pD* promoter, ubiquitously active in all cells of the disk's epithelium (weak ubiquitous GFP expression in **Figure 8 A2**). This latter, ubiquitous, activity was abolished by a deletion of the WI from the mini-*white* (**Figure 8 A3**) and was not observed with WI on its own (**Figure 8 A1**) indicating cooperation between GI and WI in mediating the *cis* activity of this enhancer onto *pD*. The Fab8 responder, similarly to the GI responder (**Figure 8 A2**), also trapped the epithelial enhancer via the *pD* promoter (**Figure 8 A4**), while the 1A2 responder did not show any activity from any of the promoters (**Figure 8 A5**).

Heterozygotes between the "blank" sender and "blank" responder transgenes produced extremely faint but visible expression of GFP in the WM, indicative of a WI-mediated *trans* interaction between *e8* and *pD* (**Figure 8 B1, B1'**). Surprisingly, this

interaction was augmented in the GI responder (WM in **Figure 8 B2**). This enhancement of WI-WI mediated transvection by GI, was confirmed by deleting one WI, which abolished GFP expression in the WM (**Figure 8 B3**). This is in contrast to our previous result where GI-WI interaction in *cis* had an inhibitory effect on the WI-mediated transvection (**Figure 2**), suggesting that the transcriptional outcome of GI-WI *cis* interaction is context-dependent. Fab8 also showed a detectable, albeit weaker, enhancement of *e8*→*pD* transvection at the WM, whereas 1A2 had no detectable effect (**Figure 8 B4, B5**). Consistent with the conclusion that this transvection was mediated by homotypic WI/WI, when we deleted WI from the sender transgene, no transvection was observed at the WM in combination with any of the responders (**Figure 8 C1-C5**). WI-mediated transvection was also weakly enhanced by adding a heterologous insulator in the sender homologue (**Figure 8 D1-G1**) and testing against the "blank" responder. Again, only *e8* was transvected, although this time both *pH* and *pD* responded, consistent with the fact that no insulator lies between the two. In this case 1A2 was able to enhance transvection comparably to GI and Fab8 (**Figure 8G1**, compare to **D1** and **F1**). Removal of WI from the GI sender abolished transvection (**Figure 8E1**).

Unlike the weak transvection effects discussed so far with the "blank" senders (**Figure 8**, column B) or responders (**Figure 8**, row 1), where only WI was present, we got very strong transvection of both *e7* and *e8*, when we combined homotypic insulators in sender and responder, i.e., GI/GI, Fab8/Fab8 or 1A2/1A2 (**Figure 8 D2, D3, E2, E3, F4, G5**). In all cases *e7* transactivated the *pH-lacZ* reporter (in the AMPs) and *e8* transactivated *pD-gfp* (in the WM), consistent with orientation-dependent

function of all three insulators. GI produced the strongest effect and 1A2 the weakest. When the WI was deleted from either the sender or the responder GI construct, no difference in transvection efficiency was seen (**Figure 8 D2, E2, D3, E3**) thus confirming, in a different context, our earlier conclusion that WIs do not affect GI-mediated transvection.

Moderately strong WM GFP expression (*e8*→*pD* transvection,) was also seen in apparently heterotypic insulator combinations, specifically Fab8 or 1A2 senders with a GI responder (**Figure 8 F2, G2**) and the reciprocal, i.e., a GI sender with Fab8 or 1A2 responders (**Figure 8 D4, D5**). Upon deleting the WI from either the GI sender or responder, however, all of these effects were abolished (**Figure 8 F3, G3, E4, E5**), consistent with being mediated via WI/WI and enhanced by the presence of the heterologous insulators, similar to the effects noted earlier with "blank" sender/responder constructs. On the other hand, the AMP *lacZ* and WM GFP expression seen in both 1A2/GI combinations (**Figure 8 G2 and D5**) was maintained, albeit much more weakly, upon WI deletion (**Figure 8 G3/G3'** and **E5/E5'**), suggesting that this results from a *trans* heterotypic interaction between GI and 1A2, which is not surprising, since both are Su(Hw)-binding insulators.

In conclusion, we provide evidence that insulator landscape can affect enhancer-promoter communication both in *cis* (enhancer trapping) and in *trans*. All homotypic insulator combinations tested support transvection. The presence of additional heterotypic insulators in one or both homologues can augment (or in other contexts suppress) this effect. We finally provide evidence that heterotypic insulators can weakly promote transvection if they belong to the same class.

694

695 Discussion

696 Transvection is the ability of an enhancer to activate transcription from an unlinked
 697 promoter located at the same locus of the homologous chromosome. Using a
 698 collection of enhancers and promoters driving GFP and LacZ reporters and targeted
 699 to specific genomic loci, we were able to study transvection and characterize
 700 important parameters influencing its outcome. The salient features of this
 701 phenomenon borne out by our results are the following: (1) We confirmed that
 702 homologue pairing is a prerequisite for transvection, as already known from classical
 703 studies. In *Drosophila*, homologue pairing in non-meiotic cells is widespread, but
 704 seems to evolve gradually during the first half of embryogenesis (Hiraoka *et al.* 1993;
 705 Fung *et al.* 1998; Gemkow *et al.* 1998) – we showed that transvection unfolds in a
 706 similarly gradual manner, being stochastic and erratic during the early stages of
 707 embryogenesis, while recapitulating precisely the *cis*-activity of the enhancer at later
 708 embryonic and larval stages, once homologue pairing has been completed. (2)
 709 Insulators are needed for transvection. At least four different insulators, GI, WI, Fab8
 710 and 1A2, are capable of mediating transvection. We focussed on GI and WI, which
 711 are commonly found in transgenesis vectors. (3) GI-mediated transvection is strong,
 712 resulting in about 20% of the expression level the same enhancer-promoter
 713 combination would give in a *cis* configuration, and can work with all enhancer-
 714 promoter combinations tested. In contrast, WI-mediated transvection is weak and is
 715 only detectable if the responding promoter is accompanied by another enhancer
 716 (**Figure S3**) or another insulator (**Figure 8 B2**). (4) While necessary, the presence of

717 GIs in both homologues is not sufficient for transvection. (4a) The most important
718 parameters in determining the transvection outcome are the number and position of
719 GIs in both homologues – the strongest effects are seen when a single GI is present
720 in each homologue, one immediately upstream of the responding promoter and
721 another adjacent to the interacting enhancer. Additional GIs at other positions of
722 homologously inserted transgenes have a detrimental effect, probably by competing
723 against the 5'GI/5'GI enhancer-promoter enabling interaction. (4b) The presence of a
724 promoter in *cis* to the enhancer reduces its effectiveness in transvection. This *cis*-
725 preference phenomenon, which has been described before (Geyer *et al.* 1990;
726 Martínez-Laborda *et al.* 1992; Hendrickson and Sakonju 1995; Casares *et al.* 1997;
727 Sipos *et al.* 1998; Morris *et al.* 1999a; b, 2004; Bateman *et al.* 2012a; Kravchuk *et al.*
728 2016), depends on the activity of the *cis*-linked promoter: we have shown that the
729 TATA element seems to play a more important role than the DPE in inhibiting a *cis*-
730 linked enhancer from acting in *trans*. (4c) The GI is highly asymmetric with 12 Su(Hw)
731 binding sites all in the same orientation: the GI orientation with respect to the
732 enhancer-promoter interacting pair is also important, although not crucial for the
733 transvection outcome. The 3' side of the GI is the optimal for promoting
734 transvection, but the 5' side is also functional but less effectively. (5) Non-homotypic
735 insulators generally do not promote transvection, with the exception of GI/1A2, both
736 of which bind Su(Hw) and showed a weak interaction. (6) Non-homotypic insulators
737 do however *cis*-influence the transvection produced by homotypic insulators in a
738 context-dependent manner. For example GIs can enable or inhibit WI/WI-mediated
739 transvection. However, WI seems unable to affect GI/GI-mediated transvection.

740 A role of insulators in transvection has been described, but the mechanism been at
 741 best unclear (Fukaya and Levine 2017). Whereas some studies propose that
 742 insulators are needed for transvection (Lim *et al.* 2018), others propose that they
 743 only have an accessory role (Kravchenko *et al.* 2005; Schoborg *et al.* 2013), or that
 744 they affect transvection by promoting homologue pairing (Fujioka *et al.* 2016). And
 745 finally, other studies ignore them altogether (Bateman *et al.* 2012a; Mellert and
 746 Truman 2012; Blick *et al.* 2016). Instead, the classical proposed role of insulators is to
 747 insulate, i.e. to inhibit enhancer-promoter communication in *cis*, although
 748 sometimes they can enable such enhancer-promoter communication, e.g. via their
 749 so-called "bypass" activity (Cai and Shen 2001; Muravyova *et al.* 2001; Kyrchanova *et*
 750 *al.* 2008a; Fujioka *et al.* 2013). These two apparently contradictory activities have
 751 been reconciled by the "looping" model (reviewed in Chetverina *et al.* 2017;
 752 Schwartz and Cavalli 2017), which is based on evidence that insulators mediate
 753 homotypic, or sometimes heterotypic, interactions (Kyrchanova *et al.* 2008a; Li *et al.*
 754 2011; Vogelmann *et al.* 2014; Bonchuk *et al.* 2015; Fujioka *et al.* 2016). Via these
 755 interactions, insulators can form chromatin loops and these loops can either bring
 756 enhancers and promoters in proximity (e.g. when both are near the loop's anchor
 757 points) or avert their proximity (e.g. when one is within one loop and the other is
 758 outside that loop). These activities occur in *cis* and shape chromosomal architecture
 759 in parallel to affecting transcriptional regulation. The same insulator-insulator
 760 interactions can occur in *trans* (Kravchenko *et al.* 2005; Fujioka *et al.* 2016; Lim *et al.*
 761 2018) and this could mediate interactions of enhancers on one homologue with
 762 promoters on the other (transvection).

763 One model proposes that insulators promote transvection by mediating homologue
764 pairing in somatic cells. This hypothesis is supported by the fact that different classes
765 of insulators are widely distributed in the *Drosophila* genome (Bartkuhn *et al.* 2009;
766 Bushey *et al.* 2009; Negre *et al.* 2010; Schwartz *et al.* 2012) and their congruent
767 matching could underlie paternal-maternal homologue alignment from end to end.
768 Alternatively, insulators may not mediate homologue pairing *per se*, rather prior
769 homologue pairing is a prerequisite for allowing insulators to mediate transvection.
770 We favour the latter model: although we did not directly assay pairing, we have
771 numerous instances where addition of extra copies of the GI has a detrimental effect
772 on transvection (**Figures 6 and 7**). This result would be hard to reconcile with a
773 model where insulators promote pairing, as we would expect pairing, and thus
774 transvection, to locally increase as more and more insulators are added. Consistent
775 with the view that homologue pairing precedes transvection is the fact that screens
776 designed to identify somatic homologue pairing factors did not reveal any of the
777 numerous insulator binding proteins (Bateman *et al.* 2012b; Joyce *et al.* 2012).
778 Moreover, a recent study imaged two homologously inserted transgenes in live
779 embryos and found the same frequency of colocalization (pairing) whether an
780 insulator was included or not (Lim *et al.* 2018) – yet transvection between these
781 genes required an insulator on both homologues, in agreement with our results.
782 Another recent study mapped DNA elements mediating pairing and transvection
783 from the *ss* locus: the two activities were found to map on two different fragments
784 (Viets *et al.* 2018).

785 If insulators do not mediate homolog pairing, their role could be to enable the
786 productive interaction of enhancers with certain promoters. Several of our
787 observations support such a more active role: (1) In order to promote *trans*-
788 interaction between *e8* and either *pH* or *p7*, the WI requires the presence of *another*
789 *enhancer* (*e7*) nearby. (2) Transvection supported by single 5' GIs is orientation-
790 dependent. Having both homologues in the FOR orientation is much more effective
791 than both homologues bearing a GI in the REV orientation (**Figure 6**); if insulator-
792 insulator interactions were the only parameter influencing transvection, having
793 congruently disposed insulators in both homologues would most likely produce an
794 identical result, whether the configuration were FOR/FOR or REV/REV. (3) When two
795 GIS are present in each homologue, they exhibit a strong bias for "vertical" trans
796 GI/GI interactions (**Figure 7**). The fact that this bias can be alleviated by promoter
797 mutations is consistent with more direct insulator-promoter communication. (4) A
798 forward GI exhibits a strong promoter preference: it transfects the *3xP3* enhancer
799 only to *pH* and not to *p7* or *p8* in a certain transgene combination (**Figure S9**).

800 Recent data agree with such a more direct role of insulators in enhancer-promoter
801 communication. A genome-wide chromatin occupancy analysis for more than 15
802 insulator binding proteins showed that a large proportion of their binding sites is
803 near a promoter or an enhancer (Cubéñas-Potts *et al.* 2016). Direct contacts
804 between insulators and nearby enhancers and promoters has been detected in
805 transgenes *via* 3C (Kyrchanova *et al.* 2013). Live imaging of two loci separated (in *cis*)
806 by >100kb has shown that homotypic insulators promote proximity between these
807 loci, but they do it much more effectively in the presence of a promoter in the one

locus that gets activated by enhancers on the other (Chen *et al.* 2018). Finally, studies replacing specific insulator elements in the Bithorax Complex with other insulators, strongly support the ability of the resident insulators, like Fab-7 and Fab-8, to interact with neighboring enhancers (the *iab* elements) to bring them in the proximity of the *AbdB* promoter (Kyrchanova *et al.* 2016; Postika *et al.* 2018). How insulators select which enhancers to pair with which promoters is an important question that still remains to be elucidated.

Why has the necessity for insulators been overlooked in some of the studies on transvection? Most probably because the fly genome and common transgenesis cloning vectors are rich in insulators. For example, in one study (Mellert and Truman 2012) all transgenes used contained mini-*white* and its associated WI, which we have shown is capable of mediating some enhancer action in *trans*, consistently transvection was observed only with a subset of enhancers in that study. Two other studies (Bateman *et al.* 2012a; Blick *et al.* 2016) used recombinase-mediated cassette exchange system which allows for transgene integration without vector sequences, however, both of the landing sites used (JB53F and JB37B) are in close proximity (200 bp and 1 kb) to endogenous insulators bound by CTCF (Negre *et al.* 2010). In our experiments we never observed transvection in the absence of an insulator in the transgenic construct, even though we tested five different *attP* landing sites. Out of these five *attP* sites, the closest endogenous insulator to the regulatory elements present in our constructs is 6 kb away in the case of VK2, while in the case of the rest of the loci (VK13, VK37, VK40 and *attP40*) this distance is no less than 20 kb, taking into consideration the vector sequences of the integrated

831 transgenes (Negre *et al.* 2010). This suggests, not only a necessity of insulator
 832 elements for transvection, but also that they need to be closely linked to the
 833 transfecting enhancer and promoter; which agrees with the emerging genome-wide
 834 association of insulators with enhancers and promoters (Bartkuhn *et al.* 2009;
 835 Cubeñas-Potts *et al.* 2016). As a corollary, one should be careful when planning to
 836 use heterozygous transgene combinations in the same landing site. If one wishes to
 837 minimize transvection, one should preferably not include GI or any other insulator in
 838 the transgenes; at the expense of losing shielding from position effects. If shielding is
 839 desired, GIs can be used in only one of the transgenes. If shielding of both
 840 transgenes is desired, GIs can be placed in different orientations and as far as
 841 possible from the transgenes' enhancers and promoters. If, on the other hand, one
 842 wishes to promote transgene transvection, one should place GIs in the forward
 843 orientation directly upstream of the transgenes' cis-regulatory elements.

844 The association between insulators and transcriptional *cis*-regulatory elements
 845 (enhancers and promoters) is not a peculiarity of *Drosophila*; it has been reported
 846 also for vertebrates (Guo *et al.* 2015; reviewed in Hnisz *et al.* 2016). On the other
 847 hand, *Drosophila* (dipterans in general) seem to be unique in establishing somatic
 848 homologue pairing early in development and maintaining it throughout life (Abed *et*
 849 *al.* 2018; Erceg *et al.* 2018). Transvection could be an epiphenomenon of these two
 850 biological processes: insulator interactions with enhancers/promoters and
 851 homologue pairing. This would explain why it is more often encountered in
 852 *Drosophila*, but is only sporadic in mammals (Apte and Meller 2012; Stratigi *et al.*
 853 2015). Does transvection also serve a role in regulating transcriptional output and

accordingly could it be positively selected in evolution? Some studies have suggested that it increases transcription from the two alleles, or that it coordinates their transcriptional on/off decisions (Goldsborough and Kornberg 1996; Johnston *et al.* 2014). How widespread this effect is across the genome and whether it contributes to organism adaptation to fluctuations in environmental conditions or response to stressful stimuli is not known. At the least, transvection would ensure robustness of gene expression levels in the face of genetic variation, specifically heterozygosity for mutations in promoters or enhancers.

Acknowledgments

We thank Michalis Averof for *pMinos{3xP3-dsRed}*, Maria Monastirioti for *pGL3-hsp70-luc*, Chrysoula Pitsouli and Norbert Perrimon for providing *attP2* and *attP40* flies; Marina Gkantia, Giorgos Samantsidis, Alexis Molfetas, Babis Galouzis and Vasilis Ntasis for their help in generating plasmid constructs and isolating transgenic flies. Many thanks to Dr Charalampos Spilianakis for critical reading of the manuscript. This work was supported by the EU Marie Curie EST program No 7295/FAMED, ARISTEIA grants (No 4436 and 1967) from the General Secretariat for Research and Technology of Greece (co-funded by the EU European Social Fund), the Fondation Santé and intramural IMBB support.

Literature Cited

Abed J. A., J. Erceg, A. Goloborodko, S. C. Nguyen, R. B. McCole, *et al.*, 2018 Highly

876 Structured Homolog Pairing Reflects Functional Organization of the Drosophila
877 Genome. *bioRxiv* 443887. <https://doi.org/10.1101/443887>

878 Adryan B., G. Woerfel, I. Birch-Machin, S. Gao, M. Quick, *et al.*, 2007 Genomic
879 mapping of Suppressor of Hairy-wing binding sites in Drosophila. *Genome Biol*
880 8: R167. <https://doi.org/10.1186/gb-2007-8-8-r167>

881 Apte M. S., and V. H. Meller, 2012 Homologue pairing in flies and mammals: gene
882 regulation when two are involved. *Genet Res Int* 2012: 430587.
883 <https://doi.org/10.1155/2012/430587>

884 Arnone M. I., I. J. Dmochowski, and C. Gache, 2004 Using Reporter Genes to Study
885 cis-Regulatory Elements, pp. 621–652 in *Methods in cell biology*,.

886 Babu P., and S. Bhat, 1980 Effect of zeste on white complementation. *Basic Life Sci.*
887 16: 35–40.

888 Barges S., J. Mihaly, M. Galloni, K. Hagstrom, M. Müller, *et al.*, 2000 The Fab-8
889 boundary defines the distal limit of the bithorax complex iab-7 domain and
890 insulates iab-7 from initiation elements and a PRE in the adjacent iab-8 domain.
891 *Development* 127: 779–90.

892 Barolo S., L. A. Carver, and J. W. Posakony, 2000 GFP and B-galactosidase
893 transformation vectors for promoter/enhancer analysis in Drosophila.
894 *Biotechniques* 29: 726–732.

895 Bartkuhn M., T. Straub, M. Herold, M. Herrmann, C. Rathke, *et al.*, 2009 Active
896 promoters and insulators are marked by the centrosomal protein 190. *EMBO J.*

897 28: 877–88. <https://doi.org/10.1038/emboj.2009.34>

898 Bateman J. R., J. E. Johnson, and M. N. Locke, 2012a Comparing enhancer action in
899 cis and in trans. *Genetics* 191: 1143–1155.
900 <https://doi.org/10.1534/genetics.112.140954>

901 Bateman J. R., E. Larschan, R. D’Souza, L. S. Marshall, K. E. Dempsey, *et al.*, 2012b A
902 Genome-Wide Screen Identifies Genes That Affect Somatic Homolog Pairing in
903 *Drosophila*. *G3: Genes|Genomes|Genetics* 2: 731–740.
904 <https://doi.org/10.1534/g3.112.002840>

905 Baxley R. M., J. D. Bullard, M. W. Klein, A. G. Fell, J. A. Morales-Rosado, *et al.*, 2017
906 Deciphering the DNA code for the function of the *Drosophila* polydactyl zinc
907 finger protein Suppressor of Hairy-wing. *Nucleic Acids Res.* 45: 4463–4478.
908 <https://doi.org/10.1093/nar/gkx040>

909 Berghammer a J., M. Klingler, and E. a Wimmer, 1999 A universal marker for
910 transgenic insects. *Nature* 402: 370–371. <https://doi.org/10.1038/46463>

911 Bischof J., R. K. Maeda, M. Hediger, F. Karch, and K. Basler, 2007 An optimized
912 transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases.
913 *Proc. Natl. Acad. Sci. U. S. A.* 104: 3312–7.
914 <https://doi.org/10.1073/pnas.0611511104>

915 Blanton J., M. Gaszner, and P. Schedl, 2003 Protein:protein interactions and the
916 pairing of boundary elements in vivo. *Genes Dev.* 17: 664–675.
917 <https://doi.org/10.1101/gad.1052003>

- 918 Blick A. J., I. Mayer-Hirshfeld, B. R. Malibiran, M. A. Cooper, P. A. Martino, *et al.*,
919 2016 The capacity to act in trans varies among drosophila enhancers. *Genetics*
920 203: 203–218. <https://doi.org/10.1534/genetics.115.185645>
- 921 Bonchuk A., O. Maksimenko, O. Kyrchanova, T. Ivlieva, V. Mogila, *et al.*, 2015
922 Functional role of dimerization and CP190 interacting domains of CTCF protein
923 in *Drosophila melanogaster*. *BMC Biol* 13: 63. [https://doi.org/10.1186/s12915-](https://doi.org/10.1186/s12915-015-0168-7)
924 015-0168-7
- 925 Bushey A. M., E. Ramos, and V. G. Corces, 2009 Three subclasses of a *Drosophila*
926 insulator show distinct and cell type-specific genomic distributions. *Genes Dev*
927 23: 1338–1350. <https://doi.org/10.1101/gad.1798209>
- 928 Byrd K., and V. G. Corces, 2003 Visualization of chromatin domains created by the
929 gypsy insulator of *Drosophila*. *J Cell Biol* 162: 565–574.
930 <https://doi.org/10.1083/jcb.200305013>
- 931 Cai H., and M. Levine, 1995 Modulation of enhancer-promoter interactions by
932 insulators in the *Drosophila* embryo. *Nature* 376: 533–536.
933 <https://doi.org/10.1038/376533a0>
- 934 Cai H. N., and P. Shen, 2001 Effects of cis arrangement of chromatin insulators on
935 enhancer-blocking activity. *Science* 291: 493–495.
936 <https://doi.org/10.1126/science.291.5503.493>
- 937 Casares F., W. Bender, J. Merriam, and E. Sánchez-Herrero, 1997 Interactions of
938 *Drosophila* Ultrabithorax regulatory regions with native and foreign promoters.

- 939 Genetics 145: 123–37.
- 940 Celis J. F. de, J. de Celis, P. Ligoxygakis, A. Preiss, C. Delidakis, *et al.*, 1996 Functional
- 941 relationships between Notch, Su(H) and the bHLH genes of the E(spl) complex:
- 942 the E(spl) genes mediate only a subset of Notch activities during imaginal
- 943 development. Development 122: 2719–28.
- 944 Chen J. L., K. L. Huisinga, M. M. Viering, S. A. Ou, C. T. Wu, *et al.*, 2002 Enhancer
- 945 action in trans is permitted throughout the Drosophila genome. Proc Natl Acad
- 946 Sci U S A 99: 3723–3728. <https://doi.org/10.1073/pnas.062447999>
- 947 Chen H., M. Levo, L. Barinov, M. Fujioka, J. B. Jaynes, *et al.*, 2018 Dynamic interplay
- 948 between enhancer–promoter topology and gene activity. Nat. Genet. 50: 1296–
- 949 1303. <https://doi.org/10.1038/s41588-018-0175-z>
- 950 Chetverina D., E. Savitskaya, O. Maksimenko, L. Melnikova, O. Zaytseva, *et al.*, 2008
- 951 Red flag on the white reporter: A versatile insulator abuts the white gene in
- 952 Drosophila and is omnipresent in mini-white constructs. Nucleic Acids Res. 36:
- 953 929–937. <https://doi.org/10.1093/nar/gkm992>
- 954 Chetverina D., M. Fujioka, M. Erokhin, P. Georgiev, J. B. Jaynes, *et al.*, 2017
- 955 Boundaries of loop domains (insulators): Determinants of chromosome form
- 956 and function in multicellular eukaryotes. BioEssays 39: 1–11.
- 957 <https://doi.org/10.1002/bies.201600233>
- 958 Cubebñas-Potts C., and V. G. Corces, 2015 Architectural proteins, transcription, and
- 959 the three-dimensional organization of the genome. FEBS Lett. 589: 2923–30.

960 <https://doi.org/10.1016/j.febslet.2015.05.025>

961 Cubeñas-Potts C., M. J. Rowley, X. Lyu, G. Li, E. P. Lei, *et al.*, 2016 Different enhancer
 962 classes in *Drosophila* bind distinct architectural proteins and mediate unique
 963 chromatin interactions and 3D architecture. *Nucleic Acids Res.*
 964 <https://doi.org/10.1093/nar/gkw1114>

965 Delidakis C., M. Monastirioti, and S. S. Magadi, 2014 E(spl): genetic, developmental,
 966 and evolutionary aspects of a group of invertebrate Hes proteins with close ties
 967 to Notch signaling. *Curr. Top. Dev. Biol.* 110: 217–62.
 968 <https://doi.org/10.1016/B978-0-12-405943-6.00006-3>

969 Doyle B., G. Fudenberg, M. Imakaev, and L. A. Mirny, 2014 Chromatin loops as
 970 allosteric modulators of enhancer-promoter interactions. *PLoS Comput Biol* 10:
 971 e1003867. <https://doi.org/10.1371/journal.pcbi.1003867>

972 Erceg J., J. A. Abed, A. Goloborodko, B. R. Lajoie, G. Fudenberg, *et al.*, 2018 The
 973 genome-wide, multi-layered architecture of chromosome pairing in early
 974 *Drosophila* embryos. *bioRxiv* 443028. <https://doi.org/10.1101/443028>

975 Fujioka M., G. Sun, and J. B. Jaynes, 2013 The *Drosophila* eve Insulator Homie
 976 Promotes eve Expression and Protects the Adjacent Gene from Repression by
 977 Polycomb Spreading. *PLoS Genet.* 9.
 978 <https://doi.org/10.1371/journal.pgen.1003883>

979 Fujioka M., H. Mistry, P. Schedl, and J. B. Jaynes, 2016 Determinants of Chromosome
 980 Architecture: Insulator Pairing in cis and in trans. *PLoS Genet.* 12: 1–25.

981 <https://doi.org/10.1371/journal.pgen.1005889>

982 Fukaya T., and M. Levine, 2017 Transvection. *Curr. Biol.* 27: R1047–R1049.

983 <https://doi.org/10.1016/j.cub.2017.08.001>

984 Fung J. C., W. F. Marshall, A. Dernburg, D. A. Agard, and J. W. Sedat, 1998

985 Homologous chromosome pairing in *Drosophila melanogaster* proceeds through

986 multiple independent initiations. *J Cell Biol* 141: 5–20.

987 Gelbart W. M., 1982 Synapsis-dependent allelic complementation at the

988 decapentaplegic gene complex in *Drosophila melanogaster*. *Proc. Natl. Acad.*

989 *Sci.* 79: 2636–2640.

990 Gemkow M. J., P. J. Verveer, and D. J. Arndt-Jovin, 1998 Homologous association of

991 the Bithorax-Complex during embryogenesis: consequences for transvection in

992 *Drosophila melanogaster*. *Development* 125: 4541–4552.

993 Georgiev P., T. Tikhomirova, V. Yelagin, T. Belenkaya, E. Gracheva, *et al.*, 1997

994 Insertions of Hybrid P Elements in the yellow Gene of *Drosophila* Cause a Large

995 Variety of Mutant Phenotypes. *Genetics* 146.

996 Gerasimova T. I., D. A. Gdula, D. V. Gerasimov, O. Simonova, and V. G. Corces, 1995 A

997 *drosophila* protein that imparts directionality on a chromatin insulator is an

998 enhancer of position-effect variegation. *Cell* 82: 587–597.

999 [https://doi.org/10.1016/0092-8674\(95\)90031-4](https://doi.org/10.1016/0092-8674(95)90031-4)

1000 Geyer P. K., M. M. Green, and V. G. Corces, 1990 Tissue-specific transcriptional

1001 enhancers may act in trans on the gene located in the homologous

1002 chromosome: the molecular basis of transvection in *Drosophila*. EMBO J 9:
1003 2247–2256.

1004 Geyer P. K., and V. G. Corces, 1992 DNA position-specific repression of transcription
1005 by a *Drosophila* zinc finger protein. Genes Dev 6: 1865–1873.

1006 Goldsborough A. S., and T. B. Kornberg, 1996 Reduction of transcription by
1007 homologue asynapsis in *Drosophila* imaginal discs. Nature 381: 807–810.
1008 <https://doi.org/10.1038/381807a0>

1009 Golovnin A., I. Biryukova, O. Romanova, M. Silicheva, A. Parshikov, *et al.*, 2003 An
1010 endogenous Su(Hw) insulator separates the yellow gene from the Achaete-
1011 scute gene complex in *Drosophila*. Development 130: 3249–3258.
1012 <https://doi.org/10.1242/dev.020263>

1013 Golovnin A., E. Melnick, A. Mazur, and P. Georgiev, 2005 *Drosophila* Su(Hw) insulator
1014 can stimulate transcription of a weakened yellow promoter over a distance.
1015 Genetics 170: 1133–1142. <https://doi.org/10.1534/genetics.104.034587>

1016 Groth A. C., M. Fish, R. Nusse, and M. P. Calos, 2004 Construction of transgenic
1017 *Drosophila* by using the site-specific integrase from phage ϕ C31. Genetics 166:
1018 1775–82.

1019 Guo Y., Q. Xu, D. Canzio, J. Shou, J. Li, *et al.*, 2015 CRISPR Inversion of CTCF Sites
1020 Alters Genome Topology and Enhancer/Promoter Function. Cell 162: 900–910.
1021 <https://doi.org/10.1016/j.cell.2015.07.038>

1022 Harrison D. A., D. A. Gdula, R. S. Coyne, and V. G. Corces, 1993 A leucine zipper

- 1023 domain of the suppressor of Hairy-wing protein mediates its repressive effect
- 1024 on enhancer function. *Genes Dev* 7: 1966–1978.
- 1025 <https://doi.org/10.1101/gad.7.10.1966>
- 1026 Hendrickson J. E., and S. Sakonju, 1995 Cis and trans interactions between the iab
- 1027 regulatory regions and abdominal-A and abdominal-B in *Drosophila*
- 1028 *melanogaster*. *Genetics* 139: 835–48.
- 1029 Heride C., M. Ricoul, K. Kieu, J. von Hase, V. Guillemot, *et al.*, 2010 Distance between
- 1030 homologous chromosomes results from chromosome positioning constraints. *J*
- 1031 *Cell Sci* 123: 4063–4075. <https://doi.org/10.1242/jcs.066498>
- 1032 Hiraoka Y., A. F. Dernburg, S. J. Parmelee, M. C. Rykowski, D. A. Agard, *et al.*, 1993
- 1033 The onset of homologous chromosome pairing during *Drosophila melanogaster*
- 1034 embryogenesis. *J Cell Biol* 120: 591–600.
- 1035 Hnisz D., D. S. Day, and R. A. Young, 2016 Insulated Neighborhoods: Structural and
- 1036 Functional Units of Mammalian Gene Control. *Cell* 167: 1188–1200.
- 1037 <https://doi.org/10.1016/j.cell.2016.10.024>
- 1038 Holdridge C., and D. Dorsett, 1991 Repression of hsp70 heat shock gene transcription
- 1039 by the suppressor of hairy-wing protein of *Drosophila melanogaster*. *Mol Cell*
- 1040 *Biol* 11: 1894–1900. <https://doi.org/10.1128/MCB.11.4.1894>.Updated
- 1041 Horn C., B. Jaunich, and E. A. Wimmer, 2000 Highly sensitive, fluorescent
- 1042 transformation marker for *Drosophila* transgenesis. *Dev. Genes Evol.* 210: 623–
- 1043 629. <https://doi.org/10.1007/s004270000111>

- 1044 Johnston R. J., C. Desplan, R. J. Johnston Jr., C. Desplan, R. J. Johnston, *et al.*, 2014
- 1045 Interchromosomal communication coordinates intrinsically stochastic
- 1046 expression between alleles. *Science* 343: 661–5.
- 1047 <https://doi.org/10.1126/science.1243039>
- 1048 Joyce E. F., B. R. Williams, T. Xie, and C. ting Wu, 2012 Identification of genes that
- 1049 promote or antagonize somatic homolog pairing using a high-throughput FISH-
- 1050 based screen. *PLoS Genet* 8: e1002667.
- 1051 <https://doi.org/10.1371/journal.pgen.1002667>
- 1052 Joyce E. F., J. Erceg, and C. ting Wu, 2016 Pairing and anti-pairing: A balancing act in
- 1053 the diploid genome. *Curr. Opin. Genet. Dev.* 37: 119–128.
- 1054 <https://doi.org/10.1016/j.jgde.2016.03.002>
- 1055 Kellum R., and P. Schedl, 1991 A position-effect assay for boundaries of higher order
- 1056 chromosomal domains. *Cell* 64: 941–950.
- 1057 Kellum R., and P. Schedl, 1992 A group of scs elements function as domain
- 1058 boundaries in an enhancer-blocking assay. *Mol Cell Biol* 12: 2424–2431.
- 1059 Klämbt C., E. Knust, K. Tietze, and J. A. Campos-Ortega, 1989 Closely related
- 1060 transcripts encoded by the neurogenic gene complex enhancer of split of
- 1061 *Drosophila melanogaster*. *EMBO J.* 8: 203–10.
- 1062 Kravchenko E., E. Savitskaya, O. Kravchuk, A. Parshikov, P. Georgiev, *et al.*, 2005
- 1063 Pairing between gypsy insulators facilitates the enhancer action in trans
- 1064 throughout the *Drosophila* genome. *Mol. Cell. Biol.* 25: 9283–91.

1065 <https://doi.org/10.1128/MCB.25.21.9283-9291.2005>

1066 Kravchuk O., M. Kim, P. Klepikov, A. Parshikov, P. Georgiev, *et al.*, 2016 Transvection
 1067 in *Drosophila*: trans-interaction between yellow enhancers and promoter is
 1068 strongly suppressed by a cis-promoter only in certain genomic regions.
 1069 Chromosoma 1–11. <https://doi.org/10.1007/s00412-016-0605-6>

1070 Kuhn E. J., M. M. Viering, K. M. Rhodes, and P. K. Geyer, 2003 A test of insulator
 1071 interactions in *Drosophila*. EMBO J. 22: 2463–2471.
 1072 <https://doi.org/10.1093/emboj/cdg241>

1073 Kutach a K., and J. T. Kadonaga, 2000 The downstream promoter element DPE
 1074 appears to be as widely used as the TATA box in *Drosophila* core promoters.
 1075 Mol. Cell. Biol. 20: 4754–4764. [https://doi.org/10.1128/MCB.20.13.4754-](https://doi.org/10.1128/MCB.20.13.4754-4764.2000)
 1076 4764.2000

1077 Kvon E. Z., 2015 Using transgenic reporter assays to functionally characterize
 1078 enhancers in animals. Genomics 106: 185–192.
 1079 <https://doi.org/10.1016/j.ygeno.2015.06.007>

1080 Kyrchanova O., D. Chetverina, O. Maksimenko, A. Kullyev, and P. Georgiev, 2008a
 1081 Orientation-dependent interaction between *Drosophila* insulators is a property
 1082 of this class of regulatory elements. Nucleic Acids Res. 36: 7019–7028.
 1083 <https://doi.org/10.1093/nar/gkn781>

1084 Kyrchanova O., S. Toshchakov, Y. Podstreshnaya, A. Parshikov, and P. Georgiev,
 1085 2008b Functional interaction between the Fab-7 and Fab-8 boundaries and the

1086 upstream promoter region in the *Drosophila* Abd-B gene. *Mol. Cell. Biol.* 28:
1087 4188–4195. <https://doi.org/10.1128/MCB.00229-08>

1088 Kyrchanova O., T. Ivlieva, S. Toshchakov, A. Parshikov, O. Maksimenko, *et al.*, 2011
1089 Selective interactions of boundaries with upstream region of Abd-B promoter in
1090 *Drosophila* bithorax complex and role of dCTCF in this process. *Nucleic Acids*
1091 *Res.* 39: 3042–3052. <https://doi.org/10.1093/nar/gkq1248>

1092 Kyrchanova O., O. Maksimenko, V. Stakhov, T. Ivlieva, A. Parshikov, *et al.*, 2013
1093 Effective blocking of the white enhancer requires cooperation between two
1094 main mechanisms suggested for the insulator function. *PLoS Genet* 9:
1095 e1003606. <https://doi.org/10.1371/journal.pgen.1003606>

1096 Kyrchanova O., V. Mogila, D. Wolle, G. Deshpande, A. Parshikov, *et al.*, 2016
1097 Functional Dissection of the Blocking and Bypass Activities of the Fab-8
1098 Boundary in the *Drosophila* Bithorax Complex. *PLoS Genet.* 12: 1–22.
1099 <https://doi.org/10.1371/journal.pgen.1006188>

1100 Lee A. M., and C. T. Wu, 2006 Enhancer-promoter communication at the yellow gene
1101 of *Drosophila melanogaster*: Diverse promoters participate in and regulate trans
1102 interactions. *Genetics* 174: 1867–1880.
1103 <https://doi.org/10.1534/genetics.106.064121>

1104 Levis R., T. Hazelrigg, and G. M. Rubin, 1985 Effects of genomic position on the
1105 expression of transduced copies of the white gene of *Drosophila*. *Science* 229:
1106 558–61.

- 1107 Lewis E. B., 1954 The Theory and Application of a New Method of Detecting
- 1108 Chromosomal Rearrangements in *Drosophila melanogaster*. *Am. Nat.* 88: 225–
- 1109 239. <https://doi.org/doi:10.1086/281833>
- 1110 Li H. B., M. Muller, I. A. Bahechar, O. Kyrchanova, K. Ohno, *et al.*, 2011 Insulators, not
- 1111 Polycomb response elements, are required for long-range interactions between
- 1112 Polycomb targets in *Drosophila melanogaster*. *Mol Cell Biol* 31: 616–625.
- 1113 <https://doi.org/10.1128/MCB.00849-10>
- 1114 Lim B., T. Heist, M. Levine, and T. Fukaya, 2018 Visualization of Transvection in Living
- 1115 *Drosophila* Embryos. *Mol. Cell* 70: 287–296.e6.
- 1116 <https://doi.org/10.1016/j.molcel.2018.02.029>
- 1117 Markstein M., C. Pitsouli, C. Villalta, S. E. Celniker, and N. Perrimon, 2008 Exploiting
- 1118 position effects and the gypsy retrovirus insulator to engineer precisely
- 1119 expressed transgenes. *Nat. Genet.* 40: 476–483. <https://doi.org/10.1038/ng.101>
- 1120 Martínez-Laborda A., A. González-Reyes, and G. Morata, 1992 Trans regulation in the
- 1121 Ultrabithorax gene of *Drosophila*: alterations in the promoter enhance
- 1122 transvection. *EMBO J.* 11: 3645–52.
- 1123 McKee B. D., 2004 Homologous pairing and chromosome dynamics in meiosis and
- 1124 mitosis. *Biochim Biophys Acta* 1677: 165–180.
- 1125 <https://doi.org/10.1016/j.bbaexp.2003.11.017>
- 1126 Mellert D. J., and J. W. Truman, 2012 Transvection is common throughout the
- 1127 *Drosophila* genome. *Genetics* 191: 1129–1141.

- 1128 <https://doi.org/10.1534/genetics.112.140475>
- 1129 Moon H., G. Filippova, D. Loukinov, E. Pugacheva, Q. Chen, *et al.*, 2005 CTCF is
- 1130 conserved from Drosophila to humans and confers enhancer blocking of the
- 1131 Fab-8 insulator. EMBO Rep. 6: 165–70.
- 1132 <https://doi.org/10.1038/sj.embor.7400334>
- 1133 Morris J. R., J. Chen, S. T. Filandrinov, R. C. Dunn, R. Fisk, *et al.*, 1999a An analysis of
- 1134 transvection at the yellow locus of Drosophila melanogaster. Genetics 151:
- 1135 633–651.
- 1136 Morris J. R., P. K. Geyer, and C. T. Wu, 1999b Core promoter elements can regulate
- 1137 transcription on a separate chromosome in trans. Genes Dev 13: 253–258.
- 1138 Morris J. R., D. A. Petrov, A. M. Lee, and C. T. Wu, 2004 Enhancer choice in cis and in
- 1139 trans in Drosophila melanogaster: Role of the promoter. Genetics 167: 1739–
- 1140 1747. <https://doi.org/10.1534/genetics.104.026955>
- 1141 Muravyova E., A. Golovnin, E. Gracheva, A. Parshikov, T. Belenkaya, *et al.*, 2001 Loss
- 1142 of insulator activity by paired Su(Hw) chromatin insulators. Science 291: 495–8.
- 1143 <https://doi.org/10.1126/science.291.5503.495>
- 1144 Negre N., C. D. Brown, P. K. Shah, P. Kheradpour, C. A. Morrison, *et al.*, 2010 A
- 1145 comprehensive map of insulator elements for the Drosophila genome. PLoS
- 1146 Genet 6: e1000814. <https://doi.org/10.1371/journal.pgen.1000814>
- 1147 Parkhurst S. M., D. A. Harrison, M. P. Remington, C. Spana, R. L. Kelley, *et al.*, 1988
- 1148 The Drosophila su(Hw) gene, which controls the phenotypic effect of the gypsy

1149 transposable element, encodes a putative DNA-binding protein. *Genes Dev.* 2:
1150 1205–1215. <https://doi.org/10.1101/gad.2.10.1205>

1151 Parnell T. J., M. M. Vierung, A. Skjesol, C. Helou, E. J. Kuhn, *et al.*, 2003 An
1152 endogenous suppressor of hairy-wing insulator separates regulatory domains in
1153 *Drosophila*. *Proc Natl Acad Sci U S A* 100: 13436–13441.
1154 <https://doi.org/10.1073/pnas.2333111100>

1155 Parnell T. J., E. J. Kuhn, B. L. Gilmore, C. Helou, M. S. Wold, *et al.*, 2006 Identification
1156 of genomic sites that bind the *Drosophila* suppressor of Hairy-wing insulator
1157 protein. *Mol Cell Biol* 26: 5983–5993. <https://doi.org/10.1128/MCB.00698-06>

1158 Pfeiffer B. D., A. Jenett, A. S. Hammonds, T.-T. B. Ngo, S. Misra, *et al.*, 2008 Tools for
1159 neuroanatomy and neurogenetics in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.*
1160 105: 9715–20. <https://doi.org/10.1073/pnas.0803697105>

1161 Pfeiffer B. D., T.-T. B. Ngo, K. L. Hibbard, C. Murphy, A. Jenett, *et al.*, 2010
1162 Refinement of tools for targeted gene expression in *Drosophila*. *Genetics* 186:
1163 735–55. <https://doi.org/10.1534/genetics.110.119917>

1164 Postika N., M. Metzler, M. Affolter, M. Müller, P. Schedl, *et al.*, 2018 Boundaries
1165 mediate long-distance interactions between enhancers and promoters in the
1166 *Drosophila* Bithorax complex, (M. C. Gambetta, Ed.). *PLOS Genet.* 14: e1007702.
1167 <https://doi.org/10.1371/journal.pgen.1007702>

1168 Ramos E., D. Ghosh, E. Baxter, and V. G. Corces, 2006 Genomic organization of gypsy
1169 chromatin insulators in *Drosophila melanogaster*. *Genetics* 172: 2337–2349.

- 1170 <https://doi.org/10.1534/genetics.105.054742>
- 1171 Ringrose L., 2009 Transgenesis in *Drosophila melanogaster*. *Methods Mol. Biol.* 561:
1172 3–19. https://doi.org/10.1007/978-1-60327-019-9_1
- 1173 Roseman R. R., V. Pirrotta, and P. K. Geyer, 1993 The su(Hw) protein insulates
1174 expression of the *Drosophila melanogaster* white gene from chromosomal
1175 position-effects. *EMBO J.* 12: 435–42.
- 1176 Roseman R. R., E. A. Johnson, C. K. Rodesch, M. Bjerke, R. N. Nagoshi, *et al.*, 1995 A P
1177 element containing suppressor of hairy-wing binding regions has novel
1178 properties for mutagenesis in *Drosophila melanogaster*. *Genetics* 141: 1061–74.
- 1179 Schoborg T., S. Kuruganti, R. Rickels, and M. Labrador, 2013 The *Drosophila* gypsy
1180 insulator supports transvection in the presence of the vestigial enhancer. *PLoS*
1181 One 8: e81331. <https://doi.org/10.1371/journal.pone.0081331>
- 1182 Schwartz Y. B., D. Linder-Basso, P. V. Kharchenko, M. Y. Tolstorukov, M. Kim, *et al.*,
1183 2012 Nature and function of insulator protein binding sites in the *Drosophila*
1184 genome. *Genome Res.* 22: 2188–2198.
1185 <https://doi.org/10.1101/gr.138156.112.2188>
- 1186 Schwartz Y. B., and G. Cavalli, 2017 Three-Dimensional Genome Organization and
1187 Function in *Drosophila*. *Genetics* 205: 5–24.
1188 <https://doi.org/10.1534/genetics.115.185132>
- 1189 Sipos L., J. Mihály, F. Karch, P. Schedl, J. Gausz, *et al.*, 1998 Transvection in the
1190 *Drosophila* Abd-B domain: extensive upstream sequences are involved in

- 1191 anchoring distant cis-regulatory regions to the promoter. *Genetics* 149: 1031–
1192 50.
- 1193 Soshnev A. A., X. Li, M. D. Wehling, and P. K. Geyer, 2008 Context differences reveal
1194 insulator and activator functions of a Su(Hw) binding region. *PLoS Genet* 4:
1195 e1000159. <https://doi.org/10.1371/journal.pgen.1000159>
- 1196 Spana C., D. A. Harrison, and V. G. Corces, 1988 The *Drosophila melanogaster*
1197 suppressor of Hairy-wing protein binds to specific sequences of the gypsy
1198 retrotransposon. *Genes Dev* 2: 1414–1423.
- 1199 Spana C., and V. G. Corces, 1990 DNA bending is a determinant of binding specificity
1200 for a *Drosophila* zinc finger protein. *Genes Dev* 4: 1505–1515.
- 1201 Stratigi K., M. Kapsetaki, M. Aivaliotis, T. Town, R. a. Flavell, *et al.*, 2015 Spatial
1202 proximity of homologous alleles and long noncoding RNAs regulate a switch in
1203 allelic gene expression. *Proc. Natl. Acad. Sci.* 1–10.
1204 <https://doi.org/10.1073/pnas.1502182112>
- 1205 Thibault S. T., M. A. Singer, W. Y. Miyazaki, B. Milash, N. A. Dompe, *et al.*, 2004 A
1206 complementary transposon tool kit for *Drosophila melanogaster* using P and
1207 piggyBac. *Nat. Genet.* 36: 283–287. <https://doi.org/10.1038/ng1314>
- 1208 Thorpe H. M., and M. C. Smith, 1998 In vitro site-specific integration of
1209 bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase
1210 family. *Proc. Natl. Acad. Sci. U. S. A.* 95: 5505–10.
- 1211 Venken K. J. T., Y. He, R. A. Hoskins, and H. J. Bellen, 2006 P[acman]: a BAC transgenic

1212 platform for targeted insertion of large DNA fragments in *D. melanogaster*.
 1213 Science 314: 1747–51. <https://doi.org/10.1126/science.1134426>

1214 Viets K., M. Sauria, C. Chernoff, C. Anderson, S. Tran, *et al.*, 2018 TADs pair
 1215 homologous chromosomes to promote interchromosomal gene regulation.
 1216 bioRxiv 445627. <https://doi.org/10.1101/445627>

1217 Vogelmann J., A. Le Gall, S. Dejardin, F. Allemand, A. Gamot, *et al.*, 2014 Chromatin
 1218 insulator factors involved in long-range DNA interactions and their role in the
 1219 folding of the *Drosophila* genome. PLoS Genet 10: e1004544.
 1220 <https://doi.org/10.1371/journal.pgen.1004544>

1221 Wei W., and M. D. Brennan, 2001 The gypsy insulator can act as a promoter-specific
 1222 transcriptional stimulator. Mol. Cell. Biol. 21: 7714–7720.
 1223 <https://doi.org/10.1128/MCB.21.22.7714-7720.2001>

1224

SUPPLEMENTARY MATERIALS AND METHODS

The role of insulators in transgene transvection in *Drosophila*

Piwko et al

Plasmid constructs

Here we present the general strategy used for plasmid construction. Detailed information regarding each construct and maps are available upon request. **Table 1** is the list of all constructs used in this study together with their general characteristics, provides numeric (No) identification for each construct and associates them with the Figures in the Results section. **Table 2** provides information of the primers used in plasmid construction.

To obtain our constructs we have used four different backbones of publicly available plasmids: *egfp*-containing *pStinger* and *lacZ*-containing *pPelican* (Barolo *et al.* 2004), *pBluescript* (SK(+), Stratagene) and *pLacZ-attB* (Bischof *et al.* 2007). *pStinger*, *pPelican* and *pBluescript* were modified by cloning *attB* sequence from *pTA-attB* plasmid (Groth *et al.* 2004) in the same orientation in each plasmid, immediately upstream to the multi-cloning site (MCS; in case of *pStinger/pPelican* it is upstream of the 5' GI).

The 7 kb genomic fragment encompassing *E(spl)m7* and *E(spl)m8*, together with their upstream (enhancers and promoters) and downstream (3' UTRs) regulatory regions, was cloned with *XhoI* and *Clal* restriction enzymes from the R3012 cosmid clone encompassing part of *E(spl)* locus (Delidakis and Artavanis-Tsakonas 1992) into *pPelican-attB*. Due to strong post-transcriptional repression of both of these genes (Lai *et al.* 2005) we have replaced *E(spl)m7* and *E(spl)m8* 3' UTRs with the SV40 (derived from pGL3 vector, Promega) and the *Adh* (derived from the *Ract-HAdh* vector, Swevers *et al.* 1996) poly A terminators,

respectively. Subsequently, unique restriction sites were introduced before the start codons of both genes by mutagenesis to allow cloning of EGFP (derived from *pCRE-d2EGFP* vector, Clontech) into the ORF of *E(spl)m7* in one construct (*GFPm7-m8*, construct **No 1** in Table 1) and of *E(spl)m8* in another (*m7-GFPm8*, construct **No 2**).

The 'short genomic' constructs (*GFPm7* and *GFPm8*) were generated by the excision of the regulatory and coding portions of the untagged genes from the 'long genomic' constructs (i.e., *GFPm7* (construct **No 3**) was generated from *GFPm7-m8* by a deletion of the downstream sequences to the *m7*'s (SV40) 3' UTR, and *GFPm8* (construct **No 4**) was generated from *m7-GFPm8* by a deletion of the *m7*'s 3' UTR and its upstream sequences). The resulting constructs contained 2.1 kb sequence upstream of the start codon of *GFPm7* (which we denote as *ep7*) and 1.3 kb sequence upstream of the start codon of *GFPm8* (*ep8*).

The GI-less *ep8-GFPm8* construct (**No 5**) was generated by cloning of the *ep8-GFPm8* module (together with its *Adh* 3' UTR) from *GFPm8* into *pPelican-attB* bearing deletion of both GIs and LacZ. This construct was subsequently used to generate four single-GI versions of *ep8-GFPm8* by cloning the 5' GI from *pPelican* into 5' (upstream of *ep8*) or 3' (downstream of *Adh* 3' UTR) positions in forward or reverse orientation (constructs **No 6-9**).

The *ep8-lacZ pPelican*-based construct (**No 10**) was made by cloning PCR-amplified *ep8* (primers **ep8F** and **ep8R**) into the MCS of *pPelican*.

The *ep7-lacZ pPelican*-based construct (**No 11**) was made by cloning PCR-amplified *ep7* (primers **ep7F** and **ep7R**) into the MCS of *pPelican*. The *pBluescript* GI-less version of *ep7-lacZ* (**No 12**) was made by (1) cloning the entire *ep7-lacZ* sequence of *pPelican*-based *ep7-lacZ* (together with its SV40 3' UTR) into *pBluescriptSK(+)* in between *KpnI* and *SpeI* restriction sites; (2) and subsequent cloning of the DNA fragment containing *3xP3-dsRed* and

attB sequences from *pMinos{3xP3-dsRed}* vector (Berghammer *et al.* 1999; this vector was beforehand modified by us by inserting *attB* sequence in the vicinity of *3xP3-dsRed*) into *KpnI* site of the *pBluescript* such that *attB* site is in between the *3xP3-dsRed* and *ep7*. The GI-less version of the *ep7-lacZ* construct with mini-*white* marker (**Nº 13**) was made by the replacement of the *3xP3-dsRed* in the latter construct with *pPelican*'s mini-*white*. The *pBluescript* GI-less version of *ep7-lacZ-3xP3-dsRed* (**Nº 12**) construct was used as a basis to generate (1) GI-containing *ep7-lacZ* constructs by cloning 5' GI from *pPelican* immediately upstream of *ep7* or/and downstream of *SV40 3' UTR* in a reverse or forward orientation (constructs **Nº 14-20**); (2) the three *ep7-lacZ* constructs containing different parts of mini-*white* gene cloned into *SpeI* site (immediately downstream of *SV40 3' UTR*): the 0.24 kb *AfeI/EcoRI* fragment encompassing the 5'/promoter region of mini-*white* (*ep7-lacZ-w5'*, construct **Nº 21**), the 2.4 kb *AflIII/EcoRV* fragment encompassing mini-*white*'s gene body (*ep7-lacZ-wB*, construct **Nº 22**) and the 0.9 kb *EcoRV/BsrGI* 3' part of the mini-*white* (*ep7-lacZ-w3'*, construct **Nº 23**). The *pBluescript ep7-lacZ* construct with two GIs in forward orientation (**Nº 14**) was used to generate two 'tester' constructs by cloning *ep8-m8-Adh 3' UTR* module from the EGFP-untagged version of the *GFPm8* construct (*ep8-m8* tester, **Nº 24**) and *pH-gfp-SV40 3' UTR* module from *pHStinger* (*pH-gfp* tester, **Nº 25**) immediately downstream of the 3' GI.

The *ep7-ΔTATA-lacZ* construct (**Nº 26**) containing deletion of 20 nt encompassing TATA box was generated by introducing an *EcoRV* restriction site by site-directed mutagenesis upstream of the TATA box of *p7* in *pPelican*-based *ep7-lacZ* construct (**Nº 11**) and subsequent excision of the sequence between *EcoRV* and *BstEII*. The intermediate *ep7-lacZ* construct with introduced *EcoRV* (but without *EcoRV/BstEII* deletion; **Nº 27**) produced LacZ expression pattern and levels indistinguishable to that of the *ep7-lacZ* without *EcoRV* and

was used as a control transgene (to *ep7-ΔTATA-lacZ* and *ep7-ΔDPE-lacZ*) containing wild-type *p7*. The *ep7-ΔDPE-lacZ* (**№ 28**) construct was generated from *pPelican*-based *ep7-lacZ* by excision of 83 nt containing INR and DPE motifs in between *BstEII* and *StuI* restriction sites. The *e7-lacZ* (promoterless construct, **№ 29**) was generated by an excision of sequence between *EcoRV* and *StuI* from the *ep7-lacZ* construct with the introduced *EcoRV* site (**№ 27**).

All *pH-gfp* constructs (**№ 30-36**) were made based on *pHStinger* (a version of *pStinger* containing *pH* fused to *egfp*; Barolo *et al.* 2000). The *pH-gfp* with two GIs in forward orientation (**№ 30**) was made by inserting *attB* sequence into *pHStinger*. Subsequent deletions of the 3' GI, 5' GI and both GIs from this construct resulted in generation of constructs **№ 31**, **32** and **33**, respectively. The 5' GI^{REV} *pH-gfp* (**№ 34**) was made by cloning GI in reverse orientation into GI-less *pH-gfp*. The 5' GI^{FOR} 3' GI^{REV} *pH-gfp* (**№ 35**) was made by cloning GI in a reverse orientation in place of the 3' GI^{FOR} of GIs^{FOR} *pH-gfp* (**№ 30**). The 5' GI^{REV} 3' GI^{FOR} *pH-gfp* (**№ 36**) was made by cloning GI in a reverse orientation in place of the 5' GI^{FOR} of the GIs^{FOR} *pH-gfp* (**№ 30**).

The 2 kb *BglII-EcoRV e7*-containing fragment derived from the *ep7-lacZ* construct bearing *EcoRV* site introduced upstream to TATA box of *p7* (**№ 27**) was used to generate (1) the *e7* construct (**№ 37**) by replacing *lacZ* in the *pPelican-attB* and (2) the *e7pH-gfp* (**№ 38**) by cloning it upstream to *pH* of the GIs^{FOR} *pH-gfp* (**№ 30**). The *ep7-gfp* construct (**№ 39**) was made by PCR-amplifying the *p7* promoter from *ep7-lacZ* (**№ 11**) with **p7F** and **p7R** primers and ligating the *NheI*/*HeaIII*-digested PCR product into *NheI* and *StuI* sites in the *e7pH-gfp* construct (**№ 38**) such that *p7* sequence replaces the sequence of *pH*. The *p7-gfp* (**№ 40**) was based on the *ep7-gfp* (**№ 39**) by excision of *e7* with *KpnI*. The sequence of *e8* was PCR-amplified with **e8F** and **e8R** primers, its product was cut with *AvrII* and *NheI* and ligated to

XbaI and *NheI* sites (replacing *e7*) in the *e7* construct (**Nº 37**) to generate *e8* construct (**Nº 41**). The sequence of *e8* (as a *MfeI*-*NheI* fragment) was cloned from *e8* construct (**Nº 41**) upstream of *p7* (*EcoRI*/*NheI*) in the *p7-gfp* construct (**Nº 40**) to generate *e8p7-gfp* (**Nº 42**). The *e7p8-gfp* construct (**Nº 43**) was made by PCR-amplifying the *p8* promoter from *ep8-lacZ* (**Nº 10**) with **p8F** and **L5R** primers and ligating the *NheI*/*HeaIII*-digested PCR product into *NheI*/*StuI* sites in the *e7pH-gfp* construct (**Nº 38**) such that *p8* sequence replaces the sequence of *pH*. The *e8* sequence was cut out from the *e8* construct (**Nº 41**) with *MfeI* and *NheI* restriction enzymes and ligated to *EcoRI* and *NheI* sites (such that the *e8* sequence replaces the sequence of *e7*) in the *e7p8-gfp* construct (**Nº 43**) to generate *ep8-gfp* construct (**Nº 44**). The *p8-gfp* construct (**Nº 45**) was generated by excision of the *e7* sequence with *NaeI* and *NheI* restriction enzymes from the *e7p8-gfp* construct (**Nº 43**).

The luc constructs (**Nº 46-48**) were generated by replacing *pH-gfp* module with the restriction fragment containing *pH-luc* module (derived from the *pGL3-hsp70-luc* construct, gift from M. Monastirioti, IMBB) in the Gl^{FOR} *pH-gfp* (**Nº 30**), Gl^{FOR} *e7pH-gfp* (**Nº 38**) and Gl-less *pH-gfp* (**Nº 33**) constructs – resulting in generation of Gl^{FOR} *pH-luc* (**Nº 46**), Gl^{FOR} *e7pH-luc* (**Nº 47**) and Gl-less *pH-luc* (**Nº 48**), respectively.

The ‘blank’ sender construct (**Nº 49**) was made by cloning (1) the *HindIII*/*BamHI*-cut PCR-amplified product of *e7* (primers **e7-2F** and **e7-2R**) into *pLacZ-attB* construct (replacing *lacZ* sequence) and (2) subsequent cloning of the *NheI*/*BglII*-cut PCR amplicon of *e8* (primers **e8-2F** and **e8-2R**). The ‘blank’ sender construct (**Nº 49**) was used to generate Gl, Fab8 and 1A2 sender constructs (**Nº 50, 51 and 52**, respectively) by cloning the PCR-amplified sequences of Gl, Fab8 or 1A2 into the *BglII* site located in between *e7* and *e8*. The 400 bp Gl sequence was amplified on the *pPelican* template with **GIF** and **GIR** primers and contains full length Gl. The

540 bp Fab8 sequence was PCR-amplified from *Drosophila* genome using **Fab8F** and **Fab8R** primers, which relates to genomic location 3R: 16918976..16919515 (*Dm6*), and contains the F8²⁵⁴ sequence and part of F8⁴⁶⁹ sequence (as defined in Kyrchanova *et al.* 2008). The 454 bp 1A2 sequence was PCR-amplified from *Drosophila* genome using primers **1A2F** and **1A2R**, and relates to the exact same region defined as 1A2 insulator in Kyrchanova *et al.* 2008a. The WI-less sender constructs, **Nº 53** and **54**, were generated from the ‘blank’ sender (**Nº 49**) and the *GI* sender (**Nº 50**) constructs, respectively, by excision of the 341 bp WI-containing fragment between *BsrGI* and *BstBI* sites.

The ‘blank’ responder construct (**Nº 55**) was made by (1) replacing the *pH* in the *pHStinger* with the the *DSCP* promoter (*pD*) derived from *pBPGUw* plasmid (Pfeiffer *et al.* 2008) and (2) ligating the resulting *pD-gfp-SV40 3' UTR* module in between the *BglII* and *NheI* sites of *pLacZ-attB*, which already contains the *pH-lacZ-SV40* module (Bischof *et al.* 2007).

Subsequently, this construct was used to generate *GI* and *Fab8* responder constructs (**Nº 56** and **57**, respectively) by cloning into the *BglII* site the sequences of *GI* and *Fab8* PCR-amplified with the same sets of primers (i.e., **GIF/GIR** and **Fab8F/Fab8R**) and in the same orientation as it was for the corresponding sender constructs. The 1A2 sequence was amplified with the **1A2F2** and **1A2R** primers and ligated into *EcoRI/BglII* sites of the ‘blank’ responder construct (**Nº 55**) resulting in generation of the 1A2 responder construct (**Nº 58**).

The WI-less *GI* responder (**Nº 59**) was generated by a deletion of the 502 bp WI-containing *BstBI/NsiI* fragment from the *GI* responder construct (**Nº 56**).

The *HH-1.5*, *HH-2.1*, *HH-5.4*, *HB-1.6*, *BB-1.8*, *BB-3.1* and *BB-2.3* constructs (**Nº 60-66**) were generated by cloning 1.5 kb-, 2.1 kb- and 5.4 kb-*HindIII-HindIII*, 1.6 kb-*HindIII-BglII*, 1.8 kb-, 3.1 kb- and 2.3 kb-*BglII-BglII* fragments, respectively, derived from the *R3007* cosmid

(Delidakis and Artavanis-Tsakonas 1992) into the MCS of *lacZ*-deficient *pPelican-attB* (not listed).

The *BB-5.4*, *BB-1.3*, *BB-7.9*, *BB-1.7*, *BB-07*, *BP-3.2* and *PB-4.7* constructs (**Nº 67-73**) were generated by cloning 5.4 kb-, 1.3 kb-, 7.9 kb-, 1.7 kb- and 0.7 kb-*BglII*-*BglII*, 3.2 kb- and 4.7 kb-*PstI*-*BglII* fragments, respectively, derived from the *R3012* cosmid (Delidakis and Artavanis-Tsakonas 1992) into the MCS of *lacZ*-deficient *pPelican-attB*.

The 0.8 kb sequence of Vestigial Quadrant Enhancer (*vgQ*, Kim *et al.* 1996) was derived from a *pBluescript*-*vgQ* vector and cloned into the MCS of *lacZ*-deficient *pPelican-attB* to generate the *VGQ* construct (**Nº 74**).

Table 1 List of constructs used in the study

Nº	schematic	Name (as in text)	Marker gene	Insulators	Backbone	Figures
1		<i>GFPm7-m8</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	1
2		<i>m7-GFPm8</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	1
3		<i>GFP-E(spl)m7</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	1
4		<i>GFP-E(spl)m8</i> , <i>ep8-GFPm8**</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	1, 2, S1
5		<i>ep8-GFPm8</i>	mini-white	WI*	<i>pPelican</i>	2, S1, S8
6		<i>ep8-GFPm8</i>	mini-white	WI*, 5' Gl ^{FOR}	<i>pPelican</i>	S1, S8
7		<i>ep8-GFPm8</i>	mini-white	WI*, 5' Gl ^{REV}	<i>pPelican</i>	S8
8		<i>ep8-GFPm8</i>	mini-white	WI*, 3' Gl ^{FOR}	<i>pPelican</i>	S8
9		<i>ep8-GFPm8</i>	mini-white	WI*, 3' Gl ^{REV}	<i>pPelican</i>	S8

10		<i>ep8-lacZ</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	S3, 5
11		<i>ep7-lacZ</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	1, 2, 3
12		<i>ep7-lacZ</i>	3xP3-dsRed	-	<i>pBluescript</i>	2, S7, S8
13		<i>ep7-lacZ</i>	mini-white	WI*	<i>pPelican</i>	2, 3
14		<i>ep7-lacZ</i>	3xP3-dsRed	Gl ^{FOR}	<i>pBluescript</i>	2, 3, S6, 6, 7, S9
15		<i>ep7-lacZ</i>	3xP3-dsRed	5' Gl ^{FOR}	<i>pBluescript</i>	6, S7, S8, 7
16		<i>ep7-lacZ</i>	3xP3-dsRed	5' Gl ^{REV}	<i>pBluescript</i>	6, S8
17		<i>ep7-lacZ</i>	3xP3-dsRed	3' Gl ^{FOR}	<i>pBluescript</i>	6, S8
18		<i>ep7-lacZ</i>	3xP3-dsRed	3' Gl ^{REV}	<i>pBluescript</i>	S8
19		<i>ep7-lacZ</i>	3xP3-dsRed	5' Gl ^{FOR} , 3' Gl ^{REV}	<i>pBluescript</i>	6
20		<i>ep7-lacZ</i>	3xP3-dsRed	5' Gl ^{REV} , 3' Gl ^{FOR}	<i>pBluescript</i>	6
21		<i>ep7-lacZ-w5'</i>	3xP3-dsRed	-	<i>pBluescript</i>	S1
22		<i>ep7-lacZ-wB</i>	3xP3-dsRed	-	<i>pBluescript</i>	S1
23		<i>ep7-lacZ-w3'</i>	3xP3-dsRed	WI	<i>pBluescript</i>	S1
24		<i>ep8-m8 tester</i>	3xP3-dsRed	Gl ^{FOR}	<i>pBluescript</i>	7
25		<i>pH-gfp tester</i>	3xP3-dsRed	Gl ^{FOR}	<i>pBluescript</i>	7
26		<i>ep7-ΔTATA-lacZ</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	5, S5, 7
27		<i>ep7-lacZ</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	5, S5

28		<i>ep7-ΔDPE-lacZ</i>	mini-white	WI*, GlS ^{FOR}	<i>pPelican</i>	5, S5, 7
29		<i>e7-lacZ</i>	mini-white	WI*, GlS ^{FOR}	<i>pPelican</i>	5, S5, 7
30		<i>pH-gfp</i>	mini-white	WI*, GlS ^{FOR}	<i>pStinger</i>	3, S2, S3, S4, 4, 5, S6, 6, 7, S9
31		<i>pH-gfp</i>	mini-white	WI*, 5' Gl ^{FOR}	<i>pStinger</i>	6, S7, 7
32		<i>pH-gfp</i>	mini-white	WI*, 3' Gl ^{FOR}	<i>pStinger</i>	6
33		<i>pH-gfp</i>	mini-white	WI*	<i>pStinger</i>	3, S7
34		<i>pH-gfp</i>	mini-white	WI*, 5' Gl ^{REV}	<i>pStinger</i>	6
35		<i>pH-gfp</i>	mini-white	WI*, 5' Gl ^{FOR} , 3' Gl ^{REV}	<i>pStinger</i>	6
36		<i>pH-gfp</i>	mini-white	WI*, 5' Gl ^{REV} , 3' Gl ^{FOR}	<i>pStinger</i>	6
37		<i>e7</i>	mini-white	WI*, GlS ^{FOR}	<i>pPelican</i>	S4, 4, 5
38		<i>e7pH-gfp</i>	mini-white	WI*, GlS ^{FOR}	<i>pStinger</i>	S3, S4, 4
39		<i>ep7-gfp</i>	mini-white	WI*, GlS ^{FOR}	<i>pStinger</i>	S3, 4
40		<i>p7-gfp</i>	mini-white	WI*, GlS ^{FOR}	<i>pStinger</i>	S3, 4, S5, S9
41		<i>e8</i>	mini-white	WI*, GlS ^{FOR}	<i>pPelican</i>	4, 5
42		<i>e8p7-gfp</i>	mini-white	WI*, GlS ^{FOR}	<i>pStinger</i>	4
43		<i>e7p8-gfp</i>	mini-white	WI*, GlS ^{FOR}	<i>pStinger</i>	S3, 4
44		<i>ep8-gfp</i>	mini-white	WI*, GlS ^{FOR}	<i>pStinger</i>	4
45		<i>p8-gfp</i>	mini-white	WI*, GlS ^{FOR}	<i>pStinger</i>	S3, 4, S5, S9

46		<i>pH-luc</i>	mini-white	WI*, Gl ^{FOR}	<i>pStinger</i>	5
47		<i>e7pH-luc</i>	mini-white	WI*, Gl ^{FOR}	<i>pStinger</i>	5
48		<i>pH-luc</i>	mini-white	WI*	<i>pStinger</i>	5
49		'blank' sender	mini-white	WI*	<i>pLacZ-attB</i>	8
50		<i>Gl sender</i>	mini-white	WI*, Gl ^{REV}	<i>pLacZ-attB</i>	8
51		<i>Fab8 sender</i>	mini-white	WI*, Fab8	<i>pLacZ-attB</i>	8
52		<i>1A2 sender</i>	mini-white	WI*, 1A2	<i>pLacZ-attB</i>	8
53		'blank' sender	mini-white- dWI	-	<i>pLacZ-attB</i>	8
54		<i>Gl sender</i>	mini-white- dWI	Gl ^{REV}	<i>pLacZ-attB</i>	8
55		'blank' responder	mini-white	WI*	<i>pLacZ-attB</i>	8
56		<i>Gl responder</i>	mini-white	WI*, Gl ^{REV}	<i>pLacZ-attB</i>	8
57		<i>Fab8 responder</i>	mini-white	WI*, Fab8	<i>pLacZ-attB</i>	8
58		<i>1A2 responder</i>	mini-white	WI*, 1A2	<i>pLacZ-attB</i>	8
59		<i>Gl responder</i>	mini-white- dWI	Gl ^{REV}	<i>pLacZ-attB</i>	8
60		<i>HH-1.5</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	S2
61		<i>HH-2.1</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	S2
62		<i>HH-5.4</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	S2
63		<i>HB-1.6</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	S2
64		<i>BB-1.8</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	S2
65		<i>BB-3.1</i>	mini-white	WI*, Gl ^{FOR}	<i>pPelican</i>	S2

66		BB-2.3	mini-white	WI*, GlS ^{FOR}	pPelican	S2
67		BB-5.4	mini-white	WI*, GlS ^{FOR}	pPelican	S2
68		BB-1.3	mini-white	WI*, GlS ^{FOR}	pPelican	S2
69		BB-7.9	mini-white	WI*, GlS ^{FOR}	pPelican	S2
70		BB-1.7	mini-white	WI*, GlS ^{FOR}	pPelican	S2
71		BB-07	mini-white	WI*, GlS ^{FOR}	pPelican	S2
72		BP-3.2	mini-white	WI*, GlS ^{FOR}	pPelican	S2
73		PB-4.7	mini-white	WI*, GlS ^{FOR}	pPelican	S2
74		VGQ	mini-white	WI*, GlS ^{FOR}	pPelican	S2

*indicates WI present in the context of the mini-white marker

**GFP-E(spl)m8 and ep8-GFPm8 refer to the same construct

Table 2 List of primers

Name	Sequence (5' to 3')
ep8F	CGTCTAGAGGGGAATCTATTTTACAGCACAATCCAATAGGGG
ep8R	GCGGTACCCGGCTTGTTGCTGCCTGCTCG
ep7F	CGGCATGCGTCGCCAGAAAAATTGTAACGGCCC
ep7R	GCCTCGAGGAACTTCTTCGATCTTTCGGAGGAGG
p7F	GAGGCTAGCAGCTATAAAAGCAGCGGTAACC
L5R	TCAGACGATTCATTGGCACC
e8F	CGGCATGCGGGGAATCTATTTTACAGCACAATCCAATAGGGG
e8R	CTGGCTAGCTCCCTGGTCCCTGAAATCC
p8F	GAGGCTAGCGGTATAAAAGGACGGGACCTC
e7-2F	GGGAAGCTTGTCGCCAGAAAAATTGTAACGG
e7-2R	GGCGGATCCGTGCCGCCGCGAGAG
e8-2F	CGTCTAGAGGGGAATCTATTTTACAGCACAATCCAATAGGGG

e8-2R	CCCGCTAGCCGCTCCCTGGTCCCTGAAATCC
GIF	GGAGATCTGCATCACGTAATAAGTGTGCGT
GIR	GGAGATCTGCCGAGCACAATTGATCG
Fab8F	GGAGATCTGGGGAGGGAATTTCTTCA
Fab8R	GGAGATCTCATCTTCCGTTTCATCCGTT
1A2F1	GGCGGATCCACTACCAGGCAAGAAAGTAGGT
1A2F2	GGTGCGAATTCATACCAGGCAAGAAAGTA
1A2R	GGCGGATCCTATATGCTTCGTCTACCGTTGTG

Literature Cited

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

Barolo S., B. Castro, and J. W. Posakony, 2004 New *Drosophila* transgenic reporters: insulated P-element vectors expressing fast-maturing RFP. *Biotechniques* 36: 436–40, 442.

Berghammer a J., M. Klingler, and E. a Wimmer, 1999 A universal marker for transgenic insects. *Nature* 402: 370–371. <https://doi.org/10.1038/46463>

Bischof J., R. K. Maeda, M. Hediger, F. Karch, and K. Basler, 2007 An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. *Proc. Natl. Acad. Sci. U. S. A.* 104: 3312–7. <https://doi.org/10.1073/pnas.0611511104>

Delidakis C., and S. Artavanis-Tsakonas, 1992 The Enhancer of split [E(spl)] locus of *Drosophila* encodes seven independent helix-loop-helix proteins. *Proc. Natl. Acad. Sci. U. S. A.* 89: 8731–5. <https://doi.org/10.1073/PNAS.89.18.8731>

Groth A. C., M. Fish, R. Nusse, and M. P. Calos, 2004 Construction of transgenic *Drosophila*

by using the site-specific integrase from phage phiC31. *Genetics* 166: 1775–82.

Kim J., A. Sebring, J. J. Esch, M. E. Kraus, K. Vorwerk, *et al.*, 1996 Integration of positional signals and regulation of wing formation and identity by *Drosophila* vestigial gene. *Nature* 382: 133–138. <https://doi.org/10.1038/382133a0>

Kyrchanova O., D. Chetverina, O. Maksimenko, A. Kullyev, and P. Georgiev, 2008a Orientation-dependent interaction between *Drosophila* insulators is a property of this class of regulatory elements. *Nucleic Acids Res.* 36: 7019–7028. <https://doi.org/10.1093/nar/gkn781>

Kyrchanova O., S. Toshchakov, Y. Podstreshnaya, A. Parshikov, and P. Georgiev, 2008b Functional interaction between the Fab-7 and Fab-8 boundaries and the upstream promoter region in the *Drosophila* Abd-B gene. *Mol. Cell. Biol.* 28: 4188–4195. <https://doi.org/10.1128/MCB.00229-08>

Lai E. C., B. Tam, and G. M. Rubin, 2005 Pervasive regulation of *Drosophila* Notch target genes by GY-box-, Brd-box-, and K-box-class microRNAs. *Genes Dev.* 19: 1067–80. <https://doi.org/10.1101/gad.1291905>

Pfeiffer B. D., A. Jenett, A. S. Hammonds, T.-T. B. Ngo, S. Misra, *et al.*, 2008 Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 105: 9715–20. <https://doi.org/10.1073/pnas.0803697105>

Swevers L., L. Cherbas, P. Cherbas, and K. Iatrou, 1996 Bombyx EcR (BmEcR) and Bombyx USP (BmCF1) combine to form a functional ecdysone receptor. *Insect Biochem. Mol. Biol.* 26: 217–21.

Figure 1

Figure 1 *ep7* and *ep8* interact in *cis* and in *trans*. (**A** and **C**) *GFPm7* and (**B** and **D**) *GFPm8* expression in wing disk, eye-antennal disk and CNS (in left, middle and right columns, respectively) from the EGFP-tagged constructs shown in the diagrams on the left panel; red dotted rectangles highlight *ep7*-specific expression in the adult muscle precursors (AMPs) of the notum and in the midline of the ventral nerve cord (VNC); blue dotted rectangles indicate *ep8*-specific expression in the wing margin (WM) and in neuroblasts of the central brain and VNC. Blue and red arrows in the constructs' panel depict, respectively, *e7* and *e8* activities which are shared between *p7* (**A**) and *p8* (**B**) in the wing disk. Note lack of WM-specific expression by *ep7* (**C**) and lack of AMP-specific expression by *ep8* (**D**). Also note that both enhancers drive expression in some common areas, e.g. the eye morphogenetic furrow. (**E** and **F**) *cis*-expression patterns in wing disks dissected from hemizygous animals carrying *ep8-GFPm8* (**E**) or *ep7-lacZ* (**F**) transgenes at the *attP40* locus. (**G-J**) *ep8-GFPm8* and *ep-lacZ* crossed in the same animal. When present as heterozygotes at the same locus (**G** at *attP40*, **H** at *attP2*) both transgenes expand their pattern to locations dictated by their homologous transgene (marked with dotted rectangles). No such transvection is observed when the *ep8-GFPm8* and *ep7-lacZ* transgenes are placed in the same animal, but at non-homologous loci (**I** and **J**).

Figure 2

Figure 2 GIs and mini-*white* mediate transvection. Wing imaginal disks from animals heterozygous for variations of the *ep8-GFP-m8* and *ep7-lacZ* transgenes. The two GIs in forward orientation, flanking the two reporter, are indicated by black triangles; marker genes are located downstream of the second (3') GI: either the mini-*white* gene ('w') or the *3xP3-dsRed* cassette ('3xP3'). An image of the same disk from each genotype is split into two channels: *p8*-driven GFP-m8 in green and *p7*-driven LacZ in red. Dotted rectangles (red and blue) indicate cells (AMPs and WM, respectively) exhibiting transvection. Each combination of transgenes (in rows) is tested in the presence of Su(Hw) (*su(Hw)*^{+/−}, first column, **A-F**) and absence of Su(Hw) (*su(Hw)*^{−/−}, second column, **A'-F'**). All transgenes are inserted in *attP40* locus.

Figure 3

Figure 3 GIs, but not WIs, mediate transvection of *e7* and *3xP3* enhancers to a heterologous, enhancerless *hsp70* promoter (*pH*). Wing disks and CNSs from heterozygotes of *pH-gfp* with GIs (**A-F**) or without GIs (**G** and **H**). The second transgene in the homologous locus (*attP40*) is depicted on the side of each panel, the '+' sign denotes absence of a second transgene. The *pH-gfp* produces no expression as a hemizygote (**A**). (**B-F** and **H**) Projections of *trans* (GFP, green) and *cis* (LacZ, red and DsRed blue, when present) expression from the same wing disk and CNS of each sample containing an *ep7-lacZ* reporter are shown. Note that the observed GFP expression in wing discs of GI-less *pH-gfp* reporters (white arrow) originates from

associated tracheal branches in the vicinity of AMPs, and is an effect of trapping a trachea-specific enhancer by the *pH* promoter in *cis* in the absence of the insulating effect of the GIs (**G** and **H**) or their ‘inactivation’ in the *su(Hw)*^{-/-} background (**E** and **F**). Only combinations of transgenes with GIs in both homologs support transvection (**B** and **C**), while congruency in *white* in the absence of GIs does not (**H**). Consequently, depletion of Su(Hw) protein abolishes transvection observed from **B** and **C** combinations (**E** and **F**, respectively). Note a dotted pattern in the CNS in **C**, manifested by the 3xP3 activity in *cis* (DsRed, in blue) and in *trans* (GFP, in green), which comes from a glial cell population. The artificial 3xP3 enhancer does not drive expression in wing disks.

Figure 4

Figure 4 Transcriptional activity of different enhancer-promoter (*e-p*) pairs interacting in *cis* or in *trans* in the presence of GIs. Each genotype is examined for GFP expression in the third instar wing disk, eye disk and CNS. (**A-E**) Samples from hemizygotes with a transgene containing a given *e-p* pair linked in *cis* show higher levels of GFP expression than samples from heterozygotes containing same *e* and *p*, in a *trans* configuration (**F-J**). All transgenes are inserted in *attP40*.

Figure 5

Figure 5 Regulation of GI-mediated transvection by *cis*-preference. (A) The chart shows levels of basal and *e7*-induced (in *cis* and in *trans*) *pH*-driven luciferase activity. Levels of luciferase activity were measured from third instar larval disk-brain complexes. Luciferase

values normalized to total protein are shown as arbitrary units (a.u.). The mean and standard deviation of 5 replicates is shown. The activity of luciferase transgenes (1st column in the construct panel) is assayed on their own (as hemizygotes; horizontal line in the second column) or in combination with a second transgene in *trans*. **(B1-B7)** Transgenes containing *e8* or *e7*, with or without a *cis*-linked promoter (depicted on the right), were placed in *trans* to *pH-gfp*. For each genotype (each row) the third instar wing disk, eye disk and CNS were examined for (1) GFP expression, reflecting *trans* activity of the enhancer-containing transgenes on *pH-gfp* (green) and (2) β -galactosidase expression, reflecting *cis* activity of the enhancer-linked promoter, when *lacZ* is present (red). All transgenes contain GIs and mini-*white* and are inserted in *attP40*.

Figure 6

Figure 6 The relative position, number and orientation of GIs determine transvection effects. (A-P) Confocal z-projection of GFP expression in third instar larval wing disk (top left panel for each genotype), eye disk (top right) and CNS (bottom left). Merged *ep7*-driven LacZ (red) and -3xP3-*pH*-driven, DsRed expression (blue) in the same CNS (bottom right); note that the 3xP3 reporter shows no expression in the third instar eye or wing disk; for patterns of *ep7-lacZ* expression in disks, see previous figures. Each genotype contains a *pH-gfp* transgene with mini-*white* in trans to an *ep7-lacZ* transgene with 3xP3-dsRed and various arrangements of GIs, as indicated. Note, that D, E and I represent different samples obtained from the same genotype. All transgenes are inserted in *attP40*.

Figure 7

Figure 7 In dual GI transgenes, both 5' and 3' GIs participate in trans-interactions. Two tester transgenes, GI^{FOR}-*ep7-lacZ*-GI^{FOR}-*ep8-m8* (left column, **A-F**), and GI^{FOR}-*ep7-lacZ*-GI^{FOR}-*pH-gfp* (right column, **B, A'-F', G-I**) were tested as hemizygotes (**A, A'**), combined *inter se* (**B**) or combined with various other GI-containing transgenes in *attP40* (**C-I**). Third instar wing imaginal disks are shown for each genotype in two channels (GFP in green, LacZ in red). Note

stronger $e7 \rightarrow pH-gfp$ transvection in **D** and **D'** compared to **C** and **C'**. Also note that the ability of the 3' GI^{FOR} of both tester constructs to interact with a 5' GI^{FOR} of the *ep7-lacZ* reporter (**E**, **E'**) is severely compromised when a second GI is added to the *ep7-lacZ* reporter (**F**, **F'**). Interestingly this suppression is relieved by disabling the *p7* promoter in the latter construct (**G-I**).

Figure 8

Figure 8 1A2 and Fab8 insulators also mediate transvection. (**A1-G5**) Merged GFP and LacZ channels of confocal z-projections from third instar wing disks. (**A1-A5**) Wing disks from animals hemizygous for responder transgenes. (**B1-G5**) Wing disks from heterozygotes for sender transgenes (as indicated in each column) with responder transgenes (as indicated in each row). All transgenes are inserted in *attP40*. **A1-G5** are imaged at the same intensity settings. **A1'**, **B1'**, **E5'** and **G3'** are higher-sensitivity images of the respective panels, to reveal very low levels of transvection. Note that the responder construct in row **3**, as well as the sender constructs in columns **C** and **E**, are deleted for WI.

Figure S1

bioRxiv preprint doi: <https://doi.org/10.1101/565846>; this version posted March 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S1 *White*-mediated uni-directional transvection relies on mini-*white*-contained Wari insulator (WI). Three fragments of mini-*white*, cloned in the *ep7-lacZ(-3xP3)* construct, were tested at the *attP40* locus in transvection assay with a *ep8-EGFPm8* transgene containing full-sized mini-*white* (**A-C**). Neither promoter region of mini-*white* (*w*^{5'}, 0.24 kb, **A**), nor its ‘gene body’ (*w*^B, 2.4 kb, **B**) mediated transvection; whereas the 3’ part of mini-*white* (*w*^{3'}, 0.9 kb, **C**) did exhibit uni-directional transvection in WM (blue dotted rectangle). This fragment is known to contain WI. Addition of one or two GIs in the *ep8* construct inhibited WI-mediated transvection (**D-E**).

Figure S2

Figure S2 GI-mediated transvection is compatible with many developmental enhancers. A map of the *E(spl)* locus is shown on top and DNA fragments are marked as green rectangles with indicated length in kb. Each of these DNA fragments ('eX' in the constructs' scheme) was inserted into a transgene between GIs, placed in *trans* to GIs-flanked *pH-gfp* transgene in *attP40*. Heterozygotes were tested for activating *pH-gfp* expression in larval CNS (top), wing disk (middle) and eye-antennal disk (bottom row). Another fragment, unrelated to *E(spl)* locus, the vestigial quadrant enhancer (*vgQ*, last column) recapitulates its *cis* activity in *trans* (GFP expression only in the wing disk). We observe little or no activity from 3 *E(spl)* fragments in *trans*.

Figure S3

Figure S3 WI-mediated *trans*-activation by the *e8* enhancer requires the presence of the *e7* enhancer on the other homolog. All transgenes contain GIs and mini-*white* and are inserted in *attP40*. (A-F) Wing disks from animals hemizygous for the indicated GFP transgenes. When GFP is driven by the *pH* promoter (A), *E(spl)m7* (*p7*) promoter (B) and *E(spl)m8* (*p8*) promoter (C), no expression is observed; addition of the *e7* enhancer to any of these promoters (D-F) results in *e7*-specific expression pattern of the GFP; note the lack of WM expression. G-L show wing disks from animals heterozygous for the same *gfp* A-F transgenes and *ep8-lacZ*. M-R show the same transgene combinations in disks derived from *su(Hw)*^{-/-} mutants. Blue and red dotted rectangles highlight cells exhibiting *trans*-activity of *e7* (in

AMPs) and *e8* (in WM), respectively. In the absence of *e7*, only GI-mediated transvection is detected (**G,H,I**), whereas in the presence of *e7* both GI- (**J,K,L**) and WI-mediated (**P,Q,R**) transvection takes place. WI-mediated transvection is unidirectional, as it is only observed in the WM (GFP) of **P, Q, R**, but not in their AMPs (LacZ); compare to bidirectional transvection in **J, K, L**.

Figure S4

Figure S4 The onset of transvection is delayed in embryogenesis. All embryos are imaged ventrally with anterior to top. A *cis*-linked *e7pH* enhancer-promoter pair (top row) drives GFP expression along the ventral midline within 3 h after egg deposition (AED). At 4-5 h AED *e7pH* is broadly active in the ventral ectoderm, whereas *e7* and *pH* separated in *trans* show interaction only in small subset of these cells. From 7h AED onwards *e7* and *trans-pH* seem to interact in most cells where *e7(cis-)pH* is active. The enhancerless *pH-gfp* reporter shows no background expression in embryos as a hemizygote (data not shown).

Figure S5

Figure S5 The effects of mutations in the *p7* promoter on the ability of a cis-linked *e7* to transfect are independent of the identity of the *trans* promoter. The *ep7-lacZ* (A and E) and its derivatives harboring deletions of *p7* (B and F), $\Delta TATA$ (C and G) and ΔDPE (D and H) are placed in *trans* to *p7-gfp*- (A - D) or *p8-gfp*-enhancerless constructs (E - H). For each genotype (each row) the third instar wing disk, eye disk and CNS are simultaneously examined for (1) GFP, reflecting *trans* activity of the enhancer-containing transgenes on *pH-gfp* (green) and (2) β -galactosidase, reflecting *cis* activity the *e7*-linked promoter (*lacZ*, in red). Although, *p7*, *p8* and *pH* are promoters of different strength, their activity seems to be

affected similarly by different mutations in *e7*-linked *trans-p7* (compare to the results obtained for *pH*, **Figure 5, D-G**). All transgenes contain GIs and mini-*white* and are inserted in *attP40*.

Figure S6

Figure S6 *pH* receives input from two enhancers in *trans*, *e7* and *3xP3*. Confocal z-projection of a third instar CNS from a heterozygote between *ep7-lacZ-3xP3* and *pH-gfp* dual-GIs^{FOR} transgenes in *attP40*. Top left panel shows *trans*-activated *pH-gfp* expression; note *e7*-specific GFP expression in the VNC midline, corresponding to the *cis*-activity of *e7* (driving expression of LacZ, in red, top right panel), and the 'dotty' *3xP3*-specific expression in central brain and VNC corresponding to glial cells with active *3xP3* enhancer (expressing DsRed, in blue, in *cis*). Merged image is shown in the bottom right panel.

Figure S7

Figure S7 No transvection is observed in the absence of GIs at five different loci. The ability to interact in *trans* was tested between *ep7-lacZ* and *pH-gfp* transgenes inserted in five genomic loci (*attP40*, *VK2*, *VK13*, *VK37*, *VK40*) in two configurations of these transgenes: carrying no GIs (**A-E**) and containing one 5' GI each (**A'-E'**). For each genotype three tissues are shown: wing disk, eye disk and CNS (left to right). Green is *pH*-driven GFP. Red is *ep7*-driven LacZ. Blue is *3xP3*-driven DsRed (CNS only). Note the activity of neighboring (trapped) enhancers in all loci except *VK40*, which is also seen in the hemizygous condition for these transgenes (not shown): (1) tracheal activity in the wing disk of the uninsulated *pH-gfp* in *attP40* and *VK2* and of the GI-insulated *pH-gfp* in *VK37* (**A**, **B** and **D'**), (2) ubiquitous activity in VNC and central brain of the uninsulated *pH* in *VK2* and *VK13* (**B** and **C**), (3) wing pouch activity in wing disk of the uninsulated *p7-lacZ* in *VK2* (**B**), (4) ubiquitous wing disk activity of uninsulated and GI-insulated *p7-lacZ*, and GI-insulated *pH* in *VK13* (**C** and **C'**). Note GFP expression in notum AMPs, retina and antenna and brain and VNC midline in **A'-E'**. All this *ep7*-driven GFP expression is absent in **A-E**.

Figure S8

Figure S8 Less strict GI position and orientation requirements for interaction between *ep7* and *ep8*. **(A1-E5)** *ep8*- and *ep7*-containing single-GI transgenes in *attP40* are combined as shown. Each panel consists of six sub-panels containing confocal z-projections of the WM (only LacZ channel, in red) or the AMPs (only GFP channel, in green) from three different wing disks for each genotype. Both of these expressions result from transvection and are abrogated when one or both homologs lack a GI (row **E** or column **5**). Note, that unlike *ep7*→*pH* transvection (**Figure 6 H**), *ep8*→*ep7* transvection (LacZ in the WM) is also mediated by the 3' GIs (**C3, C4, D3, D4**). The 3' GIs also mediate weak *ep7*→*ep8* transvection (GFP in AMPs), when congruently oriented (**C3, D4**). Note, also, that the absolute orientation of 5' GIs is less important for *ep7*→*ep8* transvection (compare GFP in **A1, A2, B1, B2**) than for *ep7*→*pH* (**Figure 6I, L**).

Figure S9

Figure S9 The *3xP3* enhancer has a stronger affinity for *pH* than for two other promoters. (**A-C**) Confocal z-projections of third instar CNSs showing GFP channel alone (green) and merged channels of GFP (green), LacZ (red) and DsRed (blue). (**A'-C'**) Enlarged regions of the central brains indicated by white rectangles in **A-C**. Out of the three enhancerless dual-GI^{FOR} *gfp* transgenes, each carrying a different promoter (**A** - *pH*, **B** - *p7*, **C** - *p8*), only the *pH*-driven reporter shows robust interaction with *3xP3* in the dual-GIs^{FOR} *ep7-lacZ-3xP3* transgene in *attP40*, giving a dotted glial GFP pattern. Notice weaker and sporadic GFP expression from *trans-3xP3* by *p7* (**B**, **B'**), even though *p7* is a stronger promoter than *pH*. In contrast, the "vertical" interaction of *e7* with the trans promoter (GFP in VNC midline and optic lobes) is stronger with *p7* (**B**) than with *p8* or *pH* (**A**, **C**).