

Pawel Piwko

Regulation of gene expression in *cis* and in *trans*



Doctoral Dissertation for the Doctor of Philosophy Degree, Department of Biology, University of Crete, performed in the Institute of Molecular Biology & Biotechnology, Foundation for Research & Technology Hellas, Heraklion, Greece, under supervision of **Christos Delidakis**

Heraklion, 2019

PhD supervisor:

Prof. **Christos Delidakis** (Department of Biology, University of Crete)

PhD advisors:

Prof. **Charalampos G. Spilianakis** (Department of Biology, University of Crete)

Prof. **Kriton Kalantidis** (Department of Biology, University of Crete)

PhD committee:

Prof. **Christos Delidakis** (Department of Biology, University of Crete)

Prof. **Charalampos G. Spilianakis** (Department of Biology, University of Crete)

Prof. **Kriton Kalantidis** (Department of Biology, University of Crete)

Prof. **Despina Alexandraki** (Department of Biology, University of Crete)

Prof. **Joseph Papamatheakis** (Department of Biology, University of Crete)

Prof. **Dimitrios Tzamarias** (Department of Biology, University of Crete)

Prof. **George A. Garinis** (Department of Biology, University of Crete)

"Pure self-recognition in absolute otherness, this Aether as such, is the ground and soil of Science or knowledge in general"

Georg Wilhelm Friedrich Hegel

Table of contents

1. Preface	1
2. Chapter 1 Negative regulation of <i>E(spl)</i> genes	1
2.1. Abstract	1
2.2. Introduction	2
2.2.1. <i>E(spl)</i> and Notch	2
2.2.2. Regulation of <i>E(spl)</i> gene expression	4
2.3. Objectives	6
2.4. Results	7
2.4.1. Expression of <i>E(spl)m7</i> and <i>E(spl)m8</i>	7
2.4.1.1. Wing disk	8
2.4.1.2. Eye-antennal disk	10
2.4.1.3. CNS	11
2.4.1.4. Embryo	13
2.4.1.5. Pupa	15
2.4.2. Analysis of enhancers of <i>E(spl)m7</i> and <i>E(spl)m8</i>	16
2.4.3. Analysis of 3' UTRs of <i>E(spl)m7</i> and <i>E(spl)m8</i>	22
2.4.4. Impact of transcriptional and translational repression on <i>E(spl)m7</i> and <i>E(spl)m8</i> expression and function	28
2.5. Discussion	45
3. Chapter 2. The role of insulators in transgene transvection in <i>Drosophila</i>	50
3.1. Abstract	50
3.2. Introduction	50
3.2.1. Homologous pairing in mammals	51
3.2.2. Homologous pairing and transvection in <i>Drosophila melanogaster</i>	52
3.2.3. Determinants and dynamics of homologue pairing	53
3.2.4. Regulators of homologous pairing	57
3.2.5. Insulators	58
3.3. Objectives	59
3.4. Results	60
3.4.1. Enhancer element analysis of <i>E(spl)m7</i> and <i>E(spl)m8</i> reveals a transvection phenomenon	60
3.4.2. Transvection is mediated by homotypic interactions between GIs or Wari insulators	61
3.4.3. GI-mediated transvection is promiscuous, whereas WI-mediated transvection is highly selective	65
3.4.4. Transvection is weaker than cis enhancer-promoter activity and is suppressed by promoter cis-preference	69
3.4.5. Relative position, number and orientation of GIs determine transvection effects.	74
3.4.6. Other insulators also mediate transvection	88
3.5. Discussion	93
4. Materials and Methods	98
4.1. Plasmid constructs	98
4.2. Fly maintenance and stocks	107
4.3. Integration of attB plasmids into attP fly lines	108
4.4. Immunostaining and microscopy	108

4.5. Luciferase assays	109
5. Acknowledgments	109
6. Literature Cited	110

List of Figures

Figure 1 Organization of the <i>E(spl)</i> locus.	2
Figure 2 <i>Cis</i> -regulatory elements of <i>E(spl)m7</i> and <i>E(spl)m8</i> .	4
Figure 3 <i>e7p7</i> and <i>e8p8</i> interact in <i>cis</i> .	7
Figure 4 <i>e8</i> but not <i>e7</i> is active in most of the PNCs of the late third instar wing disk.	10
Figure 5 <i>e7p7-lacZ</i> expression pattern in eye-antennal disk.	11
Figure 6 Expression of <i>E(spl)m7</i> and <i>E(spl)m8</i> in larval CNS.	13
Figure 7 Expression of <i>E(spl)m7</i> and <i>E(spl)m8</i> in embryo.	14
Figure 8 <i>e7</i> and <i>e8</i> activities in early pupal stages.	15
Figure 9 Evolutionary conservation of the <i>E_B</i> - and <i>E_C</i> -boxes of <i>e7</i> (B1) and <i>e8</i> (B2, C1, C2).	17
Figure 10 Validation <i>in vivo</i> of the <i>cis</i> -regulatory elements identified within enhancers of <i>E(spl)m7</i> and <i>E(spl)m8</i> .	20
Figure 11 <i>E(spl)m7</i> and <i>E(spl)m8</i> are repressed by the products of <i>E(spl)</i> locus.	21
Figure 12 Identified microRNA seed boxes within the <i>E(spl)m7</i> 's and <i>E(spl)m8</i> 's 3' UTRs.	23
Figure 13 Identified microRNA binding sites within <i>E(spl)m7</i> and <i>E(spl)m8</i> 3' UTRs mediate repression.	24
Figure 14 The 3'UTR of <i>E(spl)m7</i> , and not that of <i>E(spl)m8</i> , is under control of microRNAs.	26
Figure 15 <i>bantam</i> microRNA targets <i>E(spl)m7</i> 3'UTR.	27
Figure 16 Notch activity impacts 3' UTR-mediated regulation of <i>E(spl)m7</i> and <i>E(spl)m8</i> .	28
Figure 17 <i>E(spl)m7</i> expression is regulated by microRNAs and transcriptional repressors.	29
Figure 18 GFPm7 is detected in all PNCs of the wing disk upon mutating all <i>E_B</i> and <i>E_C</i> boxes and identified microRNA binding sites within <i>E(spl)m7</i> 3' UTR.	30
Figure 19 GFPm7 is accumulated with increasingly higher levels in the WM, AMPs and PNCs of the wing disk as the transcriptional negative regulation and microRNA regulation is compromised.	32
Figure 20 Misregulation of <i>E(spl)m7</i> and <i>E(spl)m8</i> expression impacts SOP specification.	33
Figure 21 Misregulation of <i>E(spl)m7</i> and <i>E(spl)m8</i> induces SC/DC bristle loss.	35
Figure 22 Effect of transgene dosage on SC/DC bristle loss.	36
Figure 23 <i>E(spl)</i> locus copy number affects level-dependent activity of <i>E(spl)m7</i> and <i>E(spl)m8</i> .	38
Figure 24 Heterozygous BX22 deletion does not affect level-dependent activity of <i>E(spl)m7</i> and <i>E(spl)m8</i> .	39
Figure 25 <i>E(spl)m7</i> 3' UTR-mediated regulation, but not SOP specification is affected in the <i>dcr1</i> null clones.	40
Figure 26 Loss of <i>E(spl)</i> locus affects specification and patterning of SOPs and bristles.	41
Figure 27 Loss of <i>E(spl)</i> stimulates <i>GFPm7</i> expression from genomic transgenes.	43
Figure 28 Effects of <i>E(spl)m7-E(spl)m8</i> genomic transgenes on patterning of thorax bristles in mutant clones of <i>E(spl)</i> .	45
Figure 29 <i>e7p7</i> and <i>e8p8</i> interact in <i>trans</i> .	61
Figure 30 GIs and mini- <i>white</i> mediate transvection.	63

Figure 31 <i>White</i> -mediated uni-directional transvection relies on the mini- <i>white</i> -contained Wari insulator (WI).	64
Figure 32 GIs, but not WIs, mediate transvection of <i>e7</i> and <i>3xP3</i> enhancers to a heterologous, enhancerless <i>hsp70</i> promoter (<i>pH</i>).	66
Figure 33 GI-mediated transvection is compatible with many developmental enhancers.	67
Figure 34 WI-mediated <i>trans</i> -activation by the <i>e8</i> enhancer requires the presence of the <i>e7</i> enhancer on the other homolog.	68
Figure 35 Transcriptional activity of different enhancer-promoter (<i>e-p</i>) pairs interacting in <i>cis</i> or in <i>trans</i> in the presence of GIs.	70
Figure 36 The onset of transvection is delayed in embryogenesis.	70
Figure 37 Regulation of GI-mediated transvection by <i>cis</i> -preference.	72
Figure 38 The effects of mutations in the <i>p7</i> promoter on the ability of a <i>cis</i> -linked <i>e7</i> to transfect are independent of the identity of the <i>trans</i> promoter.	74
Figure 39 <i>pH</i> receives input from two enhancers in <i>trans</i> , <i>e7</i> and <i>3xP3</i> .	75
Figure 40 Genomic maps of insulators.	77
Figure 41 GIs stimulate expression in <i>cis</i> independently of their orientation.	78
Figure 42 The relative position, number and orientation of GIs determine transvection effects.	79
Figure 43 No transvection is observed in the absence of GIs at five different loci.	80
Figure 44 Genomic maps of transgene insertions used in Figure 43.	84
Figure 45 Less strict GI position and orientation requirements for interaction between <i>e7p7</i> and <i>e8p8</i> .	85
Figure 46 In dual GI transgenes, both 5' and 3' GIs participate in <i>trans</i> -interactions.	86
Figure 47 The <i>3xP3</i> enhancer has a stronger affinity for <i>pH</i> than for two other promoters.	88
Figure 48 1A2 and Fab8 insulators also mediate transvection.	90
Figure 49 Relative orientation of GIs in <i>trans</i> impacts strength of transvection but not enhancer choice.	92
Figure 50 Parameters of transvection.	93

1. Preface

Functions which guide development, maintenance and reproduction of living organisms, together with functions which can lead to their death, are accomplished by the activity of the factors encoded in their DNA. A success of conveying these functions is decided at the interface between spatio-temporal abundance of these factors and environmental cues. DNA transcription is the first step of delimiting their abundance. Thereupon, transcribed factors are subject to (post-transcriptional) regulation, and if translated, are further (post-translationally) modified, both of which shape their functionality and stability. This profuse regulation of gene expression indicates molecular complexity and underlies diversity and robustness of living organisms.

Understanding general laws driving life requires us to study diverse species and different forms of their organismal complexity. One of the most valuable model of these studies has been a fruit fly, *Drosophila melanogaster*. This manuscript contains results from the study performed on this species *in vivo* aimed at determining DNA features through which certain aspects of gene expression and its regulation are mediated. In the first chapter I characterize enhancer, promoter and 3' untranslated (3'-UTR) elements and assess their role in driving and regulating expression of two *Enhancer of split* (*E(spl)*) genes, *Espl(m7)* and *Espl(m8)*. In the second part, I identify sequences and their attributes needed to support transcriptional activation of a promoter by an enhancer located on the homologous chromosome.

2. Chapter 1 Negative regulation of *E(spl)* genes

2.1. Abstract

bHLH-Orange proteins, like the mouse Hes and the *Drosophila* Enhancer of split [*E(spl)*] proteins are transcription factors that, among a multitude of functions, are implicated in maintaining the undifferentiated state in many stem-cell based systems. For instance, they antagonize bHLH proneural activators and suppress neural fate. Their levels have to be exquisitely regulated to achieve a precise balance between proliferation and differentiation. For this reason they are subject to multi-tiered regulation.

During my PhD research I have tried to understand how the expression of two neighboring *E(spl)* genes, *E(spl)m7* and *E(spl)m8*, is regulated and how this regulation impacts biological processes controlled by their protein products. I have provided description of their enhancer, promoter and 3' untranslated elements and scrutinized the role of these elements in specifying spatio-temporal abundance of the *E(spl)m7* and *E(spl)m8* proteins. I demonstrate that compromising transcriptional negative autoregulation and translational repression can

have dramatic consequences on the expression levels and the activity of these two genes. Thus, both regulatory modes are essential in determining level-dependent functions of *E(spl)m7* and *E(spl)m8*.

2.2. Introduction

2.2.1. *E(spl)* and Notch

Enhancer of split was identified as a spontaneous dominant mutation which enhanced duplicated bristle phenotype of the recessive *Notch* allele *split* (Welshons 1956). Subsequent analysis distinguished twelve protein-coding genes spanning 60 kb of the *E(spl)* locus (**Figure 1**) (Knust *et al.* 1987; Klämbt *et al.* 1989; Delidakis *et al.* 1991; Delidakis and Artavanis-Tsakonas 1992a; Knust *et al.* 1992). Seven of these genes *E(spl)mδ*, *E(spl)mγ*, *E(spl)mβ*, *E(spl)m3*, *E(spl)m5*, *E(spl)m7* and *E(spl)m8* constitute a group of structurally related genes within the *E(spl)*-complex (*E(spl)*-C) (Klämbt *et al.* 1989; Delidakis *et al.* 1991; Delidakis and Artavanis-Tsakonas 1992b; Knust *et al.* 1992). They encode paralogous basic helix-loop-helix (bHLH) transcription repressors with an “orange domain” (bHLH-O proteins), and together with four other *Drosophila* genes (*hairy*, *deadpan*, *similar-to-deadpan*, and *her*) form the group of Hes genes (*hairy* and *enhancer-of-split*) (Delidakis *et al.* 2014). Additionally, Hes proteins contain a stereotypic terminus tetrapeptide with the amino acids tryptophan, arginine, proline, tryptophan (WRPW) through which they interact with global co-repressor Groucho (Gro) (Paroush *et al.* 1994; Jennings *et al.* 2006), and, consequently, through which they exert their repressive function (Giebel and Campos-Ortega 1997; Giagtzoglou *et al.* 2003, 2005).

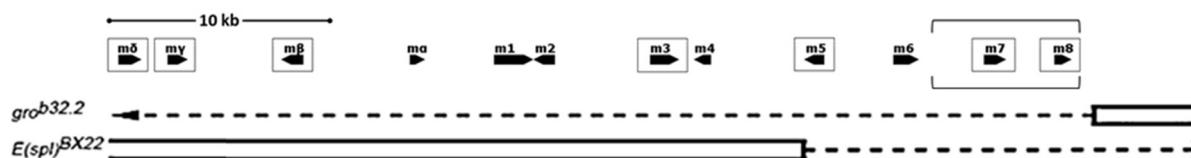


Figure 1 Organization of the *E(spl)* locus. The coding regions of *E(spl)* bHLH genes are boxed. The region of interest for this study encompassing *E(spl)m7* and *E(spl)m8* is indicated with brackets. Below, the span of two deletions of the *E(spl)* locus used in this study is indicated with dashed line.

Loss of the *E(spl)* locus results in hyperplasia of the embryonic nervous system, so called “the neurogenic phenotype”, thus indicating its important role in embryonic neurogenesis (Knust *et al.* 1987; Preiss *et al.* 1988; Delidakis *et al.* 1991; Delidakis and Artavanis-Tsakonas 1992b; Schrons *et al.* 1992; Nagel *et al.* 2004). It was later shown that this phenotype is due to the lack of the activity of the bHLH genes of *E(spl)*-C and that these genes are transcriptionally activated by Notch intracellular domain (NICD) (Bailey and Posakony 1995; Lecourtois and Schweisguth 1995). Notch signalling plays a central role during animal development endowing adjacent cells with a communication system to coordinate their fates with those of their neighbours (Artavanis-Tsakonas *et al.* 1995). In a simplistic model, the Notch transmembrane receptor is triggered in one cell via direct intercellular contact by the

binding of a ligand protein contained on the membrane of the other cell. This activation leads to the cleavage of the NICD and its translocation to the nucleus where it interacts with DNA-bound Suppressor of Hairless protein (Su(H), CSL in mammals) and the co-activator Mastermind (Mam) to stimulate transcription of target genes (Bray 2006). Thus, genes primed for activation are pre-bound by Su(H) which acts in the absence of NICD as a repressor (Furriols and Bray 2001). Genes of the *E(spl)-C* are not the only targets of Notch but by far the best-characterized and the ‘fastest’ responders of Notch activity; their induction can be detected within 10-30 min of Notch activation (Krejci and Bray 2007; Housden *et al.* 2013). The *E(spl)* genes are small, intronless and their products exhibit very short half-lives (Delidakis *et al.* 2014; Kiparaki *et al.* 2015). These features allow them to react quickly to activation and cessation of Notch signaling (Delidakis *et al.* 2014).

During the process of “lateral inhibition”, the selection of a single neural precursor from a group of equipotent cells relies upon such intercellular communication and activity of the Notch and *E(spl)* bHLH proteins (Artavanis-Tsakonas *et al.* 1995). In the developing embryo lateral inhibition singles out neural precursor cells from neuroepithelial cells (Artavanis-Tsakonas *et al.* 1999). Deficiencies of the *E(spl)* locus result in embryonic lethality due to overcommitment of neural precursor cells and concomitant loss of ventral and cephalic epidermis (Lehmann *et al.* 1983; Delidakis *et al.* 1991). This neurogenic phenotype exhibits wide range of severity which roughly depends on the size of *E(spl)* deficiency due to functional redundancy of the seven *E(spl)-C* genes (Delidakis *et al.* 1991; Schrons *et al.* 1992).

A situation analogous to that for embryonic neuroblast (and epidermal precursors) specification takes place in the developing wing imaginal disk to single out sensory organ precursor (SOP) cells, from which adult peripheral nervous system (PNS) organs arise (Ghysen *et al.* 1993; Modolell and Campuzano 1998). A single SOP is selected among a cluster of competent cells by lateral inhibition to give rise to a bristle sensillum of the imago of *Drosophila* (Huang *et al.* 1991). SOP fate depends on the expression of the proneural bHLH transcriptional activators, encoded by the “proneural” genes of the *Achaete-Sute-Complex* (AS-C). Initially, AS-C genes are expressed within the so called “proneural cluster” (PNC), a group of 20-30 cells from which one or few SOPs will be selected (Gómez-Skarmeta *et al.* 2003). Notch signalling is also active in all the cells of PNC, however, with the lowest level of expression of its target genes in the presumptive SOP (Jennings *et al.* 1994, 1995). This correlates with spatially inversed expression pattern of the proneural bHLH factors, elevated within SOP (Jennings *et al.* 1994, 1995). Proneural transcription factors trigger expression of the proneural genes as well as of *Delta* (*Dl*) - a ligand of Notch receptor (Heitzler *et al.* 1996). In this way, SOP cell activates Notch signalling (and consequently *E(spl)* genes) in the neighbouring non-SOP PNC cells (Simpson 1990; Cabrera 1990; Jennings *et al.* 1994; Bailey and Posakony 1995; Lecourtois and Schweisguth 1995; de Celis *et al.* 1996). The *E(spl)* bHLH factors prevent non-SOP PNC cells from becoming a SOP by reducing the activity of the bHLH proneural factors and by repressing transcription of the proneural genes (Lecourtois and Schweisguth 1995; Heitzler *et al.* 1996; Giagtzoglou *et al.* 2003; Miller *et al.* 2014). Thus,

acquiring neural cell fate appears to be an outcome of the interplay between the proneural bHLH activators and the “anti-neural” *E(spl)* bHLH repressors.

2.2.2. Regulation of *E(spl)* gene expression

The seven *E(spl)* bHLH genes exhibit similar spatial pattern of expression during embryogenesis and high homology in the sequence of the proteins they encode (Kläämbt *et al.* 1989; Delidakis and Artavanis-Tsakonas 1992b; Knust *et al.* 1992; Wech *et al.* 1999). Together with the observations that these proteins might functionally replace each other and more than one has to be mutated in order to achieve the complete elimination of the *E(spl)* function in neurogenesis, led to conclusion that the *E(spl)* locus originated as a result of gene duplication and that its members are functionally redundant (Kläämbt *et al.* 1989; Delidakis and Artavanis-Tsakonas 1992b; Schrons *et al.* 1992). However, the entire *E(spl)* gene complex exhibits high conservation among other *Drosophila* species, indicating that all of the genes as well as their organization are of functional importance (Maier *et al.* 1993). Although, *E(spl)* bHLH genes are transcribed in a nearly identical pattern during early embryogenesis, their expression in late embryonic, larval and imaginal stages appears to be regulated by different sets of tissue specific factors (de Celis *et al.* 1996; Wech *et al.* 1999).

Sequence analysis of the promoter-proximal upstream regions of *E(spl)* genes revealed number of *cis*-regulatory elements and their high evolutionary conservation (Nellesen *et al.* 1999). Three classes of Su(H) binding sites have been identified: (1) single sites (TATGGGAA), (2) Su(H)-paired sites (SPS) – a configuration of two single Su(H) sites in opposite orientation spaced by 17 bp (as in the promoter-proximal upstream region of *E(spl)m7*, see **Figure 2**), and (3) SPS sites with closely linked binding site for the proneural bHLH activators – the so called E_A-box (SPS+A; as in the promoter-proximal upstream region of *E(spl)m8*, see **Figure 2**) (Bailey and Posakony 1995; Nellesen *et al.* 1999). Su(H) binding sites integrate activation (via Notch) and repression (in the absence of Notch signal) of *E(spl)* genes (Furriols and Bray 2001). Consequently, the expression of *E(spl)* bHLH genes is lost upon depletion of Notch signal, while overactivity of Notch results in their elevated and ectopic expression (Jennings *et al.* 1994, 1995; Bailey and Posakony 1995; Lecourtois and Schweisguth 1995; de Celis *et al.* 1996; Wurmbach *et al.* 1999; Nellesen *et al.* 1999; Lai *et al.* 2000).

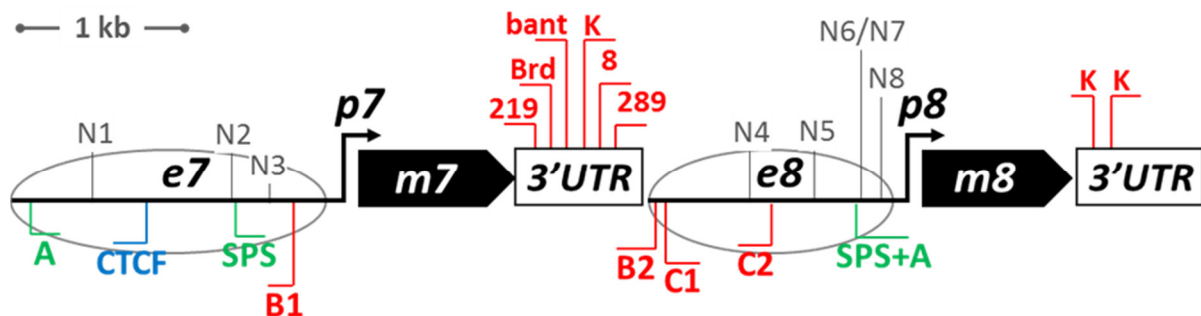


Figure 2 *Cis*-regulatory elements of *E(spl)m7* and *E(spl)m8*. Putative enhancers of *E(spl)m7* and *E(spl)m8* – *e7* and *e8*, respectively, are indicated with horizontal lines; promoters of these genes (*p7* and *p8*) are shown as bent

arrows, CDSs as black pentagons and 3'UTRs as rectangles. Identified regulatory sequences are marked with bent lines. **SPS** denotes Su(H) paired site; **A** – proneural binding site class E_A box, **B1-B2** and **C1-C2** mark two classes of bHLH repressor binding sites (class E_B box and class E_C box, respectively); CTCF denotes putative CTCF insulator; **K** and **Brd** in 3' UTRs indicate specific microRNA-binding motifs recognized by miR2, miR6, miR11 and miR13 (K-box) and miR4 and miR79 (Brd-box). **219**, **bant**, **8** and **289** denote putative seed regions recognized by miR219, bantam, miR8 and miR289, respectively.

Diverse expression patterns of various *E(spl)*-C genes in late embryonic and larval tissues depends on, in addition to Notch, other spatiotemporally restricted transcription factors acting in a context specific manner. For example, expression of *E(spl)* in adult muscle precursors in the developing wing disk requires a combination of Notch and Twist activities (Bernard *et al.* 2010). In the wing disk's PNCs, however, this expression depends also on the presence of proneural bHLH AS-C factors (Kramatschek and Campos-Ortega 1994; Hinz *et al.* 1994; Singson *et al.* 1994; Jennings *et al.* 1995; Nellesen *et al.* 1999; Cave *et al.* 2009). Consequently, most of enhancer elements of the *E(spl)* genes contain binding sites for proneural bHLH activators, CASCTG - so called E_A-boxes (as in the upstream region of *E(spl)m7* and, closely linked to SPS, of *E(spl)m8*, see **Figure 2**) (Nellesen *et al.* 1999; Cave *et al.* 2009).

Besides Notch and proneural activation, the expression of *E(spl)* genes is further modulated by a complex transcriptional and post-transcriptional regulatory network to ensure proper outcome of their level-dependent function. *E(spl)* bHLH-O proteins are thought to autorepress (Delidakis *et al.* 2014). In mice the *E(spl)* homolog *Hes7* is expressed in an oscillatory expression manner depending on negative autoregulation: the *Hes7* transcription factor represses its own gene by binding to its promoter (Bessho *et al.* 2003). This repression is short-lived due to the short half-life of the *Hes7* mRNA and protein. In this way, *Hes7* autonomously initiates oscillatory expression, which is crucial for the correct timing of the biological clock controlling the formation of somites (Bessho *et al.* 2003; Kageyama *et al.* 2007). Another *E(spl)* mouse homolog – *Hes1* also oscillates due to negative feedback and regulates cell proliferation and differentiation of the nervous system (Kageyama *et al.* 2007). There are several *E(spl)* homologs in other vertebrates showing a similar fashion of expression (zebrafish *her1* and *her7*, chick *chairy1*), suggesting a crucial role of the negative autoregulation of the *E(spl)* gene expression in controlling important developmental processes (Kageyama *et al.* 2010). Although not directly demonstrated in *Drosophila*, the dynamic pattern and the ample presence of the binding sites for bHLH-O repressors in the enhancers of *E(spl)* genes implies their autoregulation. bHLH-O factors were shown to bind with high affinity to the B-type E-box type sequences (CACGTG) and its variant E_C-box (CACGCG), and with lower affinity to E_N-boxes (CACNAG) (Tietze *et al.* 1992; Oellers *et al.* 1994; Ohsako *et al.* 1994; Van Doren *et al.* 1994; Jennings *et al.* 1999; Delidakis *et al.* 2014). Four of the high affinity binding sites are also present in the regulatory regions of *E(spl)m7* and *E(spl)m8* (one E_B-box, B1, in the promoter-proximal upstream region of *E(spl)m7*, and, one E_B-box, B1, and two E_C-boxes, C1-C2, in the upstream region of *E(spl)m8*, see **Figure 2**). However, most of the E-boxes in this region are represented by the lower affinity sites, E_N-boxes (N1-N3 upstream of *E(spl)m7* and N4-N8 upstream of *E(spl)m8*, see **Figure 2**).

In addition to their complex transcriptional regulation, *E(spl)* transcript levels are controlled by microRNAs binding to their 3' untranslated regions (UTRs) (Leviten *et al.* 1997; Lai and Posakony 1997; Lai *et al.* 1998, 2005). Sequence analysis of their 3' untranslated regions (UTRs) and subsequent mutagenesis studies revealed highly evolutionary conserved motifs, so called K-boxes (UGUGAU, recognized by miR2, miR6, miR11 and miR13; present in the 3' UTRs of *E(spl)m7* and *E(spl)m8*, see **Figure 2**) and Brd-boxes ((AGCUUUA, recognized by miR4 and miR79; present in the 3' UTRs of *E(spl)m7*, see **Figure 2**), and their important role in negative post-transcriptional regulation of Notch target genes expression (Lai *et al.* 2005; Duncan and Dearden 2010; Bejarano *et al.* 2012).

2.3. Objectives

This study aims to characterize the contribution of transcriptional and post-transcriptional negative regulatory elements in the expression of the *E(spl)* genes. Out of 12 genes residing within the *E(spl)* locus, two of them were chosen, *E(spl)m7* and *E(spl)m8*. Both of these genes exhibit a similar expression pattern; during early neurogenesis expression is confined to ventral ectoderm (neuroectoderm) and in imaginal stages they are expressed in nearly all proneural clusters in the wing discs (Knust *et al.* 1992; Kramatschek and Campos-Ortega 1994; Singson *et al.* 1994; Jennings *et al.* 1994; Bailey and Posakony 1995; de Celis *et al.* 1996; Wech *et al.* 1999; Nellesen *et al.* 1999; Cooper *et al.* 2000). This suggests that both genes are essential for lateral inhibition within proneural clusters in both embryo and larvae. *E(spl)m7* and *E(spl)m8* form a neighbouring pair of genes within the *E(spl)*-C (**Figure 1**). The fact that the *E(spl)*-C has developed by gene duplication and the similar or overlapping pattern of expression of its genes may suggest the possibility of sharing common regulatory elements.

Current knowledge about processes regulating *E(spl)* expression comes from the artificial expression systems where dynamism and sensitivity of naturally occurring *E(spl)* expression is lost. I wanted to describe naturally occurring repressional events and their importance in biological processes regulated by *E(spl)* by the use of an intact genomic fragment containing two *E(spl)*-C genes. The approach of this research is to (1) dissect and describe spatio-temporal activity of the *cis*-regulatory elements necessary for *E(spl)m7* and *E(spl)m8* expression, and, to (2) test *in vivo* the function of the *E(spl)m7* and *E(spl)m8* genes with introduced mutations within their enhancer and 3' UTR regions in the specification of sensory neural precursors of the adult peripheral nervous system (PNS).

2.4. Results

2.4.1. Expression of *E(spl)m7* and *E(spl)m8*

In order to characterize the *cis*-regulatory elements patterning expression of the *E(spl)m7* and *E(spl)m8* I 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 3 B**) and *m7-GFPm8* (EGFP fused to the *E(spl)m8* ORF, **Figure 3 C**). Due to strong post-transcriptional repression of both of these genes (Lai *et al.* 2005) I have replaced *E(spl)m7* and *E(spl)m8* 3' UTRs with the SV40 and the Adh poly A terminators, respectively. 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 3 B, C**). 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.

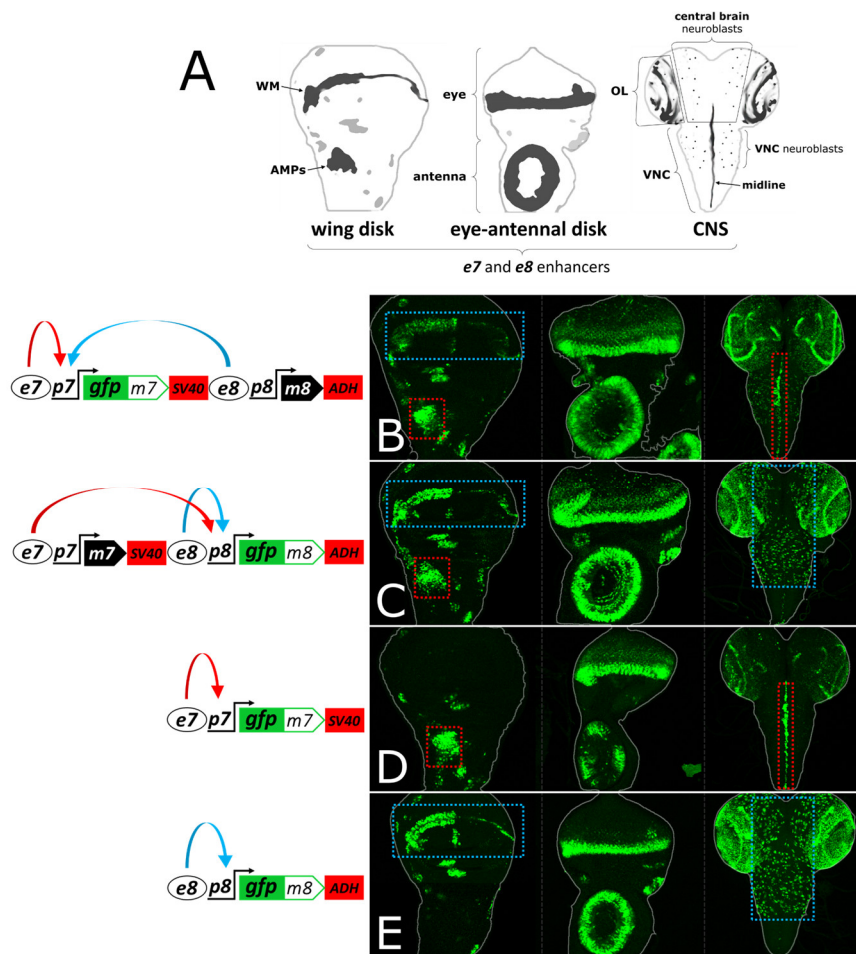


Figure 3 *e7p7* and *e8p8* interact in *cis*. (A) Schematics of a wing disk, an eye-antennal disk and two central nervous systems (CNSs) with the areas of *m7* and *m8* expression marked in shades of black. WM: wing margin;

AMPs: adult muscle precursors; OL: optic lobe; VNC: ventral nerve cord. (**B** and **D**) *GFPm7* and (**C** and **E**) *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 *e7*-specific expression in the AMPs and in the midline of the VNC; blue dotted rectangles indicate *e8*-specific expression in the WM and in neuroblasts of the central brain and VNC. Also note that both enhancers drive expression in some common areas, e.g. the eye morphogenetic furrow. In the constructs' schematics enhancers are shown as ovals (*e7* and *e8*), promoters as bent arrows (*p7* and *p8*) and insulators as triangles: black triangle: *gypsy* insulator (GI); red triangle: *Wari* insulator (WI), included in the 3' of the mini-*white* marker gene. Blue and red curved arrows in the diagrams depict, respectively, *e7* and *e8* activities which are shared between *p7* (**B**) and *p8* (**C**) in the wing disk.

I 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 3 D** and **E**, respectively). *GFPm7* contains *e7p7* – the 2.1 kb sequence upstream of *E(spl)m7* containing its putative enhancer, *e7*, and promoter, *p7*. *GFPm8* contains *e8p8* – 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 3 D**) and *GFPm8* lacked the muscle precursors (AMPs) (**Figure 3 E**). Therefore, 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 the genomic fragment encompassing both genes, *e7* and *e8* are shared between promoters of the two genes, *p7* and *p8*, 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*). Therefore, enhancer affinity for a given promoter can be modulated according to cell type.

2.4.1.1. Wing disk

I went on to characterize extensively these patterns in larval tissues. For a detailed expression analysis in the late third instar wing imaginal disk, I have used the serendipitously obtained “genomic” transgene expressing (under control of both enhancers, *e7* and *e8*) GFP fused to the ORF of *Espl(m7)* such that the ORF of *Espl(m7)* is out of frame (see construct scheme in **Figure 4**). Thus, only GFP is translated in frame and is localized in a cytoplasm as it lacks nuclear localization signal (NLS) located within the ORF of *Espl(m7)*. This transgene provided the greatest sensitivity of the *e7* and *e8* activity as it is *bona fide* GFP reporter construct of the two enhancers (lacking post-transcriptional regulation due to the use of SV40 3' UTR, and, post-translational modifications which are otherwise deposited on the fused and translated *E(spl)m7* protein and make it highly unstable (Kiparaki *et al.* 2015)).

The wing primordium expressing this transgene was immunostained against GFP and Senseless, a protein marking SOPs (Nolo *et al.* 2000). These *sens*-positive cells are singled out from clusters of cells (PNCs) through a process of lateral inhibition and will develop into the large sensory bristles (macrochaetes) of the thorax later on (Modolell and Campuzano 1998). Both, the number and arrangement of macrochaetes on the adult body and the number and

position of SOPs in the PNCs of the wing primordium are precisely spatially and temporally determined (Modolell and Campuzano 1998; Koelzer and Klein 2003).

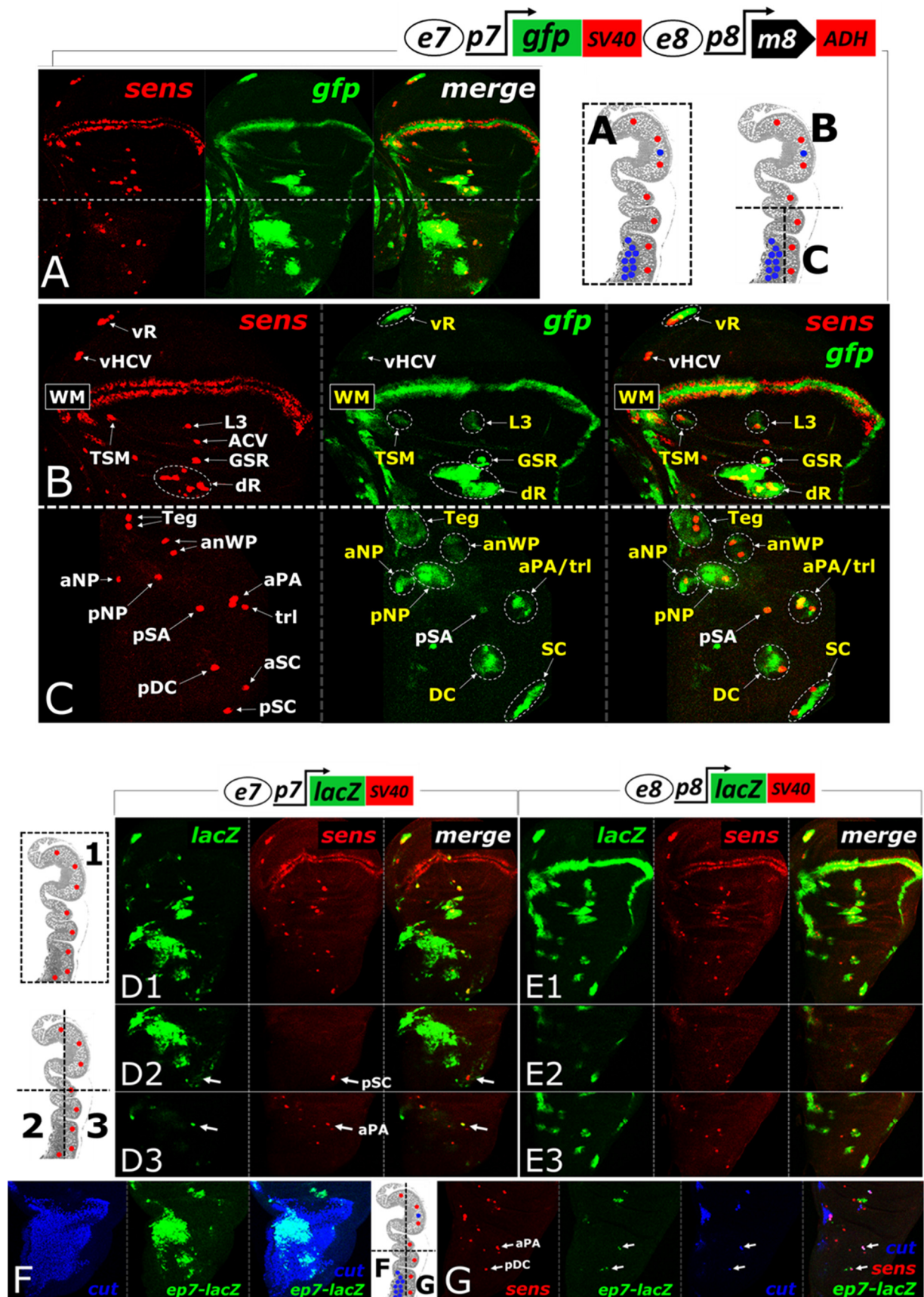


Figure 4 *e8* but not *e7* is active in most of the PNCs of the late third instar wing disk. Sens (red) and GFP (green) expression from the animals carrying *e7/e8*-GFP reporter transgene (shown top right, see also text). Schematics of the lateral sections of the wing disk (top right) show the z-stacks used to project images in **A-C**; red cells denote SOPs, a group of a blue cells at the bottom of the disk – the AMPs. **(A)** A full z-projection of the wing disk's confocal stacks. **(B)** High-resolution full z-projection of the wing pouch. **(C)** High-resolution partial z-projection of the notum and hinge. Site nomenclature and SOP assignment (white letters, in the *sens* panel of **B** and **C**) are according to (Cubas *et al.* 1991; Huang *et al.* 1991). **vR**, sensilla of the ventral radius; **vHCV**, sensillum of the ventral humeral cross vein; **WM**, chemosensory bristles of the anterior wing margin; **TSM**, twin sensilla of the wing margin; **L3**, vein L3 campaniform sensillum; **ACV**, campaniform sensillum of the anterior crossvein; **GSR**, giant sensillum of the radius; **dR**, sensilla of the dorsal radius; **Teg**, sensilla and bristles of the tegula; **anWP**, campaniform sensilla of anterior wing process; **aNP** and **pNP**, anterior and posterior notopleural; **pSA**, posterior supraalar; **aPA**, anterior postalar; **trl**, trichoid sensillum; **pDC**, posterior dorsocentral; **aSC** and **pSC**, anterior and posterior scutellar. Nomenclature and PNC assignment (yellow letters, in the *gfp* and *merge* panels of **B** and **C**) are named after their sensilla and according to (Modolell and Campuzano 1998). Note, that from all identified SOPs, GFP expression is not detected only in the vicinity of the ACV in the wing pouch. **(D1-D3)** and **(E1-E3)** show LacZ (green)/ Sens (red) patterns from *e7p7-lacZ* and *e8p8-lacZ* reporter transgenes, respectively. **(D1** and **E1)** are full z-projections, **(D2** and **E2)** show z-stacks of the notum containing AMPs, and **(D3** and **E3)** show AMPs' underlying sections where most of the SOPs are present. **(F)** is the notums' AMPs sections with the activity of the *e7p7-lacZ* immunostained against LacZ (green) and myoblast specific marker *cut* (blue). **(G)** shows the activity of *e7p7-lacZ* (green) in the AMPs underlying sections counterstained with *sens* (red) and *cut* (blue). *Cut* also stains all cells of external sensory organs. Note, that *e7*'s activity is detected weakly around pSC (**D2**) and aPA and pDC (**G**).

Full z-projection of the wing disk's confocal stacks (**Figure 4 A**) reveals GFP expression pattern characteristic to the activity of both enhancers, *e7* – in the AMPs, and *e8* – in the WM, as it was seen in **Figure 3**. A closer examination of the full z-projection of the wing pouch (**Figure 4 B**) and the sections underlying the AMPs (**Figure 4 C**) show that these enhancers are active also in most, if not all, PNCs.

Examination of the expression patterns obtained from individual enhancers (**Figure 4 D1-G**) revealed that *e7* is not active in most of the PNCs where *e8* is active; with the exception of the PNCs of dorsal and ventral radius of the wing pouch (**Figure 4**, compare **D1** to **E1**) and few cells underlying the AMPs in notum in only two PNCs (aPA/trl and DC where also *cut* is expressed **Figure 4 G**). Moreover, counterstaining with *cut* – myoblast specific marker (Blochliger *et al.* 1993) shows that *e7* (but not *e8*) is active in a subset of the AMPs (**Figure 4 F**).

2.4.1.2. Eye-antennal disk

Since both enhancers are active in virtually identical pattern, I characterized the expression pattern obtained from *e7p7-lacZ* reporter transgene (**Figure 5**). The activity of *e7* is evident in the antenna primordium, morphogenetic furrow (MF) and the photoreceptors.

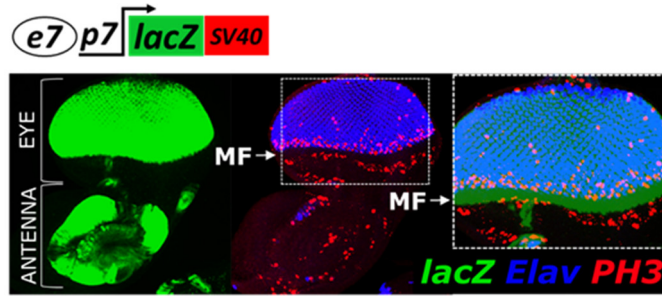


Figure 5 *e7p7-lacZ* expression pattern in eye-antennal disk. Confocal image of the eye-antennal disk expressing *e7*-driven LacZ (green), counterstained with Elav (blue) and Phosphohistone H3 (PH3, red). Elav marks photoreceptors; morphogenetic furrow (MF) is composed of cells anterior to photoreceptors which are devoid of mitotic cells expressing PH3.

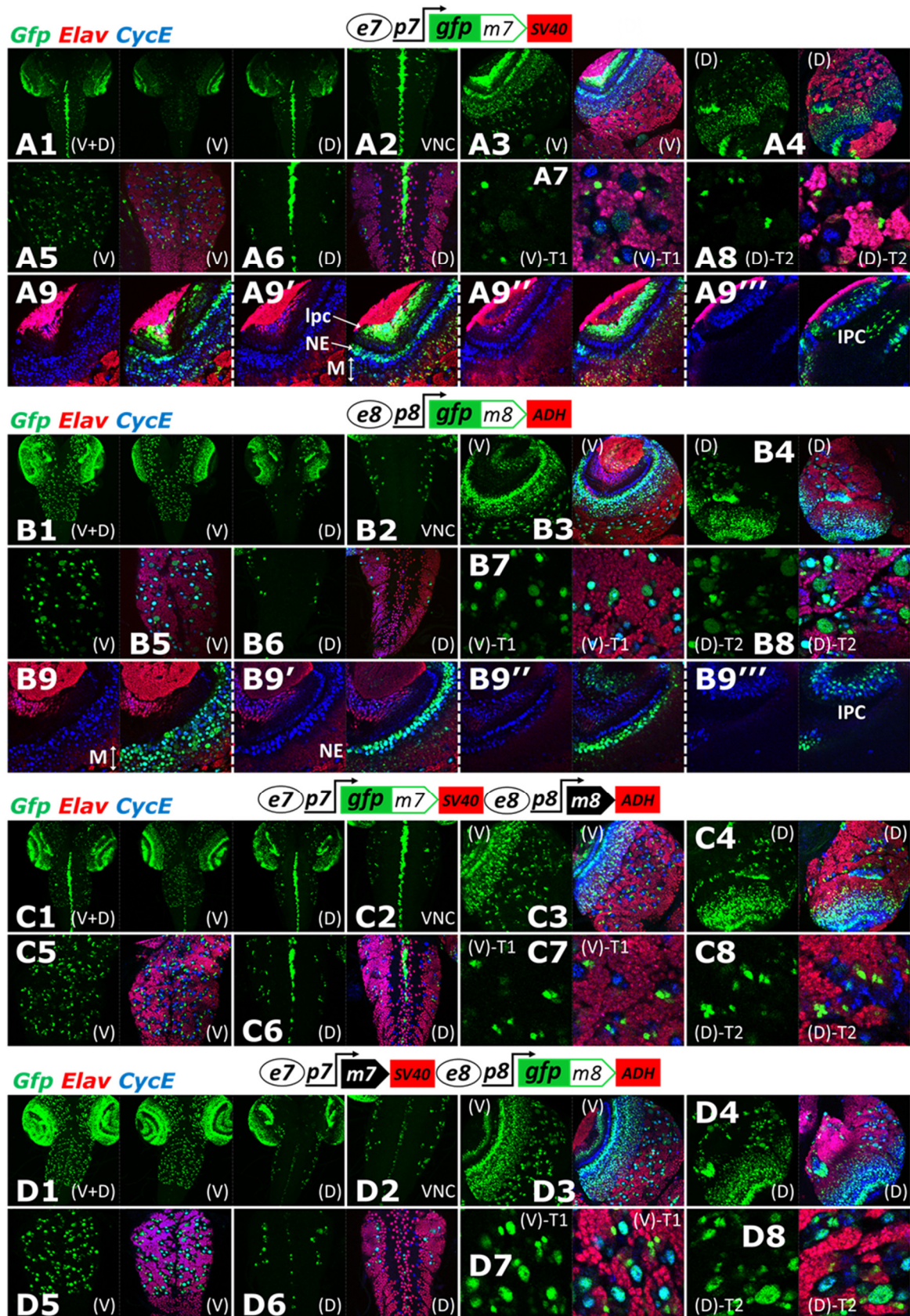
2.4.1.3. CNS

I have examined expression patterns of the genomic transgenes from **Figure 3** in the third instar larval brain. *e7* activity is evident in the midline of the VNC (**Figure 6 A1, A2, A6**), optic lobes (OLs) (**Figure 6 A1, A3, A4**), and in cells dispersed within ventral and dorsal central brain (CB) (**Figure 6 A3, A4**) and ventral thoracic ganglion (**Figure 6 A5**). The latter (dispersed) signal represent two cell types: neuroblasts (NBs) with weak expression of GFPm7 (large cells stained against Cyclin E (CycE) in blue, **Figure 6 A3-A8**) and early Hey-positive neurons (Monastirioti *et al.* 2010) with strong GFPm7 expression (small cells within population of red-stained (Elav) neurons, **Figure 6 A3-A8**, and, **Figure 6 E**, where staining against Hey, in red, reveals that these cells are a subset of Hey-positive neurons). OL's expression of *e7*-driven GFPm7 is detected in neuroectoderm (NE), lamina precursor cells, subset of medulla NBs and neurons, and inner proliferation center (IPC) (**Figure 6 A9-A9'''**).

e8p8-driven GFPm8 is detected in OLs, CB and ventral side of the thoracic ganglion (**Figure B1-B6**). The *e8* activity is strongly associated with NBs where the GFPm8 is expressed at higher levels than the *e7*-driven GFPm7 (**Figure B3-B9**). However, it is detected also in Hey-positive neurons but seems to be expressed in fewer cells, mostly surrounding NBs type 2 and at lower levels than *e7*-driven GFPm7 (**Figure B7-B8, F**). In the OLs GFPm8 is highly expressed in the NBs of medulla and NE superficially (**Figure B9-B9'**) and in the neurons of medulla (weak expression, **Figure B9'**) and the IPC in deeper sections (**Figure B9''-B9'''**).

GFPm7 and GFPm8 expression patterns from long genomic transgenes look very similar to the patterns obtained from their respective shorter transgenes (**Figure 6 C1-C6** and **D1-D6**, compare to **A1-A6** and **B1-B6**, respectively), however, some differences were noted. Hey-positive neurons express GFPm7 more frequently than it was observed with any of the short genomic constructs (**Figure 6 C7-C8**, compare to **A7-A8** and **B7-B8; F**). The latter observation suggests that *e7* and *e8* enhancers are active in different subpopulations of Hey-positive neurons. Similarly, more Hey-positive neurons express GFPm8 from the long genomic transgene (**Figure 6 D7-D8**, compare to **A7-A8** and **B7-B8; G**, compare to **E** and **F**). And, although, the *e8* enhancer (in a "long" genomic construct) does not seem to affect expression

of GFPm7 driven by *e7p7* in NBs (as it is detected equally weak as with "short" genomic *e7* transgene, **Figure 6 C7-C8**, compare to **A7-A8**), the *e7* enhancer induces weak expression of GFPm8 in the midline of the VNC (**Figure 6 D1, D2 and D6**).



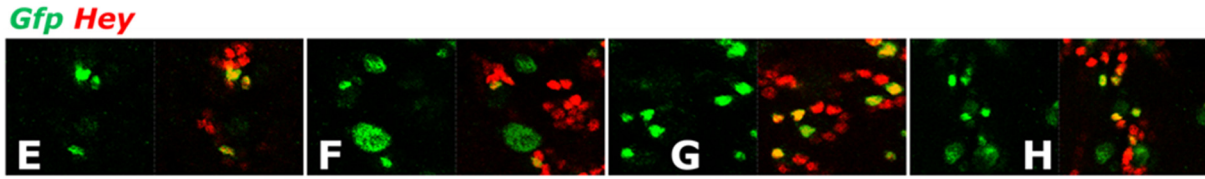


Figure 6 Expression of *E(spl)m7* and *E(spl)m8* in larval CNS. Analysis of the GFP expression patterns produced from the short and long genomic transgenes in the third instar larval brains, using anti-GFP antibody. **A to D** panels show different views and magnifications of the VNC from animals bearing indicated transgene stained against GFP (green), embryonic lethal abnormal vision (Elav, red) marking neurons and Cyclin E (CycE, blue) marking neuroblasts (NBs). (**E-H**) show staining with anti-GFP (green) and anti-Hey (red) antibodies in the area of the central brain (CB) region of the animals carrying transgenes depicted above **A-D** panels, respectively. (**A1, B1, C1, D1**) full z-projections (V+D), ventral sections (V) and dorsal sections (D) of whole CNSs. (**A2, B2, C2, D2**) magnified view of the VNC containing midline cells (D). (**A3 and A4, B3 and B4, C3 and C4, D3 and D4**) ventral (V) and dorsal (D) sections, respectively, of brain lobes (BLs). (**A5 and A6, B5 and B6, C5 and C6, D5 and D6**) ventral (V) and dorsal (D) sections, respectively, of the thoracic ganglion of the VNC. (**A7 and A8, B7 and B8, C7 and C8, D7 and D8**) magnified views showing type-1 NBs from ventral side of the CB [(V)-T1] and type-2 NBs from dorsal side of the CB [(V)-T1], respectively. (**A9-A9''' and B9-B9'''**) show sections of magnified BLs seen from ventrally from most superficial layers (**A9 and B9**) until interior of a BL (**A9''' and B9'''**); lpc - lamina precursor cells, NE - neuroectoderm, M - medulla, IPC - inner proliferation center.

2.4.1.4. Embryo

Both, *e7* and *e8*, trigger expression along the ventral midline at stage 7 (within 3 hours after egg deposition (AED)) (**Figure 7 A**, the first and the second row). From stage 7 onwards, although both enhancers are broadly active in the neuroectoderm, their activities start to diversify. *e8*, unlike *e7*, shows strong activity in dorsal procephalic neuroectoderm and in stage 13 their expression patterns are mutually exclusive: *e7* is active (superficially and laterally, in stripes), in neuroectoderm specifying primordia of PNS organs and (in deeper sections, also in stripes) in what seems to be clusters of a mesodermal origin, whereas, *e8* is active in brain and ventral nervous system (**Figure 7 A, B**). *e7* is also active in the VNC midline in stages 10 and 12 (**Figure 7 C4, C5**), whereas, *e8* is not (**Figure 7 C10, C11**). The *e7* and *e8* enhancers are not shared between *p7* and *p8* promoters in embryo as patterns obtained with long genomic transgenes are indistinguishable from their respective short versions (**Figure A, B**).

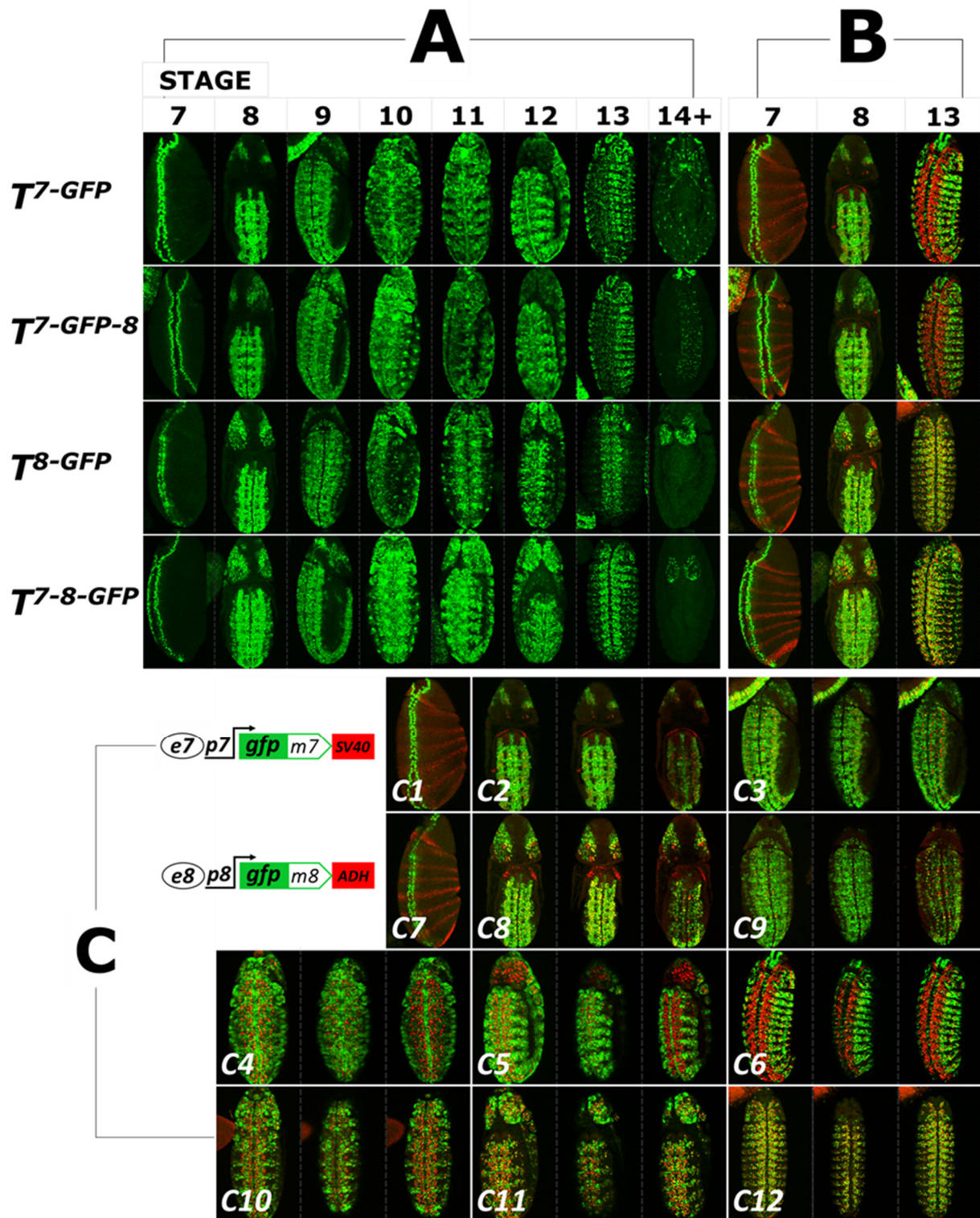
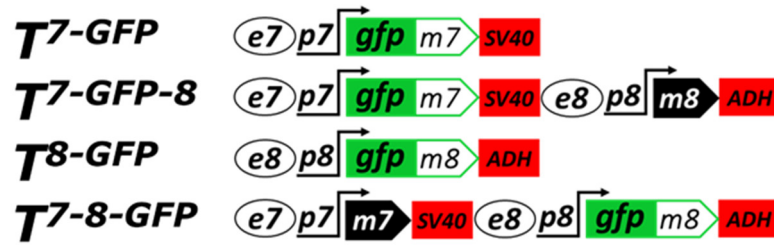


Figure 7 Expression of *E(spl)m7* and *E(spl)m8* in embryo. Analysis of the GFP expression patterns produced from the short and long genomic transgenes in the embryo. **(A)** Full z-projection of embryos of different stages (7-14+) stained with anti-GFP antibody (green) for each indicated genotype (in rows, see constructs schemes below). **(B)** Subset of these embryos (stages 7, 8, 13) are shown with overlaid deadpan (Dpn) staining (in red), marker of embryonic NBs arising below neuroectoderm from the stage 8 onward (note that, at stage 7 *dpn* expression is in stripes and is not related to NBs). **(C)** A subset of embryos from panel A (**C1**, **C7** - stage 7; **C2**, **C8** - stage 8; **C3**, **C9** - stage 9; **C4**, **C10** - stage 10; **C5**, **C11** - stage 12; **C6**, **C12** - stage 13) embryos bearing short genomic *e7*- (**C1**-**C6**)

or *e8*-transgenes (**C7-C12**) are shown as full z-projection, stack of superficial layers (neuroectoderm) and stack of deeper layers (NBs sections, Dpn, red), respectively, in each panel (except of **C1**, **C7** showing only full z-projection). All embryos are imaged ventrally with anterior to top. Constructs schemes:



2.4.1.5. Pupa

For imaging of pupal *e7* and *e8* patterns, I used *p7-gfp* transgenes fused to either of the two enhancers. Strong promoter activity of *p7* (see **Figure 35** in part 2), an enhanced version of GFP (EGFP) and, translationally unrepressed, SV40 3' UTR of SV40 allowed me to image native *gfp* expression patterns under fluorescent stereo microscope in living pupae (**Figure 8**). The two enhancers showed distinct activities. *e7* was active in dorsal longitudinal muscles (DLM) (**Figure 8 B1-B3**, and laterally **B5**), antenna, proboscis, legs and eyes (**Figure 8 B4**). *e8* was active in eyes, ocelli (triangle shape between the eyes) (**Figure C1-C3**), microchaetal stripes (**Figure C3'**) and in the presumptive Johnston's organ of the antenna (**Figure C4'**).

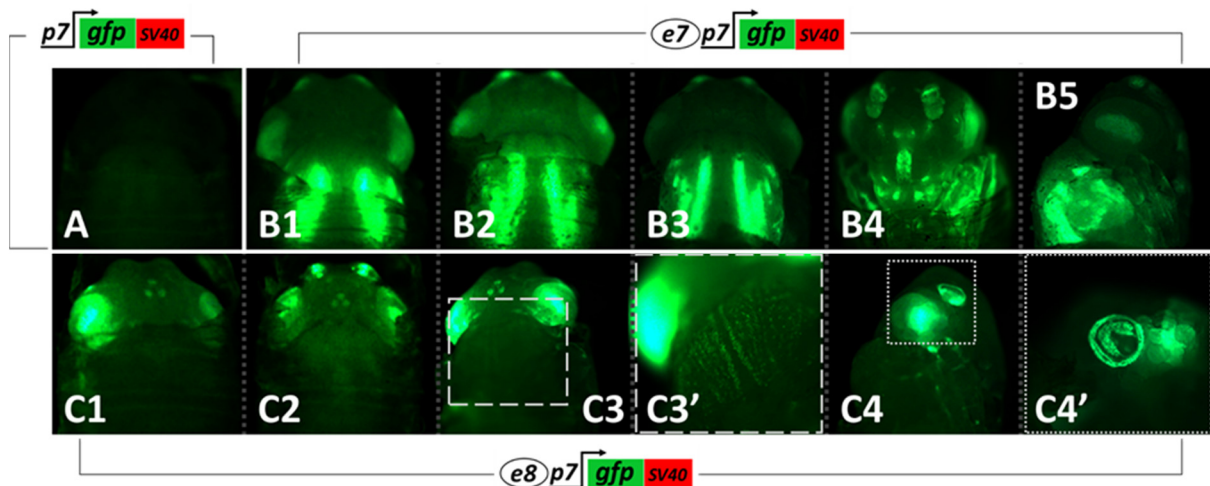


Figure 8 *e7* and *e8* activities in early pupal stages. Images were taken with the fluorescence stereo microscope of the cephalic and notal parts of pupae bearing indicated transgenes. (**A**) enhancerless *p7-gfp* transgene shows that *p7* has no basal activity or that its basal activity is not detectable. To visualize GFP expression from *e7p7-gfp* transgene, three different individuals were imaged dorsally (**B1-B3**), one ventrally (**B4**) and one laterally (**B5**). For the *e8p7-gfp*, three different individuals were imaged dorsally (**C1-C3**), a zoom was taken of the notal region from **C3** pupa (**C3'**), one pupa was imaged laterally (**C4**) with a zoom of the presumptive Johnston's organ (**C4'**).

2.4.2. Analysis of enhancers of *E(spl)m7* and *E(spl)m8*

A 7 kb genomic fragment containing *E(spl)m7* and *E(spl)m8* transcription units has been mapped for previously known regulatory elements (see **Figure 2**). The enhancers of both genes contain single paired Su(H) high-affinity site (SPS) and binding sites for proneural bHLH activator proteins, class E_A boxes (Jennings *et al.* 1999). These boxes have been shown to mediate activation and positive control of *E(spl)* genes and are necessary for their expression in the wing disk's PNCs (Bailey and Posakony 1995; Lecourtois and Schweisguth 1995; Culí and Modolell 1998; Castro 2005).

A number of bHLH repressor sites have been mapped within enhancers of *E(spl)* genes (suggestive of autorepression), although none of these sites have yet been shown to be required for any functions *in vivo* (Culí and Modolell 1998; Jennings *et al.* 1999; Maeder *et al.* 2007). The bHLH repressor sites are represented by three classes of boxes – numerous present within upstream regions of *E(spl)m7* and *E(spl)m8* class E_N-box (CACNAG); two E_B-boxes (CACGTG, B1 and B2) in each enhancers *E(spl)m7* and *E(spl)m8*, and two E_C-class boxes (CACGCG, C1 and C2) were found in *e8* (**Figure 2**). The E_N-box is bound *in vitro* with much lower affinity than E_B- and E_C-boxes, thus B-box class and C-box class seem to mostly contribute in the transcriptional negative regulation of the *E(spl)* genes (Jennings *et al.* 1999). Subsequent bioinformatic analysis of multi-species alignment between 12 *Drosophilids* indicates high evolutionary conservation of the *e7*'s B1, N1 and N2 sites, and *e8*'s C2, N5, N6 and N7 sites (**Figure 9**). The supposedly strong bHLH-O binding sites (B1 within *e7*, and B2, C1, C2 within *e8*) were chosen for point mutation to assess their functionality *in vivo*.

I have generated a series of *lacZ* reporter constructs to validate the importance of the identified *cis*-regulatory sequences of the *E(spl)m7* and *E(spl)m8*. In order to avoid genomic position effects and facilitate semi-quantification of expressed proteins all obtained constructs were inserted into a defined position in the *Drosophila* genome using the ϕ C31-based integration system (Markstein *et al.* 2008). Enhancer activities were tested *in cis* (*LacZ* reporter) and *in trans* (by transvection to an enhancerless *promoter-gfp* transgene) (see **Figure 10**).

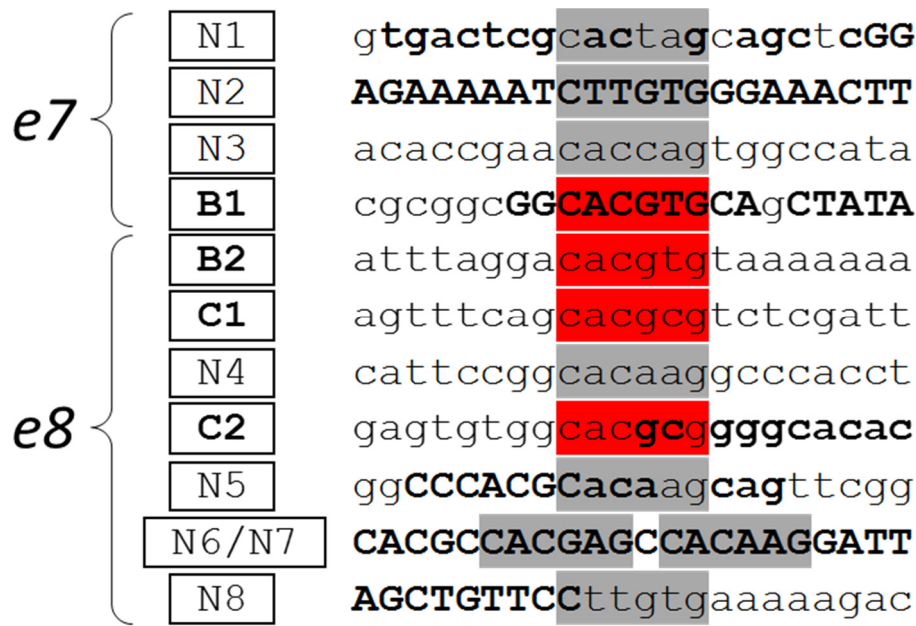


Figure 9 Evolutionary conservation of the E_B- and E_C-boxes of e7 (B1) and e8 (B2, C1, C2). DNA sequence of the two E_B-boxes (B1 and B2) and two E_C-boxes (C1 and C2) found within enhancers of *E(spl)m7* and *E(spl)m8* (their core hexamers are highlighted in red) together with their flanking sequences were processed with EvoPrinter <http://evoprinter.ninds.nih.gov/> for identification of evolutionary conservation among 12 Drosophilids. Bold capital letters represent bases in the *D.melanogaster* reference sequence that are conserved among all 12 *Drosophila* species (*D.simulans*, *D.sechellia*, *D.erecta*, *D.yakuba*, *D.ananassae*, *D.pseudoobscura*, *D.persimilis*, *D.willistoni*, *D.mojavensis*, *D.virilis* and *D.grimshawi*). Bold small letters indicate less conserved bases shared among 8 *Drosophila* species (with the exception of *D.persimilis*, *D.willistoni*, *D.virilis* and *D.grimshawi*).

Mutating B1-box did not affect expression levels obtained from the wild-type *e7*, suggesting that this box does not mediate repression (**Figure 10**, compare **B** to **A** and **B'** to **A'**). I examined functionality also of other features of the *e7p7* module. Deletion of the *p7* promoter did not completely abolished expression in *cis* (**Figure 10 C**, notice extremely faint *e7*-specific expression of LacZ), suggesting existence of a cryptic promoter within *e7*. The E_A-box-containing distal part of the *e7p7* (devoid of the the SPS and the *p7* promoter) exhibits Notch-independent activities in the wing disk's dorsal radius' PNC and in the CNS's midline (detected in *trans*, **Figure 10 D'**). In contrast, the SPS-containing proximal part of the *e7* enhancer is completely inactive, both in *cis* and in *trans* (**Figure 10 E** and **E'**). This suggests that the Notch molecule alone is not sufficient to activate *e7*. Rather, *e7* is a compound enhancer requiring synergy of both activities: E_A-box-mediated bHLH proneural activity located distally and SPS-mediated Notch activity located in the proximal part of the 2 kb long enhancer. Deletion of the putative insulator, CTCF-binding site (Negre *et al.* 2010) did not affect levels and the expression pattern of LacZ, suggesting that the binding of CTCF at this site does not play any role in *e7* function (**Figure 10**, compare **F** to **A** and **F'** to **A'**).

I have also tested functional importance of *E(spl)m7* promoter's integrity. *p7* is a multi-element promoter, containing a TATA box, an initiator (Inr) and a downstream promoter element (DPE) (Klämbt *et al.* 1989; Kutach and Kadonaga 2000). I introduced two deletions

into the *e7p7-lacZ* construct aiming to disrupt each of these activities; one, *e7p7-ΔTATA-lacZ*, removed the TATA box and another, *e7p7-ΔDPE-lacZ*, removed the Inr and DPE elements. Both of these promoter mutations retained weak yet detectable transcriptional activity in *cis* (**Figure 10 G and H**, compare to **A**; *trans* activity of these transgenes will be discussed in Chapter 2). This observation indicates that the *p7* needs both of these elements for its full activity (see summary diagram, **Figure 10 N**).

Much like it was a case for B1-box in the *e7*, B2 and C1 sites in the *e8* do not seem to mediate any repression (**Figure 10**, compare **J** to **I** and **J'** to **I'**). Mutation of the C2 site, however, resulted in increased levels of expression, both in *cis* and in *trans* (**Figure 10**, compare **K** to **I** and **K'** to **I'**). Similarly, mutation of all B/C-boxes in *e8* or a deletion encompassing these sites had comparable effect (**Figure 10, L/L', M/M'**). Therefore, out of four putative repressor binding sites, only C2 of *e8* mediates visible and robust repression.

Subsequently, I tested whether *e7* and *e8* enhancers are targeted by the products of *E(spl)* locus. To this end, I generated clones mutant for the whole *E(spl)* locus using deficiency *gro^{b32.2}* (see **Figure 1**) (Schrons *et al.* 1992) within the third larval instar wing disk of animals carrying wild-type *e7p7-lacZ* (**Figure 11 A-F**), wild-type *e8p8-lacZ* (**Figure 11 G-L**) and *e8p8-lacZ* bearing deletion of the *e8* containing all B/C-boxes (**Figure 11 M-R**). In all three cases, I observed elevated levels of LacZ expression within mutant clones centred around PNCs. As *e7* is not normally active in most of PNCs, clonal expression of LacZ in this case was ectopic. These observations mean that both, *e7* and *e8*, are under repression of the *E(spl)* proteins and that *e7*'s inactivity in most of the PNCs could be due to a strong transcriptional repression. Moreover, the fact that the *e8* lacking B/C-type repression sites still responds to the lack of *E(spl)* factors means that this repression is mediated not only through the C2-box. It is possible that in the absence of high-affinity binding site (as the C2), bHLH repressors gain access to low-affinity N-box type sites. Several N-boxes are indeed present in the proximal part of *e8* (as well as in the *e7*). This result can be explained also by the ability of the *E(spl)* bHLH proteins to be recruited on DNA via interaction with bHLH activators (through *E_A*-boxes) (Giagtzoglou *et al.* 2003, 2005). Given that bHLH proneural factors are needed for the activity of both enhancers, the latter possibility poses difficulty in separating activation from repression.

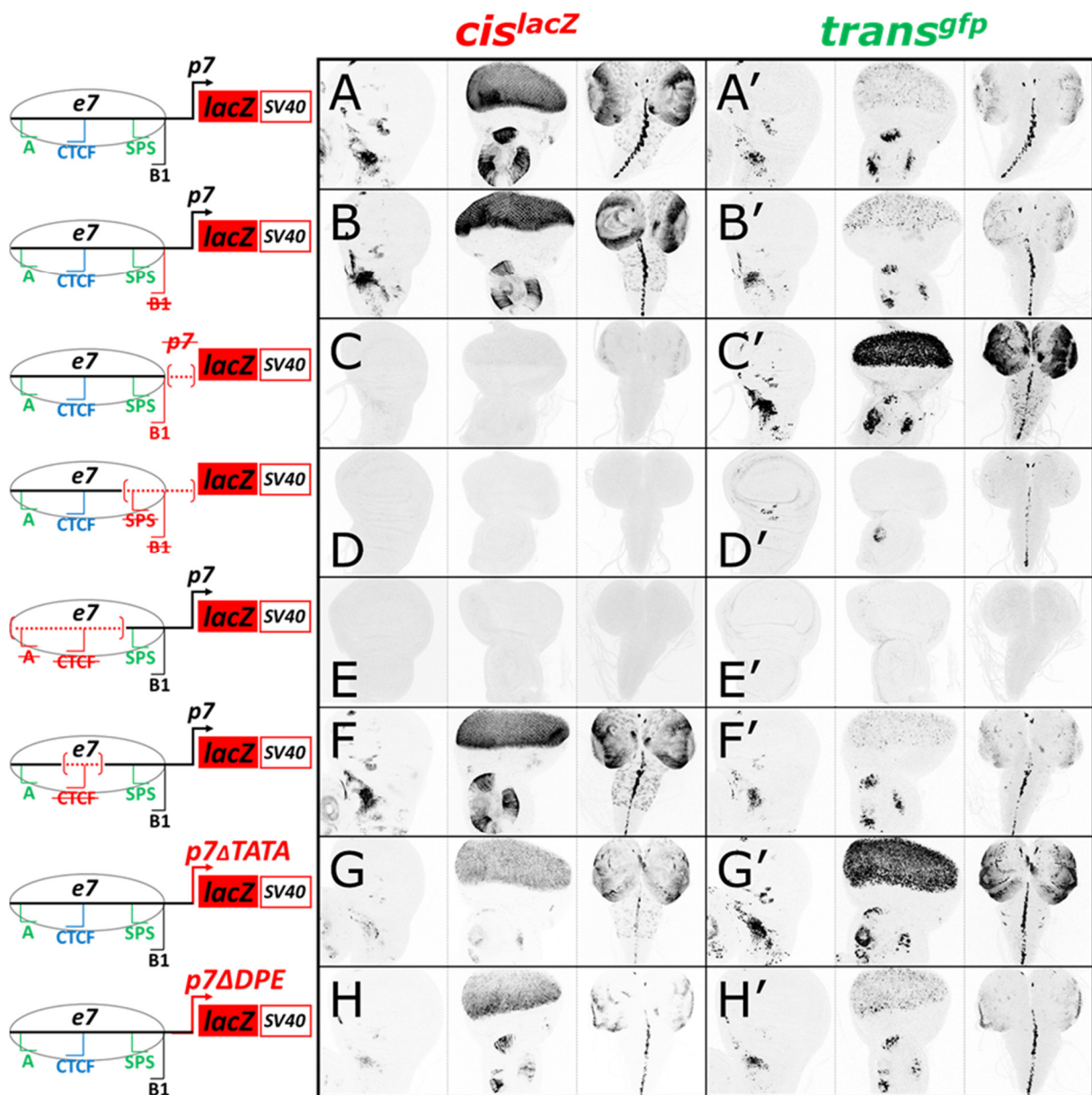


Figure 10 - continues on the next page

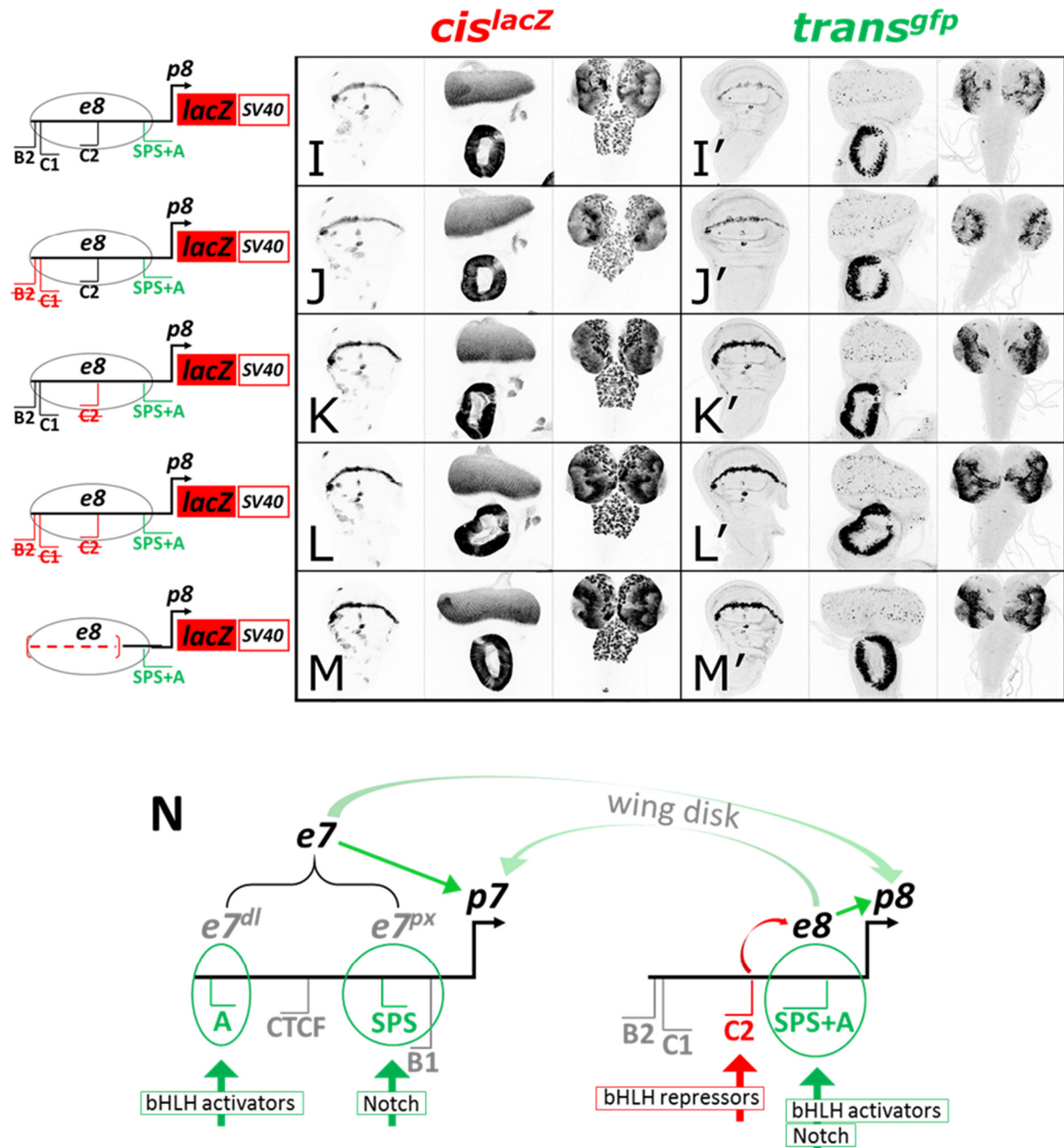


Figure 10 Validation in vivo of the cis-regulatory elements identified within enhancers of *E(spl)m7* and *E(spl)m8*. (A-M) Wing disk, eye-antennal disk and CNS from third instar larvae of animals carrying *lacZ* transgenes (shown schematically) were immunostained with anti- β -galactosidase and imaged with identical settings to allow semi-quantitative comparison of LacZ expression levels. (A'-M') Additionally, each genotype was tested for GFP expression as a transheterozygote with a chromosome containing enhancerless promoter (*heat-shock 70 promoter, pH*)-*gfp* transgene at the same locus, allowing for a detection of an enhancer activity in *trans* (see Chapter 2). (N) Summary diagram of the mapping of the *e7* and *e8* enhancers. *e7* is a composite enhancer needed its SPS- containing proximal part (*e7^{px}*, mediating Notch activity) and A-box-containing distal part (*e7^{dl}*, mediating bHLH activators' activity) for full activity. CTCF and B-box binding sites do not have any effect on the activity of *e7*. *e8* enhancer activity is located proximally to *p8* (within 0.5 kb from transcription start site, out of 2 kb of the putative *e8*) and depends on the SPS+A sites which mediates Notch (SPS) and bHLH activators (A) activity. Out of three mutated B/C boxes, only C2 is mediating repression (red arrow). In larval eye-antennal disk and CNS, *e7* and *e8* act on their downstream promoters, however, in wing disk, both enhancers act on both promoters.

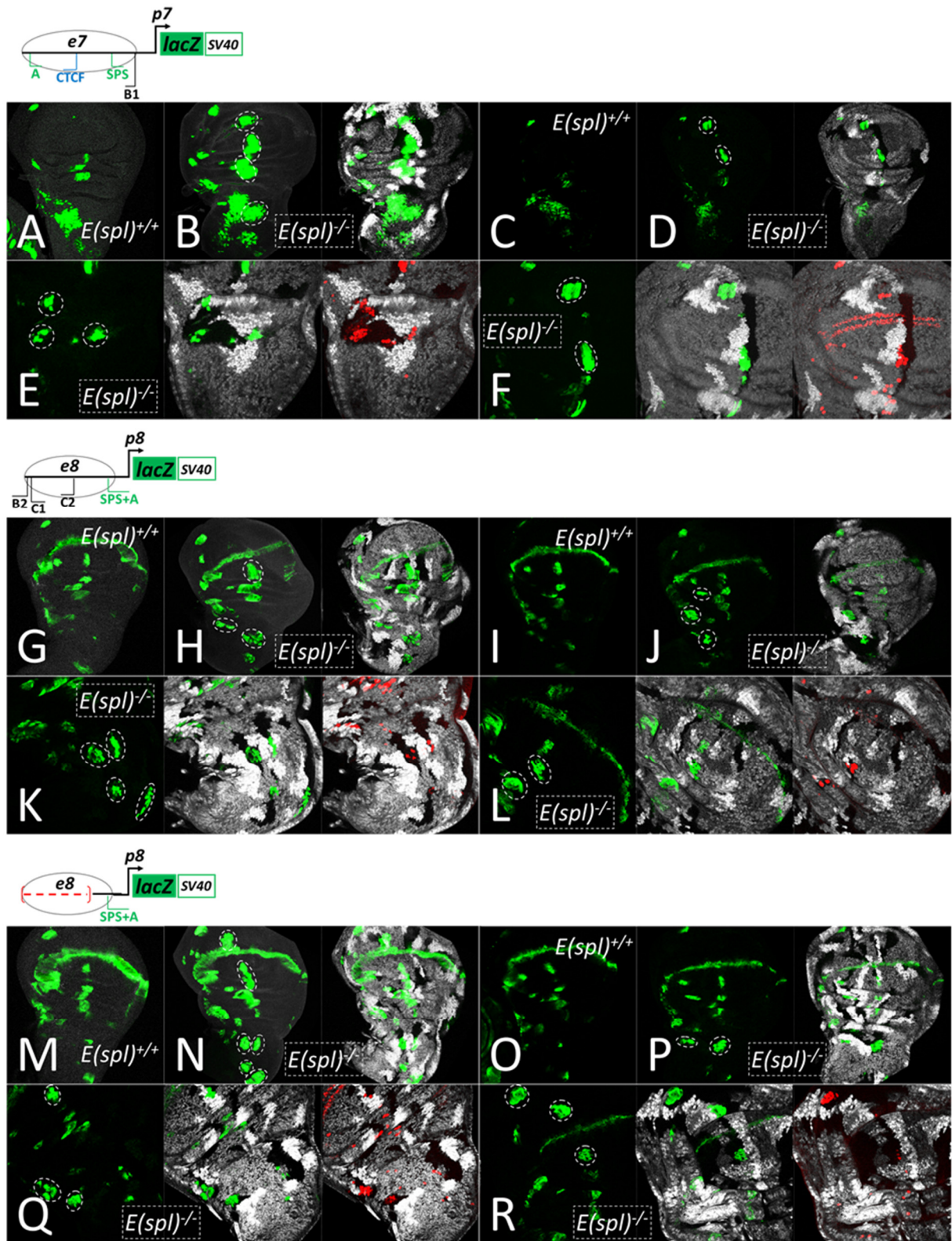


Figure 11 *E(spl)m7* and *E(spl)m8* are repressed by the products of *E(spl)* locus. Clonal analysis of the *e7p7/e8p8*-driven LacZ expression (green) using complete deficiency of the *E(spl)* locus (*Df(3R)gro*^{32.2}). Unlabeled cells (by shades of grey) are homozygous for *Df(3R)gro*^{32.2}. SOPs are also visible (*sens*, in red). Full z-projection of a wing disk with expression of the *e7p7-lacZ* (A), *e8p8-lacZ* (G) and *e8p8ΔBCC-lacZ* (M) in wild-type background (*E(spl)*^{+/+}) is compared to full z-projection of a wing disk with expression of LacZ from the same transgene in the mutant background (*E(spl)*^{-/-}), some of which arises from clones (marked with dashed line), in (B), (H) and (N), respectively. Note that clones in full z-projected disks seem smaller or are covered by overlaid sections containing

non-mutant tissue. (C), (I) and (O) with (D), (J) and (P), respectively, are similar comparisons but of a single section centered at the PNC of dR; notice ectopic and elevated levels of LacZ expression in (D), (I) and (O) (within clones marked by dashed line). (E) and (F), (K) and (L), (Q) and (R) are close-ups of notum and wing pouch, respectively. Notice that sites of ectopic expression of *e7p7-lacZ* as well as elevated expression of *e8p8-lacZ* are marked by supernumerary SOPs (red), a sign that these are areas of PNCs.

2.4.3. Analysis of 3' UTRs of *E(spl)m7* and *E(spl)m8*

Two classes microRNA responsive seed boxes have been identified within 3' UTRs of the *E(spl)-C* genes (Lai *et al.* 2005). The *E(spl)m7* 3'UTR contains one Brd-box and one K-box whilst the *E(spl)m8* 3' UTR has two K boxes, all of which are highly conserved in all *Drosophilids* (**Figure 12**). It has already been shown that ectopic expression of miR-2 and miR-11 directly regulate K-box-containing 3' UTR of *E(spl)mδ* whereas miR-4 and miR-79 regulate Brd-box-containing 3' UTRs of *E(spl)m5*, *E(spl)mδ* and *E(spl)mγ* (Lai *et al.* 2005). Using computational methods I identified four more regulatory elements within *E(spl)m7* 3'UTR: three putative target sites for miR-219, miR-8 and miR-289, and one highly conserved motif predicted to form duplexes with the bantam sequence – a microRNA which controls cell proliferation and apoptosis in *Drosophila* (Brennecke *et al.* 2003) (**Figure 12**). *E(spl)m8* 3'UTR may also be regulated by another mechanism, AU-rich element (ARE) mediated decay (Cairrao *et al.* 2009), as it bears a stretch of evolutionary conserved AT repeats immediately downstream of its two K-boxes.

In order to estimate the *in vivo* impact of microRNAs on these 3' UTRs, I have generated 3' UTR “sensor” transgenes wherein globally expressed (under a constitutive, tubulin promoter) reporter gene mRFP is coupled with a given 3' UTR. All microRNA seed boxes identified within *E(spl)m7*'s and *E(spl)m8*'s 3' UTRs were subsequently mutagenized in various combinations. Presumed de-repression of the sensor and the level of its increased expression will allow to ascertain the functional responsiveness of the mutated seed boxes to active microRNAs present *in vivo*. The SV40 3' UTR and the ADH 3' UTR were used as control 3' UTRs presumably devoid microRNA regulation. Additionally, I have generated a *E(spl)m8*'s 3' UTR sensor with the putative AREs excised and replaced by the ADH 3' UTR while preserving both K-boxes in their wild-type sequence (all constructed 3' UTRs are schematically shown in **Figure 13**).

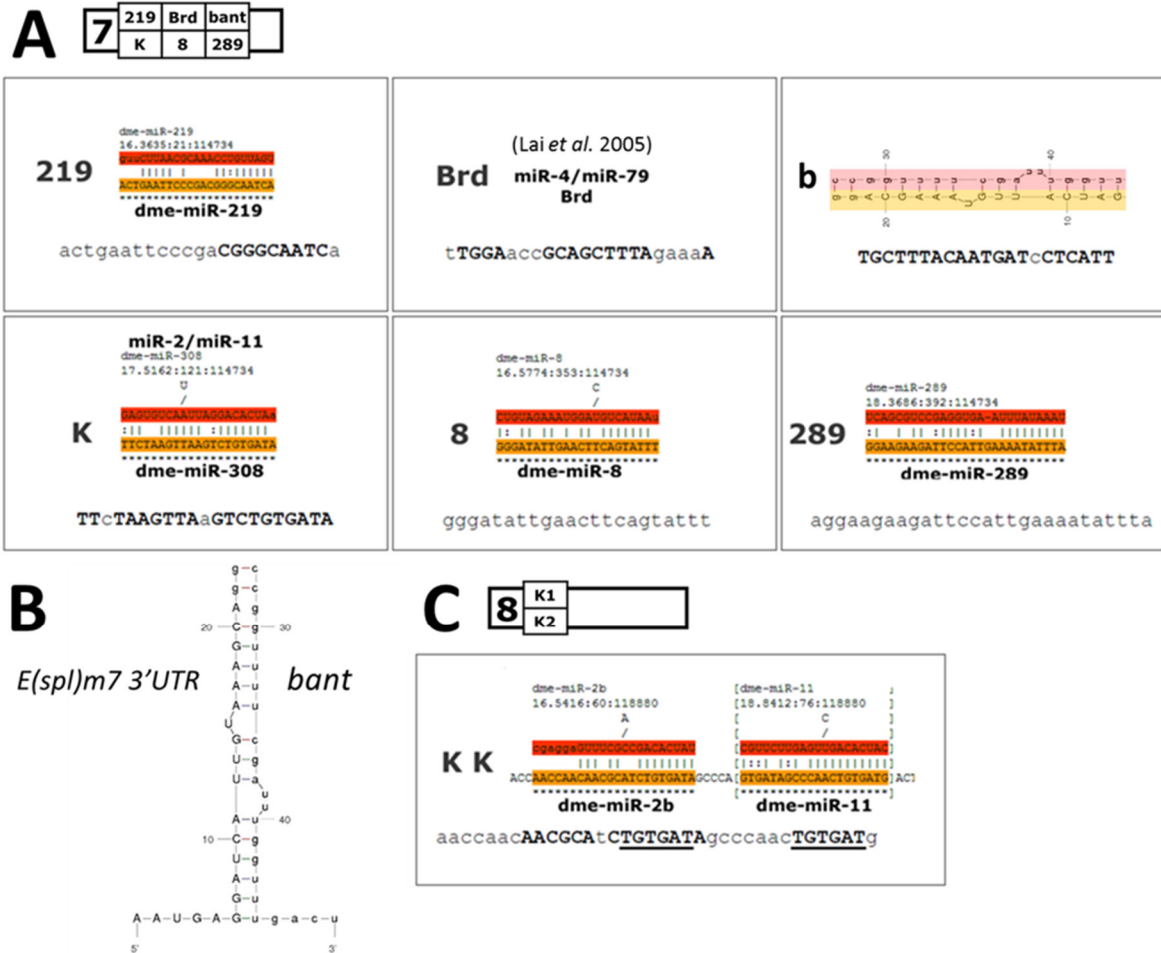


Figure 12 Identified microRNA seed boxes within the *E(spl)m7*'s and *E(spl)m8*'s 3' UTRs. (A) and (C) Schematic representation of the *E(spl)m7* 3' UTR and the *E(spl)m8* 3' UTR, respectively, containing predicted and identified microRNA binding sites: "219" – for miR-219, "Brd" – Brd-box specific for miR-4/miR-79, "bant" – for bantam, "K" – K-box specific for miR2, miR6, miR11 and miR13, "8" – for miR-8, and, "289" – for miR-289. Sequence of the respective seed box with its flanking region is shown as a read out of the EvoPrinter <http://evoprinter.ninds.nih.gov/>. Bold capital letters represent bases in the *D.melanogaster* reference sequence that are conserved among all 12 *Drosophila* species (*D.simulans*, *D.sechellia*, *D.erecta*, *D.yakuba*, *D.ananassae*, *D.pseudoobscura*, *D.persimilis*, *D.willistoni*, *D.mojavensis*, *D.virilis* and *D.grimshawi*). For the "219", "K", "8" and "289" seed boxes of the *E(spl)m7* 3' UTR, and, two K-boxes of the *E(spl)m8* 3' UTR, predicted alignments obtained with TargetScanFly 6.2 are shown (Ruby et al. 2007). (C) Predicted secondary structures of the putative binding site for bantam in the *E(spl)m7* 3' UTR and bantam microRNA using the mfold RNA folding program version 2.3 (Zuker 2003).

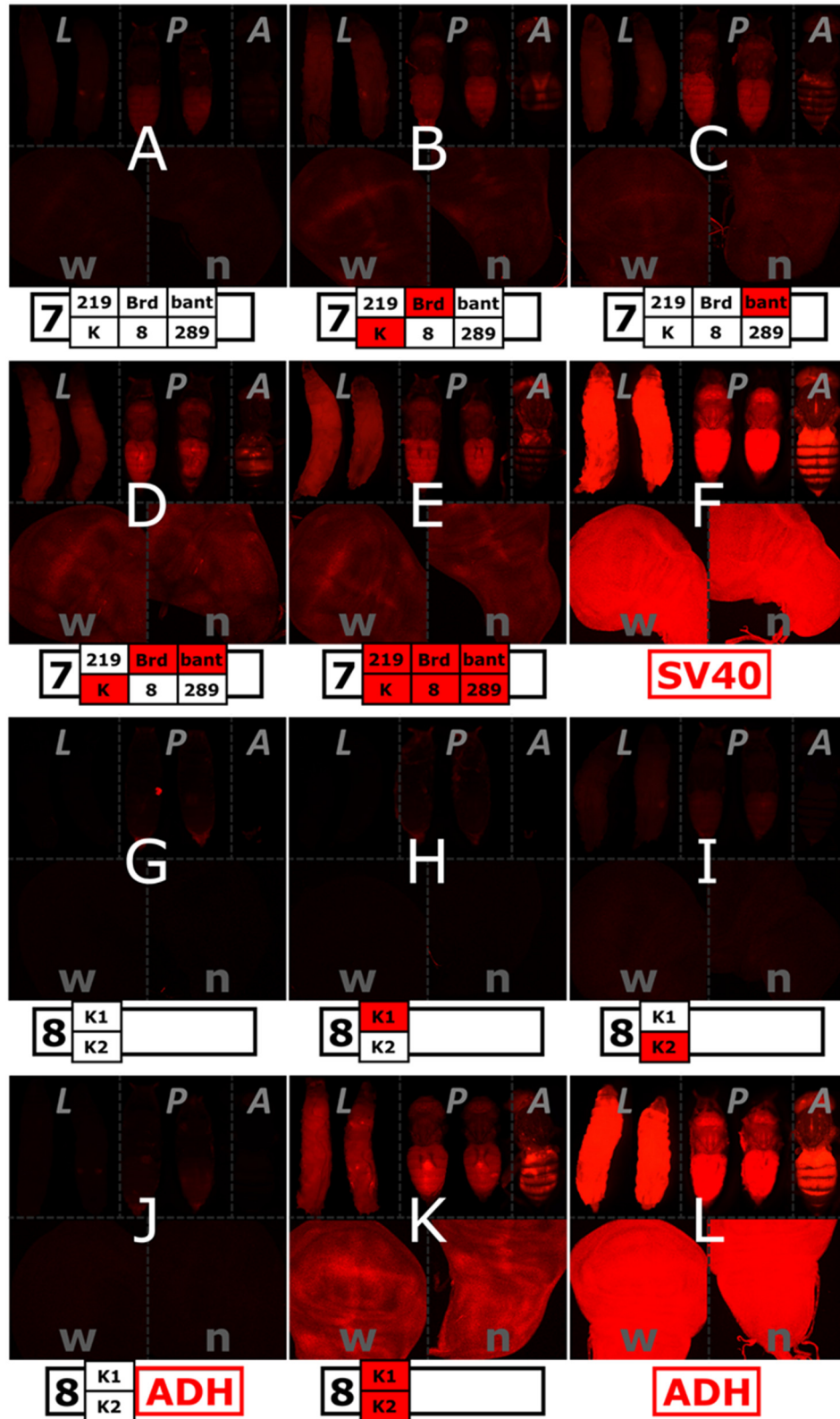


Figure 13 Identified microRNA binding sites within *E(spl)m7* and *E(spl)m8* 3' UTRs mediate repression. Semi-quantitative comparative analysis of the tubuline-mRFP sensor transgenes for wild-type and mutated 3' UTRs (shown schematically) of *E(spl)m7* (A-E), *E(spl)m8* (G-K), SV40 (F) and ADH (L). Each panel contains female and male larvae ("L"), female and male pupae ("P"), one female adult fly ("A") imaged with the same stereoscope settings; and confocal images of third instar larval wing pouch ("w") and notum ("n") image with the same confocal microscope settings. Wild-type microRNA binding sites are shown as empty boxes, whereas mutated are filled with red.

Analysis of expression levels of RFP from sensor constructs for *E(spl)m7* and *E(spl)m8* 3' UTRs (**Figure 13**) shows that identified and mutated microRNA binding sites are functional in post-transcriptional regulation. However, both 3' UTRs bearing mutations in all identified seed boxes, still exert some repression as evidenced by a comparison with SV40 and ADH 3' UTRs (**Figure 13**, compare **E** and **K** to **F** and **L**). Mutation of microRNA binding sites within the *E(spl)m7* 3' UTR has accumulative effect on de-repression. Interestingly, this is not a case for *E(spl)m8* 3' UTR where both K-boxes have to be mutated to relieve the expression of RFP from their repression. These two sites seem to be sufficient to exert almost full repression of translation and the AREs identified in the *E(spl)m8* 3' UTR seem to have little or no effect (**Figure 13, J** – in this 3' UTR, ADH 3' UTR is cloned immediately downstream of two K-boxes removing downstream AREs, compare to **G**).

DCR1 is essential for biogenesis of microRNAs (Bartel 2004). I performed clonal analysis of the third instar larval wing disks expressing transgenes shown in Figure 13 using *dcr1* null allele. Mutant cells (distinguished by lack of GFP) displayed elevated levels of RFP when the wild-type and Brd, K or/and bant deficient *E(spl)m7* 3' UTR RFP sensor transgenes were assayed (**Figure 14 A-D**), while all introduced mutations within this 3' UTR minimized this effect (**Figure 14 E**). Surprisingly, also the SV40 3' UTR (**Figure 14 F**), but not the ADH 3' UTR (**Figure 14 L**), proved to be partially repressed by microRNAs as its sensor transgene responded to the lack of DCR1. Most surprising results, however, were obtained with the *E(spl)m8* 3' UTR sensor transgenes. Although mutation of both K-boxes relieves the transcript from repression (**Figure 13 K**), lack of DCR1 had no effect on any sensor transgene containing both or any of the two K-boxes unmutated (**Figure 14 G-J**); neither did it have on the sensor with both K-boxes mutated. Thus, it seems that post-transcriptional regulation of *E(spl)m8* is *dcr1*-independent.

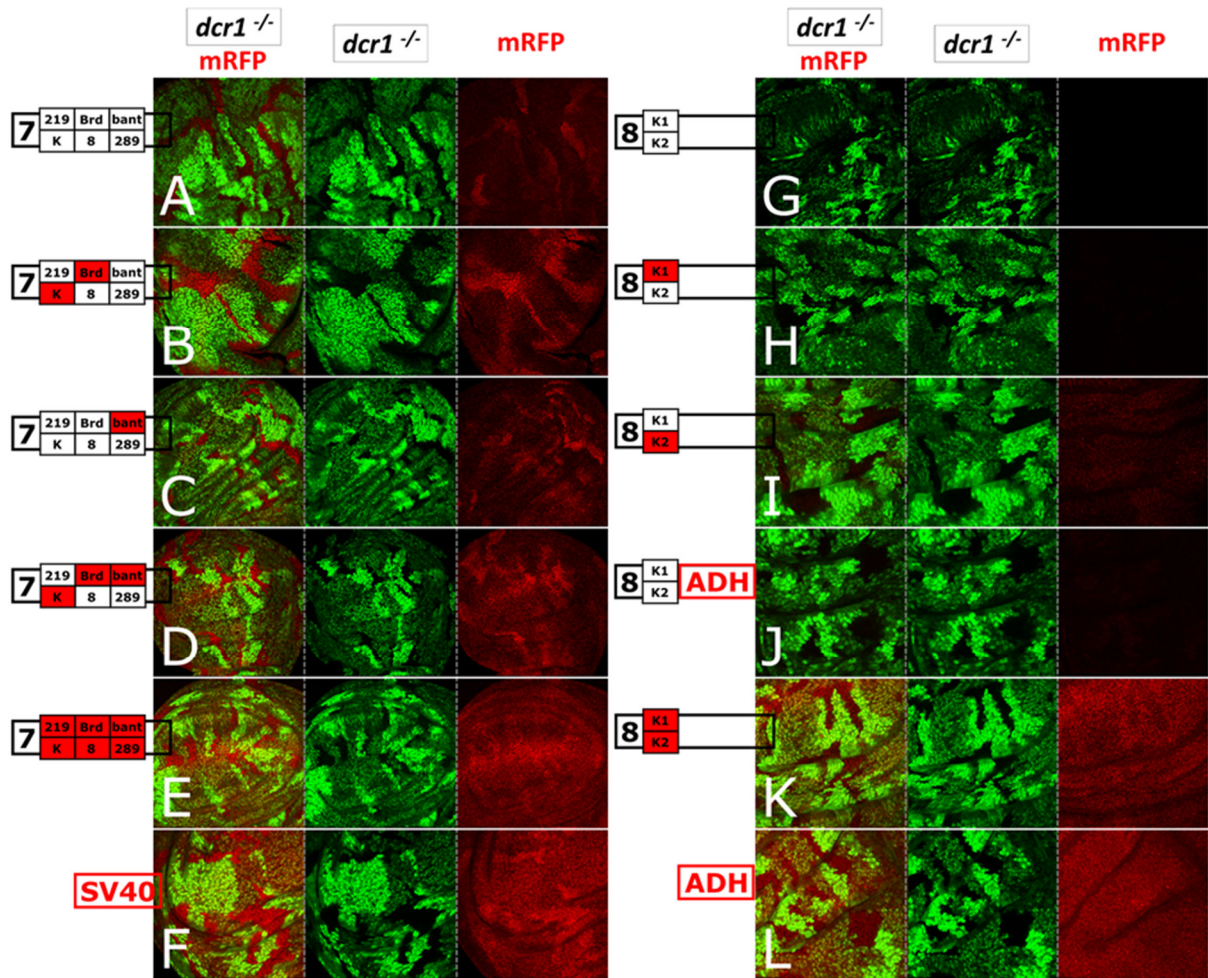


Figure 14 The 3'UTR of *E(spl)m7*, and not that of *E(spl)m8*, is under control of microRNAs. Clonal analysis of *dcr-1* (non-GFP cells, middle column) in the wing disks expressing tubuline-mRFP sensor transgenes (right column) for wild-type and mutated 3' UTRs (shown schematically) of *E(spl)m7* (A-E), *E(spl)m8* (G-K), SV40 (F) and ADH (L). The merged GFP and mRFP channels are shown in the first column.

Furthermore, I confirmed that the 3' UTR of *E(spl)m7* is targeted by bantam microRNA by testing some of the sensor constructs in a *bantam*^{-/-} mosaic background. (**Figure 15**). Clones of cells mutant for *bantam* (marked positively by GFP) showed increased levels of mRFP in the case of a transgene containing K and Brd boxes mutated (**Figure 15 B**), but not when all microRNA binding sites were wild-type (**Figure 15 A**) or mutated (**Figure 15 D**). It seems, therefore, that compromised repression by *bantam* can be compensated by microRNAs bound to K and Brd boxes.

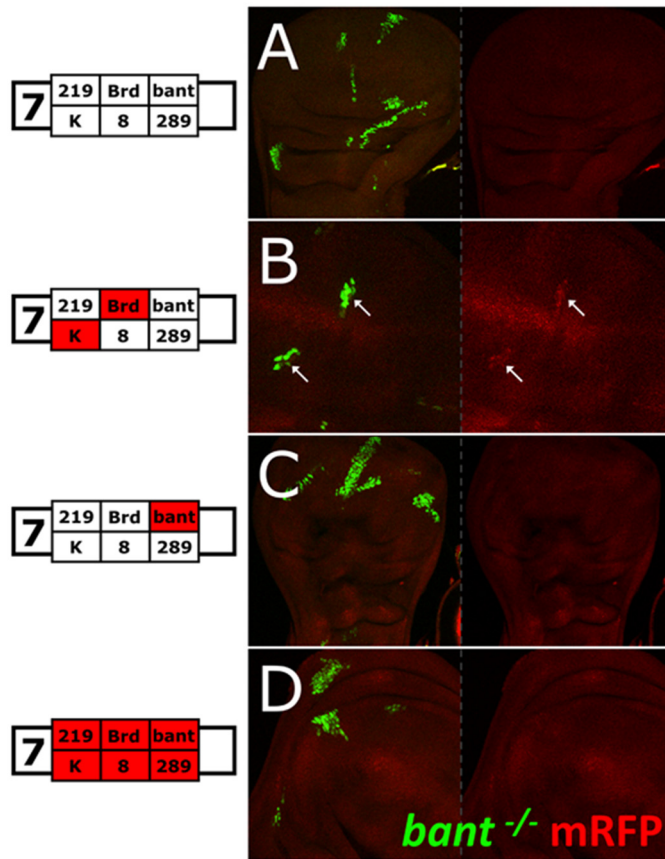


Figure 15 *bantam* microRNA targets *E(spl)m7* 3'UTR. (A-D) Bantam mutant clones are positively marked by GFP (green) in the larval wing disks expressing different tub-mRFP sensors for the 3'UTR *E(spl)m7*. Notice that the effect of derepression (indicated with white arrows) can be seen only when the two other potent microRNA binding boxes (Brd and K) are not functional, hence upregulation of mRFP mediated by the lack of bantam cannot be compensated through them.

I also addressed an impact of the Notch activity on the expression of mRFP sensors regulated through these 3' UTRs. It has been shown that Notch mediates repression of bantam microRNA and possibly other microRNAs in wing disk and CNS tissues (Becam *et al.* 2011; Wu *et al.* 2017). Expressing activated form of Notch (N^{intra}) in random clones in the third instar larval wing disks derepresses both wild-type sensors (the *E(spl)m7*'s and *E(spl)m8*'s 3'UTRs) in a manner which seems to be partially non-autonomous (**Figure 16 A and B**). This expression, however, is not triggered wherever N^{intra} is active but only in specific areas of the wing disk, that seems to be a subset of PNC pattern. This suggests that Notch acts cooperatively with other factors to derepress these 3' UTRs. I have tested whether this activity can be mediated through one of the *E(spl)* targets of Notch by expressing the *E(spl)m7* protein in the dorsal part of the wing disk (**Figure 16 C and D**). This seemed not to have any impact on these sensors (shown for wild-type and mutant 3' UTR *E(spl)m7* sensors in **Figure 16 C and D**, respectively). However, removal of all the *E(spl)* locus effected in further repression of these sensors, suggesting that Notch activity on microRNA could be in part mediated through one or several products of this locus (**Figure 16 E and F**).

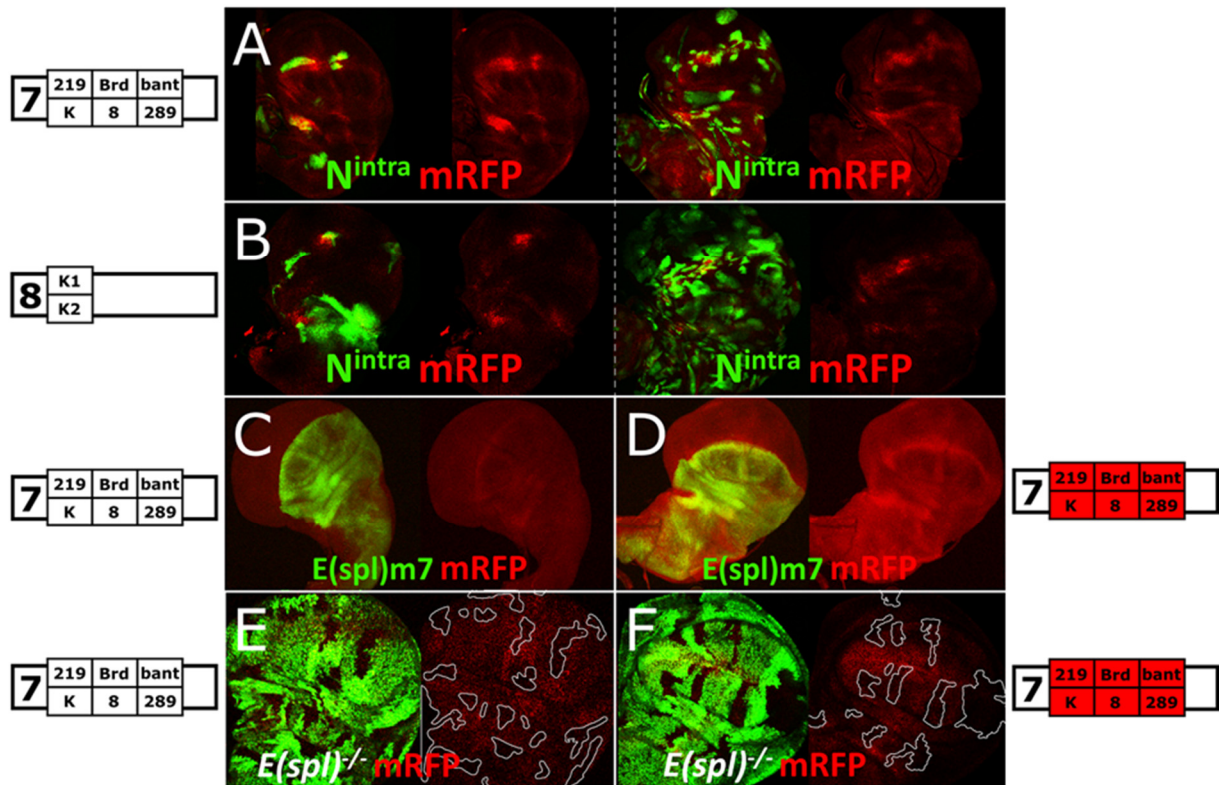


Figure 16 Notch activity impacts 3' UTR-mediated regulation of *E(spl)m7* and *E(spl)m8*. (A and B) Clonal expression of *N^{intra}*, marked by GFP in green, in the larval wing disks expressing tub-mRFP sensors for the wild-type 3'UTR of *E(spl)m7* (A) and *E(spl)m8* (B); two examples for each are shown. (C and D) *apterous*-driven expression of *E(spl)m7* (marked by GFP, in green) in the wing disks expressing tub-mRFP sensors for the wild-type (C) and mutant (D) 3'UTRs of *E(spl)m7*. (E and F) Clonal analysis of these same sensors (as in C and D, respectively) using complete deficiency of the *E(spl)* locus (*Df(3R)gro^{32.2}*), marked with non-GFP cells (outlined in the RFP, red, channel).

2.4.4. Impact of transcriptional and translational repression on *E(spl)m7* and *E(spl)m8* expression and function

To test whether the identified elements mediating transcriptional and post-transcriptional repression influence levels of expression and biological functions of the *E(spl)m7* and *E(spl)m8* I obtained fly transgenic lines carrying an intact genomic fragment encompassing *E(spl)m7* and *E(spl)m8* genes with introduced mutations of the negative regulatory elements characterized in the previous sections in their enhancers (B/C boxes) and 3'UTRs (miRNA binding sites). In a series of these transgenes, *E(spl)m7* was tagged with EGFP, what allowed me to see an effect of these mutations on the *E(spl)m7* expression (some of these were used in the analysis of expression patterns shown in **Figure 3**). Semi-quantitative comparison of expression levels in the third instar larvae wing disks revealed that the mutations in the 3'UTR of the *E(spl)m7* had the strongest effect on derepression of its expression, while the regulation mediated through *E_B/E_C* boxes had much milder effect (**Figure 17**). Much like it was shown with tub-mRFP sensor transgenes, replacement of the *E(spl)m7* 3'

UTR with SV40 3' UTR had stronger effect than mutating individual microRNA binding sites within *E(spl)m7s* 3' UTR (**Figure 17**, compare **E** to **B** to **A**). Although the effect was much weaker, mutations of the E_B/E_C boxes derepressed expression of *E(spl)m7* (**Figure 17**, compare **C** to **A**), contributed to further derepression inflicted by mutating microRNA binding sites (**Figure 17 D**), and produced the strongest levels of expression in combination with the SV40 3' UTR (**Figure 17 F**). Thus, transcriptional repression seems to have an accessory role on the regulation of *E(spl)m7* expression. Interestingly, transcriptional repression of the *E(spl)m7* seems to be mediated by the E_B/E_C boxes contained in the enhancer located immediately upstream of *E(spl)m8* (i.e., *e8*); this holds even for the AMPs, where *E(spl)m7* expression is activated by the *e7* enhancer (Figure 10A). Deletion of the *E(spl)m8* regulatory part causes upregulation of the *E(spl)m7* expression (**Figure 17 G** and **H**, compare to **A**). Similar effects are observed with mutations of the E_B/E_C boxes in the “full” genomic transgenes (**Figure 17**, compare **C** and **I** to **A**). In the latter case, mutating the E_B/E_C boxes within both enhancers have stronger effect than their mutation only within *e8* (**Figure 17**, compare **C** to **I**), suggesting that in the “genomic” context (when both enhancers are present) B1 box present in *e7* is also functional. This activity of the B1 box is not observed in the “short” genomic context (**Figure 17**, compare **G** to **H**) and the reporter gene context (**Figure 10**, compare **A** to **B**), where the *e7* is not accompanied by *e8*. This adds support to the hypothesis that the two enhancers *e7* and *e8* do not act independently of each other, they rather interact in the cells of the wing disk.

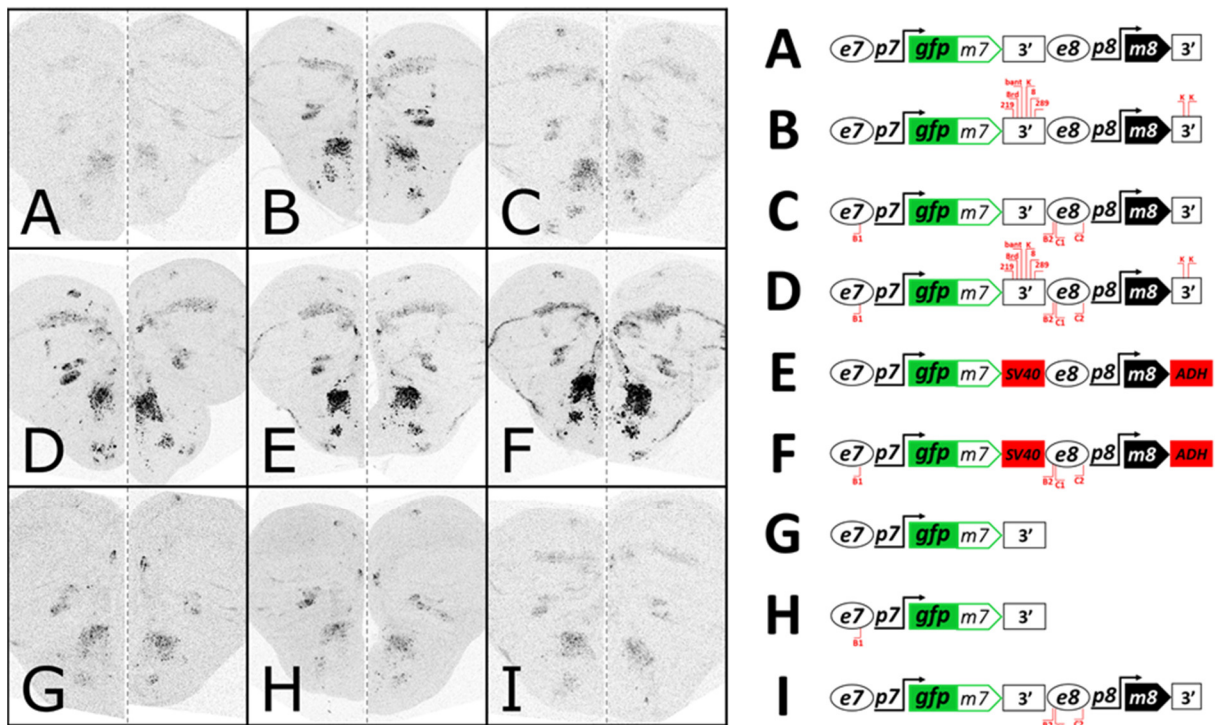


Figure 17 *E(spl)m7* expression is regulated by microRNAs and transcriptional repressors. *GFPm7* expression in the wing disks from flies carrying one copy of “genomic” transgenes schematically shown on the right. Two wing disks aligned with their anterior parts are shown for each genotype. All transgenes are inserted in the same, *attP40*, locus.

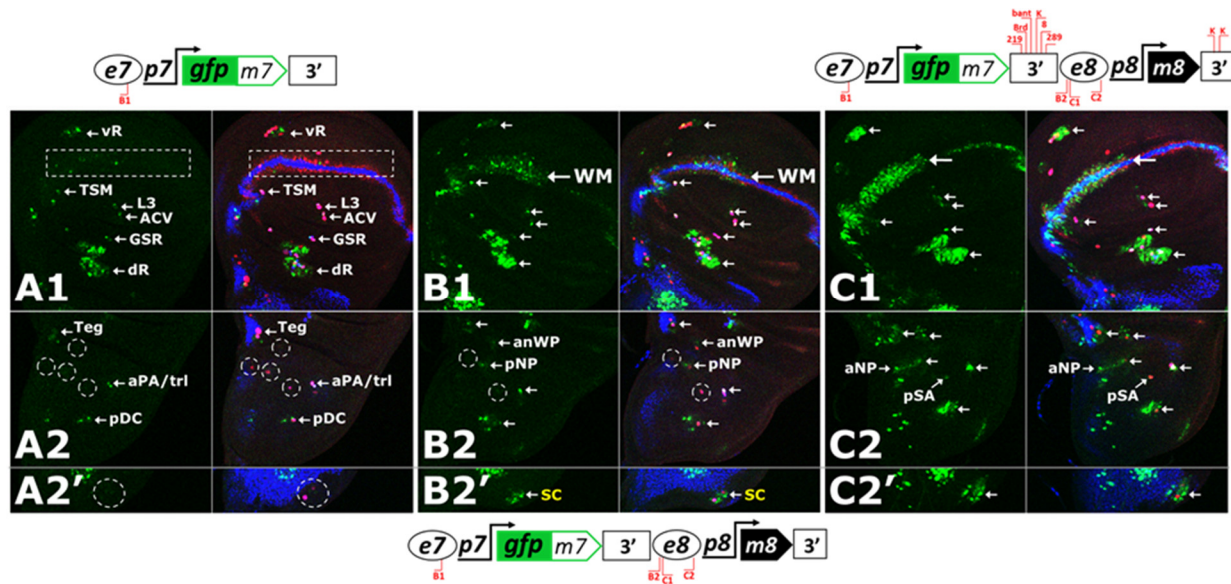


Figure 18 GFPm7 is detected in all PNCs of the wing disk upon mutating all E_B and E_C boxes and identified microRNA binding sites within $E(spl)m7$ 3' UTR. Confocal images of the wing disks from animals carrying one copy of an indicated (in columns) genomic transgene. Wing disks were immunostained against GFP (green) to detect $E(spl)m7$, Cut (blue) to detect the AMPs of the notum, the wing margin and all sensory organ cells and Sens (red) to mark SOPs. (**A1-C1**) Full projections of the wing pouch. (**A2-C2**) partial z-projection of the notum and hinge showing sections underlying the AMPs where the indicated PNCs reside. (**A2'-C2'**) partial z-projection of the most posterior part of notum showing sections containing the AMPs where the SC PNC reside. PNCs where GFPm7 is detected are indicated by arrows, whereas the lack of GFPm7 expression in them is shown by circles; The absence of the GFPm7 in the WM is indicated by a rectangle.

Analysis of the GFPm7 expression from these transgenes shows that this fusion protein is not detected in most proneural clusters of the wing disk unless mutations in the $E(spl)m7$'s 3' UTR are introduced in the transgene (**Figure 18**).

A detailed analysis of the expression pattern obtained from the $GFPm7$ transgenes harboring E_B/E_C mutations, microRNA seed boxes mutation, and SV40 and ADH 3' UTRs shows gradual increase in GFP expression in the WM, the AMPs and all PNCs (**Figure 19**). However, of note, is a visible and disproportionate increase of the GFP accumulation in the SC and DC clusters.

In order to evaluate an impact of these mutations on the biological function of $E(spl)m7$ and $E(spl)m8$ I have generated an analogous series of genomic constructs without GFP fused to $E(spl)m7$ (shown schematically in **Figure 20**), to avoid possible inhibition of the GFP-fused protein on $E(spl)$'s functions. Each of the large bristles (macrochaetes) of the adult fly's thorax arises from a specific SOP generated in the notum of the imaginal wing disk (Modolell and Campuzano 1998) (**Figure 20 A-E**). I noticed that some individuals of the flies carrying one copy of the various genomic transgenes show gain or lack of these bristles (some examples focused on DC and SC bristles are shown in **Figure 20 F**). Phenotypic quantification revealed that the increasing derepression of $E(spl)m7$ and $E(spl)m8$ expression, as estimated by the analysis of the $GFPm7$ transgenes (**Figure 17**), triggers increasingly greater loss of macrochaetes among individuals in the populations of transgenic flies (**Figure 20 G**). Interestingly, macrochaetes were affected disproportionately, with the SC and DC bristles being most impacted (**Figure 20**

G, cluster “A”), consistent with a proportionally higher accumulation of GFPm7 in these clusters **Figure 19**). Whereas point mutations within enhancers (EB/EC boxes) and 3’ UTRs (microRNA seed boxes) had little to moderate effect on the bristle loss, replacement of the *E(spl)m7* and *E(spl)m8* 3’ UTRs with SV40 and ADH 3’ UTRs augmented this effect more than 10x (from 17 and 24 bristles missing per 100 individuals for T^{BC} and T^{BC+mi} , respectively, to 226 and 280 bristles missing per 100 individuals for T^{3UTR} and $T^{BC+3UTR}$, respectively, see **Figure 20 G**).

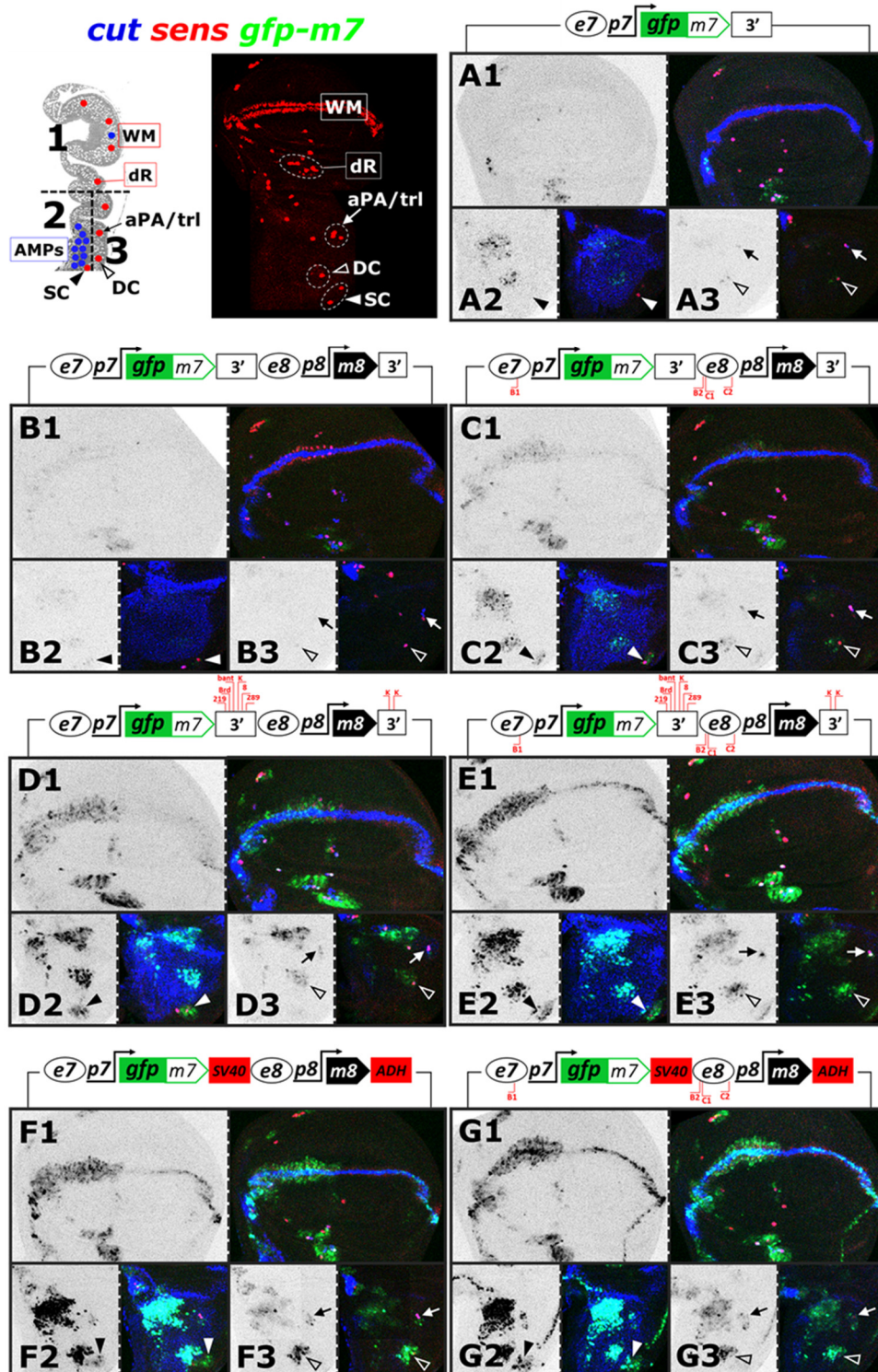


Figure 19 GFPm7 is accumulated with increasingly higher levels in the WM, AMPs and PNCs of the wing disk as the transcriptional negative regulation and microRNA regulation is compromised. Wing disks from the animals carrying one copy of a transgene indicated schematically were immunostained against GFPm7 (green), Sens (red) and Cut (blue). (A1, B1, C1, D1, E1, F1 and G1) show full z-projections of the wing pouch with GFPm7 expression on the left and merge of all three channels on the right. (A2, B2, C2, D2, E2, F2 and G2) show the partial z-projection of the notum showing sections containing the AMPs where also the SC PNC resides (indicated by white filled arrowhead). (A3, B3, C3, D3, E3, F3 and G3) show the partial z-projection of the notum showing sections underlying the AMPs where the rest of notal PNCs reside (DC and aPA/trl PNCs are indicated by white empty arrowhead and white arrow, respectively).

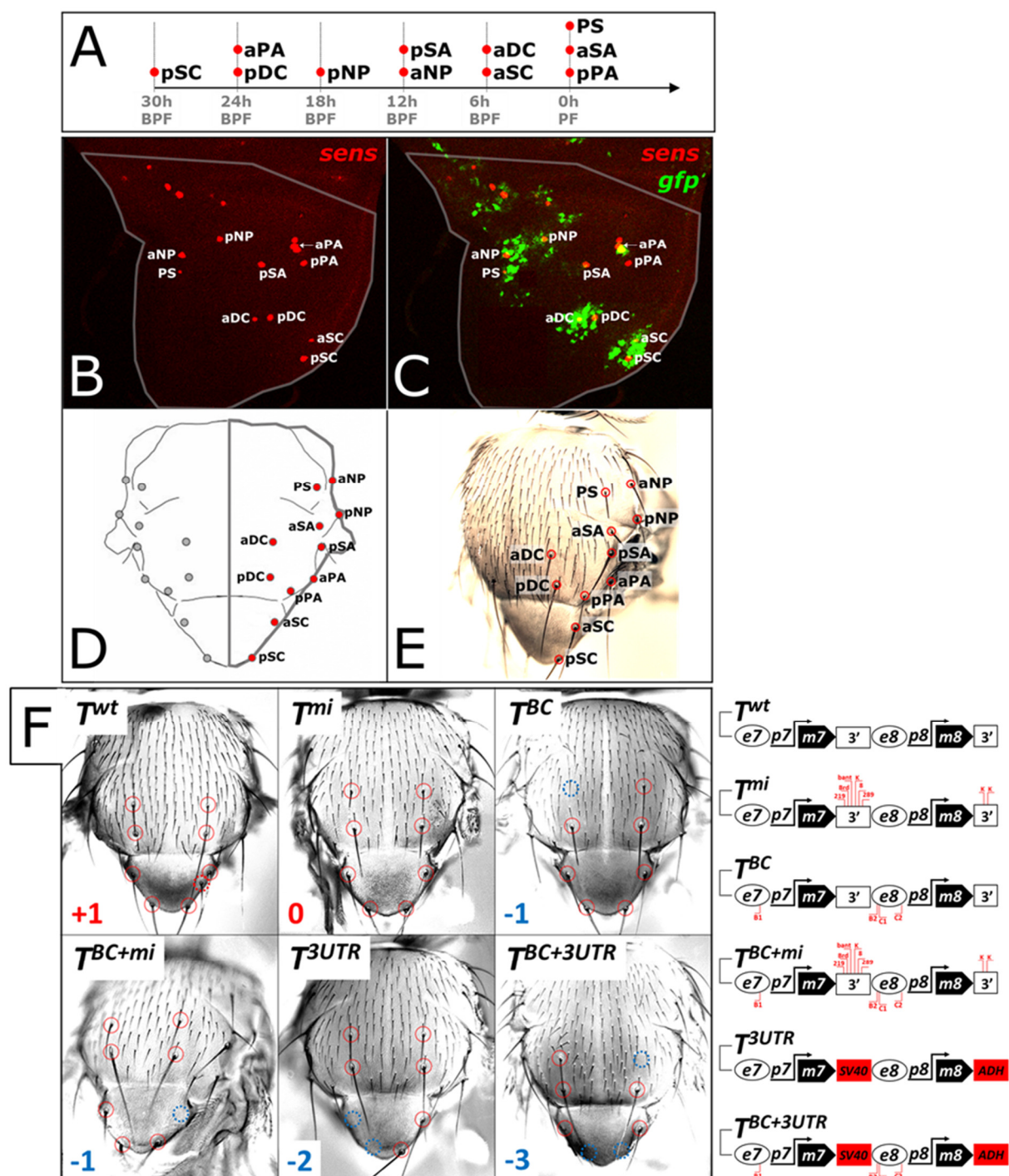
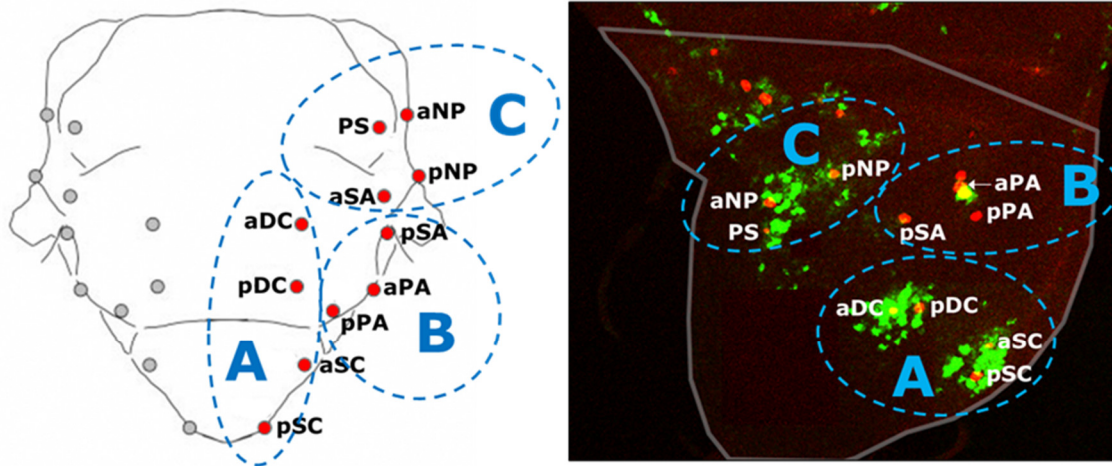


Figure 20 - continues on the next page



individuals scored		n=106	n=119	n=85	n=202	n=114	n=85	n=74
Transgene		-	T^{wt}	T^{BC}	T^{mi}	T^{BC+mi}	T^{3UTR}	$T^{BC+3UTR}$
A	pSC	0	0	0	0	0	-95	-124
	aSC	2	1	1	-6	-12	-71	-97
	pDC	0	0	0	-2	-2	-25	-30
	aDC	-1	-5	-4	-1	-5	-16	-19
sum		1	-4	-3	-9	-19	-207	-270
B	pPA	0	0	0	0	0	0	0
	aPA	0	0	0	-1	0	0	0
	pSA	0	0	0	0	0	-15	-10
sum		0	0	0	-1	0	-15	-10
C	aSA	0	0	0	0	0	-3	0
	pNP	5	0	0	0	0	0	0
	aNP	0	0	0	-7	-5	-1	0
	PS	0	0	0	0	0	0	0
sum		5	0	0	-7	-5	-5	0
total sum		6	-4	-3	-17	-24	-226	-280

Values are given per 100 individuals

Figure 20 Misregulation of *E(spl)m7* and *E(spl)m8* expression impacts SOP specification. (A) The emergence of the SOPs in the developing wing disk with time scale in hours (h) before pupae formation (BPF), modified from (Huang *et al.* 1991). (B) SOPs (*sens*, red) of the notum of the late third instar wing disk just before pupation, and (C) overlaid GFPm7 expression from a genomic transgene with *E(spl)m7* fused to the SV40 3' UTR. (D) Macrochaetes of the thorax are indicated with red dots (D) and red circles (E) on the schematic and image of the adult thorax, respectively. (F) Examples of the thoraces from individuals bearing one copy of the genomic transgenes (shown schematically); Presence of DC and SC bristles are indicated on each thorax with red circles, whereas loss of thereof is indicated by blue circles. The number under each thorax is derived from the total sum of ectopic (+n) and missing (-n) bristles. (G) Table shows quantification of all 11 pairs of macrochaetes derived from wild-type animals (-) and transgenic animals carrying one copy of indicated transgene. Macrochaetes are grouped into three clusters ("A", "B" and "C") based on the proximity of their respective SOPs in the wing disk. Notice, that the macrochaetes of the group "A" are most affected by the mutations compromising negative regulation of *E(spl)m7* and *E(spl)m8*.

I focused on the analysis of the most affected, SC and DC, macrochaetes. The SV40 and ADH 3' UTR-containing transgenes affected mainly SC bristles, with a slight bias toward posterior SC (**Figure 21**). This bias was also evident among DC bristles; the posterior DC was missing twice as often as the anterior. As the posterior SOPs in both, SC and DC, clusters appear earlier in the development than the anterior ones (see **Figure 20 A**), these observations suggest that the elevated expression of E(spl)m7 and E(spl)m8 has the most profound effect on the early developing SOPs.

The severity of the phenotype observed with transgenes containing SV40 and ADH 3' UTRs is underscored by the lethality of these transgenes. I did not observe any flies homozygous for either version of these transgenes (with E_B/E_C boxes wild-type or mutated). Moreover, the bristle-loss phenotype of the heterozygotes for these transgenes are highly (for the version with E_B/E_C boxes wild-type) and fully penetrant (for the version with mutant E_B/E_C boxes) (**Figure 21**, see the graph on top). In contrast, the wild-type and point-mutated transgenes triggered SC/DC bristle loss in no more than 20% individuals, were homozygous viable, and affected mostly anterior SC/DC bristles (**Figure 21**). An additional (second) copy of these transgenes does have an impact on the SC/DC bristle loss, however, less than two-fold (**Figure 22**). This observation suggests that the expression of E(spl) proteins from these transgenes is still efficiently buffered by transcriptional and post-transcriptional regulation so that expression levels achieved from them are not exceeding a certain threshold needed to upset phenotypic robustness.

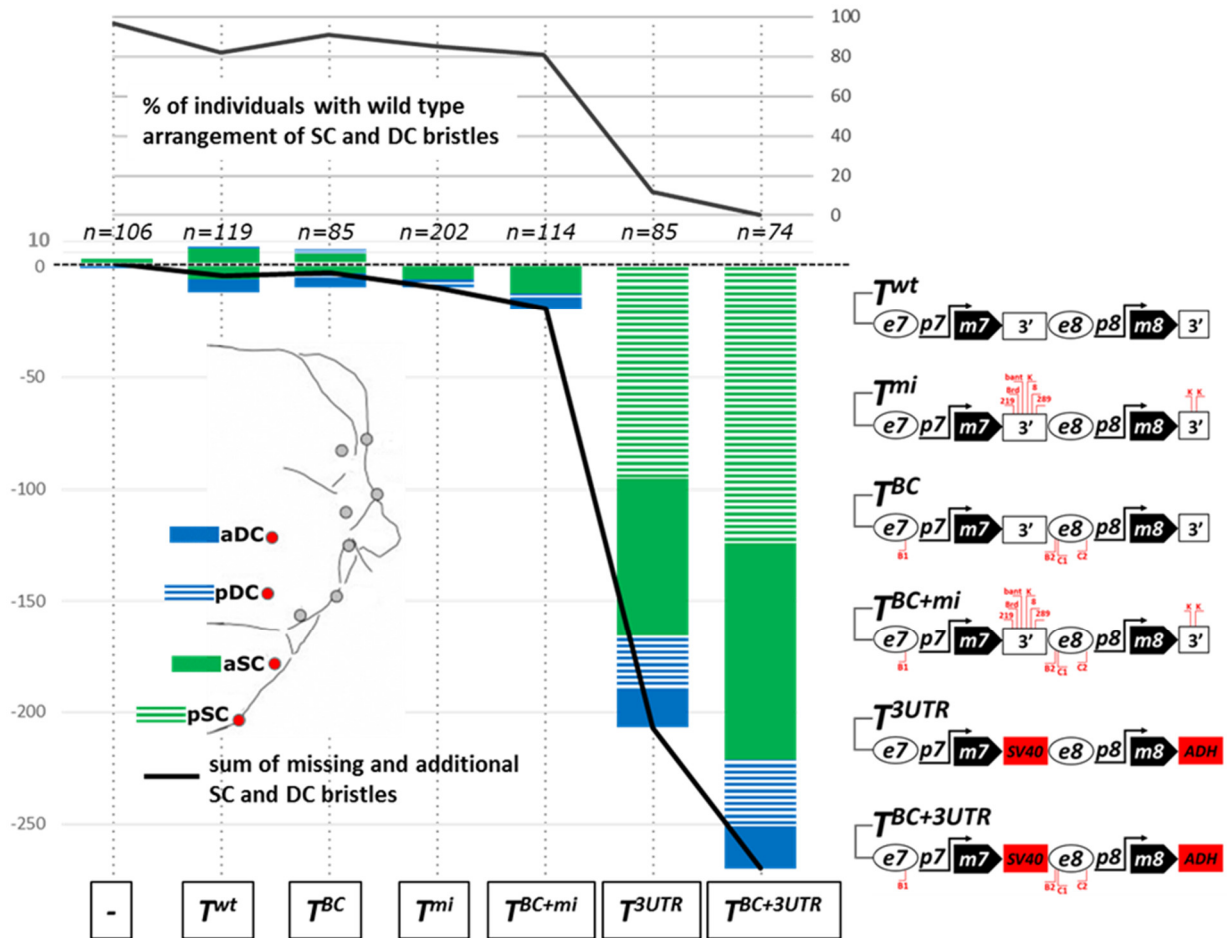


Figure 21 Misregulation of *E(spl)m7* and *E(spl)m8* induces SC/DC bristle loss. Flies containing one copy of each transgene shown schematically on the right were examined for the SC/DC bristles gain and loss. The number of gained and lost bristles is plotted for each class of bristles (aDC, pDC, aSC and pSC) as bars for each transgene. The sum of total missing and gained bristles for each transgene is plotted as a line and overlaid on the bar chart. The portion of individuals with wild-type arrangement and number of SC/DC bristles for each transgene is plotted as a curve on top. Values are given per 100 individuals. The number of individuals scored is given above the bar chart ("n").

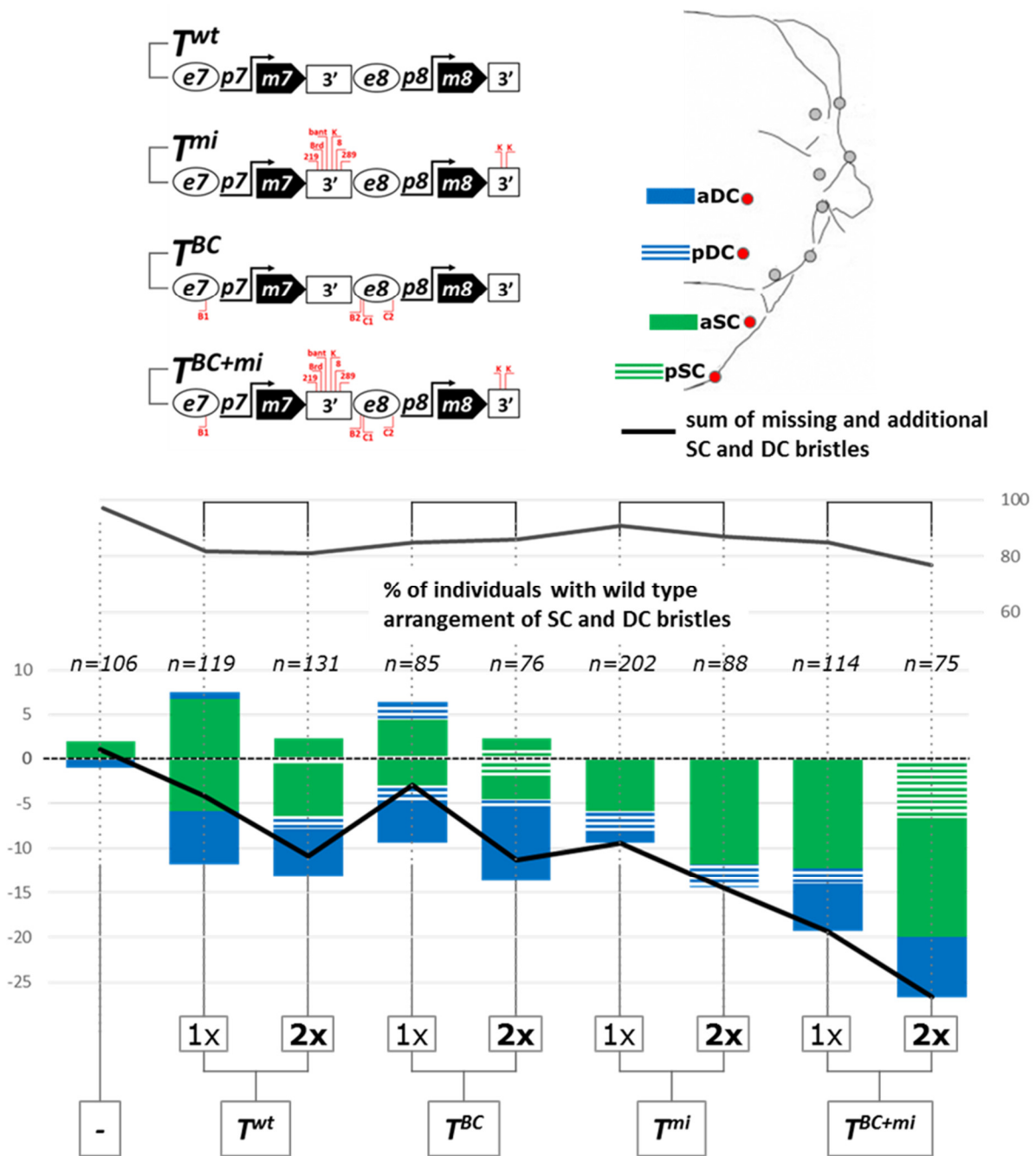


Figure 22 Effect of transgene dosage on SC/DC bristle loss. Flies heterozygous (“1x”) and homozygous (“2x”) for each transgene (shown schematically above) were examined for the SC/DC bristles gain and loss. The number of gained and lost bristles is plotted for each class of bristles (aDC, pDC, aSC and pSC) as bars for each transgene. The sum of total missing and gained bristles for each transgene is plotted as a line and overlaid on the bar chart. The portion of individuals with wild-type arrangement and number of SC/DC bristles for each transgene is plotted as a curve on top. Values are given per 100 individuals. The number of individuals scored is given above the bar chart (“n”)

I noticed that halving the *E(spl)* dosage, namely in the *E(spl)* null heterozygote animals, supports relatively normal peripheral neurogenesis, however with apparent disturbances in patterning anterior SC bristles, such that, on average, each individual exhibits a gain of one aSC bristle (**Figure 23**). Therefore, specification of this class of bristles is particularly sensitive to the expression levels of the *E(spl)* genes. The aSC bristles were also the primary targets of the elevated levels of *E(spl)m7* and *E(spl)m8* (see **Figure 21, 22**).

I asked how the expression of *E(spl)m7* and *E(spl)m8* from my transgenes would influence SC/DC bristle patterning in this sensitized background. One copy of the wild-type and *E_B/E_C*-mutant transgenes (*T^{wt}* and *T^{BC}* in **Figure 23**) had little or no effect on the aSC phenotype observed in *E(spl)* null heterozygotes. Interestingly, one copy of transgenes bearing point mutations in the *E(spl)m7* and *E(spl)m8* 3' UTRs (*T^{mi}* and *T^{BC+mi}* in **Figure 23**) brought about significant reduction of the extra aSC bristles produced by the loss of one copy of *E(spl)* locus (from 100 to less than 40 per 100 individuals) decreasing the penetrance of this phenotype by 30%. Transgenes bearing SV40 and ADH 3' UTRs (*T^{3UTR}* and *T^{BC+3UTR}* in **Figure 23**) only moderately improved what seems to be the partial rescue of aSC phenotype (to \approx 20 extra aSC bristles per 100 individuals). This modest effect and lack of the full rescue was surprising, considering strong impact of the SV40 and ADH 3' UTR-containing transgenes on SC bristle loss in wild type background (**Figure 21, and 23**, see bars marked by "+"). Strikingly, nearly all the effects on SC/DC bristle loss observed with all transgenes in the wild-type background are lost in the *E(spl)* null heterozygotes (**Figure 23**, with the most prominent reduction in bristle loss for *T^{BC+3UTR}* from total of 270 to 7 SC/DC bristles per 100 individuals). These observations imply that the level-dependent function of the *E(spl)m7* and *E(spl)m8* transgenes as a suppressor of SOP fate relies on the expression levels of an/other gene/s from *E(spl)* locus, and that both copies of this/these gene/s have to be expressed to enable higher activity of *E(spl)m7* and *E(spl)m8*.

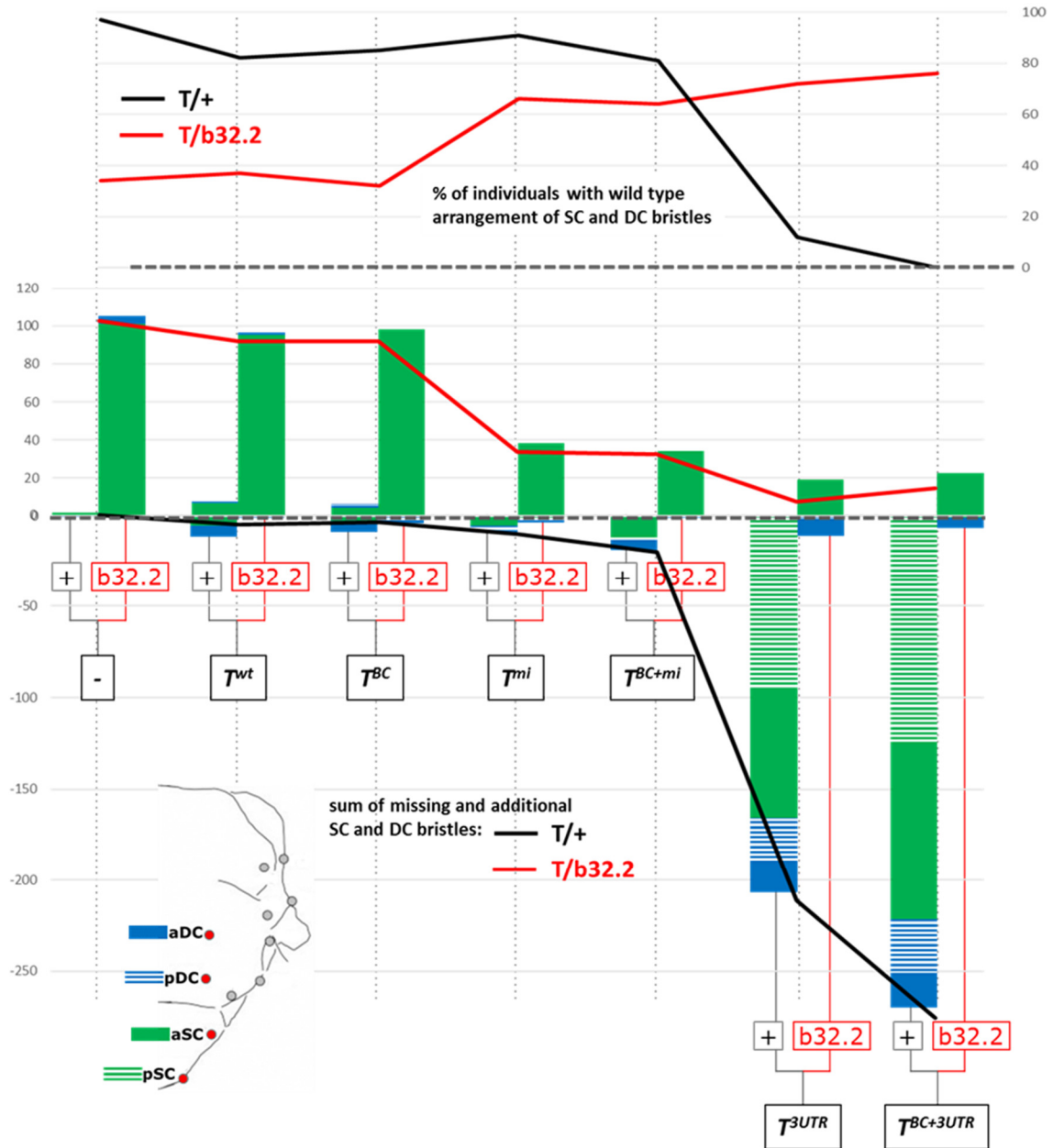


Figure 23 *E(spl)* locus copy number affects level-dependent activity of *E(spl)m7* and *E(spl)m8*. Flies bearing no transgene (first column, indicated by "-") and one of the indicated transgene (in the box below bars) were scored for SC/DC bristle gain and loss in two genetic backgrounds: wild-type (i.e., with two copies of *E(spl)* locus, indicated by "+"; contains the same data set as in Figure 21) and loss of one copy of *E(spl)* locus (heterozygotes for *Df(3R)gro*^{32.2} null allele; indicated by "b32.2"). The number of gained and lost bristles is plotted for each class of bristles (aDC, pDC, aSC and pSC) as bars for each genotype. The sum of total missing and gained bristles for each genotype in wild-type and *E(spl)* null heterozygote background is plotted as a black and red line, respectively, and overlaid on the bar chart. Values are given per 100 individuals. The portion of individuals with wild-type arrangement and number of SC/DC bristles for each genotype is plotted as black (wild-type background) and red (*E(spl)* null heterozygotes) curve on top.

Heterozygotes bearing smaller deletion of *E(spl)* locus, encompassing *E(spl)m5*, *E(spl)m6*, *E(spl)m7* and *E(spl)m8* and *gro* (see **Figure 1**), showed similar preference for

duplication of aSC bristles but 5x less frequent than heterozygotes bearing the complete deletion of *E(spl)* locus (**Figure 24**). However, contrary to the complete deletion of *E(spl)*, this partial deletion was fully rescued by one copy of the wild-type transgene (T^{wt}) and did not suppress the impact of other transgenes on aSC bristle loss (**Figure 24**). Therefore, either (a) other genes than these contained in this deletion are needed for higher activity of *E(spl)m7* and *E(spl)m8*, or (b) it is the cumulative dose of *E(spl)* genes that sets the threshold for the loss of bristles.

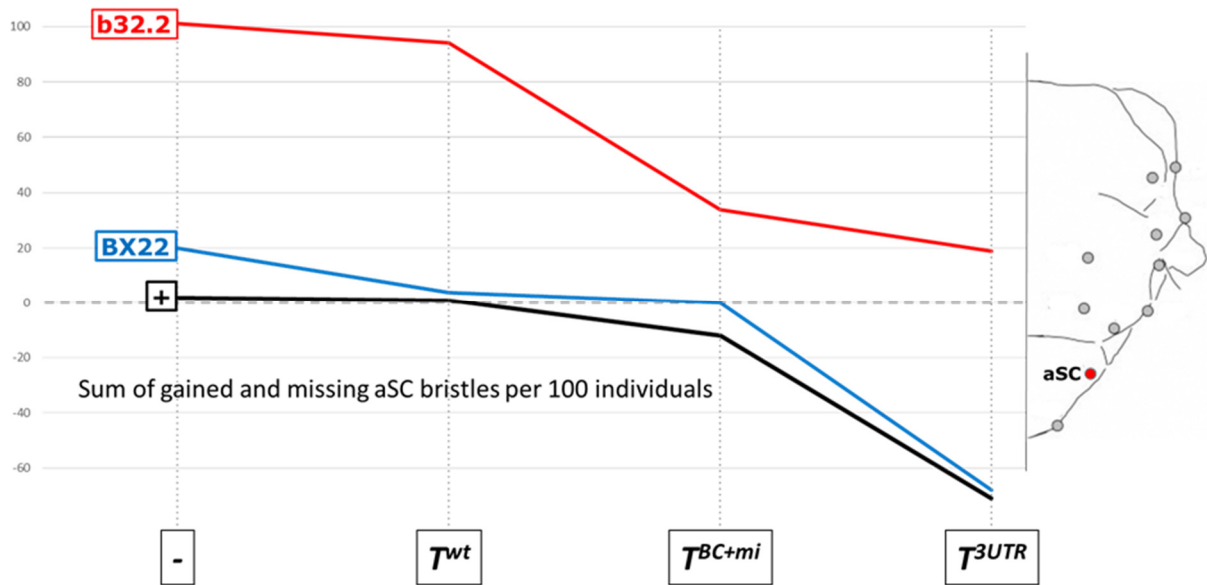


Figure 24 Heterozygous BX22 deletion does not affect level-dependent activity of *E(spl)m7* and *E(spl)m8*. Flies bearing no transgene (first column, indicated by "-") and one of the indicated transgene (in boxes below the graph) were scored for aSC bristle gain and loss in three genetic backgrounds: wild-type ("+"), heterozygous BX22 deletion of the *E(spl)* locus ("BX22") and heterozygous *E(spl)* null *Df(3R)gro^{32.2}* allele ("b32.2"). Sum of gained and missing aSC bristles is plotted for each genotype as black (for wild-type background), blue (for BX22 heterozygotes) and red (for *E(spl)* null heterozygotes) line. Values are given per 100 individuals.

Results presented thus far suggest that obstructing post-transcriptional regulation on *E(spl)m7* and *E(spl)m8* results in surplus levels of expression and activity of these genes. The effect of this activity is observable in disturbances of SOP specification and concomitant bristle loss. However, disabling *dcr-1*-dependent microRNA biogenesis in mutant clones does not affect number and position of SOPs in the larval wing disks nor bristle pattern on the adult's thorax (**Figure 25 A1-A3**). The only exception I could find was that the adult flies with induced clones during the larval stages exhibited occasional loss of a macrochaete shaft while preserving its socket (**Figure 25 A3**, red arrow). Because the socket and the shaft are siblings resulting from an asymmetric division of a single precursor (Lai and Orgogozo 2004), this implies that inhibition of *dcr-1*-dependent biogenesis of microRNAs affects bristle lineage post-SOP formation but not SOP specification itself. Despite this apparent lack of phenotype,

dcr-1 clones show elevated levels of GFPm7 fusion protein whose expression is under the control of wild-type 3' UTR of *E(spl)m7* (Figure 25 B1-B3, C1-C3)., Mutating microRNA binding sites within this 3' UTR rendered GFPm7 expression insensitive to *dcr-1* clones (Figure 25 D1-D3).

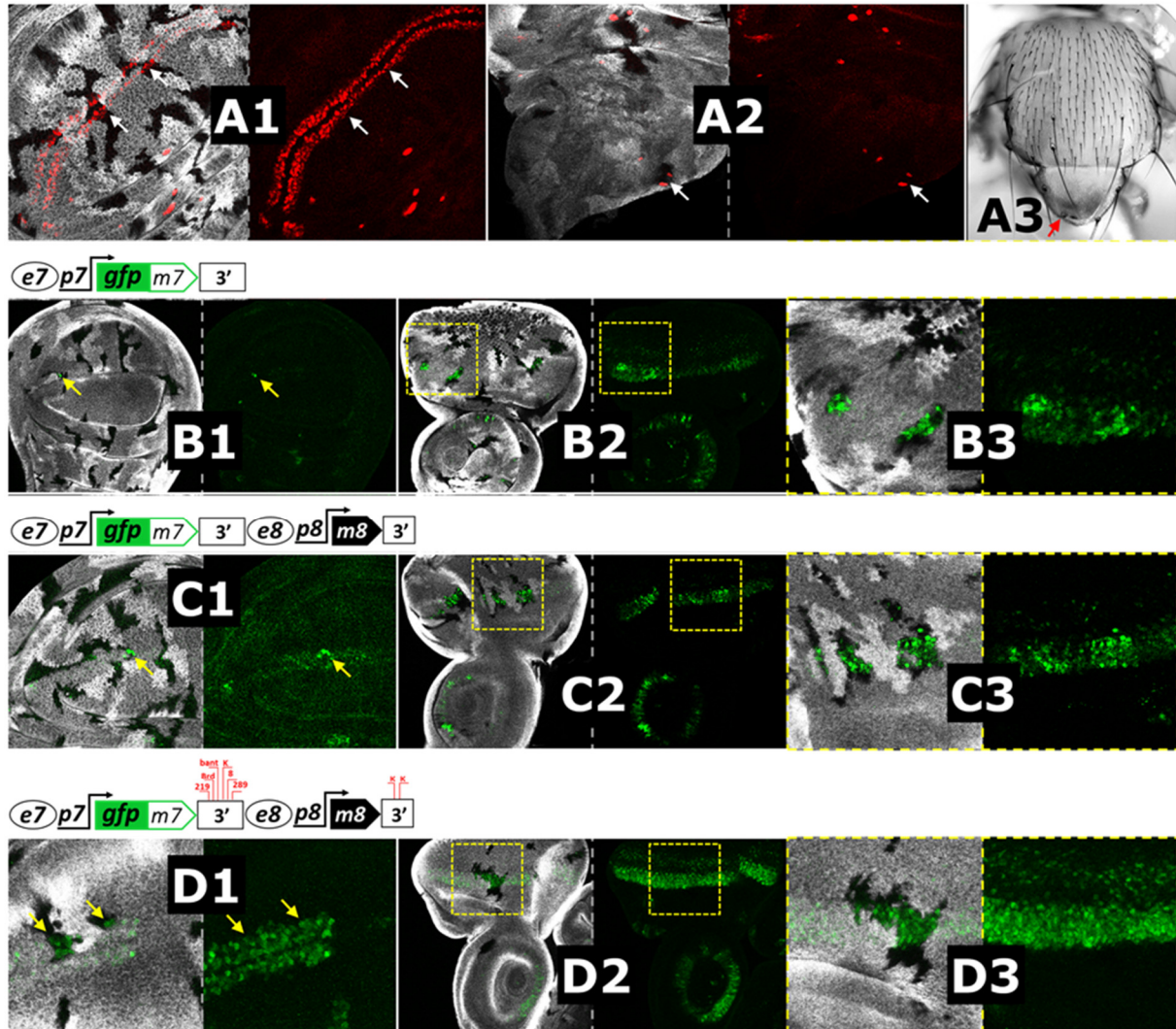


Figure 25 *E(spl)m7* 3' UTR-mediated regulation, but not SOP specification is affected in the *dcr1* null clones. (A1-A2) *dcr1* null clones (unlabeled cells by shades of grey) induced in the wing pouch (A1) and notum (A2) of the third instar larval wing disk; SOPs are stained with anti-Sens antibody (red); arrows indicate SOPs in the clones of the WM; (A3) Thorax of the adult fly with induced *dcr1* null clones in the early larval stage shows wild-type pattern of bristles with the occasional loss of bristle shaft (red arrow). *dcr1* null clones in the area of the wing pouch (B1, C1, D1) and in the eye-antennal disks (B2, C2, D2; with a magnification of the respective indicated areas B3, C3 and D3) of the animals bearing expressing GFPm7 (green) from genomic transgenes shown schematically above.

Unlike *dcr-1*, induction of mutant clones for *E(spl)* locus causes disorder of the WM's SOPs and supernumerary SOPs within PNCs of the larval wing disk, disturbances of microchaetal pattern and gain of micro- and macro-chaetes on the adult's thorax (Figure 26).

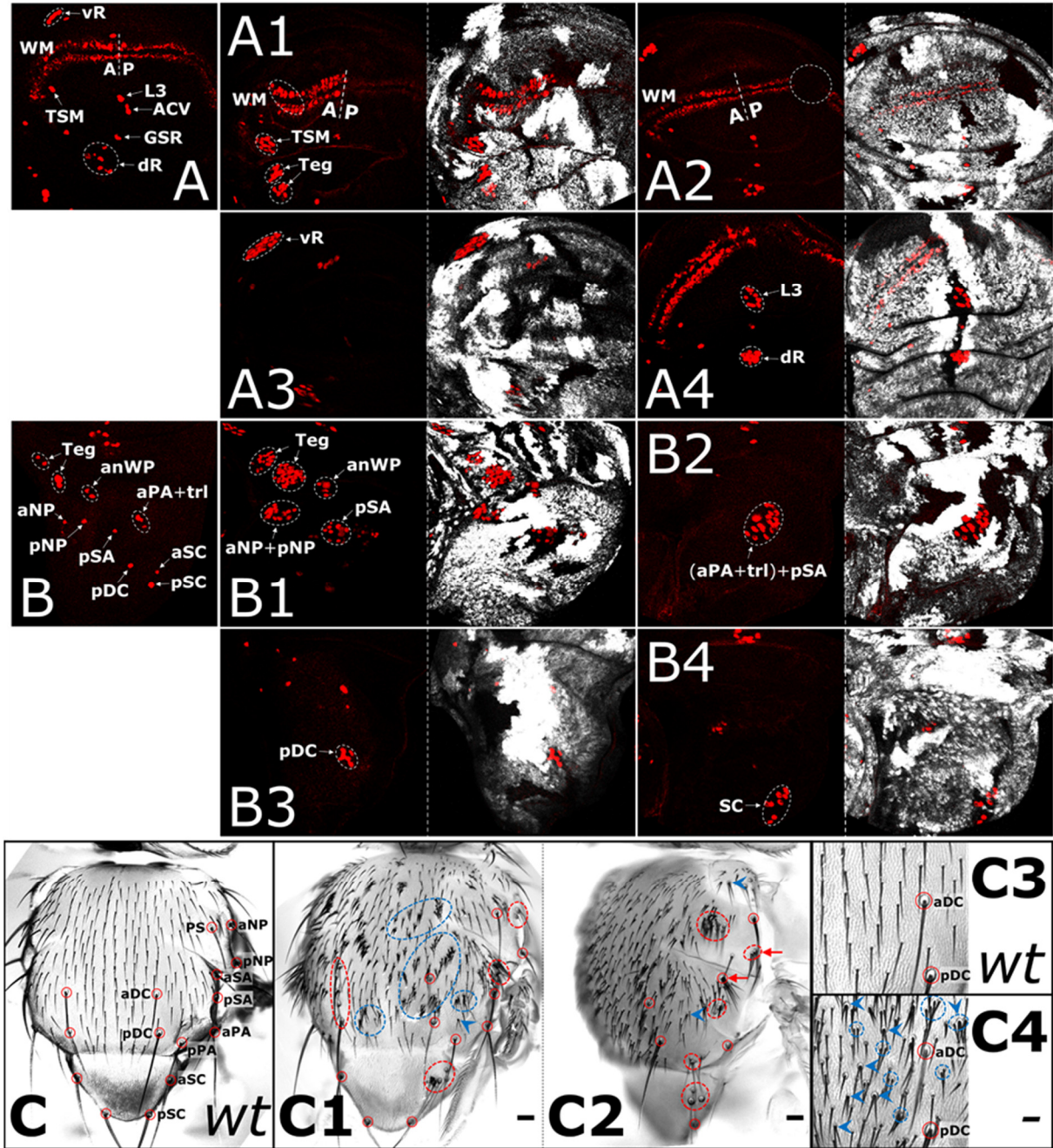
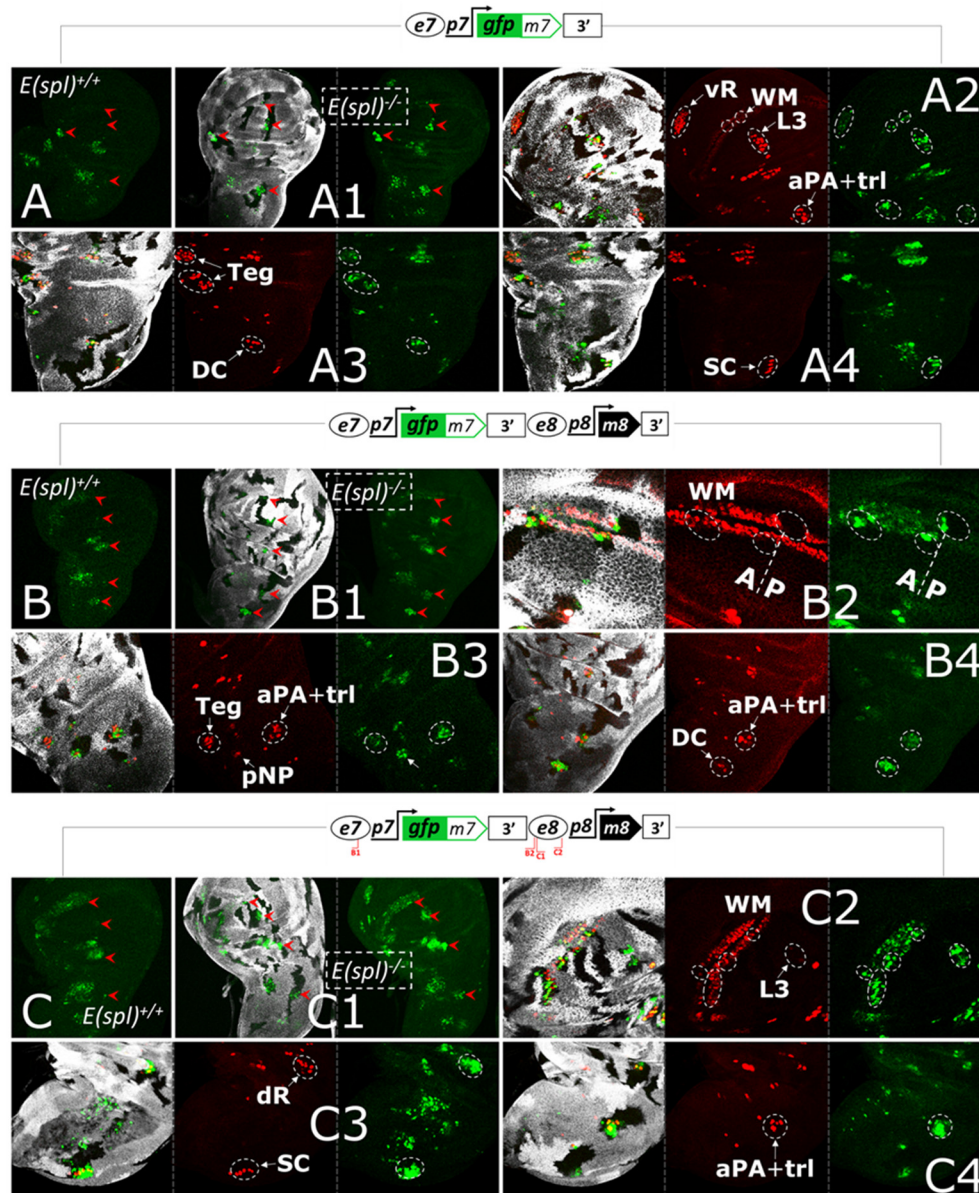


Figure 26 Loss of *E(spl)* locus affects specification and patterning of SOPs and bristles. *E(spl)* null clones in the third instar wing disks are marked by lack of shades of grey, SOPs are labeled with red (*sens*). (A) and (B) show wing pouch and notum, respectively, of the wild-type, fully *E(spl)*^{+/+} wing disk. (A1-A4) and (B1-B4) show examples of wing pouches and nota, respectively, with induced random *E(spl)* null clones. Note an excess of SOPs in a clone of anterior part of the WM (A1) and absence of *sens* staining in a clone of posterior part of the WM (A2). Note, that each time a clone falls within an area of a given PNC, the SOP/s of this PNC is amplified (encircled by dotted lines, compare to their equivalent wild-type SOPs). Thorax (C) and magnification of its DC area (C3) of the wild-type fly with indicated macrochaete (red circles). Two examples of thoraxes (C1-C2) and magnification of the DC area (C4) from animals with induced *E(spl)* null clones; red dotted circles and blue dotted circles indicate supernumerary macrochaetes and microchaetes, respectively; red arrows and blue arrowheads indicate duplications of, respectively, the macrochaetes' and microchaetes' shafts within single socket.

These clones show also elevated levels of GFPm7 when induced in wing disks of transgenic animals bearing genomic *GFPm7* transgenes (**Figure 27**, note, however, that the difference between expression levels of the *GFPm7* controlled by SV40 3' UTR within and outside the clones is much less pronounced). Interestingly, all *GFPm7* transgenes with mutated and replaced 3' UTRs fully restored normal arrangements of SOPs within *E(spl)* clones (**Figure 27**, **D2-D4**, **E2-E4**). This means that the two copies of *E(spl)* genes provided by a hemizygous transgene are sufficient to provide the function of all 14 copies (2x7) of the endogenous *E(spl)* genes, but only when relieved of miRNA suppression. Whereas, *GFPm7-m8* transgenes with wild-type 3' UTRs and wild-type or mutated E_B/E_C -boxes did not prevent *E(spl)* null-induced disorder of the WM (**Figure 27** **A2**, **B2**, **C2**) and supernumerary SOPs within PNCs (**Figure 27** **A3-4**, **B3-4**, **C3-4**). Of note is pronounced GFP derepression seen in *E(spl)* mutant clones expressing E_B/E_C -mutant *GFPm7-m8* transgene (**Figure 27** **C1-C4**). This demonstrates, in agreement with my previous results with lacZ reporter transgenes, that *E(spl)* factors mediate repression of *e7* and *e8* enhancers not (only) through E_B/E_C -boxes.



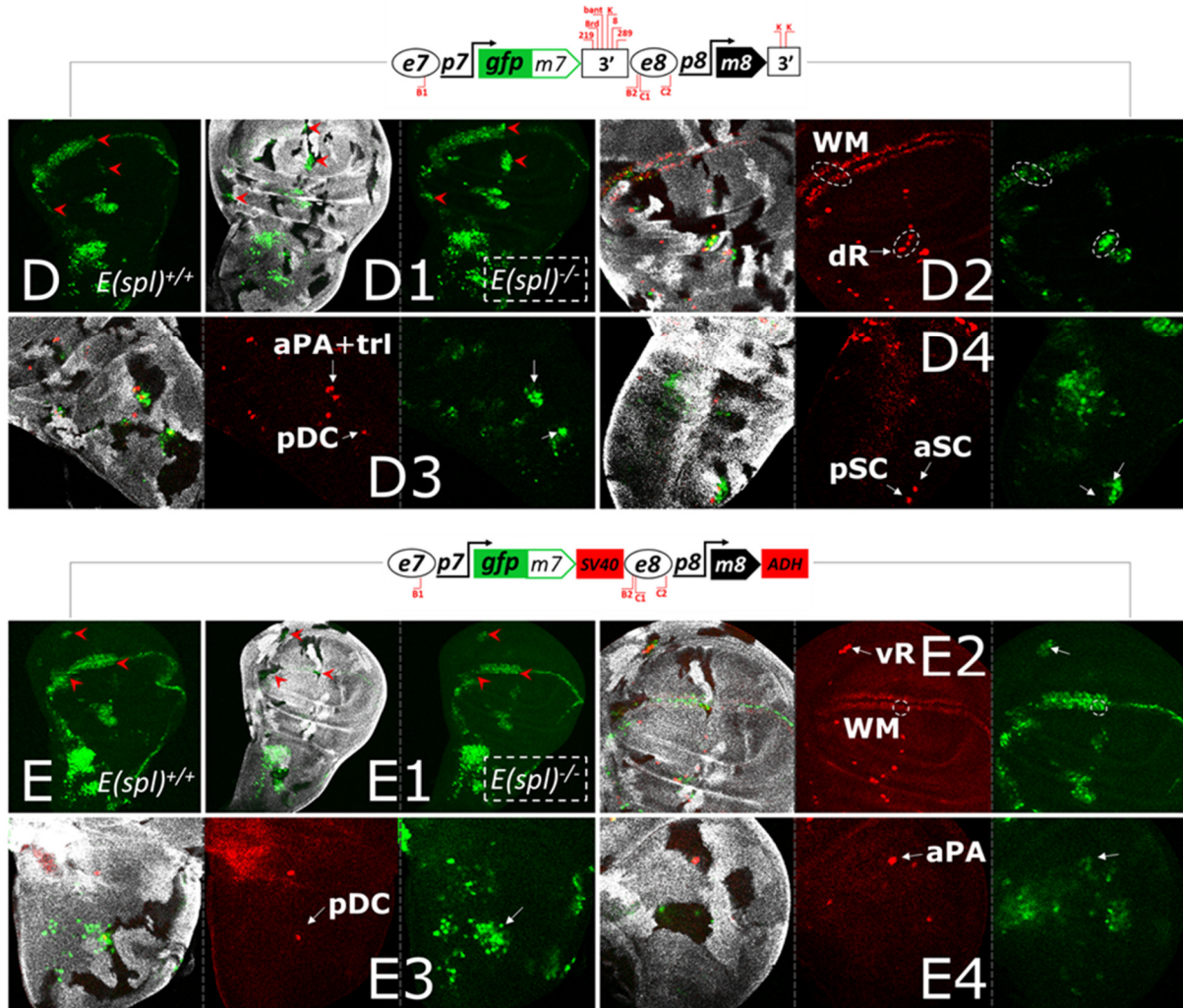


Figure 27 Loss of *E(spl)* stimulates *GFPm7* expression from genomic transgenes. Clonal analysis of the *GFPm7* expression (in green, driven from genomic transgenes depicted above each panel) using complete deficiency of the *E(spl)* locus (*Df(3R)gro^{32.2}*). (A, B, C, D and E) Fully z-projected wing disks expressing *GFPm7* in wild-type background (*E(spl)^{+/+}*) are compared to fully z-projected wing disks with expression of *GFPm7* from the same, respective, transgenes but in the mutant background (*E(spl)^{-/-}*) (A1, B1, C1, D1 and E1). Unlabeled cells (by shades of grey) are homozygous for *Df(3R)gro^{32.2}*. Red arrowheads mark areas of GFP clonal expression (in A1, B1, C1, D1 and E1) and in analogous areas of GFP pattern in the wild-type background (in A, B, C, D and E). (A2, B2, C2, D2 and E2) show one example (for each transgene) of clones induced in the wing pouch and (A3-A4, B3-B4, C3-C4, D3-D4 and E3-E4) two from examples of clones induced in the notum. SOPs are visible in red (*sens*). Abnormal SOPs formations and corresponding GFP expression are indicated with arrows and encircled with dashed line; normal SOPs formations are indicated just with arrows.

This shows that *GFPm7-m8* genomic transgenes provide functional *E(spl)m7* proteins. However, flies homozygous for the SV40/ADH 3' UTR *GFPm7-m8* transgenes, unlike their GFP-less analogues, are not fully lethal, suggesting that the *E(spl)m7* fused to GFP is less potent. Therefore, for the examination of the *E(spl)* null bristle rescue phenotypes in adult flies I have used genomic GFP-untagged transgenes (i.e., *T^{wt}*, *T^{BC}*, *T^{BC+mi}* and *T^{BC+3UTR}*). Random clones of the null *E(spl)* locus were induced at early larval stage, in the same way as it was done in the presence of GFP transgenes. The entire population of flies arising from these larvae were subsequently screened for aberrations in the bristle patterning of the thorax. **Figure 28** shows

representative examples of the observed bristle patterns. Similarly to their GFP analogues, transgenes with wild-type 3' UTRs (T^{wt} and T^{BC}) did not prevent formation of extra macrochaetes (red dotted circles in **Figure 28 C** and **D**, compare to **A** - wild-type arrangement of bristles, and, **B** - thorax with $E(spl)$ null clones induced in the non-transgenic animal). Neither did transgene with microRNA-boxes and E_B/E_C -boxes mutated (T^{BC+mi} , **Figure 28 E**). However, all three transgenes seemed to partially rescue the supernumerary phenotype of macrochaetes. In these cases, I did not observe so numerous accumulation of bristles within one PNC-bristle cluster position as in the non-transgenic control thoraces with $E(spl)$ null clones (**Figure 28**, compare **C-E** to **B**). The T^{wt} seemed to be less effective than the other two transgenes.

Notably, striking differences occur in the patterning of microchaetes. These finer bristles do not originate from SOPs in the larval stages but later on from SOPs specified during early pupal development and are organized in more or less regular rows on the fly's thorax (Corson *et al.* 2017) (**Figure 26 A**, **Figure 28 A**). Induction of $E(spl)$ null clones results in their multiplication and disorganized pattern (**Figure 26 C1-C4**, **Figure 28 B**, some of the presumed clones are marked with blue ovals, duplicated bristles are indicated with blue arrowheads). Less severe distortion in microchaete patterning is observed on thoraces of clonally mutant animals bearing transgenes with wild-type 3' UTRs (T^{wt} and T^{BC}), though disorganization, duplication and gain of bristles are still clearly visible (**Figure 28 C** and **D**). None of these distortions were observed with the transgene containing microRNA-boxes and E_B/E_C -boxes mutated (T^{BC+mi} , **Figure 28 E**), indicating that this transgene produces enough $E(spl)m7$ and $E(spl)m8$ proteins to fully rescue micro- but only partially rescue macro-chaete patterning upon the loss of $E(spl)$ locus. Remarkably, the use of transgene with E_B/E_C -boxes and SV40/ADH 3' UTRs ($T^{BC+3UTR}$) brought about what seems to be an "over-rescue" of the $E(spl)^{-/-}$ microchaete phenotype (**Figure 28 F**). This is manifested by thorax balding in patches presumably outlining a territory of an induced $E(spl)$ null clone (blue ovals, **Figure 28 F**). Additionally, the thoraces of this genotype lack frequently macrochaetes (green circles, **Figure 28 F**). It is difficult to say whether this is a case of the $E(spl)^{-/-}$ "over-rescue" because the $T^{BC+3UTR}$ transgene triggers macrochaete loss also in the wild-type background ($E(spl)^{+/+}$) (see

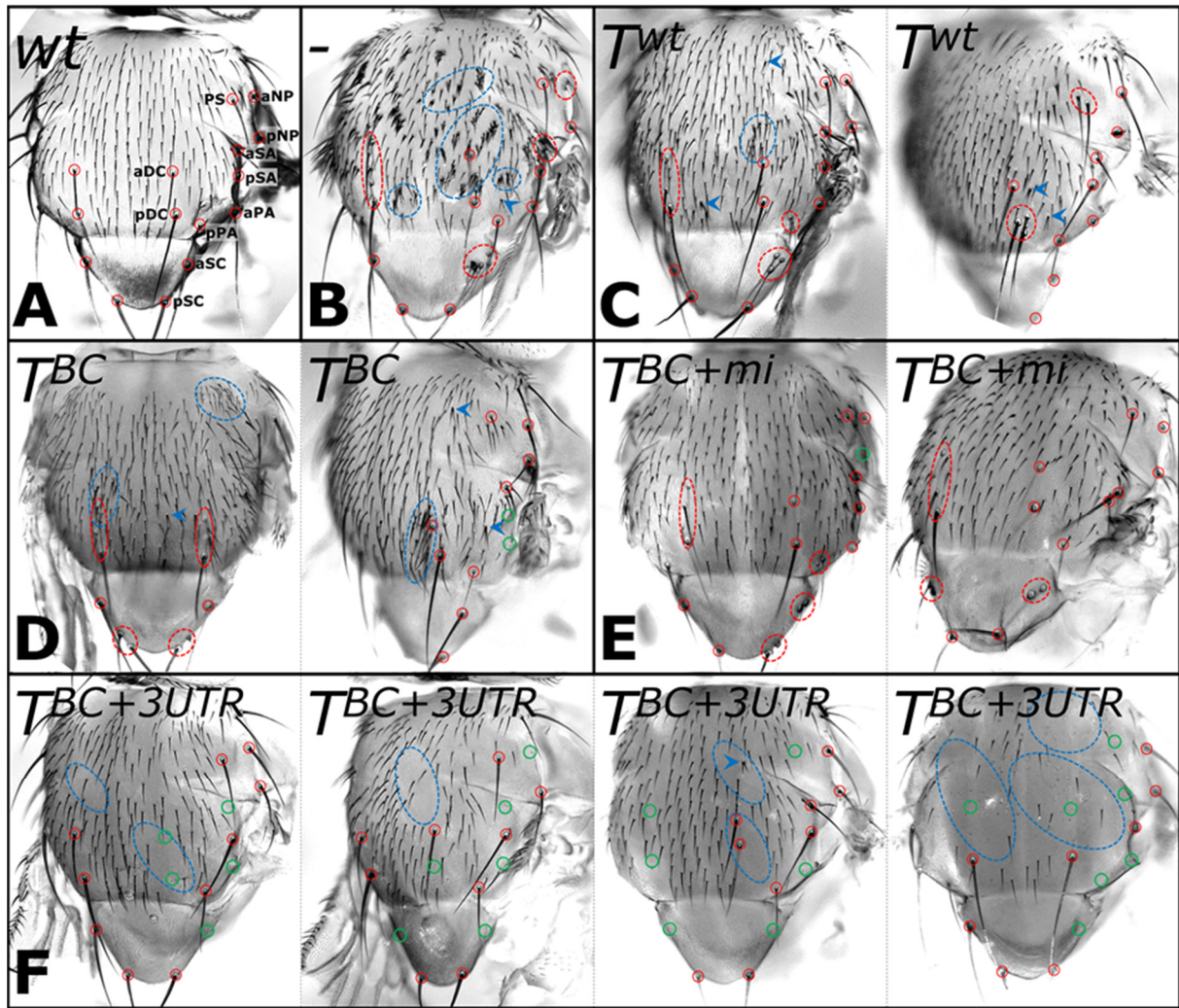


Figure 20 and 21).

Figure 28 Effects of *E(spl)m7-E(spl)m8* genomic transgenes on patterning of thorax bristles in mutant clones of *E(spl)*. (A) Wild-type thorax. (B) Thorax of a non-transgenic ("") fly with induced *E(spl)* null clones. (C-F) Examples of thoraces obtained from animals bearing indicated GFP-less genomic transgene with induced *E(spl)* null clones. Single macrochaetes in their correct positions are marked with red circles, lack thereof is marked with green circle; red dotted circles and blue dotted circles indicate areas of the *E(spl)* null clones presumed by a presence of supernumerary macrochaetes and microchaetes, respectively (in B-E), or lack thereof (in F); blue arrowheads indicate duplicated microchaetes' shafts within single socket.

2.5. Discussion

Complex living systems employ diverse mechanisms to achieve their phenotypic stability in the face of environmental and genetic perturbations (Hartman *et al.* 2001). Gene redundancy, a result of gene duplication, provides robustness in case of a gene's failure (Kitano 2004). *E(spl)* genes expressing bHLH factors are a result of gene duplication, have overlapping functions and are often co-expressed (Delidakis *et al.* 2014). This multiplication of a gene must be aided by parallel control mechanisms in order to attain proper dosage of its

expression and its function. One such strategy is a diversification of the spatio-temporal expression of the individual duplicates. An example of this can be seen in different CNS expression patterns of *E(spl)m7* and *E(spl)m8* reported in this manuscript. However, *E(spl)* genes are often co-expressed what may reflect sharing the same enhancers (as is the case for *E(spl)m7* and *E(spl)m8* in the wing disk) or dependence of their expression on the same transcription factors (e.g., Notch). Thus, their expression must be further buffered to achieve precise levels of their activity. Because they are potent and crucial regulators of development, the seven bHLH genes of the *E(spl)* complex must be provided with efficient, adjustable and multi-tiered gene control. I have examined two modes of gene control: negative transcriptional auto-regulation and 3' UTR-mediated translational repression. I focused on the pair of neighboring *E(spl)* genes, *E(spl)m7* and *E(spl)m8*, and on the consequences of perturbing their regulation on animal viability and on patterning external sensory organs of the adult fly. The results presented in this manuscript show that both forms of the addressed regulation cooperate to sculpt the wild-type expression of these two genes.

The elevated expression from *e7* and *e8*-containing reporter and genomic GFP transgenes in null *E(spl)* clones reveals that the activity of these enhancers is moderated by the *E(spl)* bHLH factors expressed in the PNCs of the wing disk. Therefore, bHLH *E(spl)* genes are capable of self-regulating their expression levels via negative transcriptional feedback, at least in case of *E(spl)m7* and *E(spl)m8*. B- and C-type motifs were proposed to mediate this regulation due to their high affinity toward bHLH *E(spl)* factors demonstrated *in vitro* (Tietze *et al.* 1992; Oellers *et al.* 1994; Ohsako *et al.* 1994; Van Doren *et al.* 1994; Jennings *et al.* 1999). Mutating these sites in the *e7* and *e8*, however, did not (in case of *e7*) or only partially (in case of *e8*) compromised their transcriptional repression *in vivo*. Moreover, both mutated enhancers still exhibited higher activity in cells lacking the *E(spl)* locus, proving that repression of these enhancers is conducted through other sites or in addition to *E_B*- and *E_C*-boxes. These could be represented by the N-type bHLH repressor binding sites and proneural activators-bound *E_A*-boxes. The latter possibility stems from the ability of the bHLH *E(spl)* repressors to interact (with their N-terminal and Orange domains) with (the transactivation domains of) the proneural proteins, and in this way be recruited to the DNA at sites bound by proneural factors - i.e., *E_A*-boxes (Giagtzoglou *et al.* 2003, 2005; Zarifi *et al.* 2012). The N-boxes show lower specificity for the bHLH *E(spl)* factors *in vitro* but are numerous present in both enhancers and most of them exhibit high degree of sequence conservation between *Drosophila* species. Out of the four B- and C-type E-boxes analyzed in this study only one was visibly functional and had moderate sequence conservation, while the most conserved B1-box seemed not to mediate any repression. Thus, the situation *in vitro* can be different from that *in vivo* and the N boxes might as well be bound by the bHLH repressors with strong affinity *in vivo*. The *in vitro* results were obtained with *in vitro* transcribed and translated proteins. *E(spl)* factors *in vivo* might undergo modification/s which can change their binding affinity.

Alternatively, the bHLH *E(spl)* factors might be effectively recruited to all those sites (i.e., *E_B*/*E_C*-, N- and *E_A*-boxes) but with non-linear relationship between the number of sites

occupied and the degree of gene repression, such that mutation (inactivation) of a subset of these sites could be compensated by stronger affinity and/or activity of the bHLH factors at the non-mutated sites. A similar relation, although concerning post-transcriptional regulation, seems to be exhibited by microRNA binding sites within the 3' UTR of *E(spl)m7*. This 3' UTR contains binding site for *bantam*, however, null *bantam* clones do not display derepression unless two other potent microRNA-binding sites (K- and Brd-boxes) are mutated (**Figure 15 A-B**). This indicates that the inactivity of *bantam*-binding box (in absence of *bantam* expression) can be compensated through other (functional) microRNA-binding sites. A good quality ChIP-seq data for E(spl) proteins would be useful to distinguish between these and other possibilities, however, up until now (due to the lack of a good anti-E(spl) antibody) such data is not available.

Analysis of the *E(spl)m7* and *E(spl)m8* 3' UTRs revealed that these genes are also under a strong translational repression. Mutating the putative microRNA binding sites compromised this regulation causing derepression of the mRFP sensors controlled by the mutant 3' UTRs (**Figure 13**). While the 3' UTR of *E(spl)m7* with all six putative microRNA seed boxes mutated had a cumulative effect on the mRFP derepression, it retained some of the capacity to be controlled by *dcr1*-derived microRNAs via other, unidentified binding sites. On the contrary, mutation of the two K-boxes in the 3' UTR of *E(spl)m8* had strong, synergistic impact on derepression and showed no effect when were individually mutated (**Figure 13**). The latter 3' UTR seems also not to be controlled by *dcr1*-derived microRNAs (**Figure 14**), suggesting that the two K-boxes must be targeted by some non-*dcr1*-dependent microRNAs or other repressive mechanism. Dcr1 is a canonical enzyme required for biogenesis of most metazoan microRNAs and small interfering RNAs (siRNAs) and alternative ways of generating these RNAs are thought to be unusual (Bartel 2018). However, examples of Dcr1-independent microRNAs (vertebrate mir-451) and siRNAs are known (Herrera-Carrillo and Berkhout 2017). Of note is a peculiarity of the *E(spl)m8* 3' UTR's K-boxes: (1) the fact that the mutation of one can fully complement mutation of another and (2) their striking proximity; they are separated by a (unlike them, nonconserved) sequence of only 8 nucleotides, making it unlikely to be targeted concurrently by two microRNAs. Therefore, it seems that both of these boxes are equally strong mediators of repression, and the activity of one can be fully compensated by the other. Of note is also the remarkable difference between derepression resulting from mutating all the putative microRNA seed boxes in *E(spl)m7* and *E(spl)m8* 3' UTRs and derepression brought about by replacing them with SV40 and ADH 3' UTRs (**Figure 13**). It is unlikely that this difference reflects residual microRNA control of the mutated 3' UTRs; the 3' UTR of *E(spl)m7* bearing mutations in all identified microRNA seed boxes only weakly relieves repression in *dcr1*-mutant clones, implying that these derepressed levels of mRFP are lower than the mRFP levels observed with SV40 and ADH 3' UTR sensors (**Figure 14 E, Figure 13 E, compare to F and L**). This suggests that *E(spl)m7* and *E(spl)m8* 3' UTRs comprise other (than microRNA-specific) regulatory elements suppressing mRNA stability. These may be sequences forming secondary structures (Rabani *et al.* 2008; Geisberg *et al.* 2014), binding sites of RNA-binding proteins (Olivas and Parker 2000; Shalgi *et al.* 2005; Duttagupta *et al.* 2005; Hogan *et al.* 2008; Hasan

et al. 2014) and AU-rich elements (AREs) mediating mRNA decay (Cairrao *et al.* 2009; Helfer *et al.* 2012). I identified putative AREs in the *E(spl)m8* 3' UTR, however their deletion did not significantly affect repression mediated by this 3' UTR (**Figure 13 J**).

Interestingly, Notch, in addition to acting as a direct transcriptional activator of *E(spl)m7* and *E(spl)m8*, alleviates translational repression mediated by the 3' UTRs of these two genes (**Figure 16 A, B**). This regulation seems not to be mediated by bHLH *E(spl)* factors (*E(spl)m7* overexpression does not reproduce Notch activity on these 3' UTRs, **Figure 16 C-D**) and may be transduced further by other systemic signaling pathways to produce non-autonomous derepression of these 3' UTRs (**Figure 16 A, B**). Two studies identified Notch as a direct transcriptional activator of *bantam* microRNA in the fly brain (Wu *et al.* 2017) and an indirect repressor of the same microRNA in the wing margin of the wing disk (Becam *et al.* 2011). Therefore, Notch can also regulate translational repressive mechanisms in a context-dependent manner.

I show that *E(spl)m7* and *E(spl)m8* are expressed in PNCs of the larval wing disk and this expression depends on the *e8* enhancer. These are not the only *E(spl)* genes expressed in the PNCs. *E(spl)mδ* and *E(spl)mγ* were also reported to be expressed in a subset of PNCs (Singson *et al.* 1994; Bailey and Posakony 1995; de Celis *et al.* 1996; Nellesen *et al.* 1999; Cooper *et al.* 2000). These and/or other bHLH *E(spl)* factors seem to limit the abundance of *E(spl)m7* and *E(spl)m8* on one hand, and cooperate with them in order to restrict cells adopting the SOP fate on the other. Depletion of all *E(spl)* factors from PNCs triggers upregulation of transgenic *GFP-E(spl)m7* within PNCs (**Figure 27**). However, halving the *E(spl)* gene dose reduces the activity of transgene-derived *E(spl)m7* and *E(spl)m8* proteins evidenced by their lower effect on the bristle loss (**Figure 23**). These results indicate complexity of interactions between *E(spl)* proteins. One important feature of the bHLH *E(spl)* has to be considered: they act (as repressors) as homodimers and heterodimers with other bHLH *E(spl)* proteins. Different heterodimers could have distinct context-dependent properties, e.g., binding specificity, potency of inducing repression. However, *E(spl)m7* and *E(spl)m8* can also act on their own. Even the wild-type version of the genomic transgene (*T^{wt}*) is capable of (partially) rescuing supernumerary macrochaete in *E(spl)* null clones. Moreover, the same transgene in the wild-type background (i.e., as an extra copy of *E(spl)m7* and *E(spl)m8* genes) has virtually no effect on bristle loss, even less so on survival of the animals. This demonstrates plasticity of the *E(spl)* locus.

The wild-type background (two copies of the *E(spl)* locus) provides substantial repression for the *E(spl)m7* and *E(spl)m8* genes, but it is not sufficient, on its own, to maintain proper (low) levels of the proteins expressed by these genes. It is most clearly manifested by the frequent macrochaete loss caused by the genomic transgene in which *E(spl)m7* and *E(spl)m8* genes are devoid of the regulation mediated through their cognate 3' UTRs (i.e., replaced by SV40 and ADH 3' UTRs) (**Figure 21**). Similarly, when transcriptional repression is compromised (in *E(spl)* null clones), modifications of the 3' UTRs result in even greater activity of *E(spl)m7* and *E(spl)m8*. This is best manifested in microchaete patterns. The full rescue of

the supernumerary *E(spl)*^{-/-} microchaete phenotype is observed only with the transgene expressing *E(spl)m7* and *E(spl)m8* at intermediate levels (the transgene with microRNA-boxes and E_B/E_C-boxes mutated, *T^{BC+mi}*, **Figure 28 E**). Lower expression levels of these genes (from transgenes bearing wild-type 3' UTRs) do not rescue, whereas, higher expression levels (from a transgene bearing SV40 and ADH 3' UTRs) lead to a loss of microchaete (**Figure 28**). The lack of rescue by the transgenes containing wild-type 3' UTRs indicates that the compromised transcriptional repression is not sufficient to yield levels of *E(spl)m7* and *E(spl)m8* needed for normal patterning of microchaetes when translational repression is unaffected. The loss of microchaetes caused by the transgene containing SV40 and ADH 3' UTRs in *E(spl)* null clones and the fact that this same transgene does not affect patterning of microchaete in the wild-type background (*E(spl)*^{+/+}) reveal that transcriptional repression can efficiently buffer against deficient post-transcriptional control. In sum, these and other results presented in this manuscript demonstrate that transcriptional and post-transcriptional repression are prominent regulatory modes utilized by *E(spl)* genes. Their profound impact on the *E(spl)* gene expression suggests that these mechanisms are important constituents of the plastic and robust trait of the multiple redundant components of the *E(spl)* locus.

3. Chapter 2. The role of insulators in transgene transvection in *Drosophila*

3.1. 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. I 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. I 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*. I 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 identity of enhancers and promoters in the vicinity of a trans-interacting insulator pair is also important, indicative of complex insulator-enhancer-promoter interactions.

3.2. Introduction

Classical studies of changes in gene expression caused by alterations of genomic order pointed to a close link between the organization of chromatin and the robustness of gene regulation. Parallel observations of nonrandom distribution of chromosomes in interphase nuclei through microscopy seemed to confirm this link (Cremer and Cremer 2010), which was further only reinforced by methods such as Hi-C (Lieberman-Aiden *et al.* 2009; Hou *et al.* 2012; Sexton *et al.* 2012; Feng *et al.* 2014), uncovering a high degree of accuracy with which genome assembles. Highly conserved associations within single chromosomes led to distinction of higher-order chromatin modules, so called 'topologically associating domains' (TADs) (Dixon *et al.* 2012; Sexton *et al.* 2012), within which associations are favored. It has long been demonstrated that physical proximity between distant enhancer and a promoter can bridge their activities into an operative transcriptional unit (Dunaway and Droge 1989; Müller *et al.* 1989; Mahmoudi *et al.* 2002; Vieira *et al.* 2004), and instances of analogous cases *in vivo* have

been reported since then in vast number among all taxa (for example (Ronshaugen and Levine 2004; Spilianakis *et al.* 2005; Lomvardas *et al.* 2006)).

Structuring chromatin fibers is complex and involves an array of multilayered organization. The lowest level of chromatin structure consists a chromosome comprising one molecule of DNA. The functional unity of a diploid genome relies not only on intrachromosomal, but also on interhomologous and interheterologous interactions between chromosomes. Homologous pairing is one of the key aspects of nuclear organization and an intrinsic property of all sexually reproducing Eukaryotes; it is a general attribute of meiosis where it has a well-established role in homologue segregation (McKee 2004). However, homologous pairing in somatic interphase nuclei occurs to different extents among species (McKee 2004; Apte and Meller 2012). In most studied cases, including plant species (Bollmann *et al.* 1991), fungi (Aramayo and Metzenberg 1996) and mammals (Apte and Meller 2012), nonmeiotic chromosome pairing is observed only at sporadic loci, or at centromeres as is the case of diploid budding yeast (Apte and Meller 2012). In Dipteran insects, however, homologous pairing seems to be a general feature of nuclear organization (Stevens 1908; Metz 1916; McKee 2004). In *Drosophila*, homologous chromosomes are intimately associated from end to end in virtually all the interphase nuclei of the organism (McKee 2004). The polytene chromosomes of the salivary glands are an extreme example of such an association where up to thousands of chromatids are joined together with close synapsis along their entire lengths (Duncan 2002).

3.2.1. Homologous pairing in mammals

The ubiquitous presence of homologous pairing in interphase nuclei of *Drosophila* is in striking contrast to that observed in mammals, where such a general somatic pairing has not been observed. Homologous chromosomes have been shown to exhibit a low frequency of joint damage after laser-UV-microirradiation and caffeine treatment in Chinese hamster lung cells (Cremer *et al.* 1982) and to rearrange infrequently (via homologous exchange) after irradiation (Boei *et al.* 2006). Moreover, numerous accounts of homologous pairing have been associated with human cancer cells (Haaf and Schmid 1989; Lewis *et al.* 1993; Brown *et al.* 1994; Williams *et al.* 1995; Atkin and Jackson 1996; Zhang *et al.* 1997; Koeman *et al.* 2008) and diseases (LaSalle and Lalande 1996; Stout *et al.* 1999; Thatcher *et al.* 2005). The human chromosomes are segregated in distinct domains of the nucleus (Cremer and Cremer 2010) and seem to rarely come into contact; or at least, the inter-homologue distances are observed to be larger than inter-heterologue distances (Heride *et al.* 2010). It is thought that maintaining a spatial separation between homologous chromosomes in human cells is an important condition for genome stability and proper gene expression (Heride *et al.* 2010). Most somatic pairing events occurring in normal human cells are attributed to regions involved in establishment or maintenance of monoallelic expression (Hu *et al.* 1997; Duvillie *et al.* 1998). Mutations that affect the competence for pairing of either allele were shown to have an impact on expression of the second allele (Hu *et al.* 1997; Duvillie *et al.* 1998; Tsai *et*

al. 2003; Landers *et al.* 2005). In the most prominent example, at the onset of X-chromosome inactivation, the interaction between two X chromosomes is needed to break symmetry and destine one homolog for silencing (Anguera *et al.* 2006). The well-studied homologous associations in mammals are transient and tightly regulated (Williams *et al.* 2010). In the latter case, the two X's pair at the *X inactivation center* (*Xic*) locus (Sun *et al.* 2010); *Xic* pair parts after 30 minutes from its formation when the X chromosomes assume different fates (Xu *et al.* 2007); its initiation requires active transcription at the locus (Xu *et al.* 2007), binding of CCCTC-binding factor (CTCF) (Xu *et al.* 2007) and the pluripotency factor, Oct4 (Donohoe *et al.* 2009). Examination of this and other cases of interchromosomal associations in mammals led to observation that these interactions happen preferentially between transcriptionally active genes (Lieberman-Aiden *et al.* 2009; Yaffe and Tanay 2011). This correlation was further examined and the model was proposed whereby homologous pairing is facilitated by (in addition to other requirements) large-scale active features comprising the *trans*-interacting regions. In this view, specific homologous alleles happen to come into proximity within specialized transcription factories where they have an opportunity for transcriptional regulation in *trans* (Krueger *et al.* 2012). These observations indicate that homologous pairing in somatic mammalian cells, although biologically significant, is infrequent, transient and localized.

3.2.2. Homologous pairing and transvection in *Drosophila melanogaster*

In contrast, the omnipresent homologous pairing in *Drosophila* seems to have a minor biological importance (Wu 1993). A homologue asynapsis caused by gross chromosomal rearrangements, including deletions, generally are viable, and haploid patches of tissue survive to adulthood (Santamaria 1983; Wu 1993). There are just a few observations where homologous pairing is important for wild-type levels of gene expression. It has been demonstrated for *Ultrabitorax* (*Ubx*) gene product (Goldsborough and Kornberg 1996) and for *spineless* (*ss*), where in the latter case, homologous pairing is proposed to coordinate the expression state between stochastically expressed alleles of *ss* (Johnston *et al.* 2014). Despite the apparent rarity of pairing-dependent need for gene regulation, a wealth of pairing-sensitive intragenic complementations between mutant alleles are documented at numerous loci (for example: (Lewis 1954; Gelbart 1982; Geyer *et al.* 1990; Leiserson *et al.* 1994), reviewed in (Duncan 2002)). These intragenic complementations are a result of a case of gene regulation termed *transvection* (Lewis 1954), whereby an enhancer's activity is deposited on a promoter located in the second allele of the same locus. Transvection describes all phenomena of gene regulation underlying somatic homologue pairing, including gene activation and silencing (Sigrist and Pirrotta 1997; Chen *et al.* 2002; Mellert and Truman 2012). However, transvection effects, often exhibit a variegated pattern of gene expression, suggesting stochastic interactions between the homologs (Bateman *et al.* 2012a; Mellert and Truman 2012). Importantly, the observed variegation is not an effect of increased distance between enhancer and promoter apposed in *trans*, as distancing enhancer-promoter pair in *cis* reduces and delays transcription but does not lead to variegation (Bateman *et al.* 2012a).

This could mean that homologous pairing in *Drosophila* is particularly dynamic. Indeed, precise measurements, assaying activity of malic enzyme (MEN), , revealed that *trans*-regulation in the *men* locus is more susceptible to environmental changes and genetic background, compared to more robust *cis*-regulation of *men* (Bing *et al.* 2014). Moreover, levels of transvection are repeatedly reported to vary between cell types (Fedorova *et al.* 2008; Mellert and Truman 2012), loci (Chen *et al.* 2002; Fedorova *et al.* 2008) and fly lines (Fedorova *et al.* 2008). In one study *trans*-activated transcription was measured to be, depending on the locus, 1,7 to 2,2% of the total levels of *cis*-activated transcript (Bateman *et al.* 2012a).

The current thinking about homologous pairing in *Drosophila* proposes a central structural role in organizing the genome. However, its role in gene regulation remains mysterious. Dynamic behavior and modest influence on gene function implicate transvection in augmenting phenotypic plasticity (Bing *et al.* 2014). It is postulated to affect buffering the copy number variation (CNV) (Zhou *et al.* 2011), early dosage compensation (Lavery *et al.* 2011), position–effect variegation (PEV) (Dernburg *et al.* 1996), but its role may be subsidiary or compensatory to more direct regulatory mechanisms of gene dosage (Apte and Meller 2012).

Maybe the most compelling adaptive advantage of extensive homologue pairing in *Drosophila* is borne out from studies on homologous recombinational (HR) repair pathway (Gloor 2002; Wyman *et al.* 2004). It was shown that the homologous chromosome is frequently used in gene conversion (GC), a mode of HR, for repairing the double-strand breaks (DBSs), like those generated by the *I-SceI* endonuclease (Rong and Golic 2003). This observation suggested that allelic recombination is a prominent repair pathway in *Drosophila* and can effectively compete with intrachromosomal recombination (Rong and Golic 2003). In contrast, the homologous chromosome is rarely used for HR repair in mammalian cells where nonhomologous end joining (NHEJ) repair outcompetes allelic recombination (Stark and Jasin 2003).

It is curious why homologous pairing is generally prohibited in mammals and yet so robust in *Drosophila*. Conceivably, *Drosophila* and mammals developed fundamentally different strategies to organize their genomes and gene expression. The most obvious manifestation of this difference is careful allocation of chromosomes in discrete territories of mammalian nuclei, compared to overall insensitivity to mutual cross-regulation by closely associated homologs in *Drosophila*. It may imply a shift in the mechanisms controlling gene dosage linked to a one of many alternative ways in which genome can be organized.

3.2.3. Determinants and dynamics of homologue pairing

The pairing kinetics of FISH-labeled loci in the blastoderm (pre nc14) embryos suggests that the capacity of homologous chromosomes to pair is proportional with the duration of the cell cycle. Following fertilization, *Drosophila* embryo progresses through 13 synchronous mitotic cycles to create a syncytial blastoderm with thousands of nuclei arrayed below the

surface of the outer membrane (Foe and Odell 1993). Initially, mitotic divisions are a few minutes long and then slow down during cycles 11–13 before finally pausing for at least 60 minutes during interphase 14, at which time the syncitial blastoderm cellularizes (Foe and Odell 1993). Accordingly, pairing between homologs is not observed in cell cycle 12, that takes 13 minutes to complete; the first paired loci emerge during the 25 minute-long cycle 13, and the high levels of pairing are apparent during cycle 14, which can last up to 175 minutes (Fung *et al.* 1998; Gilbert 2000). This correlation led to presumption that the rise of pairing in cycle 14 is due to the extended time of this cycle, and not due to a newly acquired competence for pairing during this cycle (Dernburg *et al.* 1996; Fung *et al.* 1998; Gemkow *et al.* 1998).

Pairing interactions begin to form during the nuclear cycle 13 (nc13) of *Drosophila* embryogenesis which foreshadows vast changes in embryonic development, cellularization and the maternal-to-zygotic transition (MZT) (Anderson and Lengyel 1979; Edgar and Schubiger 1986; Gilbert 2000). At the beginning of each mitotic cycle prior to nc14, unpaired homologs adopt the extended and highly nonrandom 'Rabl' conformation that is polarized along the nucleus with centromeres at one nuclear pole and telomeres at the other (Foe and Alberts 1985; Hiraoka *et al.* 1990). Each of the chromosomes at this stage occupies a distinct, largely nonoverlapping territory in the nucleus (Marshall *et al.* 1996). The fluorescent *in situ* hybridization (FISH) single-particle tracking of chromatin motion revealed that the displacement of chromatin results from diffusion, rather than from active motility (Marshall *et al.* 1997; Vazquez *et al.* 2001). This movement is confined within a limited subregion for a given locus (Marshall *et al.* 1997; Vazquez *et al.* 2001), and its dynamics is independent of the size of a chromosome (Gemkow *et al.* 1998). Colocalization of twin FISH foci linearly situated on the homologous chromosomes does not progress sequentially, implying that the stable homologous association along the length of chromosomes are formed by multipoint recognition (Fung *et al.* 1998). These observations lead to a model whereby the pairing of homologous chromosomes initiates at the start of each interphase (after nc13) and propagates by random diffusional contacts between multiple independent sites of spatially constrained homologs.

Observable effects of pairing-sensitive gene regulation are often taken as an indication of paired state between homologs. One such assay relies on the ability for complementation between two alleles of a given gene, one lacking an enhancer and another lacking a promoter. Many such phenotypic complementations are observed when the structure of the chromosomes on which these alleles reside is largely preserved (Lewis 1954; Duncan 2002). However, when the same mutant alleles are apposed on two structurally different homologs, e.g. one intact and one containing rearrangements, transvection can be suppressed (Lewis 1954). Elimination of Transvection in this case echoes the disruption of pairing, which is brought about by structural asymmetry of homologs (Hiraoka *et al.* 1993; Gemkow *et al.* 1998). The level of this inhibition usually reflects the size and the type of chromosomal rearrangements; for example, a large pericentric inversion (containing the centromere) typically inhibits complementation of an allele which is proximal to the breakpoint, while a

smaller paracentric inversion (within the chromosome's arm), even if it encompasses the allele, may still support transvection, albeit with a smaller degree (Lewis 1954; Gelbart 1982; Leiserson *et al.* 1994; Golic and Golic 1996; Bing *et al.* 2014). Nevertheless, pairing interactions between the alleles contained in these rearrangement heterozygotes can be restored in mutant backgrounds which slow the rate of cell division, indicating that *Drosophila* chromosomes retain the ability for homologous pairing even if they are faced with structural rearrangements (Golic and Golic 1996). It can explain, at least in part, the cell specificity of transvection effects and why the pairing between certain alleles is relatively sensitive to many rearrangements, while in case of the others is not; for example, structural heterozygosity easily disrupts transvection of *eyes absent* (*eya*), but not of the *brown* (*bw*) alleles, although both of these genes function in the same tissue, in the eye; however, *eya* functions in cells just anterior to the morphogenetic furrow where the cells are either actively dividing or have recently divided, while *brown* is expressed in cells that completed their final mitotic division several days earlier (Golic and Golic 1996). Thus, it seems that the association of homologous sequences is only delayed by a structural heterozygosity and it can be prevented when given insufficient time to occur in the cells with relatively fast rate of mitotic divisions (Golic and Golic 1996).

These observations suggest uniform kinetics of diffused chromatin and a common mechanism of homology search for all chromosomes. Such mechanism of sensing and pairing was postulated to rely on sequence-directed protein-independent assembly of similar dsDNA molecules (Inoue *et al.* 2007; Danilowicz *et al.* 2009) and on the DNA sequence-based selective association between nucleosomes with identical DNA sequence (Nishikawa and Ohyama 2013), both principles substantiated experimentally *in vitro*. In this simplistic view, homologous pairing is deterministic and requires just enough time to take place. Thus, the onset of pairing is derived from the correlation of the distance between the homologs and the time for chromatin to self-assemble given by the length of a cell cycle. The regulation of homologous pairing would therefore involve mechanisms that control anchoring chromosomes to the attachment sites of nuclear skeleton and constrain chromatin motion by affecting its diffusibility and propensity for self-recognition.

However, local homology is not the sole determinant for chromosome pairing. A study on pairing kinetics between structurally rearranged chromosomes showed that the same FISH-labeled homologous loci converge with lower frequency in homozygotes for rearranged homologs than in homozygotes for wild-type chromosomes (Hiraoka *et al.* 1993). Therefore, it seems that chromosomal position and surroundings of a given locus can determine kinetics of homology search for this locus.

This locus-sensitive competence for pairing raises the possibility that pairing between different loci is regulated by distinct mechanisms, in addition to global determinants of pairing. Particular loci were accordingly found to colocalize with specific features of nuclear structure. 7 out of 42 FISH probes targeted to chromosome 2L were consistently coinciding with nuclear lamina in early embryos (Marshall *et al.* 1996). Similar observation revealed

preferential association of the X-linked heterochromatic 359 base pair (bp) repeat sequence to nuclear scaffold attachment sites (Kas and Laemmli 1992; Marshall *et al.* 1997). The existence of specific interactions of discrete loci situated along the chromatin fiber with nuclear envelope and nuclear matrix suggests that these interactions are effectively impacting the range and the rate of chromosome motility (Vazquez *et al.* 2001). Moreover, the state of chromatin also has a consequence on the variable kinetics of pairing among different loci across homologs. Regions in or near centric heterochromatin pair significantly less or pair more slowly than euchromatic regions (Williams *et al.* 2007), implying possible mechanistic differences between heterochromatic and euchromatic pairing (Joyce *et al.* 2012). Heterochromatin is thought to be less prone to homologue pairing, possibly due to its repeated nature and propensity for intramolecular associations (Williams *et al.* 2007).

The correlation of the onset of pairing with the initiation of zygotic transcription raises the possibility, on one hand, that pairing is induced by some zygotic gene product/s (Hiraoka *et al.* 1993), and on the other, that the nucleation of pairing may be coupled with transcriptional state of a given locus, e.g. whereby transcriptionally competent 'apposed' alleles would colocalize or compete for the same transcriptional factories (Cook 1997). However, no particular zygotic gene product was found essential for pairing to initiate in the cell cycles 13 and 14 of embryogenesis, implying that the early inclination of homologs to pair may be a part of the developmental program encoded by the maternal genome (Bateman and Wu 2008). Dependence of homologous pairing on transcriptional state of paired loci also seems unlikely. For instance, the pairing frequencies between FISH-tracked alleles of the Bithorax-Complex (BX-C) are comparable in all cells of the embryo, independently of the BX-C transcriptional activity (Gemkow *et al.* 1998).

There are significant differences between levels of pairing depending on locus, cell type, time of development and stage of the cell cycle (Csink and Henikoff 1998; Fung *et al.* 1998; Gemkow *et al.* 1998; Sass and Henikoff 1999; Vazquez *et al.* 2002; Ronshaugen and Levine 2004; Fritsch *et al.* 2006). For example, a great majority of cells during the first 20 minutes of the 14th embryonic mitotic cycle shows complete pairing at the repetitive histone locus which is not interrupted until anaphase (Hiraoka *et al.* 1993; Fung *et al.* 1998). In contrast, the frequency of homologous pairing in the bulk of hypotonically treated nuclei from squashed larval central nervous system (CNS) is partially disrupted during the course of S-phase and decreases from 85% to 75%, or 75% to 60%, depending on the locus (Csink and Henikoff 1998). Because of the several observations that entry into S-phase and G2/M transition may affect pairing, a link was suggested between cell cycle and homologous pairing (Csink and Henikoff 1996; Fung *et al.* 1998; Williams *et al.* 2007).

Further implications come from the generally observed stochasticity of pairing events. In the case of the BX-C locus, the homologous alleles coincide in 60-70% of the postmitotic cells of the embryo and are never observed to associate with a higher frequency than 70% (Gemkow *et al.* 1998). Other studies also report that pairing frequencies rarely attain 100%, and a characteristic frequency of pairing is attributed to corresponding tissue, time of

development, stage of the cell cycle and locus (Dernburg *et al.* 1996; Fung *et al.* 1998; Gemkow *et al.* 1998). These variations in pairing between cells but with reproducible frequencies among the type of tissue suggested that association of the homologs stays in equilibrium with a dissociated state. It was speculated that unpaired states represent local 'breathing' of homologous chromosomes and reflect the need of creating regions with sufficient number of open regions without homologous interactions to allow *cis* and nonhomologous *trans*-looping, and movements required during transcriptional processes (Gemkow *et al.* 1998). However, a kinetic model of homologous pairing implies that the dissociation of paired homologous regions of chromatin is energetically disfavored and therefore induction of the unpaired state has to be actively promoted (Fung *et al.* 1998).

3.2.4. Regulators of homologous pairing

Two independent studies used *Drosophila* cultured cells, RNAi knockdown and high-throughput FISH technology (Hi-FISH), to identify potential regulators of homologous pairing (Bateman *et al.* 2012b; Joyce *et al.* 2012). These screens revealed in total 44 candidate 'pairing' genes (whose inhibition decreased the frequency of single-signal nuclei) and 65 'unpairing' genes (whose inhibition increased the frequency of single-signal nuclei), with significant overlap between hits in both screens (Bateman *et al.* 2012b; Joyce *et al.* 2012). In total, 70% of the 'pairing' genes have known roles in cell division, while the function of the half of the 'unpairing' genes is linked to the cell cycle progression, with another third of the 'unpairing' genes associated with transcription or transcription-related processes (Joyce *et al.* 2012). Of significant meaning is a large number of the 'unpairing' genes whose activity is during interphase (while the most 'pairing' factors are involved in the mitotic division), thereby suggesting that the paired state of homologs is actively regulated in the interphase nuclei. Moreover, a large number of inhibitors of pairing is involved in cell cycle regulation, indicating that the disruption of pairing is functionally coupled to the progression of the cell cycle. This in turn may suggest that the discrete stages of cell cycle are sensitive to genetic interactions that rely on pairing.

In another study, overexpression of a component of condensin II complex, Cap-H2, was shown to induce disassembly of salivary gland polytene chromosomes and inhibit transvection effects in larval wing tissue, and reciprocally, the loss of Cap-H2 stimulated transvection (Hartl *et al.* 2008). This outcome was particularly intriguing as Condensin II is proposed to mediate chromosome condensation which results in assembling metaphase chromosomal structure (Hirano 2006), suggesting that chromosome compaction also takes place in interphase nuclei and antagonizes homologue alignment (Hartl *et al.* 2008; Smith *et al.* 2013). In line with this finding, the RNAi knockdown of Cap-H2, or any of the several components of the Condensin II complex recovered in the Hi-FISH screen, enhanced homologous pairing (Joyce *et al.* 2012). In addition, one third of the 'pairing' genes were found to act in a Cap-H2 dependent manner, as their depletion combined with the loss of 'unpairing' Cap-H2 activity could restore pairing to wild-type levels (Joyce *et al.* 2012). Taken together, these findings imply the existence of a

regulatory network of genes that control somatic pairing along the duration of the cell cycle, and that some of these genes may facilitate pairing indirectly by inhibition of condensin II activity. Moreover, what appears here, is the emergence of a close relationship between cell cycle and condensed state of chromosomes in regulating homologous pairing. Finally, although a link between regulators of cell cycle and homologous pairing can account for changes in pairing levels observed throughout the cycle, the striking result of these findings is the number of more than 100 genes involved directly or indirectly in regulation of pairing. This includes components of the Anaphase-Promoting Complex (APC), protein turnover machinery, chromatin proteins among others, and suggests a still more complex layer of pairing regulation.

The relationship between homologous pairing and transvection is not clear. Several chromatin proteins were identified to regulate transvection, including Zeste, Polycomb proteins and Suppressor of hairy wing [Su(Hw)] (Jack and Judd 1979; Babu and Bhat 1980; Pirrotta *et al.* 1985); however, none of those were recovered in the Hi-FISH screens for regulators of homologous pairing (Bateman *et al.* 2012b; Joyce *et al.* 2012), suggesting that homologous pairing and transvection are mechanistically unrelated.

3.2.5. Insulators

Insulators were initially discovered as boundaries of loop domains because of their ability to establish inter-insulators' regions of independent regulatory activities (Grosveld *et al.* 1987; Kellum and Schedl 1991). Subsequently, two seemingly opposing functions were ascribed to insulators: they blocked or promoted enhancer-promoter interactions depending on the relative position, orientation and arrangement of these elements (Bell *et al.* 1999; Cai and Shen 2001; Muravyova *et al.* 2001; Melnikova *et al.* 2004; Gruzdeva *et al.* 2005; Comet *et al.* 2006; Kyrchanova *et al.* 2008b; Ali *et al.* 2016; Chetverina *et al.* 2017). These activities of insulators are currently reconciled in a model which postulates that orientation-dependent pairing between compatible insulators restricts interactions between regulatory elements located on the opposite sides of the paired insulator conformation, while promotes the interactions which are found on the same side to the paired insulators (Kyrchanova *et al.* 2008b; Fujioka *et al.* 2016; Chetverina *et al.* 2017). Thus, this model implies autonomy of the interaction between insulators, and that this interaction determines transcriptional activities indirectly by increasing or decreasing spatial proximity between competent enhancers and promoters. A requirement for mediating transcriptional activation would be a close association of insulators with enhancers and promoters. Indeed, *Drosophila* insulator proteins are enriched at or near active enhancers (Cubeñas-Potts *et al.* 2016) and active promoters (Bartkuhn *et al.* 2009; Bushey *et al.* 2009; Negre *et al.* 2010), and examples of this configuration have been demonstrated to mediate transcriptional activation in *Drosophila* and mammals (Sigrist and Pirrotta 1997; Muller *et al.* 1999; Kravchenko *et al.* 2005; Gruzdeva *et al.* 2005; Vazquez *et al.* 2006; Xu *et al.* 2011; Guo *et al.* 2012, 2015). There is growing evidence, however, that the specific interactions between proteins bound to insulators with those

bound to promoters and/or enhancers constitute an additional layer providing specificity to the enhancer-promoter interactions. Indirect association of insulator proteins (i.e., recruited via protein-protein interaction and not through DNA sequence specific binding) is frequently detected on *Drosophila* active promoters (Liang *et al.* 2014), Su(Hw)-bound *gypsy insulator* (GI) was shown to interact with an enhancer (Kyrchanova *et al.* 2013), and, to physically and genetically interact with promoters (Cai and Levine 1997; Wei and Brennan 2000, 2001; Golovnin *et al.* 2005; Kurshakova *et al.* 2007; Chopra *et al.* 2009; Erokhin *et al.* 2011). More recently it has been shown that the *Drosophila* bithorax complex (BX-C) *Fab-7* and *Fab-8* insulators themselves direct their neighboring enhancers to their appropriate gene promoter (Postika *et al.* 2018). These observations suggest that insulators-bound proteins can interact with enhancer- and promoter-bound factors, and that these interactions confer specificity to the long-distance communication between the regulatory domains and their respective target genes.

Su(Hw) is one of many insulator-bound proteins found in *Drosophila*. Insulator elements bound by Su(Hw) and other proteins have been sporadically correlated with transvection, but their role in this process remains obscure. Whereas some studies propose that insulators are needed for transvection (Lim *et al.* 2018), others propose that they only have an accessory role (Kravchenko *et al.* 2005; Schoborg *et al.* 2013), or that they affect transvection by promoting homologue pairing (Fujioka *et al.* 2016).

3.3. Objectives

Transgenic reporter genes are powerful tools for studying gene regulation. However, a transgene is often susceptible to interactions with surrounding chromatin leading to significant variations in expression patterns and levels, depending on the site of insertion in the genome (Levis *et al.* 1985). In order to factor-out these "position effects" multiple lines of transgenic animals must be studied. Recently, this problem has been tackled by two strategies now widely used in *Drosophila*. One relies on the ability to guide integration of a transgene to a specific location in the genome via the use of the Φ C31-mediated integration system (Thorpe and Smith 1998; Groth *et al.* 2004; Pfeiffer *et al.* 2008; Kvon 2015; Markstein *et al.* 2008). By targeting all reporters to the same "landing site", they are directly comparable and only a single transgenic line per construct needs to be analyzed. The second way to minimize genomic position effects utilizes insulator sequences in transgenesis vectors.

I started using both of these approaches to characterize enhancer modules of two neighboring *Drosophila* genes *E(spl)m7* and *E(spl)m8* (see Chapter 1 of this thesis). To that end I generated a series of reporter constructs flanked by two copies of the insulator sequence from the *gypsy* transposon (*gypsy insulator*, GI) and integrated each construct into the same *attP* locus. When I tested two different reporters in a heterozygous configuration, I noted that they markedly affected each other's expression in *trans*.

Using the series of reporter transgenes, I decided to search for sequence determinants of transvection.

3.4. Results

3.4.1. Enhancer element analysis of *E(spl)m7* and *E(spl)m8* reveals a transvection phenomenon

In a previous chapter of this thesis I characterized two upstream enhancers of *E(spl)m7* and *E(spl)m8* genes, *e7* and *e8*, which drive distinct expression patterns in the CNS and in the wing disk and similar patterns in the eye disk (**Figure 3**). In the context of the genomic fragment encompassing both genes, *e7* and *e8* are shared between promoters of the two genes, *p7* and *p8*, in the WM and AMPs of the wing disk (both genes expressed), but they act exclusively on their downstream gene in the VNC midline (only *m7*) and the neuroblasts (only *m8*).

When I made heterozygous animals containing both transgenes (*e7p7-lacZ* and *e8p8-GFPm8*), I noticed that *e7p7-lacZ* displayed novel expression in the wing margin (characteristic of *e8*), while *e8p8-GFP* was expressed in the AMPs (characteristic of *e7*) (**Figure 29 C**, compare to **Figure 29 A** and **B**). This effect was observed when both transgenes were inserted into the same attP landing site, either *attP40* (chromosome 2) or *attP2* (chromosome 3) (**Figure 29 C** and **D**). No such inter-transgene interaction was observed when one transgene was inserted in *attP40* and the other in *attP2* (**Figure 29 E** and **F**), suggesting that homologue pairing is required for this interaction; in other words, we are observing a transvection phenomenon.

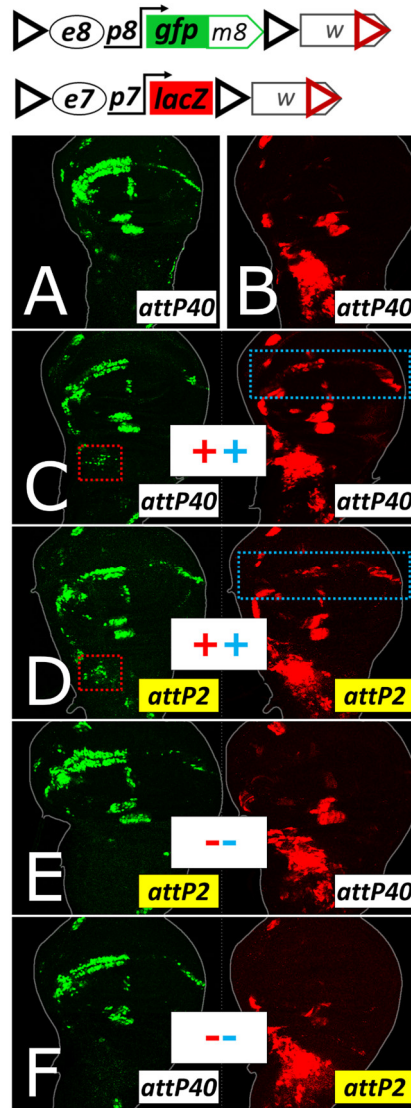


Figure 29 *e7p7* and *e8p8* interact in *trans*. (A and B) *cis*-expression patterns in wing disks isolated from hemizygous animals carrying *e8p8-GFPm8* (A) or *e7p7-lacZ* (B) transgenes in the *attP40* locus. (C-F) *e8p8-GFPm8* and *e7p7-lacZ* crossed in the same animal. When present as heterozygotes in the same locus (C in *attP40*, D in *attP2*) both transgenes expand their pattern to locations dictated by their homologous transgene (marked with dotted rectangles). Occurrence or not of transvection is marked by a + or – symbol, respectively, in the middle of the panel. Red symbols refer to *e7*→*p8* transvection, also marked by red dotted rectangles. Blue symbols refer to *e8*→*p7* transvection, also marked by blue dotted rectangles. No transvection is observed when the *e8p8-GFPm8* and *e7p7-lacZ* transgenes are placed in the same animal, but in non-homologous loci (E and F).

3.4.2. Transvection is mediated by homotypic interactions between GIs or Wari insulators

Having encountered robust bi-directional gene activation in *trans* (*e8p8* activated by *e7* in the AMPs and *e7p7* activated by *e8* at the WM), I decided to dissect the sequences that mediate this phenomenon. Firstly, I asked whether the observed transvection is a general phenomenon requiring only an enhancer paired to a promoter, or whether it depends on

additional transgene sequences. The constructs in **Figure 29** were based on the *pPelican* vector (Barolo *et al.* 2000), which contains a *pUC8* backbone, the mini-*white* gene, two *gypsy* insulators (GI) flanking a multiple cloning site (MCS), as well as a Φ C31 attB site inserted by us. In order to systematically address the role of vector sequences I re-cloned the *e7p7-lacZ* reporter construct into a minimal *pBluescript* backbone to which we added a Φ 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 Pax6 (Toy/Eyeless) responsive enhancer (Berghammer *et al.* 1999; Horn *et al.* 2000). All constructs for this analysis were introduced into the *attP40* locus.

Both *pBluescript*-based *e7p7-lacZ* constructs (mini-*white* and *3xP3-dsRed*) expressed LacZ in the same pattern as the original *pPelican*-based reporter. However, when they were tested in *trans* to the original *e8p8-GFPm8*, no transvection was observed, namely we detected neither LacZ in the WM nor GFP in the AMPs (**Figure 30 B and C**, compare to **A**). I concluded that transvection at *attP40* does not happen whenever an enhancer-promoter pair is placed in *trans*, but needs additional sequences from the vectors. Upon flanking the *e7p7-lacZ-3xP3-dsRed* reporter with two GIs in the same orientation as in *pPelican*, bidirectional transvection (*e8*→*p7*, blue dotted rectangle and *e7*→*p8*, red dotted rectangle in **Figure 30 D**) was restored implicating the paired GIs as mediators of the effect. Because the GI is known to be bound by the Su(Hw) zinc-finger protein, which mediates its insulator activity (Parkhurst *et al.* 1988; Spana *et al.* 1988; Spana and Corces 1990; Holdridge and Dorsett 1991; Geyer and Corces 1992; Gerasimova *et al.* 1995; Ramos *et al.* 2006), I tested my transgene pairs in a *su(Hw)* mutant background. Indeed, *e7*→*p8* (AMP) transvection was lost in this background, but, surprisingly, *e8*→*p7* (WM) transvection was now apparent in all transgene combinations that had a mini-*white* marker in both homologues (**Figure 30 A', B' and E'**), even the one that had not displayed this effect in wild type (wt) genetic background (**Figure 30 B vs B'**). I then modified the *pPelican*-based *e8p8-GFPm8* construct by removing its GIs. This construct was capable of supporting *e8*→*p7* transvection with the mini-*white* but not the *3xP3-dsRed* version of the *pBluescript-e7p7-lacZ* reporter in both wt and *su(Hw)* genetic backgrounds (**Figure 30 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*.

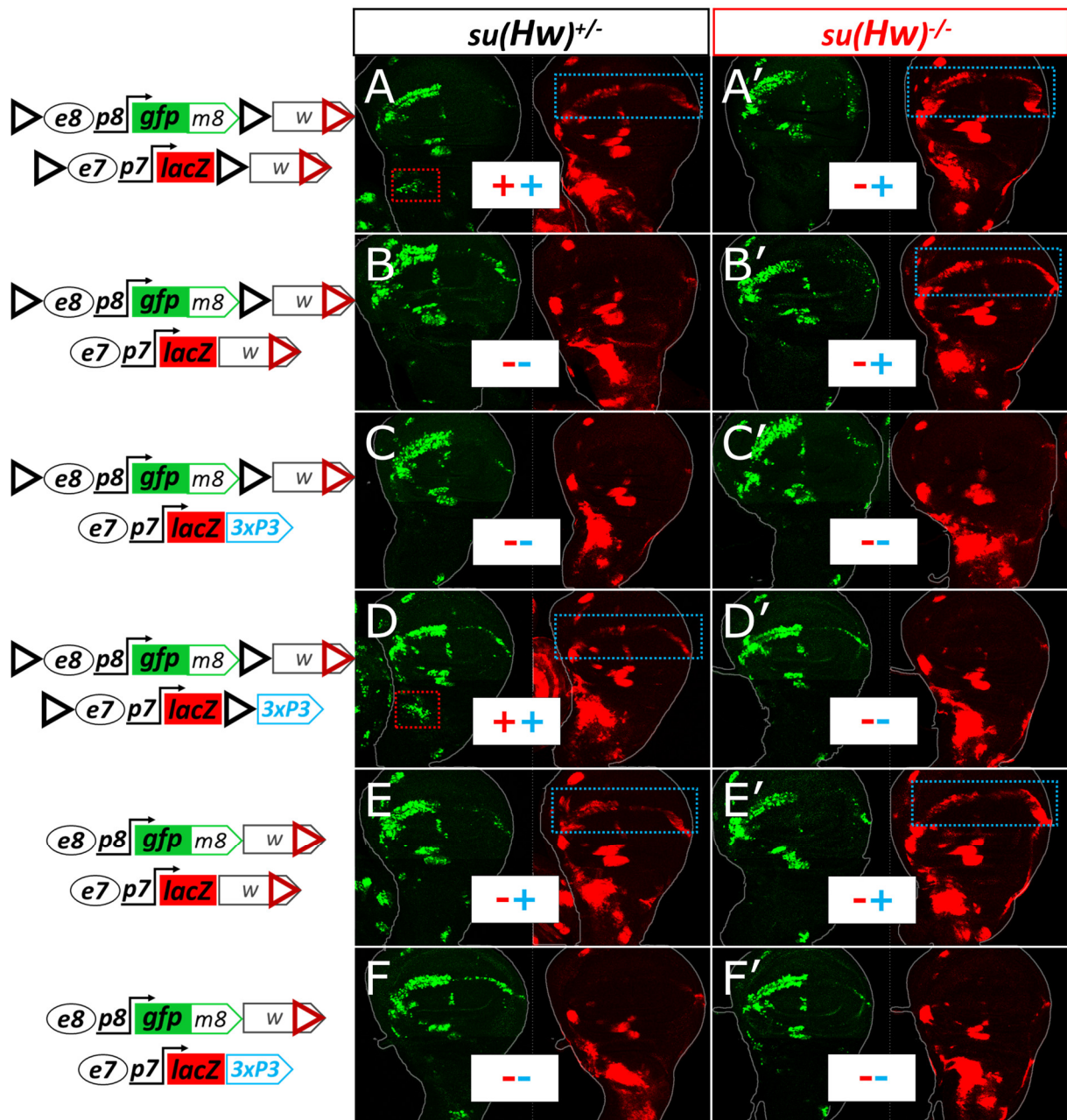


Figure 30 GIs and mini-white mediate transvection. Wing imaginal disks from animals heterozygous for variations of the *e8p8-GFP-m8* and *e7p7-lacZ* transgenes. The two GIs in forward orientation, flanking the two reporters, are indicated by black triangles; marker genes are located downstream of the 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 – also indicated by + or – in the same color (red or blue). 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 the *attP40* locus.

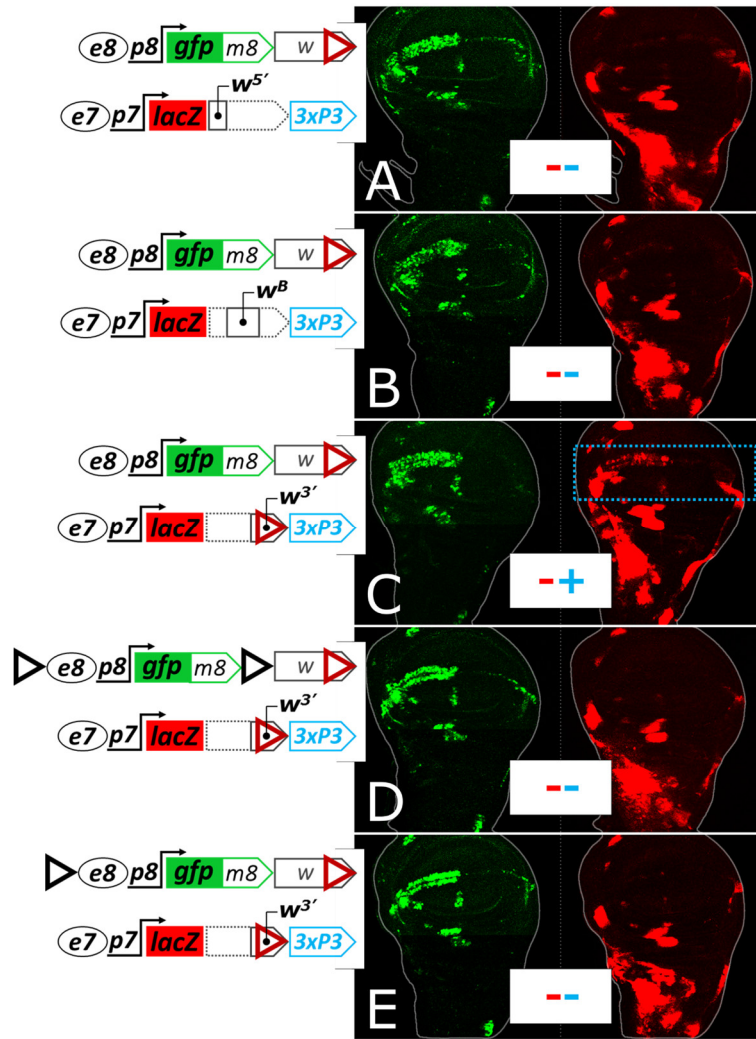


Figure 31 *White*-mediated uni-directional transvection relies on the mini-*white*-contained Wari insulator (WI). Three fragments of mini-*white*, cloned in the *e7p7-lacZ-3xP3* construct, were tested in the *attP40* locus in a transvection assay against an *e8p8-EGFPm8* transgene containing full-sized 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; blue + sign). This fragment is known to contain WI. Addition of one or two GIs in the *e8p8* construct inhibited WI-mediated transvection (D-E). + and – signs refer to presence and absence of transvection, respectively; red refers to *e7*→*p8* transvection (GFP in AMPs) and blue refers to *e8*→*p7* transvection (LacZ in the WM).

I mapped the transvection-mediating element of the *white* locus within its 3' part: out of 3 subfragments derived from mini-*white* in **Figure 31 A-C**, only its 3'most 0.9 kb recapitulated *white-white*-mediated transvection (**Figure 31 C**). It was previously reported that this sequence contains the 3'-UTR as well as an insulator element, dubbed Wari (hereafter referred to as 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 I confronted a WI-containing *e7p7-lacZ* with a WI-containing (mini-*white*) *e8p8-GFPm8* flanked by GIs (**Figure 31**

D), transvection was abolished, consistent with what I had observed earlier with the entire mini-*white* (**Figure 3B**). I hypothesized that this inhibition could result from presumptive insulation imposed by the GI located between *e8* and WI in *e8p8-GFPm8*, thereby restricting access of WI to *e8*. This was not the case, as the inhibition of WI-mediated transvection was sustained even when I deleted the 3' GI, leaving only the 5' GI intact (**Figure 31 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, I 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 48**).

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

Because *e7p7* and *e8p8* interact in *cis* in their native context (**Figure 1A, B**), it raises the possibility that the transvection I observed is tied to the specificity of these two regulatory modules for each other. Thus, I sought to determine whether GIs and WIs could mediate transvection of *e7p7* and *e8p8* to an unrelated, heterologous promoter. To address this question, I 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 32 A**). When crossed to an *e7p7-lacZ* construct, congruent in two flanking GIs and mini-*white*, GFP was detected with the *e7*-specific pattern (**Figure 32 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 32 C**), while removal of GIs from one of the transgenes (with congruent mini-*white* markers) abolished the effect (**Figure 32 D**). Unlike the *e8p8*→*e7p7* transvection, removing GIs entirely, but keeping congruent mini-*white* markers did not support transvection (**Figure 32 H**). Therefore, paired WIs cannot support transvection between *e7* and *pH*.

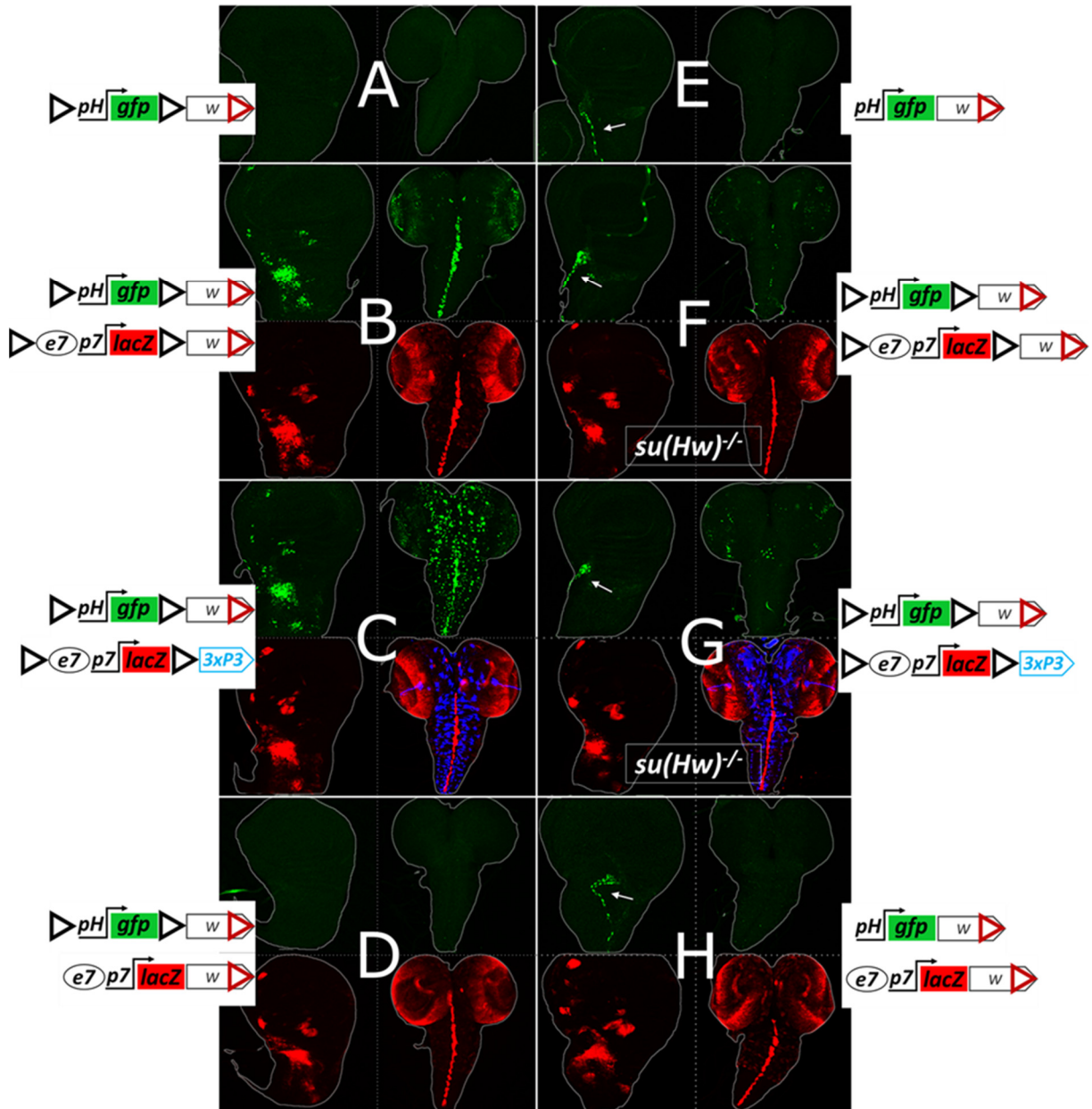


Figure 32 GIs, but not WIs, mediate transvection of *e7* and *3xP3* enhancers to a heterologous, enhancerless *hsp70* promoter (*pH*). Representative late larval wing disks and CNSs from the indicated genotypes are shown. (A-H) Projections of *trans* (GFP, green) and *cis* (LacZ red and DsRed blue), when present, expression from the same wing disk or CNS of each sample. The *ph-gfp* produces no (*cis*) expression as a hemizygote (A, E), other than tracheal branch in the uninsulated version (white arrow in E), probably originating from enhancer trapping. The same is observed in the insulated version upon GI inactivation in the *su(Hw)*^{-/-} background (F and G). Only combinations of transgenes with GIs in both homologs support transvection (GFP in AMPs and CNS in B and C), while congruency in *white* in the absence of GIs in one or both homologs does not (D, H). Depletion of Su(Hw) protein abolishes transvection (F and G, which bear the same transgenes as B and C, 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 (see blue pattern in Figure 1 A). The artificial *3xP3* enhancer does not drive expression in wing disks.

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 32 C). Depletion of Su(Hw) suppressed transvection of both enhancers, *e7* (Figure

32 F, G) and *3xP3* (CNS in **Figure 32 G**), and allowed *pH* to trap a tracheal enhancer (in *cis*) in the vicinity of *attP40* (white arrows in **Figure 32 F, G**). 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 me to use the Gl-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 Gl's activated specific patterns of expression in *trans* from the Gl-flanked *pH-gfp* transgene (**Figure 33**) – a full 15 out of these 18 fragments displayed robust enhancer activity in third instar larval disk/CNS tissue. Additionally, using this system, I was able to recapitulate in *trans* the *cis*-pattern of another enhancer unrelated to the *E(spl)* locus, the *vestigial* quadrant enhancer (last column, **Figure 33**).

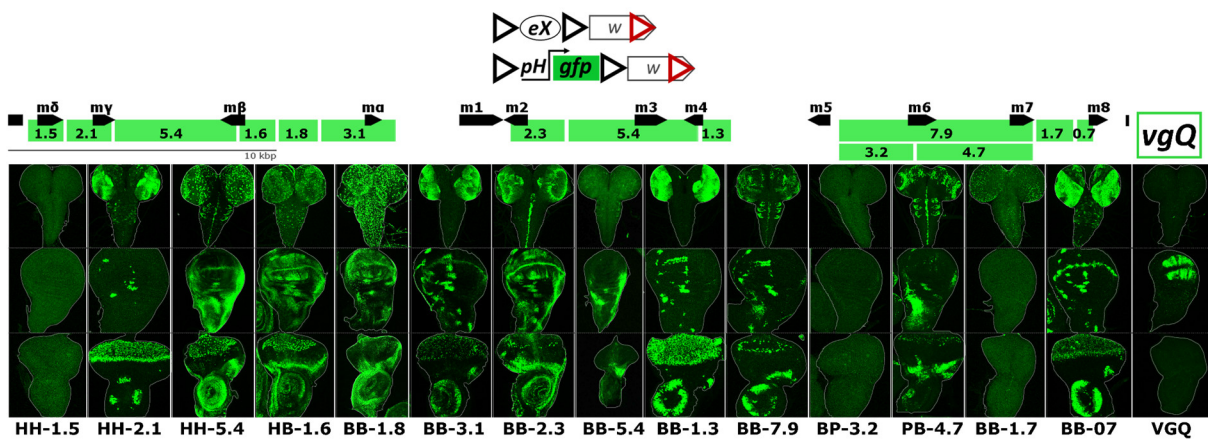


Figure 33 GI-mediated transvection is compatible with many developmental enhancers. A map of the *E(spl)* locus is shown on top and DNA fragments are marked as green rectangles with indicated length in kb. Each of these DNA fragments ('eX' in the constructs' scheme) was inserted into a transgene between Gl's, placed in *trans* to Gl's-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 the *E(spl)* locus, the *vestigial* quadrant enhancer (*vgQ*, last column) recapitulates its *cis* activity in *trans* (GFP expression only in the wing disk).

Unlike GI, WI-mediated transvection was specific for the *e8p8/e7p7* 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*), I tested various minimal GFP reporters driven by different basal promoters, *pH*, *p7* and *E(spl)m8* promoter (*p8*). I confronted these enhancerless non-expressing reporters with an *e8p8-lacZ* transgene. I made sure that all combinations were congruent for both Gl's and mini-*white*, which enabled me to simultaneously test GI-mediated transvection in a wt background and WI-mediated transvection in the absence of Su(Hw) (**Figure 34**). All three basal reporters responded to *e8p8* enhancer in a *su(Hw)^{+/-}* background (GI-dependent transvection); the *p7* was by far the

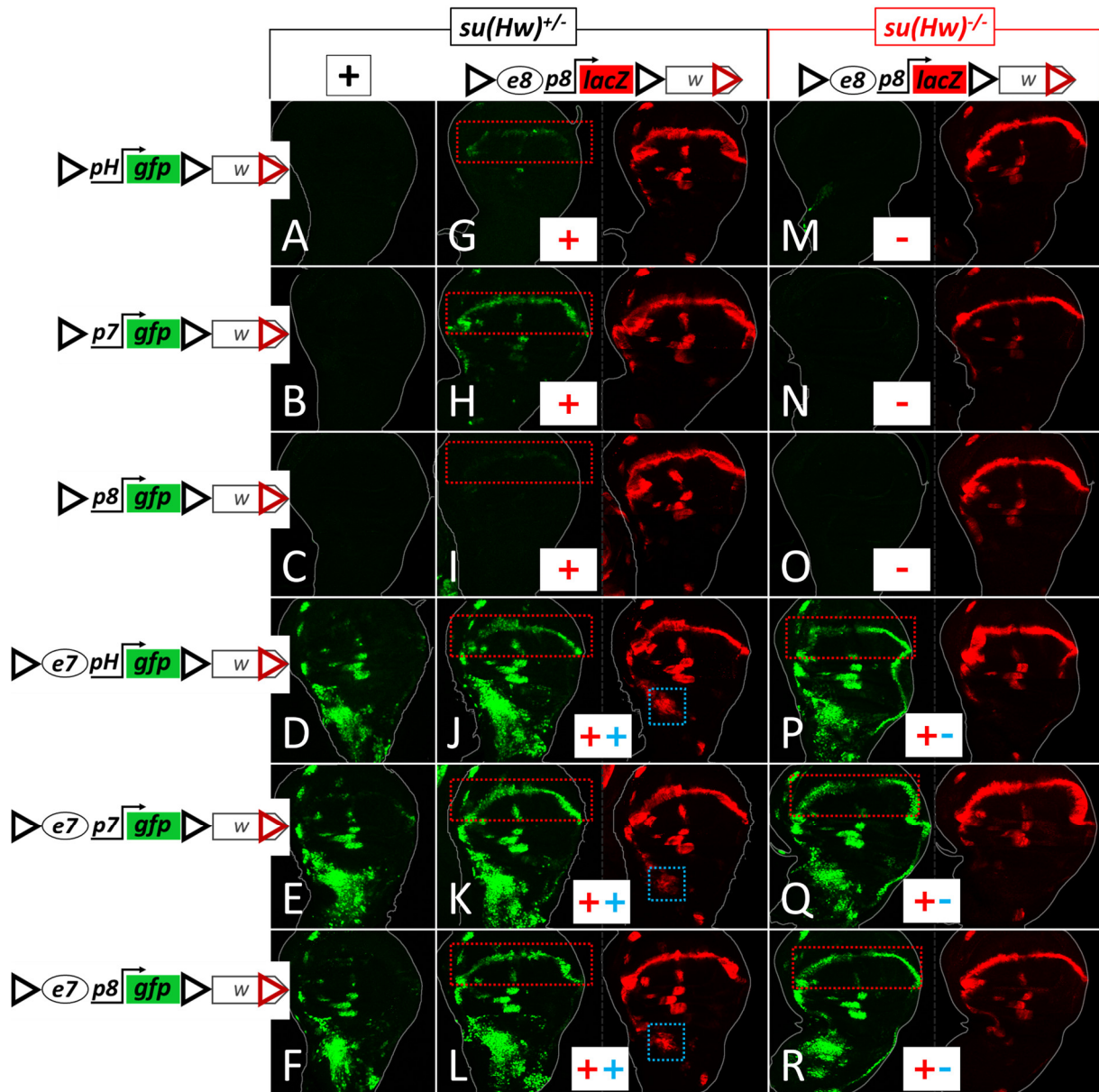


Figure 34 WI-mediated *trans*-activation by the *e8* enhancer requires the presence of the *e7* enhancer on the other homolog. All transgenes contain GIs and mini-*white* and are inserted in *attP40*. (A-F) Wing disks from animals hemizygous for the indicated GFP transgenes. When GFP is driven by the *pH* promoter (A), *E(spl)m7* (*p7*) promoter (B) or *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 *e8p8-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. These are also summarized by blue and red signs, respectively. + indicates transvection, - no transvection.

strongest responder, with *pH* following and *p8* showing a very weak activation (Figure 34 A-C). However, in the *su(Hw)^{-/-}* background (WI-dependent transvection), the *e8* enhancer did not transfect to any of the three promoters (Figure 34 M-O, compare to Figure 34 G-I). Therefore, the unidirectional *e8p8*→*e7p7* transvection supported by WIs was not due to a selectivity of WI for *p7*. When the three basal promoters were fused to the *e7* enhancer (Figure 34 D-F), these reporters started expressing (as hemizygotes) in the AMPs and some

proneural clusters, but not in the WM, consistent with the activity of the *e7* enhancer (**Figure 34 D-F**). Once confronted in *trans* with *e8p8-lacZ*, these *e7*-bearing reporters were able to robustly express GFP in the WM (**Figure 34 J-L**) and this expression was retained in the *su(Hw)* mutant genetic background (**Figure 34 P-R**). Thus, the transvection mediated by the interaction between WIs was not promoter context specific, but rather enhancer context specific, with the responding gene requiring the presence of the *e7* enhancer in *cis* in order to sustain WI-mediated transvection. In conclusion, whereas *trans*-paired GIs mediated transvection between any enhancer-promoter pair tested, *trans*-paired WIs were more selective and mediated only *e8*→*e7* transvection, an effect that was dominantly suppressed by the presence of GI elements. This *e8*→*e7* transvection may reflect some intrinsic affinity of these two enhancers for each other, but it still requires the presence of GIs or WIs in order to materialize.

3.4.4. Transvection is weaker than *cis* enhancer-promoter activity and is suppressed by promoter *cis*-preference

To gauge the relative strength of transvection compared to *cis* enhancer-promoter (e-p) interaction I generated e-p pairs driving GFP expression in *cis* and compared them to the same e-p pairs driven in *trans*. All constructs designed for this purpose were based on the backbone of *pPelican/pStinger* vectors (Barolo *et al.* 2000) which contain two GIs to enable GI-mediated transvection, and subsequently inserted into the *attP40* locus. These flanking GIs also provided efficient insulation: all enhancerless promoter-reporter constructs had undetectable levels of expression in all three larval tissues tested (wing disk, eye disk and CNS; data not shown). In all cases we observed lower GFP levels from transvection than from the *cis*-combination (**Figure 35**). I tested three promoters, *p7*, *p8* and *pH*. Regardless of the enhancer assayed (*e7* or *e8*) the strongest expression levels, both in *cis* and in *trans*, were produced by *p7*, whereas *p8* was the weakest out of the three promoters. This suggests that a promoter's strength for driving transcription is its intrinsic property and does not depend on the enhancer activating it, at least for the two enhancers tested.

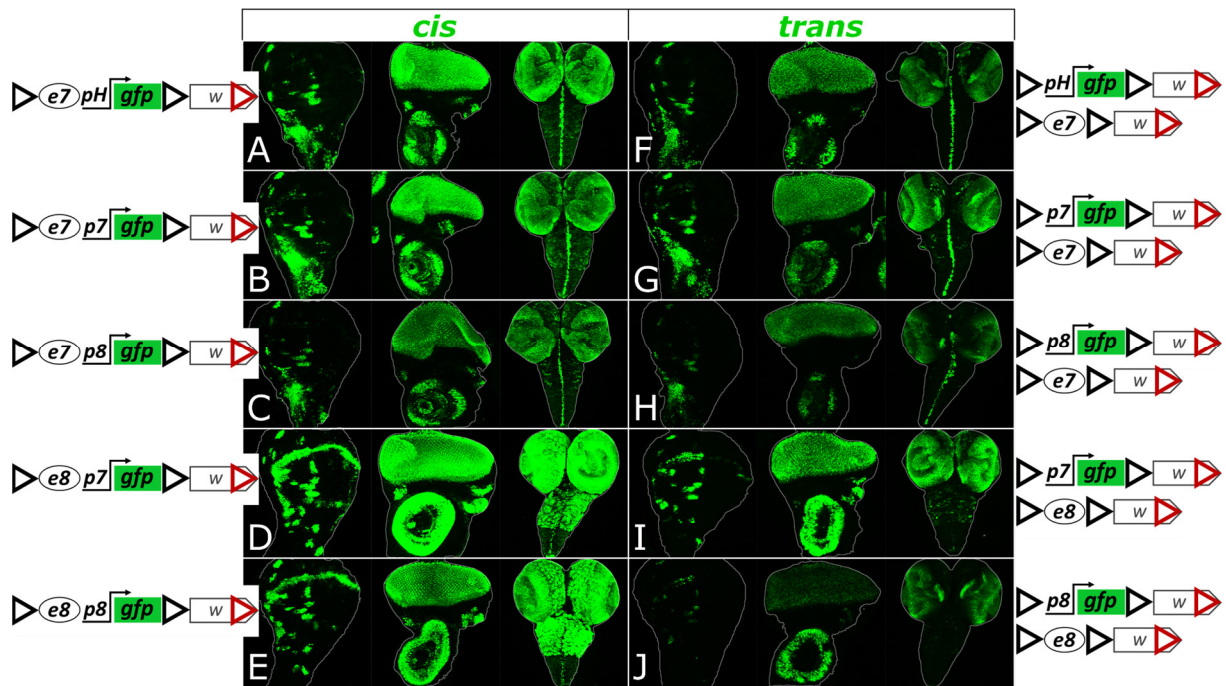


Figure 35 Transcriptional activity of different enhancer-promoter (*e-p*) pairs interacting in *cis* or in *trans* in the presence of Glis. Each genotype is examined for GFP expression in the third instar wing disk, eye disk and CNS. (A-E) Samples from hemizygotes of a transgene containing a given *e-p* 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*.

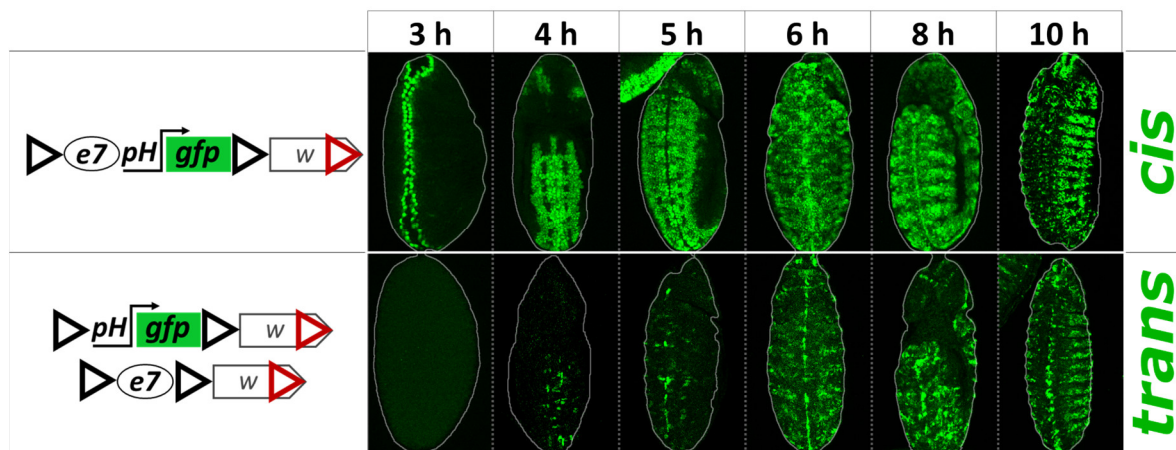


Figure 36 The onset of transvection is delayed in embryogenesis. All embryos are imaged ventrally with anterior to top. A *cis*-linked *e7pH* enhancer-promoter pair (top row) drives GFP expression along the ventral midline within 3 h after egg deposition (AED). At 4-5 h AED *e7pH* is broadly active in the ventral ectoderm, whereas *e7* and *pH* separated in *trans* show interaction only in a 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). All transgenes are inserted in *attP40*.

I also studied the embryonic *cis* versus *trans* expression of *e7* using the *pH* promoter constructs. In *cis*, *e7* displayed very dynamic expression, starting at the mesectoderm in stage

7 (3 h after egg laying, AEL), then in a neuroectodermal cluster pattern up to stage 10 (5 h AEL) and later in the VNC midline and the epidermis in a complex segmentally repeated pattern. Interestingly, the earlier patterns could not be transected or were transected only in sporadic cells. From stage 11 (6 h AEL) onward, this sporadic transvection gave way to a more complete one, where the *e7 trans* pattern recapitulated the *cis* pattern (**Figure 36**). 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, I 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 37 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).

In the above experiments I noticed that confronting an enhancerless reporter in *trans* to a solitary enhancer gave more robust expression compared to all my previous experiments, where the transfecting enhancer was linked in *cis* to a promoter (**Figure 37**, compare **B1** to **B2** for *e8*, and, **B3** and **B4** to **B5** for *e7*). This came as no surprise, since numerous earlier studies on transvection have indicated that an enhancer's action in *trans* is suppressed by the presence of a promoter in *cis* (Geyer *et al.* 1990; Martínez-Laborda *et al.* 1992; Hendrickson and Sakonju 1995; Casares *et al.* 1997; Sipos *et al.* 1998; Morris *et al.* 1999a; b, 2004; Bateman *et al.* 2012a; Kravchuk *et al.* 2016). Different promoters of varying core element composition have been reported to display *cis*-preference (i.e. to attenuate transvection). As a general rule, mutations compromising transcriptional strength of a *cis*-promoter usually release the

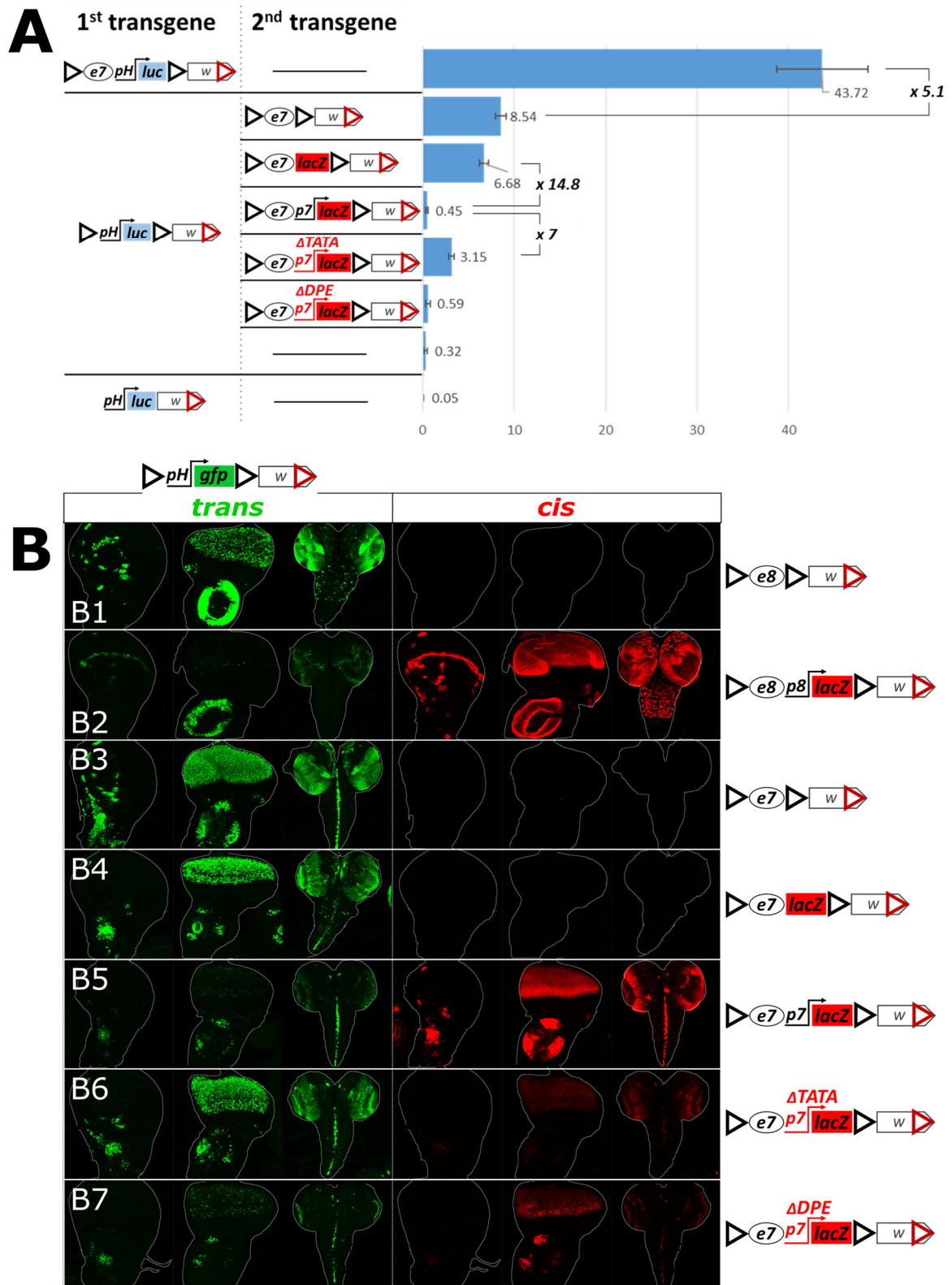


Figure 37 Regulation of GI-mediated transvection by *cis*-preference. (A) The chart shows levels of basal and *e7*-induced (in *cis* and in *trans*) *pH*-driven luciferase activity. Levels of luciferase activity were measured from third instar larval disk-brain complexes. Luciferase values normalized 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 (green), reflecting *trans* activity of the enhancer-containing transgenes on *pH-gfp* and (2) β -galactosidase expression (red), reflecting *cis* activity of the enhancer-linked promoter, when *lacZ* is present. All transgenes contain GIs and mini-*white* and are inserted in *attP40*.

enhancer towards *trans* action (Morris *et al.* 1999b, 2004; Lee and Wu 2006). I made two mutations on the *p7* promoter in an attempt to compromise its strength without completely inactivating it. *p7* is a multi-element promoter, containing a TATA box, an initiator (Inr) and a downstream promoter element (DPE) (Klämbt *et al.* 1989; Kutach and Kadonaga 2000). I introduced two deletions into the *e7p7-lacZ* construct aiming to disrupt each of these activities; one, *e7p7- Δ TATA-lacZ*, removed the TATA box [deletion of -41 to -22 bp relative to the transcription start site (TSS)] and another, *e7p7- Δ DPE-lacZ*, removed the Inr and DPE elements (deletion of -16 to +67 bp relative to TSS). Both of these promoter mutations retained weak yet detectable transcriptional activity (**Figure 37, B5-B7, *cis* column**). Even though the reduction in *cis* promoter activity was comparable between Δ TATA and Δ DPE, the two had dramatically different effects on transvection of the linked *e7* enhancer (**Figure 37, B5-B7, *trans* column**). *p7- Δ TATA* partially relieved *e7* from *cis*-preference inhibition, leading to much higher *trans*-activation of *pH-gfp* than that observed with *p7- Δ DPE*. 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 38**). When the various *e7p7-lacZ* versions were confronted with the *pH-luc* reporter, I 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 *e7p7* or the *e7p7 Δ DPE*) (**Figure 37 A**). Interestingly, in this assay the intact *e7p7* (and *e7p7- Δ DPE*) 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. I 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, I can only measure the resultant luciferase activity in the whole brain-disk extract with no cell-to-cell resolution, so I cannot test this scenario. Regardless, these luciferase constructs enabled me 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.

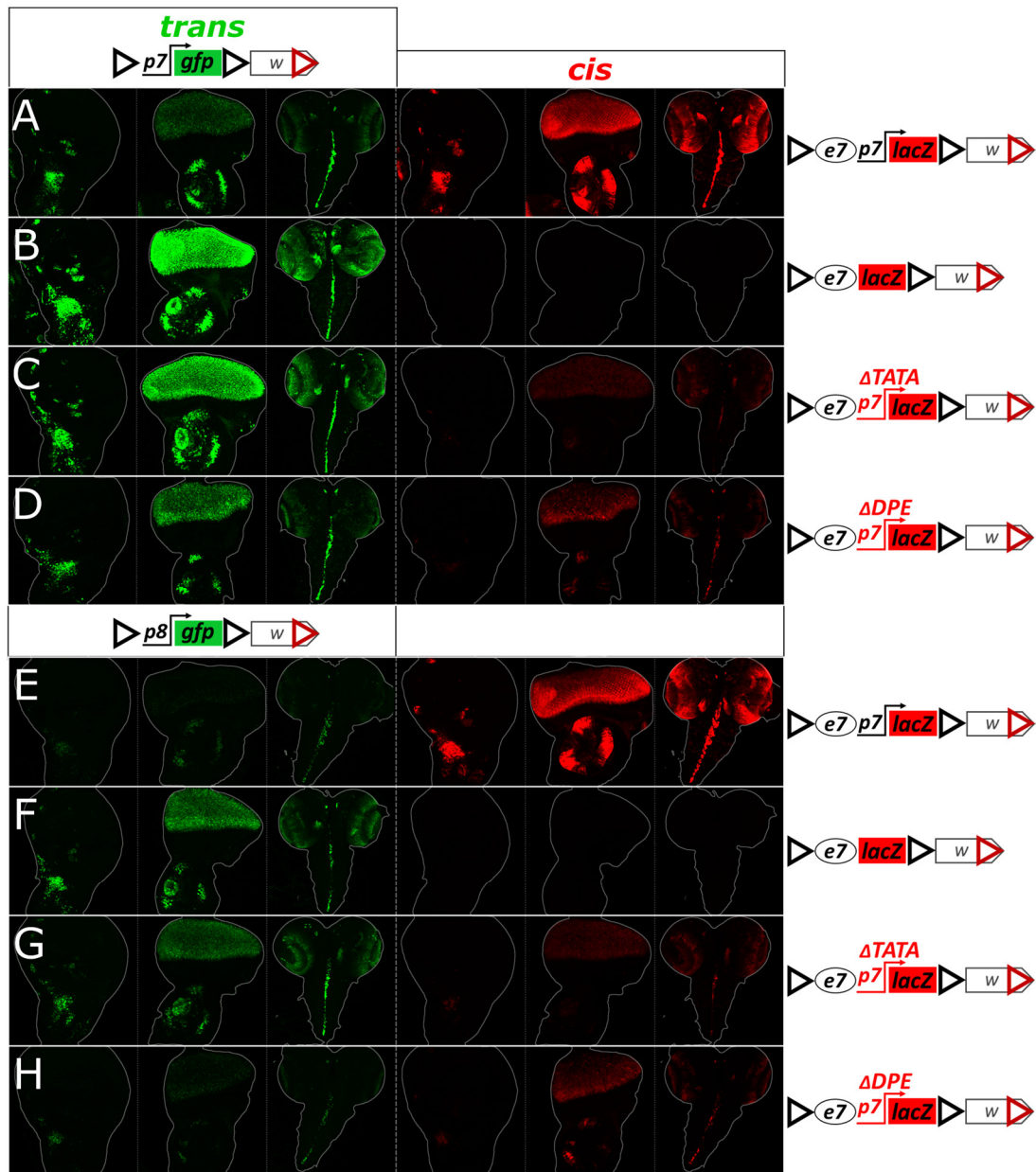


Figure 38 The effects of mutations in the *p7* promoter on the ability of a *cis*-linked *e7* to transvect are independent of the identity of the *trans* promoter. The *e7p7-lacZ* (A and E) and its derivatives harboring deletions of *p7* (B and F), Δ TATA (C and G) and Δ DPE (D and H) are placed in *trans* to *p7-gfp* (A - D) or *p8-gfp* enhancerless reporters (E - H). For each genotype (each row) the third instar wing disk, eye disk and CNS are examined for (1) GFP, reflecting *trans* activity of *e7* on *pH-gfp* (green) and (2) β -galactosidase, reflecting *cis* activity the *e7*-linked promoter (red). Although, *p7*, *p8* and *pH* are promoters of different strength, their activity seems to be affected similarly by different mutations in *e7*-linked *p7* (compare to the results obtained for *pH*, Figure 5, D-G). All transgenes contain GIs and mini-white and are inserted in *attP40*.

3.4.5. Relative position, number and orientation of GIs determine transvection effects

In all the previous experiments, all my 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 39**). GI is 367 bp long and its "forward" orientation is the same as the one found in the original *gypsy* transposon, where GI is located shortly downstream of the 5' LTR (Spana *et al.* 1988, **Figure 40**). The two transgenes (*e7p7-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 32** and **Figure 39**). 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}. I have introduced a series of modifications in the configuration of the GIs within these two constructs in order to understand how their relative number, position and orientation influence transvection. All resultant constructs for this analysis were introduced into the *attP40* locus.

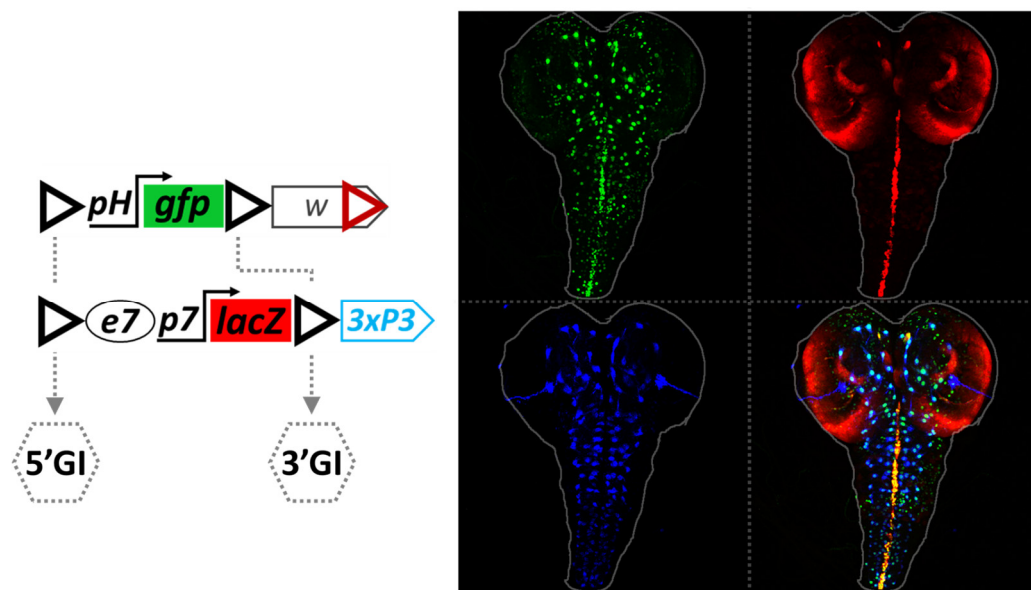
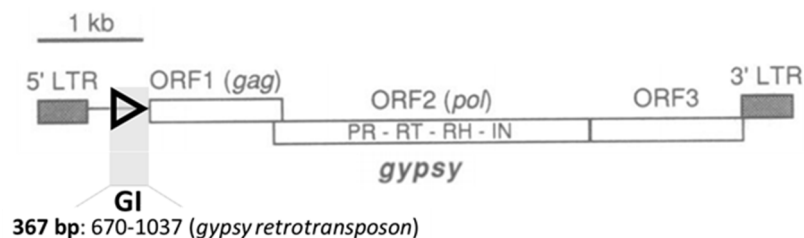


Figure 39 *pH* receives input from two enhancers in *trans*, *e7* and *3xP3*. Confocal z-projection of a third instar CNS from a heterozygote between *e7p7-lacZ-3xP3* and *pH-gfp* dual-GIs^{FOR} transgenes in *attP40*. Top left panel shows *trans*-activated *pH-gfp* expression; note *e7*-specific GFP 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*, bottom left panel). Merged image is shown in the bottom right panel.



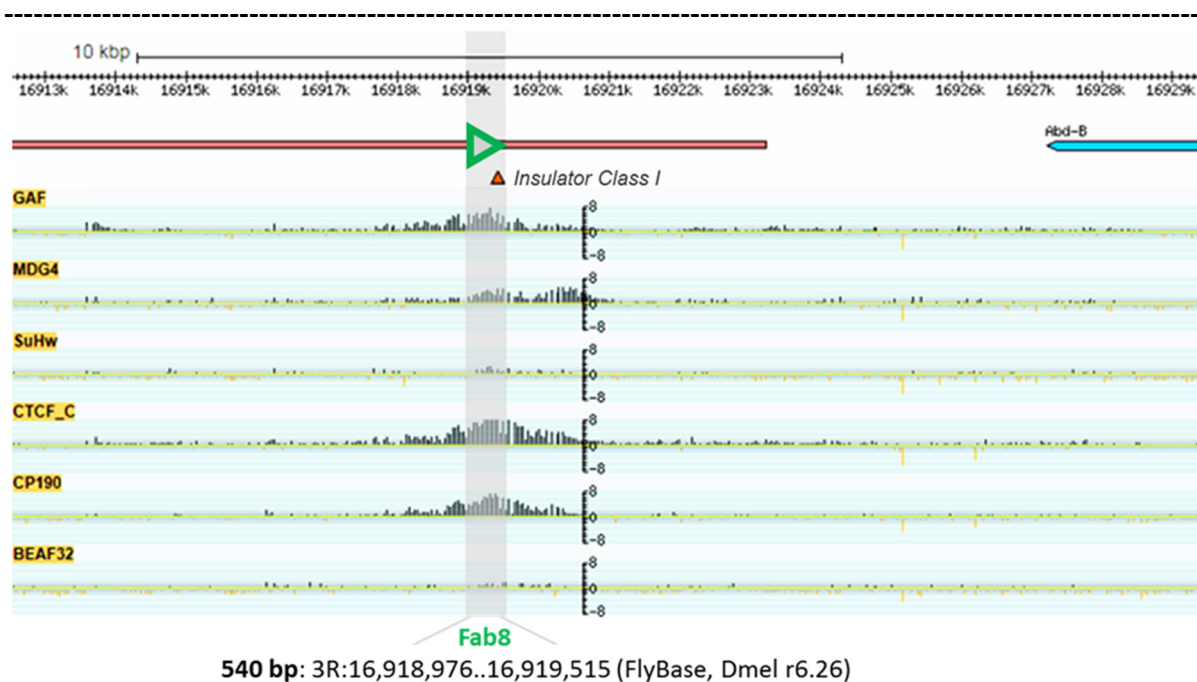
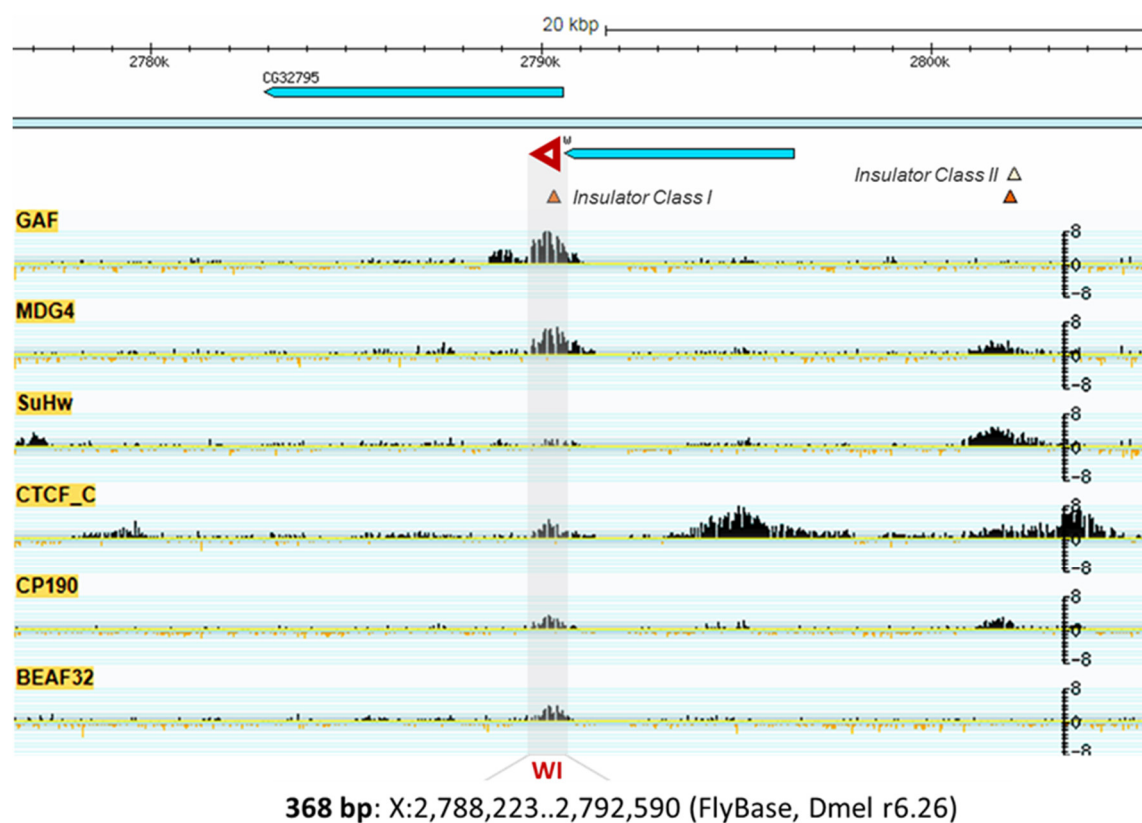


Figure 40 Continues on the next page

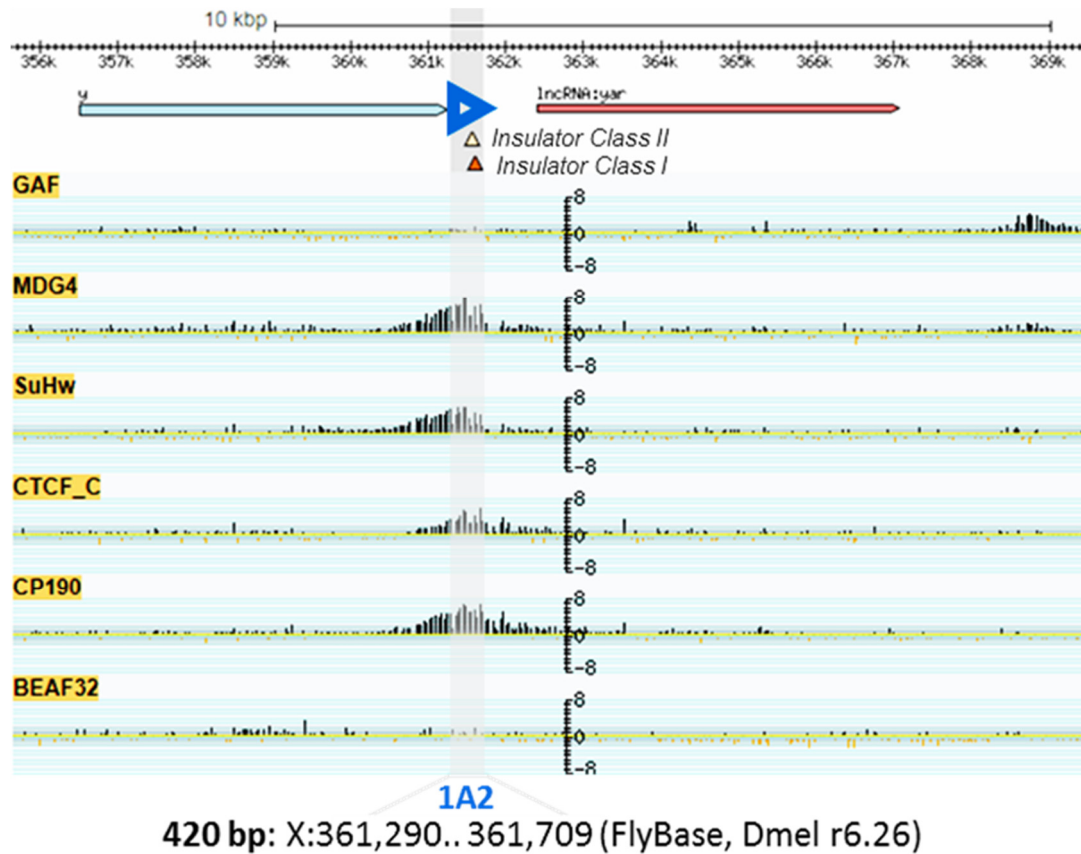


Figure 40 Genomic maps of insulators. Insulators are shown as triangles in the orientation used in the various construct schematics. The position of each insulator is marked in the genome using the genome browser of Flybase (Thurmond *et al.* 2019). Below each genomic map are shown six chromatin-IP tracks for insulator binding proteins derived from modENCODE (Celniker *et al.* 2009) and a peak-calling track for putative insulators (Negre *et al.* 2010), as well as the exact sequence coordinates of the fragment cloned in our constructs. Genomic maps and Ch-IP tracks are not possible for GI, since it comes from a repetitive element; instead the sequence coordinates come from the *gypsy* retrotransposon sequence in Spana *et al.* 1988 and Kim *et al.* 1994.

First, I examined how differently positioned and oriented GIs influence expression in *cis* of the *e7p7*-driven LacZ and *3xP3*-driven dsRed in larval brains using whole series of the obtained *e7p7-lacZ-3xP3* transgenes (**Figure 41**). Presence of GI^{FOR} upstream and adjacent to any of these *e-p* pairs resulted in elevated levels of expression driven by the 3'-linked *e-p* pair (note elevated levels of DsRed, blue, in **Figure 41 A2** as compared to **A1**; and elevated levels of LacZ, red, in **A3** as compared to **A1**). The 5' GI^{FOR} stimulated also the expression of DsRed at a distance, but weaker than the *3xP3*-adjacent (3') GI^{FOR} (compare **Figure 41 A3** to **A1** and **A2**). Consequently, dual-GIs^{FOR} *e7p7-lacZ-3xP3* expressed, both, LacZ and DsRed at higher levels (**Figure 41 A4**). Interestingly, in both cases where GI is not present between *e7p7-lacZ* and *3xP3-dsRed* modules, LacZ is detected in *3xP3*-specific glial cells, revealing that *3xP3* enhancer interacts with *p7* promoter if not separated by a GI (**Figure 41 A1** and **A3**). Accordingly, interposing a GI abolishes interaction between them (**Figure 41 A2** and **A4**). The absolute orientation of a standalone (5' or 3') GI and the relative orientation of the 5' and 3' GIs are irrelevant for the GI's impact on gene expression in *cis* (**Figure 41 B1-B8**). What matters is the

position of a GI relative to orientation of a promoter (or direction of transcription), or, proximity to an enhancer. Whenever GI is positioned at the 5' (back) side of a promoter (or, what is the same case, adjacent to an enhancer) transcription of a gene is stimulated. A convincing demonstration of this rule is observed with a transgene producing elevated levels of both, LacZ and DsRed, from two *e-p* modules which "share" one GI at their 5' sides and which are transcribed outwardly (**Figure 41 B9**).

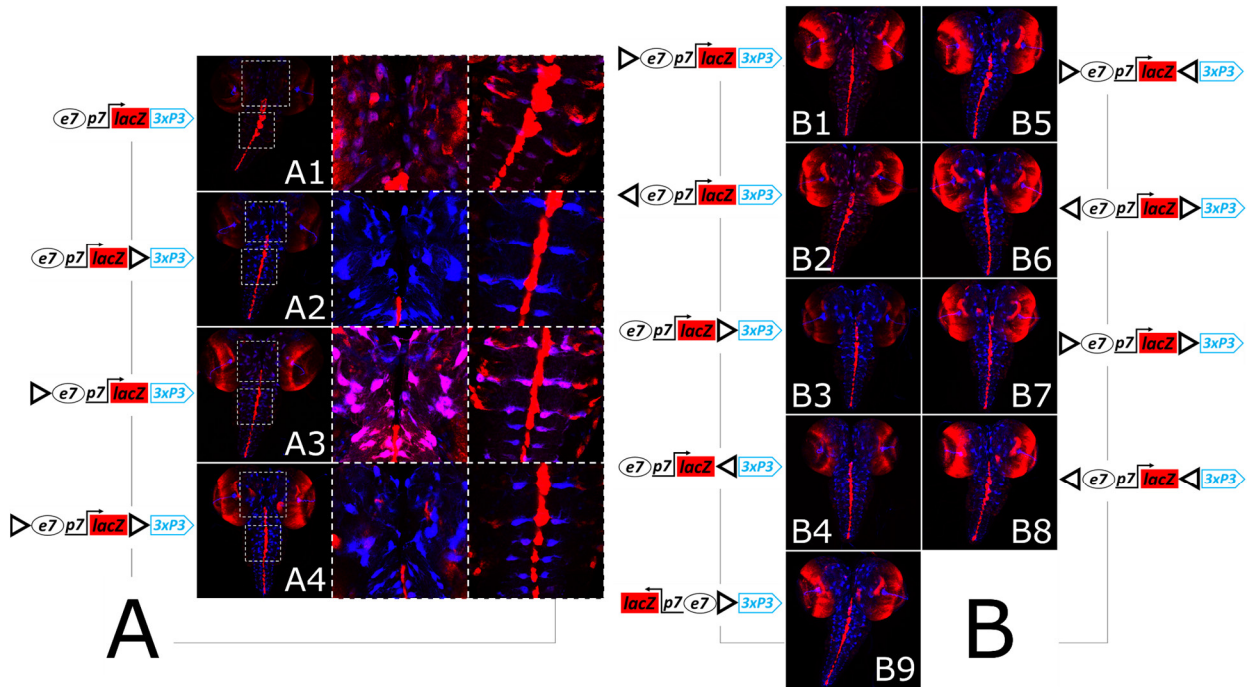


Figure 41 GIs stimulate expression in *cis* independently of their orientation. Confocal z-projections of third instar CNSs showing merge of LacZ (red) and DsRed (blue) channels. Panel **A** shows CNSs from animals bearing *e7p7-lacZ-3xP3* transgenes GI-less (**A1**) or containing GI/*s*^{FOR} (**A2-A4**) with enlarged regions of the central brain and the VNC indicated by white rectangles. Panel **B** shows CNSs from animals bearing whole set of constructs with all combinations of GI/s orientation (FOR/REV) and position (5' or/and 3') (**B1-B8**) and a construct which was made by inverting *e7p7-lacZ* module in **B3** construct.

Deletion of the 3' GI^{FOR} in the transvection-receiving construct, *pH-gfp*, while preserving the dual-GIs^{FOR} configuration in the 'sending' *e7p7-lacZ* construct, caused a reduction in transvection of *e7*, with concomitant increase of *3xP3* transvection (compare **Figure 42 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 42 C to A**). Finally, deletion of the 3' GIs^{FOR} in both constructs led to augmented GFP expression with an *e7* pattern and an almost undetectable *3xP3* pattern (compare **Figure 42 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' GIs^{FOR} of the two transgenes. This interaction seems to be weakened by the presence of a 3' GI^{FOR} in either or both of the interacting transgenes. However, the 3' GI^{FOR} in the sending construct is required for effective transvection of the

3xP3, suggesting that the 3xP3 enhancer, like *e7*, needs an adjacent GI in order to robustly act in *trans*.

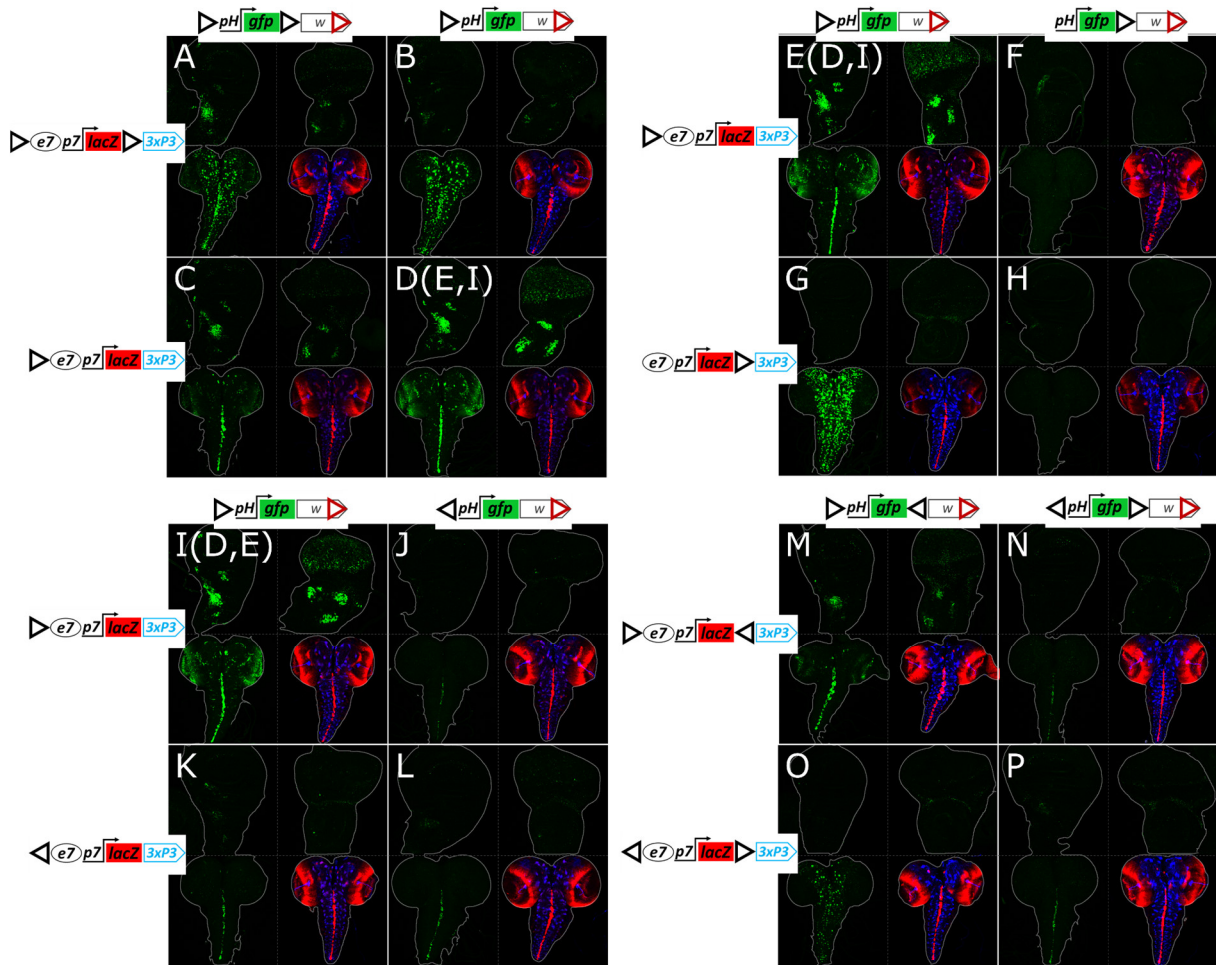


Figure 42 The relative position, number and orientation of GIs determine transvection effects. (A-P) Confocal z-projection of GFP expression in third instar larval wing disk (top left panel for each genotype), eye disk (top right) and CNS (bottom left). Bottom right shows the merged *e7p7*-driven LacZ (red) and *3xP3-pH*-driven, DsRed expression (blue) in the same CNS as the bottom left panel; for patterns of *e7p7-lacZ* expression in disks, see previous figures (Figures 2, 3, 4, 6); *3xP3-pH-dsRed* shows no expression in third instar imaginal disks. Each genotype contains a *pH-gfp* transgene with mini-white in *trans* to an *e7p7-lacZ* transgene with *3xP3-dsRed* and various arrangements of GIs (black triangles), as indicated. Note that **D**, **E** and **I** represent different samples obtained from the same genotype. All transgenes are inserted in *attP40*.

Consistently, when the dual-enhancer construct (*e7p7-lacZ-3xP3-dsRed*) contained a single GI^{FOR}, only its downstream enhancer was transvected to a GI^{FOR}-preceded *pH-gfp*: *e7* (with weak sporadic activity of *3xP3*) (**Figure 42 D, E, I**) or *3xP3* (**Figure 42 G**). Moreover, the presence of the GI^{FOR} upstream of the *pH* promoter was essential for its activation by a *trans*-enhancer, as its deletion abolished transvection altogether, even when another GI^{FOR} was present at the 3' position (**Figure 42 F and H**). It therefore seems that the *trans*-activity of both enhancers (*e7* and *3xP3*) obtained from the dual-GIs^{FOR} *e7p7-lacZ* transgene (**Figure 42 A and B**) resulted from the interaction between the 5' GI^{FOR} preceding *pH* with the two GIs^{FOR} of *e7p7-lacZ* in *trans*, each upstream of each enhancer.

I next addressed the significance of the orientation of GIs in mediating transvection. Surprisingly, the presence of a reversed GI in the 5' position (5' GI^{REV}) in either of the two transgenes strongly reduced transvection effects, even when both transgenes contained 5' GI^{REV} (**Figure 42 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 42 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 me 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 42 M** compared to **A**) and less than the combination where 3' GIs are absent altogether (**Figure 42 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 42 P** to **L**), nor were incongruent 5' GI configurations improved by a 3' GI (**Figure 42 N, O**) (note, however, that *3xP3* was efficiently transvected to *pH* in **Figure 42 O** as both *3xP3* and *pH* are preceded by GI^{FOR}).

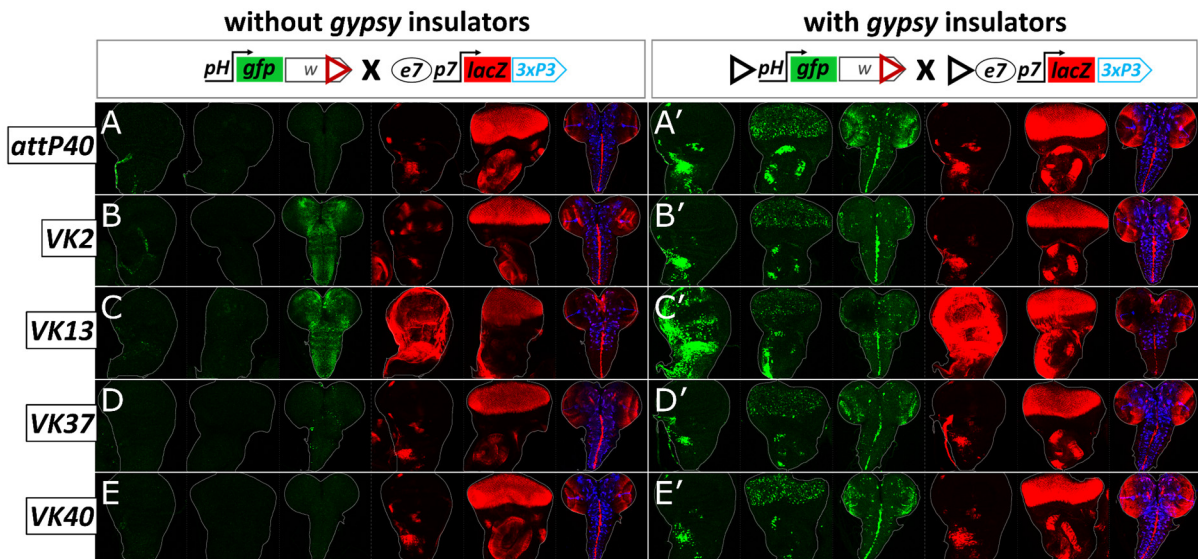
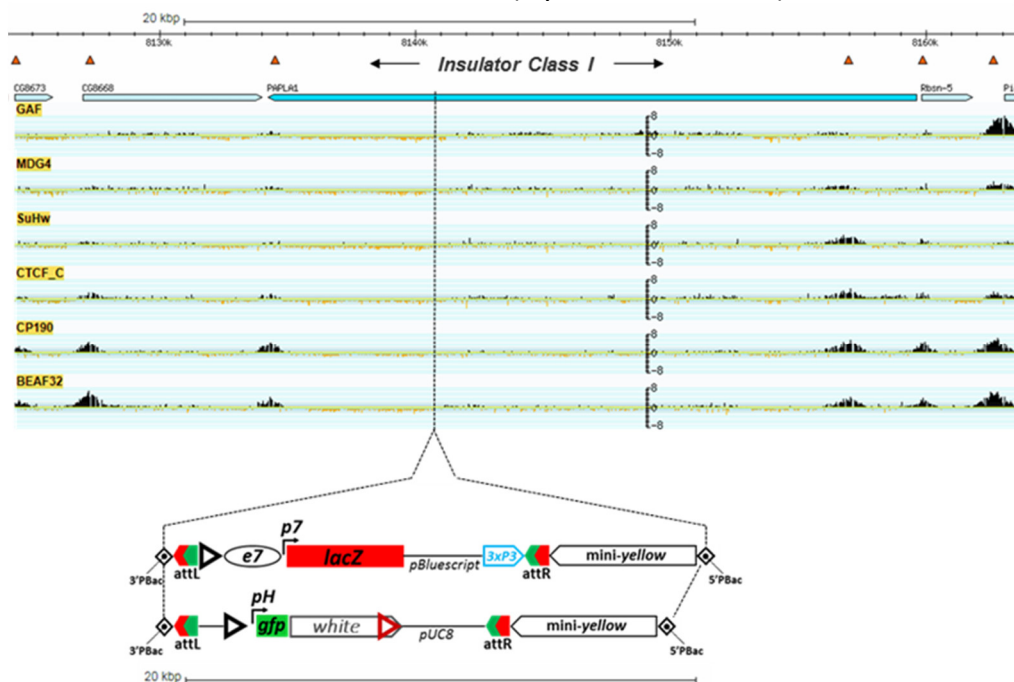


Figure 43 No transvection is observed in the absence of GIs at five different loci. The ability to interact in *trans* was tested between *e7p7-lacZ* and *pH-gfp* transgenes inserted in five genomic loci (*attP40*, *VK2*, *VK13*, *VK37*, *VK40*) in two configurations of these transgenes: carrying no 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 *e7p7*-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 expression in the wing disk of the uninsulated *pH-gfp* in *attP40* and *VK2* and of the GI-

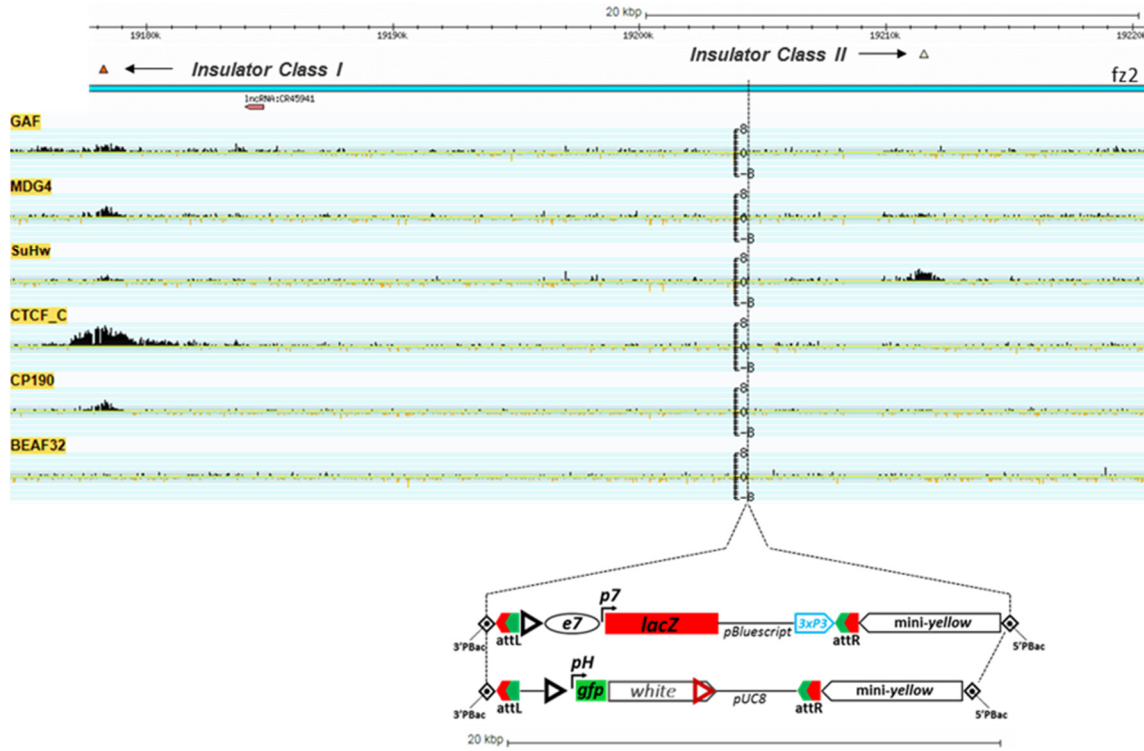
insulated *pH-gfp* in VK37 (**A, B and D'**), (2) ubiquitous activity in VNC and central brain of the uninsulated *pH-gfp* in VK2 and VK13 (**B and C**), (3) wing pouch activity of the uninsulated *p7-lacZ* in VK2 (**B**), (4) ubiquitous wing disk activity of uninsulated and GI-insulated *p7-lacZ*, and GI-insulated *pH* in VK13 (**C and C'**). Note GFP expression in notum AMPs, eye, antenna, optic lobes and VNC midline in **A'-E'**. All this *e7*-driven GFP expression is absent in **A-E**.

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? I tested constructs with a single 5' GI^{FOR} in four more *attP* loci and we got robust transvection in all cases (**Figure 43**). 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, I 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; I was able to detect some non-*e7* dependent patterns of expression of my transgenes in four out of the five loci tested (all except VK40; see **Figure 43**). Why nearby insulators are unable to support transvection in the GI-less constructs is not clear. Although the chromatin occupancy of many insulator-binding proteins has been described, only a fraction of these bound sites act as insulators in functional assays (Soshnev *et al.* 2008; Schwartz *et al.* 2012; Van Bortle *et al.* 2014). With this caveat in mind, the putative insulator landscape of each of the landing sites used is shown in **Figure 44**: the closest putative insulator could map anywhere from 1 kb (VK40) to 25 kb (VK37) away from my GI-less transgenic reporters and yet no transvection is observed.

VK2 2L:8,140,863 [-] (FlyBase, Dmel r6.26):



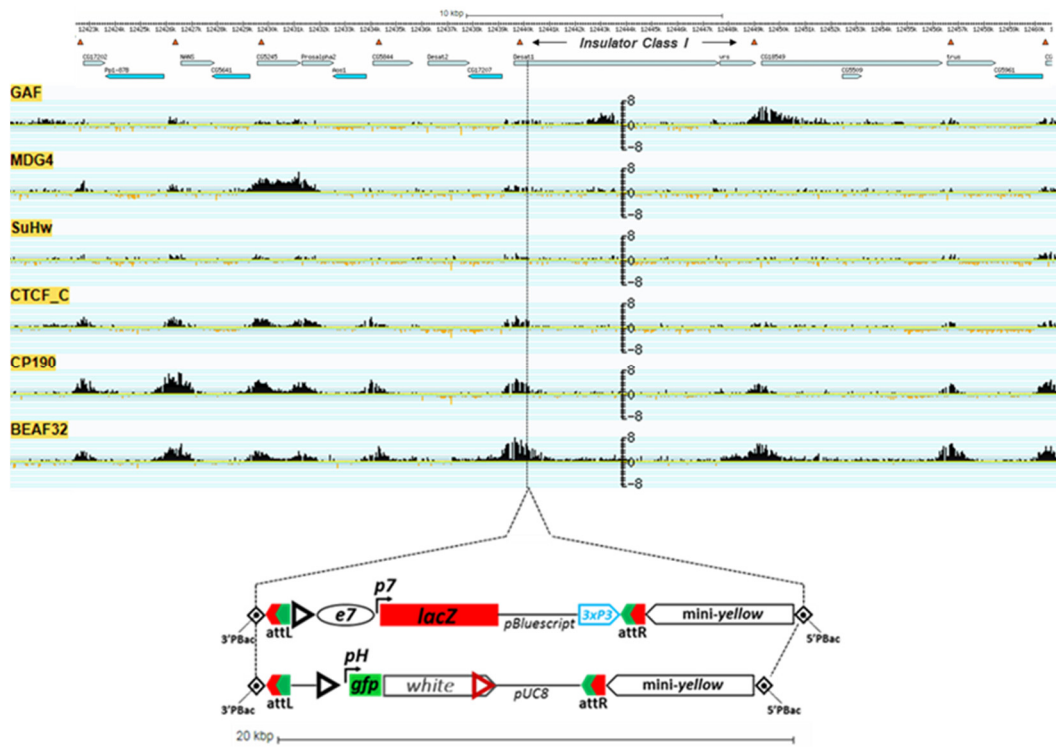
VK13 3L:19,204,358 [-] (FlyBase, Dmel r6.26):



VK37 2L:1,582,820 [+] (FlyBase, Dmel r6.26)



VK40 3R:12,440,193 [-] (FlyBase, Dmel r6.26)



attP40 2L:5,108,448 [-] (FlyBase, Dmel r6.26)

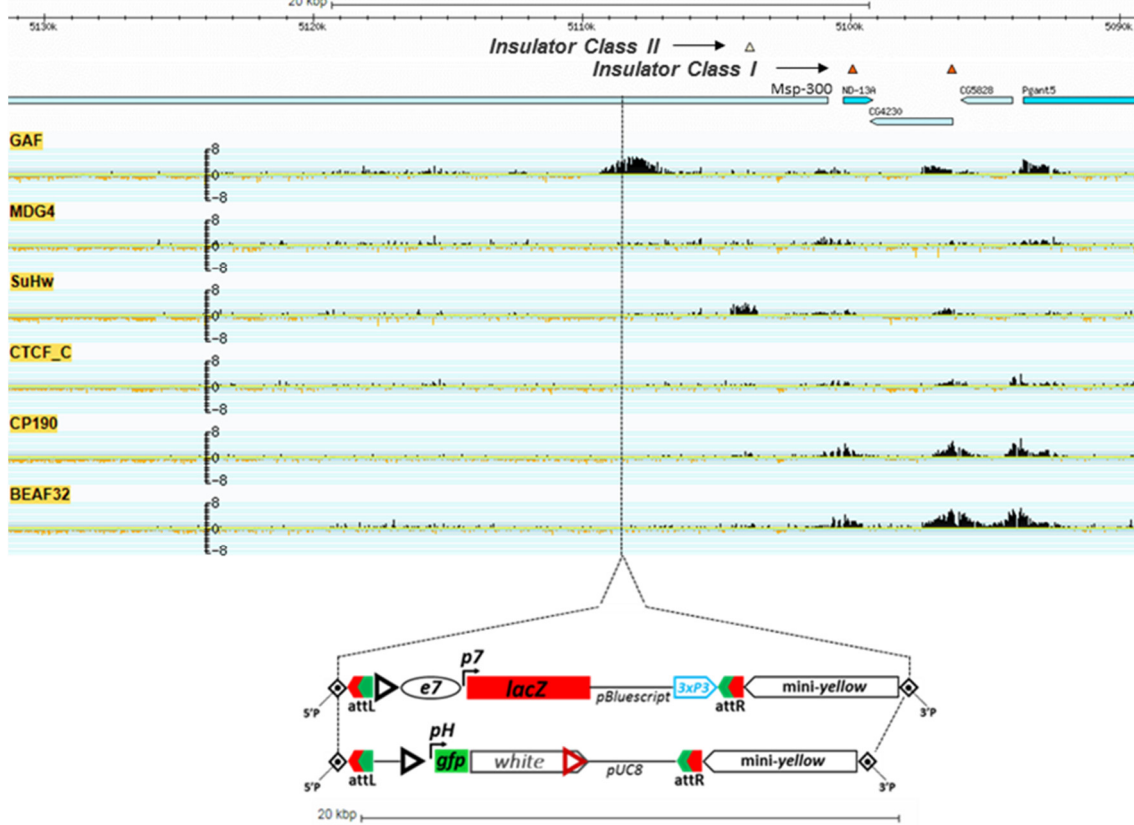


Figure 44 Genomic maps of transgene insertions used in Figure 43. The inserted transgenes are shown in the same scale as the genomic map. Note that only the GI-containing version of the two transgenes is shown. Our constructs are flanked by *attL* and *attR*, generated by ϕ C31-integrase-mediated recombination to the different *attP* landing sites. mini-yellow and the two *Piggybac* (3'PBac, 5'PBac) or *P-element* (5'P, 3'P) ends come from the landing site. The remaining symbols are the same used in all other construct schematics in this work. Information is given on the chromatin occupancy of six insulator binding proteins (Celniker *et al.* 2009) and on the position of putative insulators (Negre *et al.* 2010) in the neighborhood of the insertion sites. For VK37 and *attP40*, we have inverted the genomic map (sequence coordinates decreasing left to right) in order to depict the inserts in their correct orientation.

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 *e7p7-lacZ-3xP3* \rightarrow *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 *e7p7-lacZ* series of transgenes was tested in *trans* to an *e8p8-m8GFP* series (instead of the *pH-GFP*), the *e7p7* \rightarrow *e8p8* transvection largely obeyed the above rules, but the *e8p8* \rightarrow *e7p7* transvection was detectable even with a single 3' GI, regardless of orientation (**Figure 45**). Still, the single 5' GI^{FOR} configuration gave the strongest *trans*-effect even with this transgene combination. The removal of GI from either homologue completely abolished the effect, as expected.

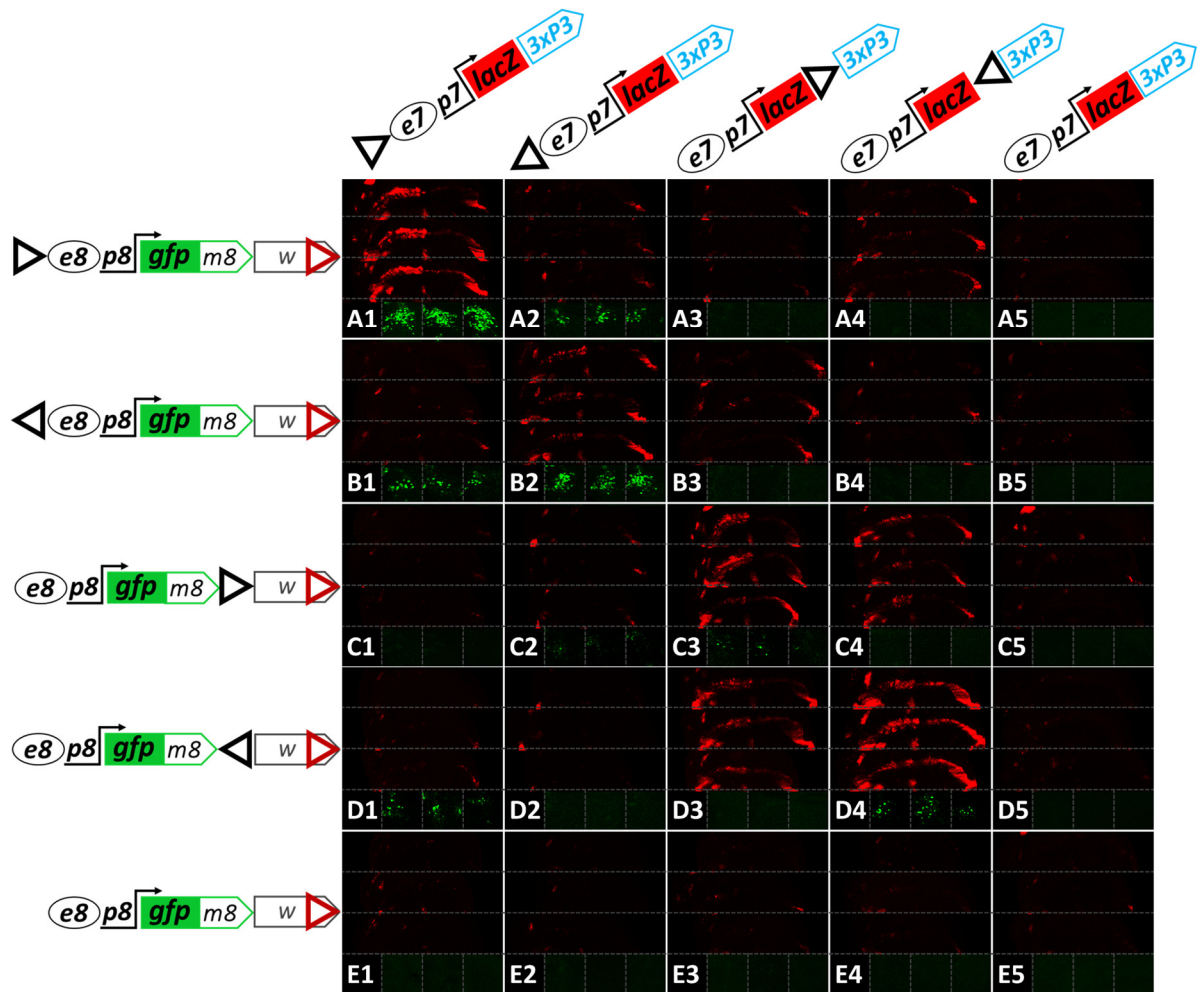


Figure 45 Less strict GI position and orientation requirements for interaction between *e7p7* and *e8p8*. (A1-E5) *e8p8-GFPm8* and *e7p7-lacZ* transgenes bearing a single GI are combined as shown in *attP40*. Each panel consists of six sub-panels containing confocal z-projections of the WM (only LacZ 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 and column 5). Note, that unlike *e7p7*→*pH* transvection (Figure 7 H), *e8p8*→*e7p7* transvection (LacZ in the WM) is also mediated by the 3' GIs (C3, C4, D3, D4). The 3' GIs also mediate weak *e7p7*→*e8p8* transvection (GFP in AMPs), when congruently oriented (C3, D4). Note, also, that the absolute orientation of 5' GIs is less important for *e7p7*→*e8p8* transvection (compare GFP in A1, A2, B1, B2) than for *e7p7*→*pH* (Figure 7 I, L).

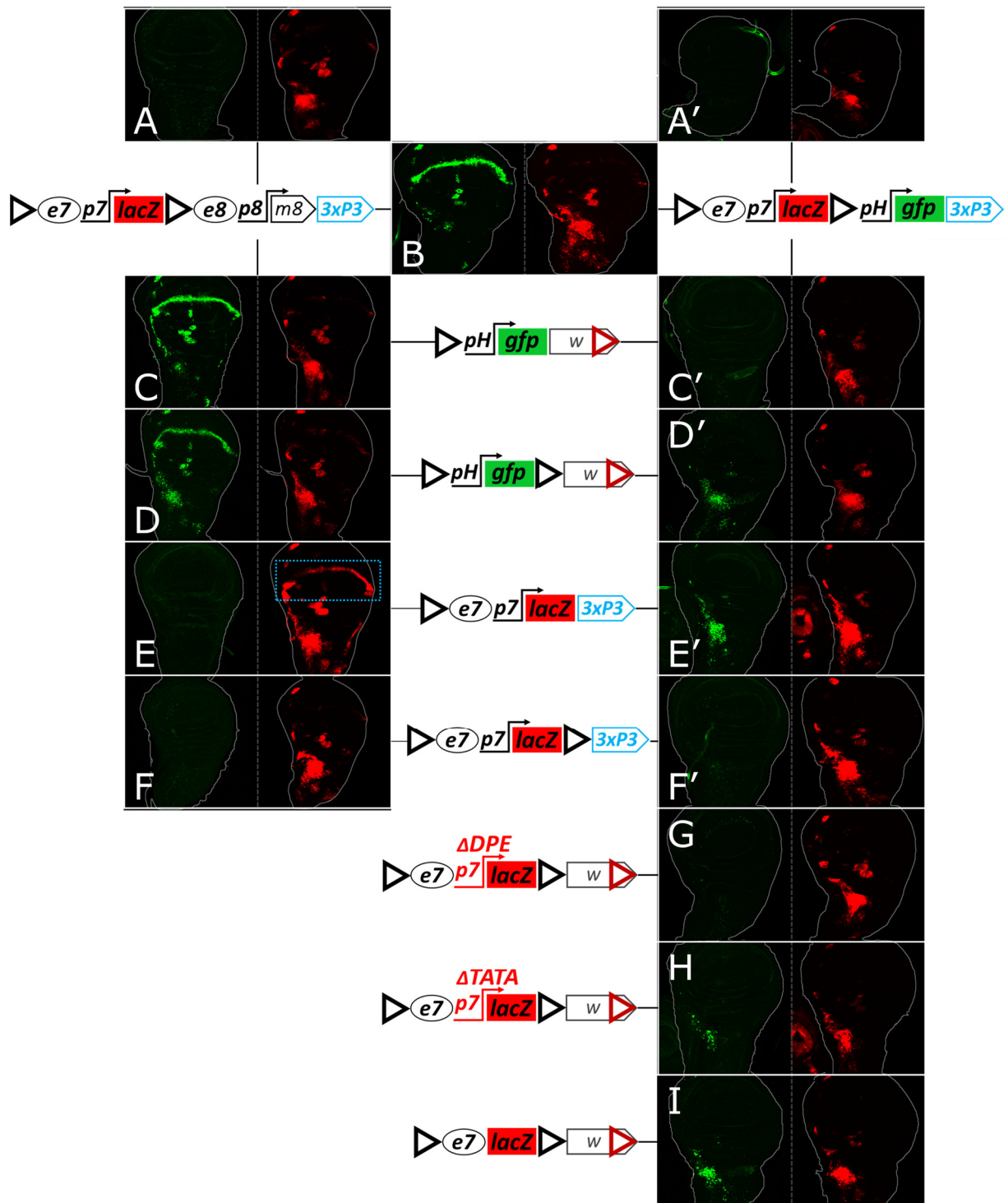


Figure 46 In dual GI transgenes, both 5' and 3' GIs participate in *trans*-interactions. Two tester transgenes, $GI^{FOR-e7p7-lacZ-GI^{FOR-e8p8-m8}}$ (left column, A-F), and $GI^{FOR-e7p7-lacZ-GI^{FOR-pH-gfp}}$ (right column, A'-F', G-I) were tested as hemizygotes (A, A'), combined *inter se* (B) or 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). Any GFP expression is caused by transvection, since the hemizygotes of all GFP constructs used express no GFP, other than a piece of trachea in A', probably due to enhancer trapping. LacZ expression in the AMPs may come from *cis* or *trans*, but WM LacZ expression, marked by a blue dotted rectangle in E, is triggered by the *e8* enhancer in *trans*.

Why was the addition of a 3' GI^{FOR} detrimental to *e7p7*→*pH* transvection (**Figure 42 A-D, M-P**)? Could it potentially engage in homotypic interactions with either the *cis* or the *trans* 5' GI that might compete with the ability of these 5' GIs to support transvection? To gain insight on the activity of this 3' GI^{FOR} I appended a "tester" module to our GI^{FOR}-*e7p7-lacZ*-GI^{FOR}-3xP3-*dsRed* transgene. Two different modules were cloned immediately 3' of the 3' GI^{FOR}: (1) a 'receiving' *pH-gfp* module or (2) a 'sending', *e8p8-m8* module (*e8p8* driving an untagged *E(spl)m8* CDS). **Figure 46** presents the results obtained in wing disks. The tester modules did not influence *e7p7-lacZ* expression: both transgenes, on their own, expressed LacZ in the AMPs, as expected (**Figure 46 A/A'**); an *e8*-specific wing margin LacZ pattern was not detected, consistent with insulation of *e7p7-lacZ* from *e8* (**Figure 46 A**); similarly, the transgene containing the insulated enhancerless *pH-gfp* module showed no GFP expression (**Figure 46 A'**), suggesting that the 3' GI in this construct insulates *e7p7* from *pH*, instead of enabling their interaction. It therefore seems that the two GIs^{FOR} in these tester constructs do not productively interact in *cis* (see also Cai and Shen 2001; Kyrchanova *et al.* 2008a). When tested against each other, I observed a robust activity of the *pH-gfp* module in the WM, indicative of a *trans* interaction between the two 3' GIs^{FOR} resulting in *e8p8-m8*→*pH-gfp* transvection (**Figure 46 B**). Therefore, 3' GIs prefer to homotypically interact in *trans*. The presence of only few GFP positive AMPs and no apparent WM LacZ expression in this combination (**Figure 46 B**) demonstrate that the "diagonal" interactions between GIs^{FOR} (i.e., 5' GIs^{FOR}-3' GI^{FOR}) are less favored than the "vertical" ones (5'-5' and 3'-3').

Placing the 5' GI^{FOR} *pH-gfp* transgene in *trans* to the testers gave strong transvection of *e8* (via the 3' GI^{FOR}; **Figure 46 C**) but only weak or no transvection of *e7* (via the 5' GI^{FOR}; **Figure 46 C, C'**). Interestingly, *e7*→*pH* transvection with both testers was augmented when a 3' GI^{FOR} was added to the responding *pH-GFP* transgene, as evidenced by broadly expressed GFP in the AMPs (**Figure 46 D, D'**). Therefore, when confronted with dual GIs^{FOR}, a single GI^{FOR} preferentially interacts with the *trans* 3' GI^{FOR}, but this shifts to the 5' GI^{FOR} when a second GI^{FOR} is added resulting in a dual/dual configuration. This conclusion was supported by confronting the tester transgenes with a single vs dual GI^{FOR} configured *e7p7-lacZ* reporter. Only the single 5' GI^{FOR} was able to interact in *trans* with the two tester modules, as evidenced by *e8p8-m8*→*e7p7-lacZ* and *e7p7-lacZ*→*pH-gfp* transvection (**Figure 46 E, E'**), whereas both of these transvection effects were lost in the dual/dual configuration (**Figure 46 F, F'**). Surprisingly, the *e7p7-lacZ*→*pH-gfp* transvection was regained when the *p7* promoter was compromised or (better) deleted in the dually GI^{FOR} flanked *e7p7-lacZ* transgene (**Figure 46 G-I**). I therefore conclude that GIs sample different homotypic interactions in *trans*, both vertical and diagonal. The latter are disfavored, but can still occur and their ability to support transvection is influenced by the enhancer-promoter interactions in their vicinity, in agreement with recent live imaging data that show that insulator-insulator interactions (both *cis* and *trans*) are stabilized when accompanied by productive enhancer-promoter interactions (Lim *et al.* 2018; Chen *et al.* 2018). For example, the *e7p7-pH* interaction is not sufficient to sustain diagonal transvection (**Figure 46 F'**), unless relieved from *cis* promoter preference (**Figure 46 H, I**). As another example, a similar diagonal GI-GI interaction can sustain

transvection of *e8* (from the *e8p8-m8* tester) to *pH-gfp* (**Figure 46 D**), but not to *e7p7-lacZ* (**Figure 46 F**). Such alternative *trans*-interactions in dual-GI combinations probably compete with the more favorable 5'GI - 5'GI interactions and could underlie the suppression of transvection produced by the addition of 3' GIs in **Figure 42 A-D**.

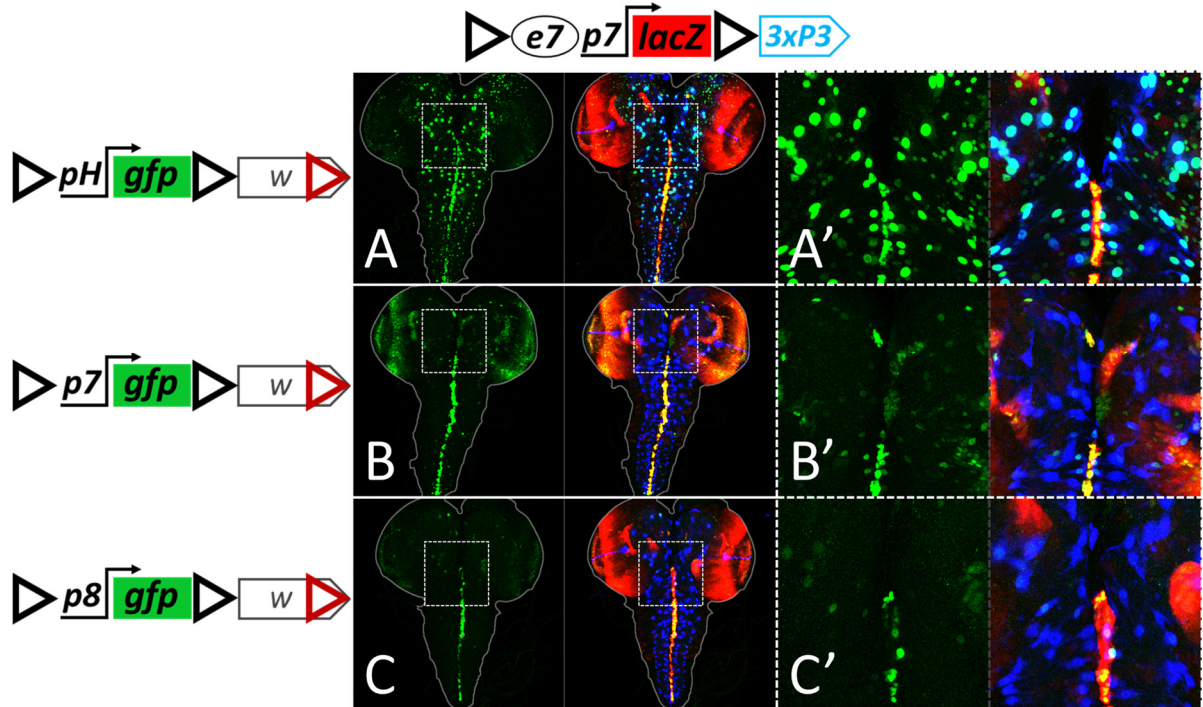


Figure 47 The 3xP3 enhancer has a stronger affinity for *pH* than for two other promoters. (A-C) Confocal z-projections of third instar CNSs showing GFP channel alone (green) and merged channels of GFP (green), LacZ (red) and DsRed (blue). (A'-C') Enlarged regions of the central brains indicated by white 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} *e7p7-lacZ-3xP3* transgene in *attP40*, giving a dotted glial GFP pattern (A, A'). 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).

Finally, I note that the *3xP3*→*pH* transvection in **Figure 42 A** is also mediated by a 5' – 3' (diagonal, less favored) GI^{FOR} interaction. I 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 47**). Therefore, when alternative GI-GI interactions are possible, they can be biased positively or negatively by the affinity that their nearby enhancer/promoter elements have for each other.

3.4.6. Other insulators also mediate transvection

To determine whether other *Drosophila* insulators also mediate transvection, I 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 the 367 bp GI, I tested two new insulators: (1) the 540 bp Fab8 insulator isolated from the *Abdominal-B* region of the *bithorax* complex (**Figure 40**); Fab8 insulator activity depends on dCTCF, the ortholog of the vertebrate CTCF protein (Barges *et al.* 2000; Moon *et al.* 2005; Kyrchanova *et al.* 2008a, 2016), and (2) the 454 bp 1A2 insulator located downstream of the *yellow* gene (**Figure 40**), containing two Su(Hw) binding sites (Golovnin *et al.* 2003; Parnell *et al.* 2003). Fab8 and 1A2 exemplify two major classes of endogenous insulators [centered around 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; Schwartz *et al.* 2012; 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 48 A1**), similar to the expression of the uninsulated *pH-gfp* reporter at the *attP40* locus (**Figure 32 E**). 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 48 A2**). This latter, ubiquitous, activity was abolished by a deletion of the WI from the mini-*white* (**Figure 48 A3**) 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 48 A2**), also trapped the epithelial enhancer via the *pD* promoter (**Figure 48 A4**), while the 1A2 responder did not show any activity from any of the promoters (**Figure 48 A5**).

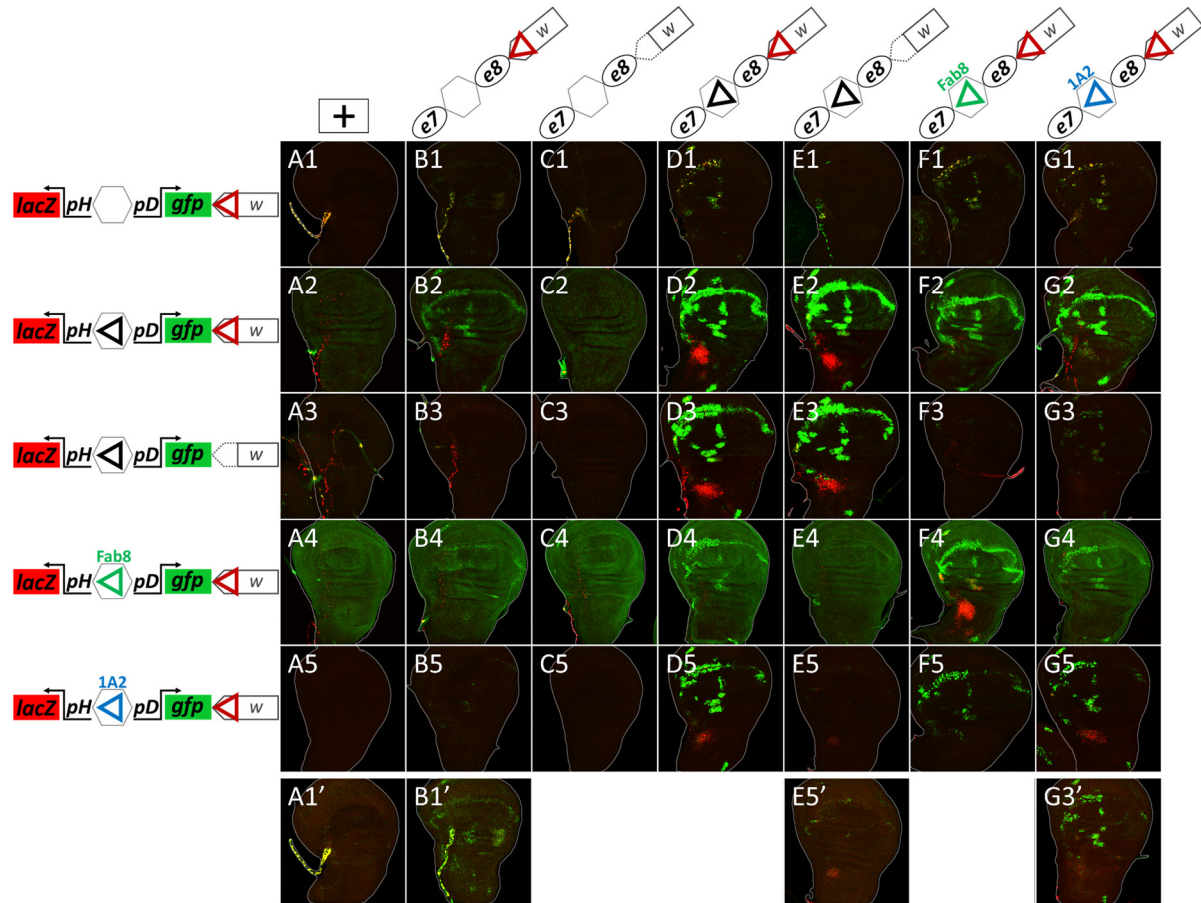


Figure 48 1A2 and Fab8 insulators also mediate transvection. All panels show merged GFP (green) and LacZ (red) channels of confocal z-projections from third instar wing disks. **(A1-A5)** Wing disks from animals hemizygous for responder transgenes; insulators are shown as triangles. **(B1-G5)** Wing 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** were 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 (red triangle).

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 48 B1, B1'**). This interaction was augmented in the GI responder (WM in **Figure 48 B2**). This enhancement of WI-WI mediated transvection by GI, was confirmed by deleting one WI, which abolished GFP expression in the WM (**Figure 48 B3**). This is in contrast to my previous result where GI-WI interaction in *cis* had an inhibitory effect on the WI-mediated transvection (**Figure 30**), 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 48 B4, B5**). Consistent with the conclusion that this transvection was mediated by homotypic WI/WI, when I deleted WI from the sender transgene, no transvection was observed at the WM in combination with any of the responders (**Figure 48 C1-C5**). WI-

mediated transvection was also weakly enhanced by adding a heterologous insulator in the sender homologue (**Figure 48 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 48 G1**, compare to **D1** and **F1**). Removal of WI from the GI sender abolished transvection (**Figure 48 E1**).

Unlike the weak WI-mediated transvection effects discussed so far with the "blank" senders (**Figure 48**, column B) or responders (**Figure 48**, row 1), I got very strong transvection of both *e7* and *e8*, when I combined homotypic insulators in sender and responder, i.e., GI/GI, Fab8/Fab8 or 1A2/1A2 (**Figure 48 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 48 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 48 F2, G2**) and the reciprocal, i.e., a GI sender with Fab8 or 1A2 responders (**Figure 48 D4, D5**). Upon deleting the WI from either the GI sender or responder, however, all of these effects were abolished (**Figure 48 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 48 G2 and D5**) was maintained, albeit much more weakly, upon WI deletion (**Figure 48 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.

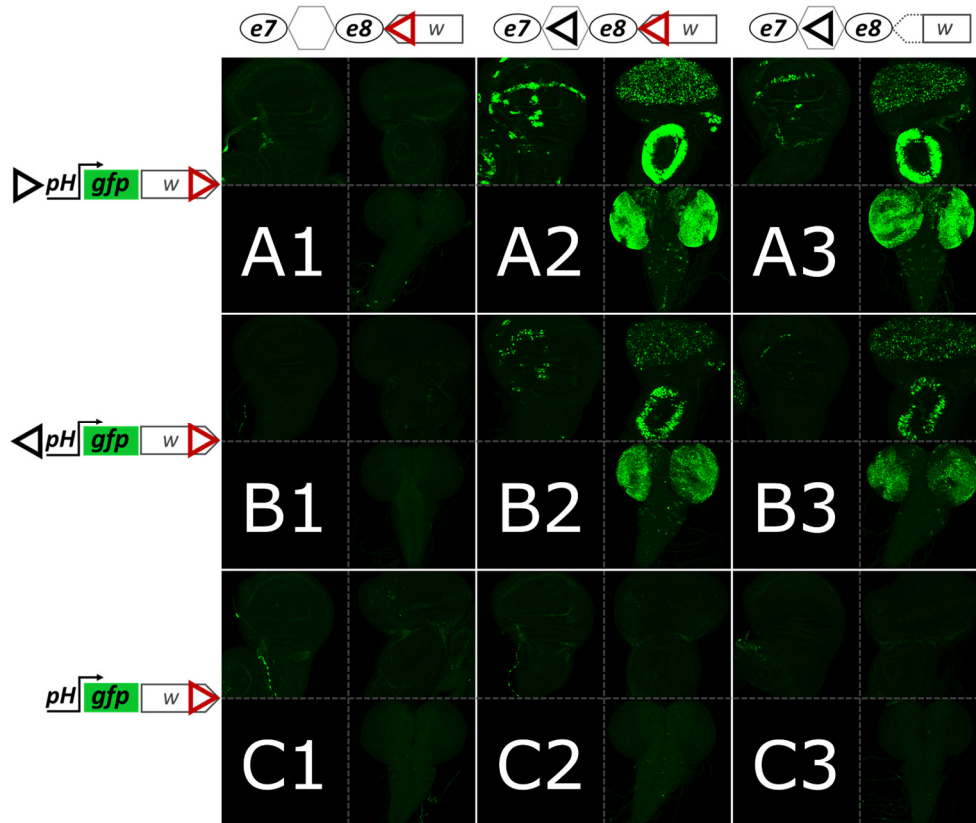


Figure 49 Relative orientation of GIs in *trans* impacts strength of transvection but not enhancer choice. All panels show merged GFP (green) channels of confocal z-projections from third instar wing disks, eye disks and CNSs for each genotype. Tissues were extracted from heterozygotes for blank and GI sender transgenes (indicated in each column) with 5' GI (FOR and REV) and GI-less *pH-gfp* transgenes (indicated in each row).

Finally, I asked whether the relative orientation of GIs in *trans* can determine which enhancer (from two located on both sides of a GI) will be transvected. For this purpose I used a GI-sender having *e7* on the 3' end and *e8* on the 5' end of a standalone GI^{REV} in combination with two *pH-gfp* transgenes having FOR or REV oriented 5' GI (**Figure 49 A2, B2**). If GIs interact in congruent fashion (head-to-head), one could expect that the *pH-gfp* will be activated by the enhancer that is on the same side of a GI as the *pH* promoter is. However, it is not the case. In both cases I observed robust expression of the *e8* enhancer, irrespectively of the relative orientation of the interacting GIs but weaker when the two GIs were incongruent (**Figure 49**, compare **B2** to **A2**). Very weak GFP expression in the VNC midline in the incongruent GI combination show (**Figure 49 A2**) that congruent GI interaction, although much weaker, also takes place in this heterozygote. Therefore, the relative orientation of GIs affects strength of transvection but not enhancer choice. This interaction necessitates GIs on both homologs (see **Figure 49 A1, B1, C1-C3**) and is not affected by WIs (**Figure 49**, compare **A2** to **A3** and **B2** to **B3**; note, however, that also in here, WI of the sender transgene augments the *e8-pH* interaction in WM of the wing disk).

In conclusion, I 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. I finally provide evidence that heterotypic insulators can weakly promote transvection if they belong to the same class.

3.5. Discussion

Transvection is the ability of an enhancer to activate transcription from an unlinked promoter located at the same locus of the homologous chromosome. Using a collection of enhancers and promoters driving GFP and LacZ reporters and targeted to specific genomic loci, I was able to study transvection and characterize parameters influencing its outcome (summarized in **Figure 50**).

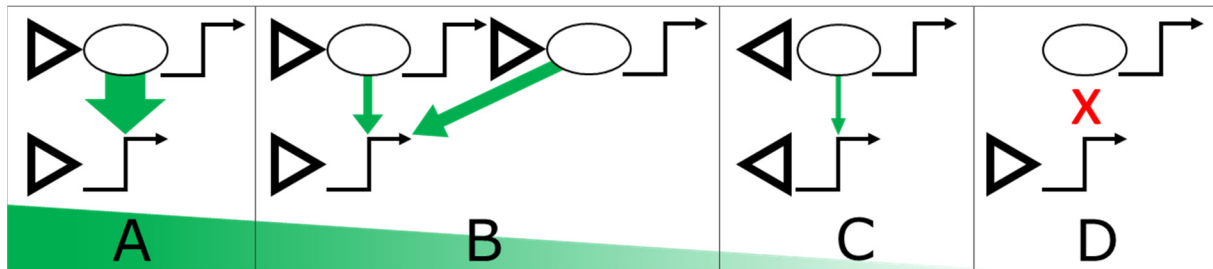


Figure 50 Parameters of transvection. A schematic summary of the parameters determining the ability of enhancers (ovals) to communicate with promoters (bent arrows) in *trans*. The top and bottom schematics in each panel depict elements present in the two paired homologs. Triangles represent *gypsy* insulators. A graded series of examples is shown with transvection ranging from high (A) to undetectable (D). For more details see Discussion.

The salient features of this phenomenon borne out by our results are the following: (1) I confirmed that homologue pairing is a prerequisite for transvection, as already known from classical studies. In *Drosophila*, homologue pairing in non-meiotic cells is widespread, but seems to evolve gradually during the first half of embryogenesis (Hiraoka *et al.* 1993; Fung *et al.* 1998; Gemkow *et al.* 1998) – I showed that transvection unfolds in a similarly gradual manner, being stochastic and erratic during the early stages of embryogenesis, while recapitulating precisely the *cis*-activity of the enhancer at later embryonic and larval stages, once homologue pairing has been completed. (2) Insulators are needed for transvection. At least four different insulators, GI, WI, Fab8 and 1A2, are capable of mediating transvection. I focussed on GI and WI, which are commonly found in transgenesis vectors. (3) GI-mediated transvection is strong, resulting in about 20% of the expression level the same enhancer-promoter combination would give in a *cis* configuration, and can work with all enhancer-promoter combinations tested. In contrast, WI-mediated transvection is weak and is only detectable if the responding promoter is accompanied by another enhancer (**Figure 34**) or another insulator (**Figure 48 B2**). (4) While necessary, the presence of GIs in both homologues

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

Based on these observations, one should be careful when planning to use heterozygous transgene combinations in the same landing site. If one wishes to minimize transvection, one should preferably not include GI or any other insulator in the transgenes, at the expense of losing shielding from position effects. If shielding is desired, GIs can be used in only one of the transgenes. If shielding of both transgenes is desired, GIs can be placed in different orientations and as far as possible from the transgenes' enhancers and promoters. If, on the other hand, one wishes to promote transgene transvection, one should place GIs in the forward orientation directly upstream of the transgenes' *cis*-regulatory elements.

A role of insulators in transvection has been described, but the mechanism been at best unclear (Fukaya and Levine 2017). Whereas some studies propose that insulators are needed for transvection (Lim *et al.* 2018), others propose that they only have an accessory role (Kravchenko *et al.* 2005; Schoborg *et al.* 2013), or that they affect transvection by promoting homologue pairing (Fujioka *et al.* 2016). And finally, other studies ignore them altogether (Bateman *et al.* 2012a; Mellert and Truman 2012; Blick *et al.* 2016). Instead, the classical proposed role of insulators is to insulate, i.e. to inhibit enhancer-promoter communication in *cis*, although sometimes they can enable such enhancer-promoter communication, e.g. via their so-called "bypass" activity (Cai and Shen 2001; Muravyova *et al.* 2001; Kyrchanova *et al.* 2008b; Fujioka *et al.* 2013). These two apparently contradictory activities have been reconciled by the "looping" model (reviewed in Chetverina *et al.* 2017;

Schwartz and Cavalli 2017), which is based on evidence that insulators mediate homotypic, or sometimes heterotypic, interactions (Kyrchanova *et al.* 2008b; Li *et al.* 2011; Vogelmann *et al.* 2014; Bonchuk *et al.* 2015; Fujioka *et al.* 2016). Via these interactions, insulators can form chromatin loops and these loops can either bring enhancers and promoters in proximity (e.g. when both are near the loop's anchor points) or avert their proximity (e.g. when one is within one loop and the other is outside that loop). These activities occur in *cis* and shape chromosomal architecture in parallel to affecting transcriptional regulation. The same insulator-insulator interactions can occur in *trans* (Kravchenko *et al.* 2005; Fujioka *et al.* 2016; Lim *et al.* 2018) and this could mediate interactions of enhancers on one homologue with promoters on the other (transvection).

One model proposes that insulators promote transvection by mediating homologue pairing in somatic cells. This hypothesis is supported by the fact that different classes of insulators are widely distributed in the *Drosophila* genome (Bartkuhn *et al.* 2009; Bushey *et al.* 2009; Negre *et al.* 2010; Schwartz *et al.* 2012) and their congruent matching could underlie paternal-maternal homologue alignment from end to end. Alternatively, insulators may not mediate homologue pairing *per se*, rather prior homologue pairing is a prerequisite for allowing insulators to mediate transvection. I favor the latter model: although I did not directly assay pairing, I have encountered numerous instances where addition of extra copies of the GI has a detrimental effect on transvection (**Figures 7 and 8**). This result would be hard to reconcile with a model where insulators promote pairing, as we would expect pairing, and thus transvection, to locally increase as more insulators are added. Consistent with the view that homologue pairing precedes transvection is the fact that screens designed to identify somatic homologue pairing factors did not reveal any of the numerous insulator binding proteins (Bateman *et al.* 2012b; Joyce *et al.* 2012). Moreover, a recent study imaged two homologously inserted transgenes in live embryos and found the same frequency of colocalization (pairing) whether an insulator was included or not (Lim *et al.* 2018) – yet transvection between these genes required an insulator on both homologues, in agreement with my results. Another recent study mapped DNA elements mediating pairing and transvection from the *ss* locus: the two activities were found to map on two different fragments (Viets *et al.* 2018).

If insulators do not mediate homolog pairing, their role could be to enable the productive interaction of enhancers with certain promoters. Several of my observations support such a more active role: (1) In order to promote *trans*-interaction between *e8* and either *pH* or *p7*, the WI requires the presence of *another enhancer* (*e7*) nearby. (2) Transvection supported by single 5' GIs is orientation-dependent: the FOR orientation is much more effective than the REV orientation (**Figure 42**); if insulator-insulator interactions were the only parameter influencing transvection, having congruently disposed insulators in both homologues would most likely produce an identical result, whether the configuration were FOR/FOR or REV/REV. (3) When two GIs are present in each homologue, they exhibit a strong bias for "vertical" trans GI/GI interactions (**Figure 46**). The fact that this bias can be alleviated

by promoter mutations is consistent with more direct insulator-promoter communication. (4) A forward GI exhibits a strong promoter preference: it transfects the 3xP3 enhancer only to *pH* and not to *p7* or *p8* in a certain transgene combination (**Figure 47**).

Recent data agree with such a more direct role of insulators in enhancer-promoter communication. A genome-wide chromatin occupancy analysis for more than 15 insulator binding proteins showed that a large proportion of their binding sites is near a promoter or an enhancer (Cubeñas-Potts *et al.* 2016). Direct contacts between insulators and nearby enhancers and promoters has been detected in transgenes *via* 3C (Kyrchanova *et al.* 2013). Live imaging of two loci separated (in *cis*) by >100kb has shown that homotypic insulators promote proximity between these 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-6* and *iab-7* 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 selectively mediating 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 which allows for transgene integration without vector sequences, making these instances of transvection harder to reconcile with the need for an insulator. One possible hypothesis would be that the inserted transgenes trapped nearby insulators. My GI-less transgenes were never able to trap nearby insulators, but I used five landing sites (*attP40*, *VK2*, *VK13*, *VK37*, *VK40*) distinct from those used in the above two studies (*JB53F* and *JB37B*), so my results cannot be compared. Given the strong dependence of transvection on insulator position, orientation and the nature of the interacting enhancers and promoters, it is likely that all of these factors will also influence insulator trapping. The large diversity of insulators in *Drosophila* (currently more than 15 binding factors identified; Maksimenko *et al.* 2015; Chetverina *et al.* 2017) is suggestive of a potentially high degree of selectivity in their interactions both with each other and with enhancer and promoter elements in their vicinity. Further work is needed to characterize these interactions and their importance in transcriptional regulation.

The association between insulators and transcriptional *cis*-regulatory elements (enhancers and promoters) is not a peculiarity of *Drosophila*; it has been reported also for vertebrates (Guo *et al.* 2015; reviewed in Hnisz *et al.* 2016). On the other hand, *Drosophila* (dipterans in general) seem to be unique in establishing somatic homologue pairing early in

development and maintaining it throughout life (Abed *et al.* 2018; Erceg *et al.* 2018). Transvection could be an epiphenomenon of these two biological processes: insulator interactions with enhancers/promoters and homologue pairing. This would explain why it is more often encountered in *Drosophila*, but is only sporadic in mammals (Apte and Meller 2012; Stratigi *et al.* 2015). Does transvection also serve a role in regulating transcriptional output and accordingly could it be positively selected in dipteran 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.

4. Materials and Methods

4.1. Plasmid constructs

Here I present the general strategy used for plasmid construction. Detailed information regarding each construct and maps I deposited in the laboratory as Vector NTI files and database in MS Excel. **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. All critical features in all the constructs used in this study were verified by sequencing.

To obtain my constructs I 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* and *pPelican* it is upstream of the 5' GI, see **A** and **B** in **Table 1**). *pBluescript* was additionally modified by cloning the *e7p7-lacZ* sequence downstream of *attB* (in the MCS) and *3xP3* upstream of *attB* (shown in **Table 1 C**). The latter construct was used as a base to generate all *3xP3*-containing plasmids.

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 1992b) into *pPelican-attB*. Due to strong post-transcriptional repression of both of these genes (Lai *et al.* 2005) I 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 *e7p7*) and 1.3 kb sequence upstream of the start codon of *GFPm8* (*e8p8*).

The GI-less *e8p8-GFPm8* construct (**No 5**) was generated by cloning of the *e8p8-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 *e8p8-GFPm8* by cloning the 5' GI from *pPelican* into 5' (upstream of *e8p8*) or 3' (downstream of *Adh* 3' UTR) positions in forward or reverse orientation (constructs **No 6-9**).

The *e8p8-lacZ* *pPelican*-based construct (**No 10**) was made by cloning PCR-amplified *e8p8* (primers **e8p8F** and **e8p8R**) into the MCS of *pPelican-attB*.

The *e7p7-lacZ* *pPelican*-based construct (**No 11**) was made by cloning PCR-amplified *e7p7* (primers **e7p7F** and **e7p7R**) into the MCS of *pPelican*. The *pBluescript* GI-less version of *e7p7-lacZ* (**No 12**, see also **C** in **Table 1**) was made by (1) cloning the entire *e7p7-lacZ* sequence of *pPelican*-based *e7p7-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 *e7p7*. The *pBluescript* GI-less version of the *e7p7-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 *e7p7-lacZ-3xP3-dsRed* (**No 12**) construct

was used as a basis to generate (1) GI-containing *e7p7-lacZ* constructs by cloning 5' GI from *pPelican* immediately upstream of *e7p7* or/and downstream of *SV40 3' UTR* in a reverse or forward orientation (constructs **No 14-20**); (2) the three *e7p7-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* (*e7p7-lacZ-w5'*, construct **No 21**), the 2.4 kb *AflIII/EcoRV* fragment encompassing mini-*white*'s gene body (*e7p7-lacZ-wB*, construct **No 22**) and the 0.9 kb *EcoRV/BsrGI* 3' part of the mini-*white* (*e7p7-lacZ-w3'*, construct **No 23**). The *pBluescript e7p7-lacZ* construct with two GIs in forward orientation (**No 14**) was used to generate two 'tester' constructs by cloning *e8p8-m8-Adh 3' UTR* module from the EGFP-untagged version of the *GFPm8* construct (*e8p8-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 *e7p7-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 *e7p7-lacZ* construct (**No 11**) and subsequent excision of the sequence between *EcoRV* and *BstEII*. The intermediate *e7p7-lacZ* construct with introduced *EcoRV* (but without *EcoRV/BstEII* deletion; **No 27**) produced LacZ expression pattern and levels indistinguishable to that of the *e7p7-lacZ* without *EcoRV* and was used as a control transgene (to *e7p7-ΔTATA-lacZ* and *e7p7-ΔDPE-lacZ*) containing wild-type *p7*. The *e7p7-ΔDPE-lacZ* (**No 28**) construct was generated from *pPelican*-based *e7p7-lacZ* by excision of 83 nt containing INR and DPE motifs in between *BstEII* and *StuI* restriction sites. The *e7-lacZ* (promoterless construct, **No 29**) was generated by an excision of sequence between *EcoRV* and *StuI* from the *e7p7-lacZ* construct with the introduced *EcoRV* site (**No 27**).

All *pH-gfp* constructs (**No 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 (**No 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 **No 31, 32** and **33**, respectively. The 5' GI^{REV} *pH-gfp* (**No 34**) was made by cloning GI in reverse orientation into GI-less *pH-gfp*. The 5' GI^{FOR} 3' GI^{REV} *pH-gfp* (**No 35**) was made by cloning GI in a reverse orientation in place of the 3' GI^{FOR} of GIs^{FOR} *pH-gfp* (**No 30**). The 5' GI^{REV} 3' GI^{FOR} *pH-gfp* (**No 36**) was made by cloning GI in a reverse orientation in place of the 5' GI^{FOR} of the GIs^{FOR} *pH-gfp* (**No 30**).

The 2 kb *BglIII-EcoRV e7*-containing fragment derived from the *e7p7-lacZ* construct bearing *EcoRV* site introduced upstream to TATA box of *p7* (**No 27**) was used to generate (1) the *e7* construct (**No 37**) by replacing *lacZ* in the *pPelican-attB* and (2) the *e7pH-gfp* (**No 38**) by cloning it upstream to *pH* of the GIs^{FOR} *pH-gfp* (**No 30**). The *e7p7-gfp* construct (**No 39**) was made by PCR-amplifying the *p7* promoter from *e7p7-lacZ* (**No 11**) with **p7F** and **p7R** primers and ligating the *NheI/HeaIII*-digested PCR product into *NheI* and *StuI* sites in the *e7pH-gfp* construct (**No 38**) such that *p7* sequence replaces the sequence of *pH*. The *p7-gfp* (**No 40**) was based on the *e7p7-gfp* (**No 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 *e8p8-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 *e8p8-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 GIs^{FOR} *pH-gfp* (**No 30**), GIs^{FOR} *e7pH-gfp* (**No 38**) and GI-less *pH-gfp* (**No 33**) constructs – resulting in generation of GIs^{FOR} *pH-luc* (**No 46**), GIs^{FOR} *e7pH-luc* (**No 47**) and GI-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 (**D** in **Table 1**) replacing *lacZ* sequence and (2) subsequent cloning of the *NheI/BglIII*-cut PCR amplicon of *e8* (primers **e8-2F** and **e8-2R**). The 'blank' sender construct (**No 49**) was used to generate GI, Fab8 and 1A2 sender constructs (**No 50, 51** and **52**, respectively) by cloning the PCR-amplified sequences of GI, Fab8 or 1A2 into the *BglIII* site located in between *e7* and *e8*. The 400 bp GI sequence was amplified on the *pPelican* template with **GIF** and **GIR** primers and contains full-length 367 bp GI (see **Figure**

S7). The 540 bp Fab8 sequence was PCR-amplified from *Drosophila* genome using **Fab8F** and **Fab8R** primers and contains the F8²⁵⁴ sequence and part of F8⁴⁶⁹ sequence (as defined in Kyrchanova *et al.* 2008; see **Figure S7**). The 420 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 (see **Figure S7**). The WI-less sender constructs, **No 53** and **54**, were generated from the 'blank' sender (**No 49**) and the *Gl* 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 *Gl* and *Fab8* responder constructs (**No 56** and **57**, respectively) by cloning into the *BglII* site the sequences of *Gl* 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 *Gl* responder (**No 59**) was generated by a deletion of the 502 bp WI-containing *BstBI/NsiI* fragment from the *Gl* 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 1992b) into the MCS of *lacZ*-deficient *pPelican-attB*.

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 1992b) 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**).

e7p7^{B1}-lacZ (**No 75**) is based on *e7p7-lacZ* (**No 11**) where B1-box of *e7* was mutated into *EcoRV* site using primer **B1-XC-EcoRV**. *e7^{dl}p7-lacZ* (**No 76**) was obtained by deletion of *EcoRI-XhoI* (and treated with Klenow fragment) proximal *e7* fragment from *e7p7-lacZ* (**No 11**). *e7^{px}p7-lacZ* (**No 77**) was obtained by deletion of *SphI-EcoRI* (and treated with Klenow fragment) distal *e7* fragment from *e7p7-lacZ* (**No 11**). *e7p7ΔCTCF-lacZ* (**No 78**) was obtained by deletion of *BstBI-EcoRI* (and treated with Klenow fragment) CTCF-binding region of *e7 e7p7-lacZ* (**No 11**). *e8p8^{B2C1}-lacZ* (**No 79**), *e8p8^{B2C1C2}-lacZ* (**No 80**) and *e8p8^{C2}-lacZ* (**No 82**) are based on *e8p8-lacZ* (**No 10**) where B2- and C1-boxes were mutated using primers **B2-XC-EcoRV** and **C1-XC-XhoI** to obtain construct **No 79**, C2-box was mutated using primer **C2-XC-XmaI** to obtain construct **No 82**, and all of above primers were used to obtain construct **No 80**. *e8^{px}p8-lacZ* (**No 80**) was created by deleting proximal fragment of *e8* with *BglII* from *e8p8-lacZ* (**No 10**).

T^{GFP-wt} (**No 83**) was made by cloning 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 1992b) into *pPelican-attB*, and tagging *E(spl)m7* with EGFP similar as it was done with construct **No 1**. Subsequently, the latter construct was used to generate *T^{GFP-mi}* (**No 84**) by mutating the putative microRNA binding sites in the 3' UTRs of *E(spl)m7* and *E(spl)m8* with mutagenic primers **BANT73wtSacN**, **m7-289-H3**, **m7-219-Nae**, **m7-8-Aat**, **m7KXho**, **m7BrdSmaI** and **m8KKmut**. *T^{GFP-BC}* (**No 85**) was made also based on *T^{GFP-wt}* (**No 83**) where all B/C-boxes were mutated (primers **B1-XC-EcoRV**, **B2-XC-EcoRV**, **C1-XC-XhoI** and **C2-XC-XmaI**). *T^{GFP-BC+mi}* (**No 86**) construct combines all microRNA-binding boxes and B/C-boxes point mutations. *T^{GFP-BC+3UTR}* (**No 86**) is made based on construct **No 1**, where the mutation within B/C-boxes were introduced (same as in **No 83**). *Short T^{GFP-wt}* and *short T^{GFP-B1}* were obtained by cloning *XhoI* of construct **No 83** and **No 85** (respectively) containing *e7p7-EGFPm7-3' UTR* unit into *dPelican-attB*.

T^{wt} (**No 91**) is untagged version of construct **No 83** and its derivatives (*T^{BC}* - **No 92**, *T^{mi}* - **No 93**, *T^{BC+mi}* - **No 94**) were generated in an analogous fashion as similar derivatives of construct **No 83**. *T^{3UTR}* (**No 95**) is untagged version of construct **No 1** and the *T^{BC+3UTR}* (**No 95**) is derived from it by mutating all B/C-boxes (as in **No 85**).

To obtain sensor *m7^{wt}* (**No 97**), 1014bp *E(spl)m7* 3'UTR was PCR amplified with **m73'UTRforXba** and **m73'UTRrevXba** and cloned into *XbaI* of construct **E**. Subsequently, the sensor *m7^{wt}* (**No 97**) was used to generate


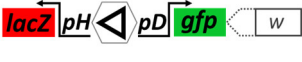















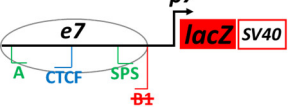
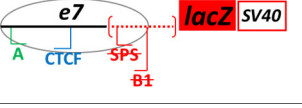
sensor $m7^{Brd,K}$ (№ 98, by mutating Brd and K boxes with **m7BrdSmaI** and **m7KXho**), sensor $m7^{bant}$ (№ 99, by mutating bantam box with **BANT73wtSacN**), sensor $m7^{Brd,K,b}$ (№ 100, by using primers used to generate constructs № 98 and № 99), sensor $m7^{ALL}$ (№ 101, by using primers used to generate constructs №100 and primers **m7-289-H3**, **m7-219-Nae** and **m7-8-Aat**). To obtain sensor $m8^{wt}$ (№ 103), 854 bp *E(spl)m8* 3'UTR was PCR amplified with **m83'UTRforXba** and **m83'UTRrevXho** and cloned into *XbaI-XhoI* of construct E. Subsequently, the sensor $m8^{wt}$ (№ 103) was used to generate sensor $m8^{K1}$ (№ 103, by mutating K1 box with primer **M8K1**), sensor $m8^{K2}$ (№ 104, by mutating K2 box with primer **M8K2**) and sensor $m8^{KK}$ (№ 107, by mutating both K boxes with primer **m8KKmut**).

Table 1 List of constructs used in the study

№	schematic	Name (as in text)	Marker gene	Insulators	Backbone
A		<i>pStinger-attB</i>	mini-white	WI, Gl ^{FOR}	<i>pStinger</i>
B		<i>pPelican-attB</i>	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
C		<i>e7p7-lacZ-attB</i>	3xP3	-	<i>pBluescript</i>
D		<i>pLacZ-attB</i>	mini-white	WI	<i>pLacZ-attB</i>
E		<i>mRFP sensor</i>	mini-white	WI	<i>pPelican</i>
1		<i>GFPm7-m8</i>	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
2		<i>m7-GFPm8</i>	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
3		<i>GFP-E(spl)m7</i>	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
4		<i>GFP-E(spl)m8</i>	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
5		<i>e8p8-GFPm8</i>	mini-white	WI	<i>pPelican</i>
6		<i>e8p8-GFPm8</i>	mini-white	WI, 5' Gl ^{FOR}	<i>pPelican</i>
7		<i>e8p8-GFPm8</i>	mini-white	WI, 5' Gl ^{REV}	<i>pPelican</i>
8		<i>e8p8-GFPm8</i>	mini-white	WI, 3' Gl ^{FOR}	<i>pPelican</i>
9		<i>e8p8-GFPm8</i>	mini-white	WI, 3' Gl ^{REV}	<i>pPelican</i>
10		<i>e8p8-lacZ</i>	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
11		<i>e7p7-lacZ</i>	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
12		<i>e7p7-lacZ</i>	3xP3-dsRed	-	<i>pBluescript</i>

13		<i>e7p7-lacZ</i>	mini-white	WI	<i>pPelican</i>
14		<i>e7p7-lacZ</i>	<i>3xP3-dsRed</i>	GIs ^{FOR}	<i>pBluescript</i>
15		<i>e7p7-lacZ</i>	<i>3xP3-dsRed</i>	5' GI ^{FOR}	<i>pBluescript</i>
16		<i>e7p7-lacZ</i>	<i>3xP3-dsRed</i>	5' GI ^{REV}	<i>pBluescript</i>
17		<i>e7p7-lacZ</i>	<i>3xP3-dsRed</i>	3' GI ^{FOR}	<i>pBluescript</i>
18		<i>e7p7-lacZ</i>	<i>3xP3-dsRed</i>	3' GI ^{REV}	<i>pBluescript</i>
19		<i>e7p7-lacZ</i>	<i>3xP3-dsRed</i>	5' GI ^{FOR} , 3' GI ^{REV}	<i>pBluescript</i>
20		<i>e7p7-lacZ</i>	<i>3xP3-dsRed</i>	5' GI ^{REV} , 3' GI ^{FOR}	<i>pBluescript</i>
21		<i>e7p7-lacZ-w5'</i>	<i>3xP3-dsRed</i>	-	<i>pBluescript</i>
22		<i>e7p7-lacZ-wB</i>	<i>3xP3-dsRed</i>	-	<i>pBluescript</i>
23		<i>e7p7-lacZ-w3'</i>	<i>3xP3-dsRed</i>	WI	<i>pBluescript</i>
24		<i>e8p8-m8 tester</i>	<i>3xP3-dsRed</i>	GIs ^{FOR}	<i>pBluescript</i>
25		<i>pH-gfp tester</i>	<i>3xP3-dsRed</i>	GIs ^{FOR}	<i>pBluescript</i>
26		<i>e7p7-ΔTATA-lacZ</i>	mini-white	WI, GIs ^{FOR}	<i>pPelican</i>
27		<i>e7p7-lacZ</i>	mini-white	WI, GIs ^{FOR}	<i>pPelican</i>
28		<i>e7p7-ΔDPE-lacZ</i>	mini-white	WI, GIs ^{FOR}	<i>pPelican</i>
29		<i>e7-lacZ</i>	mini-white	WI, GIs ^{FOR}	<i>pPelican</i>
30		<i>pH-gfp</i>	mini-white	WI, GIs ^{FOR}	<i>pStinger</i>
31		<i>pH-gfp</i>	mini-white	WI, 5' GI ^{FOR}	<i>pStinger</i>
32		<i>pH-gfp</i>	mini-white	WI, 3' GI ^{FOR}	<i>pStinger</i>
33		<i>pH-gfp</i>	mini-white	WI	<i>pStinger</i>
34		<i>pH-gfp</i>	mini-white	WI, 5' GI ^{REV}	<i>pStinger</i>

35		<i>pH-gfp</i>	mini-white	WI, 5' GI ^{FOR} , 3' GI ^{REV}	<i>pStinger</i>
36		<i>pH-gfp</i>	mini-white	WI, 5' GI ^{REV} , 3' GI ^{FOR}	<i>pStinger</i>
37		<i>e7</i>	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
38		<i>e7pH-gfp</i>	mini-white	WI, Gl ^{FOR}	<i>pStinger</i>
39		<i>e7p7-gfp</i>	mini-white	WI, Gl ^{FOR}	<i>pStinger</i>
40		<i>p7-gfp</i>	mini-white	WI, Gl ^{FOR}	<i>pStinger</i>
41		<i>e8</i>	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
42		<i>e8p7-gfp</i>	mini-white	WI, Gl ^{FOR}	<i>pStinger</i>
43		<i>e7p8-gfp</i>	mini-white	WI, Gl ^{FOR}	<i>pStinger</i>
44		<i>e8p8-gfp</i>	mini-white	WI, Gl ^{FOR}	<i>pStinger</i>
45		<i>p8-gfp</i>	mini-white	WI, Gl ^{FOR}	<i>pStinger</i>
46		<i>pH-luc</i>	mini-white	WI, Gl ^{FOR}	<i>pStinger</i>
47		<i>e7pH-luc</i>	mini-white	WI, Gl ^{FOR}	<i>pStinger</i>
48		<i>pH-luc</i>	mini-white	WI	<i>pStinger</i>
49		'blank' sender	mini-white	WI	<i>pLacZ-attB</i>
50		GI sender	mini-white	WI, GI ^{REV}	<i>pLacZ-attB</i>
51		<i>Fab8</i> sender	mini-white	WI, Fab8	<i>pLacZ-attB</i>
52		1A2 sender	mini-white	WI, 1A2	<i>pLacZ-attB</i>
53		'blank' sender	mini-white-dWI	-	<i>pLacZ-attB</i>
54		GI sender	mini-white-dWI	GI ^{REV}	<i>pLacZ-attB</i>
55		'blank' responder	mini-white	WI	<i>pLacZ-attB</i>
56		GI responder	mini-white	WI, GI ^{REV}	<i>pLacZ-attB</i>
57		<i>Fab8</i> responder	mini-white	WI, Fab8	<i>pLacZ-attB</i>

58		1A2 responder	mini-white	WI, 1A2	<i>pLacZ-attB</i>
59		GI responder	mini-white-dWI	GI ^{REV}	<i>pLacZ-attB</i>
60		HH-1.5	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
61		HH-2.1	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
62		HH-5.4	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
63		HB-1.6	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
64		BB-1.8	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
65		BB-3.1	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
66		BB-2.3	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
67		BB-5.4	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
68		BB-1.3	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
69		BB-7.9	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
70		BB-1.7	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
71		BB-07	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
72		BP-3.2	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
73		PB-4.7	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
74		VGQ	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
75		<i>e7p7^{B1}-lacZ</i>	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>
76		<i>e7^{dl}p7-lacZ</i>	mini-white	WI, Gl ^{FOR}	<i>pPelican</i>

77		$e7^{px}p7-lacZ$	mini-white	WI, Gl ^{FOR}	pPelican
78		$e7p7\Delta CTCT-lacZ$	mini-white	WI, Gl ^{FOR}	pPelican
79		$e8p8^{B2C1}-lacZ$	mini-white	WI, Gl ^{FOR}	pPelican
80		$e8p8^{B2C1C2}-lacZ$	mini-white	WI, Gl ^{FOR}	pPelican
81		$e8^{px}p8-lacZ$	mini-white	WI, Gl ^{FOR}	pPelican
82		$e8p^{C2}-lacZ$	mini-white	WI, Gl ^{FOR}	pPelican
83		T^{GFP-wt}	mini-white	WI, Gl ^{FOR}	pPelican
84		T^{GFP-mi}	mini-white	WI, Gl ^{FOR}	pPelican
85		T^{GFP-BC}	mini-white	WI, Gl ^{FOR}	pPelican
86		$T^{GFP-BC+mi}$	mini-white	WI, Gl ^{FOR}	pPelican
87		$T^{GFP-BC+3UTR}$	mini-white	WI, Gl ^{FOR}	pPelican
88		short T^{GFP-wt}	mini-white	WI, Gl ^{FOR}	pPelican
89		short T^{GFP-B1}	mini-white	WI, Gl ^{FOR}	pPelican
90		$T^{GFP-B2C1C2}$	mini-white	WI, Gl ^{FOR}	pPelican
91		T^{wt}	mini-white	WI, Gl ^{FOR}	pPelican
92		T^{BC}	mini-white	WI, Gl ^{FOR}	pPelican
93		T^{mi}	mini-white	WI, Gl ^{FOR}	pPelican
94		T^{BC+mi}	mini-white	WI, Gl ^{FOR}	pPelican
95		T^{3UTR}	mini-white	WI, Gl ^{FOR}	pPelican

96		$T^{BC+3UTR}$	mini-white	WI, Gl ^{FOR}	pPelican
97		sensor m7 ^{wt}	mini-white	WI, Gl ^{FOR}	pPelican
98		sensor m7 ^{Brd,K}	mini-white	WI, Gl ^{FOR}	pPelican
99		sensor m7 ^{bant}	mini-white	WI, Gl ^{FOR}	pPelican
100		sensor m7 ^{Brd,K,b}	mini-white	WI, Gl ^{FOR}	pPelican
101		sensor m7 ^{ALL}	mini-white	WI, Gl ^{FOR}	pPelican
102		sensor SV40	mini-white	WI, Gl ^{FOR}	pPelican
103		sensor m8 ^{wt}	mini-white	WI, Gl ^{FOR}	pPelican
104		sensor m8 ^{K1}	mini-white	WI, Gl ^{FOR}	pPelican
105		sensor m8 ^{K2}	mini-white	WI, Gl ^{FOR}	pPelican
106		sensor m8 ^{ADH}	mini-white	WI, Gl ^{FOR}	pPelican
107		sensor m8 ^{KK}	mini-white	WI, Gl ^{FOR}	pPelican
108		sensor ADH	mini-white	WI, Gl ^{FOR}	pPelican

Table 2 List of primers

Name	Sequence (5' to 3')
e8p8F	CGTCTAGAGGGGAATCTATTTTACAGCACAATCCAATAGGGG
e8p8R	GCGGTACCCGGCTTGTGCTGCCTGCTCG
e7p7F	CGGCATGCGTCGCCAGAAAAATTGTAACGGCCC
e7p7R	GCCTCGAGGAACCTTCTTCGATCTTTCGGAGGAGG
p7F	GAGGCTAGCAGCTATAAAAGCAGCGGTAACC
L5R	TCAGACGATTCATTGGCACC
e8F	CGGCATGCGGGGAATCTATTTTACAGCACAATCCAATAGGGG
e8R	CTGGCTAGCTCCCTGGTCCCTGAAATCC
p8F	GAGGCTAGCGGTATAAAAGGACGGGACCTC
e7-2F	GGGAAGCTTGTGCGCCAGAAAAATTGTAACGG
e7-2R	GCGGATCCGTGCCGCCGCGAGAG
e8-2F	CGTCTAGAGGGGAATCTATTTTACAGCACAATCCAATAGGGG
e8-2R	CCCGCTAGCCGCTCCCTGGTCCCTGAAATCC
GIF	GGAGATCTGCATCACGTAATAAGTGTGCGT
GIR	GGAGATCTGCCGAGCACAATTGATCG
Fab8F	GGAGATCTGGGGAGGGAATTTTCTTCA
Fab8R	GGAGATCTCATCTTCCGTTTCATCCGTT
1A2F1	GGCGGATCCACTACCAGGCAAGAAAGTAGGT

1A2F2	GGTGCGAATTCCTACTACCAGGCAAGAAAGTA
1A2R	GGCGGATCCTATATGCTTCGTCTACCGTTGTG
B1-XC-EcoRV	GCTCTCGCGGGCGGGATATCCAGCTATAAAAGCAGC
B2-XC-EcoRV	GTGCTAGCTTATTTAGGATATCTGTAAAAAATGGACTTCCC
C1-XC-XhoI	CATACAAATGAGTTTCAGCTCGAGTCTCGATTACTTATC
C2-XC-XmaI	CGAGTGTGGCCCGGGGGGCACACTTTC
BANT73wtSacN	CCGCAGCTTTAGAAAAAATGCTTTACAATGAGCTCTCTTTATTTGTTACGTTAC C
m7-289-H3	GGAAGAAGATTCCATTGAAAGCTTTTAATGTTGGTATGCGCG
m7-219-Nae	CCCGACGGGCCGGGCACGCAATTGGAACC
m7-8-Aat	GGGGATATTGAACTTCAGGCCTTGTCTTTAAATTCC
m7KXho	GTTTTCTAAGTTAAGTCCTCGAGAGGGCACGCTTTATTAACTTTTGTC
m7BrdSmaI	CGCAATTGGAACCGCAGCCCCGGGAAAAAATGCTTTACAATGATCC
m8KKmut	CCAACCAACAACGCATCCTCGAGAGCCCAACCCGGGGGCAACAAAATAGC
m73UTRforXba	GCATCTAGATAAACAAGATCTACTTGGAAGG
m73UTRrevXba	GGGTCTAGATACGTAATGCTATTGCAACCC
m83UTRforXba	GCATCTAGACAAGGGGTTAAGTGGCAGG
m83UTRrevXho	GGGCTCGAGTCGCTTCAGCGGCGCAG
M8K1	CCAACCAACAACGCATCCTCGAGAGCCCAACTGTGATGGC
M8K2	CGCATCTGTGATAGCCCAACCCCGGGGGCAACAAAATAGC

4.2. 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), *attP2* (RRID:BDSC_25710) (Groth *et al.* 2004; Markstein *et al.* 2008), *VK2* (RRID:BDSC_9723), *VK13* (RRID:BDSC_24864), *VK37* (RRID:BDSC_24872), *VK40* (RRID:BDSC_35568) (Venken *et al.* 2006); each carrying, and if not - crossed to, a chromosome expressing $\Phi C31$ integrase under the control of *nanos* derived from the *attP40* stock (RRID:BDSC_25709) (Bischof *et al.* 2007). For clonal analysis, the following mutant alleles were used: *bantam*^{Δ1} (Brennecke *et al.* 2003), *FRT82B P[gro⁺]*Df(3R)E(spl)*^{b32.2}* (Schrons *et al.* 1992) (RRID:BDSC_52011), *FRT82B Dcr-1^{Q1147X}* (Lee *et al.* 2004) (RRID:BDSC_32066). The BX22 *E(spl)* deficiency was derived from the *Df(3R)E(spl)22*, *tx¹/TM6B*, *Tb¹* stock (RRID:BDSC_3345). The following *Drosophila* genotypes (supplemented with appropriate transgene on the second chromosome) were used to generate loss-of-function clones by the FLP/FRT system (Xu and Rubin 1993):

yw hs-FLP; M(3L) ubi-GFP FRT80B/bantam^{Δ1} *FRT80B* (Figure 15)

*yw hs-FLP; FRT82B P[gro⁺]*Df(3R)E(spl)*^{b32.2}/FRT82B ubi-GFP* (Figures 11, 16, 26, 28)

*yw hs-FLP; FRT82B P[gro⁺]*Df(3R)E(spl)*^{b32.2}/FRT82B arm-lacZ* (Figure 27)

yw hs-FLP; FRT82B ubi-GFP/FRT82B Dcr-1^{Q1147X} (Figure 14)

yw hs-FLP; FRT82B arm-LacZ/FRT82B Dcr-1^{Q1147X} (Figure 25)

Heat-shock was induced 3 days after egg laying and late third instar wing disks were dissected 3 days after heat shock induction. The clones were visualized in disks by either loss of GFP or LacZ expression or expression of GFP. The *su(Hw)* mutant effects were assayed in the animals transheterozygous for *su(Hw)*^{e04061} null allele

(RRID:BDSC_18224) (Thibault *et al.* 2004) and ***su(Hw)*²** strong hypomorphic allele resulting in the ten times decreased *su(Hw)* expression (RRID:BDSC_983) (Parkhurst *et al.* 1988; Georgiev *et al.* 1997).

4.3. Integration of attB plasmids into attP fly lines

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

4.4. Immunostaining and microscopy

Fixation and immunohistochemistry of embryos and larval tissues was performed according to standard protocols. CNSs and imaginal disks were dissected from late third instar larvae, fixed with 4% paraformaldehyde and labeled with one or combination of the following primary antibodies: rabbit polyclonal **anti-GFP** (Minotech Biotechnology), mouse **anti-β-galactosidase** (Promega Cat# Z3781, RRID:AB_430877), rabbit **anti-β-galactosidase** (ICN Pharmaceuticals Cat# 55976, RRID:AB_2313707), guinea pig **anti-Deadpan** (Caygill and Brand 2017), mouse **anti-Hey** (Monastirioti *et al.* 2010) (Monastirioti M; Development. 2010 Cat# Hey, RRID:AB_2568888), rabbit **anti-PH3** (Millipore Cat# 06-570, RRID:AB_310177), rat **anti-Elav** (DSHB Cat# Rat-Elav-7E8A10 anti-elav, RRID:AB_528218), mouse **anti-Cut** (Bodmer *et al.* 1987) (DSHB Cat# 2b10, RRID:AB_528186), guinea pig **anti-Sens** (gift from Hugo Bellen) (Nolo *et al.* 2000). For their detection, I used the following secondary antibodies: goat anti-rabbit, Alexa488-conjugated (Thermo Fisher Scientific Cat# A-11034, RRID:AB_2576217) (**GFP**, **β-galactosidase** in **Figure 4** and **5**), goat anti-mouse, Alexa633-conjugated (Thermo Fisher Scientific Cat# A-21052, RRID:AB_2535719) or donkey anti-mouse, Alexa647-conjugated (Thermo Fisher Scientific Cat# A-31571, RRID:AB_162542) (**β-galactosidase**, **Cut**), Cy5-AffiniPure Donkey Anti-Guinea Pig (Jackson ImmunoResearch Labs Cat# 706-175-148, RRID:AB_2340462) (**Dpn**), Alexa Fluor® 555 Donkey Anti-Mouse, (Thermo Fisher Scientific Cat# A-31570, RRID:AB_2536180) (**Hey**), Cy3-AffiniPure Goat Anti-Rabbit (Jackson ImmunoResearch Labs Cat# 111-165-144, RRID:AB_2338006) (**PH3**), Cy3-AffiniPure F(ab')₂ Fragment Donkey Anti-Rat, (Jackson ImmunoResearch Labs Cat# 712-166-153, RID:AB_2340669) (**Elav**), Goat anti-Guinea Pig, Alexa Fluor 555 (Thermo Fisher Scientific Cat# A-21435, RRID:AB_2535856) (**Sens**). Samples were imaged on a Leica SP8 confocal platform using a 20x Oil Immersion Objective with fixed zoom levels for each tissue type (CNS, wing and eye disks), except of close-ups which were imaged with 63x objective. All samples within each figure (except of **Figures 7, 14-16**) were fixed and immunostained at the same time. Scanning of all figure samples was performed using identical microscope and software settings and, when possible, completed within one imaging session to enable semi-quantitative comparison. Where scanning of all figure samples within one session was not possible, replicates of the samples from two chosen genotypes were re-scanned together with the remaining samples in the next scanning session ensuring that the replicated samples are comparable. For each genotype, at least 10 wing disks, 5 CNSs and 3 eye-antennal disks were scanned. Images were manipulated using ImageJ (pseudocoloring, rotation and maximum intensity projection z-stacks) and arranged into data sets using Adobe Photoshop CC 2017 and Microsoft PowerPoint. Note that I used two different z-projections for some wing disk images in Part 2 of this manuscript. For the top part, containing the wing pouch, a full z-projection of the sample was done, while the bottom part, containing the notum and hinge, encompassed only sections containing the adult muscle precursors (AMPs). This was done to avoid confusing AMP expression with expression in the

overlying tegula, a sensory organ primordium. Whereas AMPs and tegula can be easily distinguished in the 3D confocal stacks, they 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.

4.5. 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 μ l 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 for normalization. 20 μ l of each homogenized sample was mixed with 50 μ l 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 μ l of each homogenized sample was mixed with 200 μ l BCA Working Reagent on clear-bottomed 96-well plates (Costar) and incubated at 37°C for 1 hr. The plates were allowed to equilibrate to room temperature for 10 min before measuring absorbance on an 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.

5. Acknowledgments

I want to thank Christos Delidakis for his inspiring guidance and boundless patience, nothing in this manuscript would be possible without his support. My PhD advisors, Kriton Kalantidis and Charalampos Spilianakis for their support and patience. My students Electra Vitsaki, Vasilis Ntasis, Marina Gkantia, Konstantina Kalodimou, Babis Galouzis, Alexis Molfetas, Eleni Kabrani, Giorgos Samantsidis and Giannis Rallis, who contributed to various parts of this work, for their excellent work and great patience. Members of Delidakis' laboratory for their cooperativity (and patience). Giannis Livadaras for injecting hundreds of my constructs into fly embryos. Michalis Averof for *pMinos{3xP3-dsRed}*, Maria Monastirioti for *pGL3-hsp70-luc*, Chrysoula Pitsouli and Norbert Perrimon for providing *attP2* and *attP40* flies. Biology Department of University of Crete and Institute of Molecular Biology & Biotechnology of ForTH (or people who run and make up the community of these institutions) for providing fantastic environment for doing science. And sponsors; 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.

6. Literature Cited

- Abed J. A., J. Erceg, A. Goloborodko, S. C. Nguyen, R. B. McCole, *et al.*, 2018 Highly Structured Homolog Pairing Reflects Functional Organization of the Drosophila Genome. *bioRxiv* 443887. <https://doi.org/10.1101/443887>
- Adryan B., G. Woerfel, I. Birch-Machin, S. Gao, M. Quick, *et al.*, 2007 Genomic mapping of Suppressor of Hairy-wing binding sites in Drosophila. *Genome Biol* 8: R167. <https://doi.org/10.1186/gb-2007-8-8-r167>
- Ali T., R. Renkawitz, and M. Bartkuhn, 2016 Insulators and domains of gene expression. *Curr. Opin. Genet. Dev.* 37: 17–26. <https://doi.org/10.1016/j.gde.2015.11.009>
- Anderson K. V. and J. A. Lengyel, 1979 Rates of synthesis of major classes of RNA in Drosophila embryos. *Dev Biol* 70: 217–231.
- Anguera M. C., B. K. Sun, N. Xu, and J. T. Lee, 2006 X-chromosome kiss and tell: how the Xs go their separate ways. *Cold Spring Harb Symp Quant Biol* 71: 429–437. <https://doi.org/10.1101/sqb.2006.71.012>
- Apte M. S., and V. H. Meller, 2012 Homologue pairing in flies and mammals: gene regulation when two are involved. *Genet Res Int* 2012: 430587. <https://doi.org/10.1155/2012/430587>
- Aramayo R., and R. L. Metzzenberg, 1996 Meiotic transvection in fungi. *Cell* 86: 103–113.
- Arnone M. I., I. J. Dmochowski, and C. Gache, 2004 Using Reporter Genes to Study cis-Regulatory Elements, pp. 621–652 in *Methods in cell biology*.
- Artavanis-Tsakonas S., K. Matsuno, and M. E. Fortini, 1995 Notch signaling. *Science* 268: 225–32.
- Artavanis-Tsakonas S., M. D. Rand, and R. J. Lake, 1999 Notch signaling: cell fate control and signal integration in development. *Science* 284: 770–6.
- Atkin N. B., and Z. Jackson, 1996 Evidence for somatic pairing of chromosome 7 and 10 homologs in a follicular lymphoma. *Cancer Genet Cytogenet* 89: 129–131.
- Babu P., and S. Bhat, 1980 Effect of zeste on white complementation. *Basic Life Sci.* 16: 35–40.
- Bailey A. M., and J. W. Posakony, 1995 Suppressor of hairless directly activates transcription of enhancer of split complex genes in response to Notch receptor activity. *Genes Dev.* 9: 2609–22. <https://doi.org/10.1101/gad.9.21.2609>
- Barges S., J. Mihaly, M. Galloni, K. Hagstrom, M. Müller, *et al.*, 2000 The Fab-8 boundary defines the distal limit of the bithorax complex iab-7 domain and insulates iab-7 from initiation elements and a PRE in the adjacent iab-8 domain. *Development* 127: 779–90.
- 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.
- Bartel D. P., 2004 MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* 116: 281–297. [https://doi.org/10.1016/S0092-8674\(04\)00045-5](https://doi.org/10.1016/S0092-8674(04)00045-5)
- Bartel D. P., 2018 Metazoan MicroRNAs. *Cell* 173: 20–51. <https://doi.org/10.1016/j.cell.2018.03.006>
- Bartkuhn M., T. Straub, M. Herold, M. Herrmann, C. Rathke, *et al.*, 2009 Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J.* 28: 877–88. <https://doi.org/10.1038/emboj.2009.34>
- Bateman J. R., and C. T. Wu, 2008 A genome-wide survey argues that every zygotic gene product is dispensable for the initiation of somatic homolog pairing in Drosophila. *Genetics* 180: 1329–1342. <https://doi.org/10.1534/genetics.108.094862>
- Bateman J. R., J. E. Johnson, and M. N. Locke, 2012a Comparing enhancer action in cis and in trans. *Genetics* 191: 1143–1155. <https://doi.org/10.1534/genetics.112.140954>
- Bateman J. R., E. Larschan, R. D'Souza, L. S. Marshall, K. E. Dempsey, *et al.*, 2012b A Genome-Wide Screen Identifies Genes That Affect Somatic Homolog Pairing in Drosophila. *G3* 2: 731–740. <https://doi.org/10.1534/g3.112.002840>
- Baxley R. M., J. D. Bullard, M. W. Klein, A. G. Fell, J. A. Morales-Rosado, *et al.*, 2017 Deciphering the DNA code for the function of the Drosophila polydactyl zinc finger protein Suppressor of Hairy-wing. *Nucleic Acids Res.* 45: 4463–4478. <https://doi.org/10.1093/nar/gkx040>
- Becam I., N. Rafel, X. Hong, S. M. Cohen, and M. Milan, 2011 Notch-mediated repression of bantam miRNA contributes to boundary formation in the Drosophila wing. *Development* 138: 3781–3789. <https://doi.org/10.1242/dev.064774>
- Bejarano F., D. Bortolamiol-Becet, Q. Dai, K. Sun, A. Saj, *et al.*, 2012 A genome-wide transgenic resource for conditional expression of Drosophila microRNAs. *Development* 139: 2821–2831. <https://doi.org/10.1242/dev.079939>
- Bell A. C., A. G. West, and G. Felsenfeld, 1999 The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98: 387–396. [https://doi.org/10.1016/S0092-8674\(00\)81967-4](https://doi.org/10.1016/S0092-8674(00)81967-4)
- 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>
- Bernard F., A. Krejci, B. Housden, B. Adryan, and S. J. Bray, 2010 Specificity of Notch pathway activation: Twist controls the transcriptional output in adult muscle progenitors. *Development* 137: 2633–2642. <https://doi.org/10.1242/dev.053181>
- Bessho Y., H. Hirata, Y. Masamizu, and R. Kageyama, 2003 Periodic repression by the bHLH factor Hes7 is an essential mechanism for the somite segmentation clock. *Genes Dev.* 17: 1451–6. <https://doi.org/10.1101/gad.1092303>
- Bing X., T. Z. Rzezniczak, J. R. Bateman, and T. J. S. Merritt, 2014 Transvection-based gene regulation in Drosophila is a complex and plastic trait. *G3 (Bethesda)*. 4: 2175–87. <https://doi.org/10.1534/g3.114.012484>
- 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>
- Blick A. J., I. Mayer-Hirshfeld, B. R. Malibiran, M. A. Cooper, P. A. Martino, *et al.*, 2016 The capacity to act in trans varies among drosophila enhancers. *Genetics* 203: 203–218. <https://doi.org/10.1534/genetics.115.185645>
- Blochlinger K., L. Y. Jan, and Y. N. Jan, 1993 Postembryonic patterns of expression of cut, a locus regulating sensory organ identity in Drosophila. *Development* 117: 441–50.
- Bodmer R., S. Barbel, S. Sheperd, J. W. Jack, L. Y. Jan, *et al.*, 1987 Transformation of sensory organs by mutations of the cut locus of D. melanogaster. *Cell* 51: 293–307. [https://doi.org/10.1016/0092-8674\(87\)90156-5](https://doi.org/10.1016/0092-8674(87)90156-5)
- Boei J. J., J. Fomina, F. Darroudi, N. J. Nagelkerke, and L. H. Mullenders, 2006 Interphase chromosome positioning affects the spectrum of

- radiation-induced chromosomal aberrations. *Radiat Res* 166: 319–326. <https://doi.org/10.1667/RR3541.1>
- Bollmann J., R. Carpenter, and E. S. Coen, 1991 Allelic interactions at the *nivea* locus of *Antirrhinum*. *Plant Cell* 3: 1327–1336. <https://doi.org/10.1105/tpc.3.12.1327>
- Bonchuk A., O. Maksimenko, O. Kyrchanova, T. Ivlieva, V. Mogila, *et al.*, 2015 Functional role of dimerization and CP190 interacting domains of CTCF protein in *Drosophila melanogaster*. *BMC Biol* 13: 63. <https://doi.org/10.1186/s12915-015-0168-7>
- Bortle K. Van, M. H. Nichols, L. Li, C. T. Ong, N. Takenaka, *et al.*, 2014 Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol* 15: R82. <https://doi.org/10.1186/gb-2014-15-5-r82>
- Bray S. J., 2006 Notch signalling: a simple pathway becomes complex. *Nat. Rev. Mol. Cell Biol.* 7: 678–689. <https://doi.org/10.1038/nrm2009>
- Brennecke J., D. R. Hipfner, A. Stark, R. B. Russell, and S. M. Cohen, 2003 bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila*. *Cell* 113: 25–36.
- Brown J. A., A. Alcaraz, S. Takahashi, D. L. Persons, M. M. Lieber, *et al.*, 1994 Chromosomal aneusomies detected by fluorescent in situ hybridization analysis in clinically localized prostate carcinoma. *J Urol* 152: 1157–1162.
- Bushey A. M., E. Ramos, and V. G. Corces, 2009 Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes Dev* 23: 1338–1350. <https://doi.org/10.1101/gad.1798209>
- Cabrera C. V., 1990 Lateral inhibition and cell fate during neurogenesis in *Drosophila*: the interactions between scute, Notch and Delta. *Development* 110: 733–42.
- Cai H. N., and M. Levine, 1997 The gypsy insulator can function as a promoter-specific silencer in the *Drosophila* embryo. *EMBO J* 16: 1732–1741. <https://doi.org/10.1093/emboj/16.7.1732>
- Cai H. N., and P. Shen, 2001 Effects of cis arrangement of chromatin insulators on enhancer-blocking activity. *Science* 291: 493–495. <https://doi.org/10.1126/science.291.5503.493>
- Cairrao F., A. S. Halees, K. S. A. Khabar, D. Morello, and N. Vanzo, 2009 AU-rich elements regulate *Drosophila* gene expression. *Mol. Cell Biol.* 29: 2636–43. <https://doi.org/10.1128/MCB.01506-08>
- Casares F., W. Bender, J. Merriam, and E. Sánchez-Herrero, 1997 Interactions of *Drosophila* Ultrabithorax regulatory regions with native and foreign promoters. *Genetics* 145: 123–37.
- Castro B., 2005 Lateral inhibition in proneural clusters: cis-regulatory logic and default repression by Suppressor of Hairless. *Development* 132: 3333–3344. <https://doi.org/10.1242/dev.01920>
- Cave J. W., L. Xia, and M. A. Caudy, 2009 The Daughterless N-terminus directly mediates synergistic interactions with Notch transcription complexes via the SPS+A DNA transcription code. *BMC Res. Notes* 2: 65. <https://doi.org/10.1186/1756-0500-2-65>
- Caygill E. E., and A. H. Brand, 2017 miR-7 Buffers Differentiation in the Developing *Drosophila* Visual System. *Cell Rep.* 20: 1255–1261. <https://doi.org/10.1016/j.celrep.2017.07.047>
- Celis J. F. de, J. de Celis, P. Ligoxygakis, A. Preiss, C. Delidakis, *et al.*, 1996 Functional relationships between Notch, Su(H) and the bHLH genes of the E(spl) complex: the E(spl) genes mediate only a subset of Notch activities during imaginal development. *Development* 122: 2719–28.
- Celniker S. E., L. A. L. Dillon, M. B. Gerstein, K. C. Gunsalus, S. Henikoff, *et al.*, 2009 Unlocking the secrets of the genome. *Nature* 459: 927–930. <https://doi.org/10.1038/459927a>
- Chen J. L., K. L. Huisinga, M. M. Viering, S. A. Ou, C. T. Wu, *et al.*, 2002 Enhancer action in trans is permitted throughout the *Drosophila* genome. *Proc Natl Acad Sci U S A* 99: 3723–3728. <https://doi.org/10.1073/pnas.062447999>
- Chen H., M. Levo, L. Barinov, M. Fujioka, J. B. Jaynes, *et al.*, 2018 Dynamic interplay between enhancer–promoter topology and gene activity. *Nat. Genet.* 50: 1296–1303. <https://doi.org/10.1038/s41588-018-0175-z>
- Chetverina D., E. Savitskaya, O. Maksimenko, L. Melnikova, O. Zaytseva, *et al.*, 2008 Red flag on the white reporter: A versatile insulator abuts the white gene in *Drosophila* and is omnipresent in mini-white constructs. *Nucleic Acids Res.* 36: 929–937. <https://doi.org/10.1093/nar/gkm992>
- Chetverina D., M. Fujioka, M. Erokhin, P. Georgiev, J. B. Jaynes, *et al.*, 2017 Boundaries of loop domains (insulators): Determinants of chromosome form and function in multicellular eukaryotes. *BioEssays* 39: 1600233. <https://doi.org/10.1002/bies.201600233>
- Chopra V. S., J. Cande, J. W. Hong, and M. Levine, 2009 Stalled Hox promoters as chromosomal boundaries. *Genes Dev* 23: 1505–1509. <https://doi.org/10.1101/gad.1807309>
- Comet I., E. Savitskaya, B. Schuettengruber, N. Nègre, S. Lavrov, *et al.*, 2006 PRE-Mediated Bypass of Two Su(Hw) Insulators Targets PcG Proteins to a Downstream Promoter. *Dev. Cell* 11: 117–124. <https://doi.org/10.1016/j.devcel.2006.05.009>
- Cook P. R., 1997 The transcriptional basis of chromosome pairing. *J Cell Sci* 110 (Pt 9): 1033–1040.
- Cooper M. T. D., D. M. Tyler, M. Furriols, A. Chalkiadaki, C. Delidakis, *et al.*, 2000 Spatially restricted factors cooperate with notch in the regulation of Enhancer of split genes. *Dev. Biol.* 221: 390–403. <https://doi.org/10.1006/dbio.2000.9691>
- Corson F., L. Couturier, H. Rouault, K. Mazouni, and F. Schweisguth, 2017 Self-organized Notch dynamics generate stereotyped sensory organ patterns in *Drosophila*. 7407. <https://doi.org/10.1126/science.aai7407>
- Cremer T., C. Cremer, T. Schneider, H. Baumann, L. Hens, *et al.*, 1982 Analysis of chromosome positions in the interphase nucleus of Chinese hamster cells by laser-UV-microirradiation experiments. *Hum Genet* 62: 201–209.
- Cremer T., and M. Cremer, 2010 Chromosome territories. *Cold Spring Harb Perspect Biol* 2: a003889. <https://doi.org/10.1101/cshperspect.a003889>
- Csink A. K., and S. Henikoff, 1996 Genetic modification of heterochromatic association and nuclear organization in *Drosophila*. *Nature* 381: 529–531. <https://doi.org/10.1038/381529a0>
- Csink A. K., and S. Henikoff, 1998 Large-scale chromosomal movements during interphase progression in *Drosophila*. *J Cell Biol* 143: 13–22.
- Cubas P., C. D. B. Molecular, C. Superior, and D. I. Cientificas, 1991 Proneural clusters of achaete-scute expression and the generation of sensory organs in the *Drosophila* imaginal wing disc. 996–1008.
- Cubeñas-Potts C., M. J. Rowley, X. Lyu, G. Li, E. P. Lei, *et al.*, 2016 Different enhancer classes in *Drosophila* bind distinct architectural proteins and mediate unique chromatin interactions and 3D architecture. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkw1114>
- Culí J., and J. Modolell, 1998 Proneural gene self-stimulation in neural precursors: an essential mechanism for sense organ development that is regulated by Notch signaling. *Genes Dev.* 12: 2036–47. <https://doi.org/10.1101/gad.12.13.2036>
- Danilowicz C., C. H. Lee, K. Kim, K. Hatch, V. W. Coljee, *et al.*, 2009 Single molecule detection of direct, homologous, DNA/DNA pairing. *Proc Natl Acad Sci U S A* 106: 19824–19829. <https://doi.org/10.1073/pnas.0911214106>
- Delidakis C., A. Preiss, D. A. Hartley, and S. Artavanis-Tsakonas, 1991 Two genetically and molecularly distinct functions involved in early neurogenesis reside within the Enhancer of split locus of *Drosophila melanogaster*. *Genetics* 129: 803–23.

- Delidakis C., and S. Artavanis-Tsakonas, 1992a 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>
- Delidakis C., and S. Artavanis-Tsakonas, 1992b 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>
- Delidakis C., M. Monastirioti, and S. S. Magadi, 2014 E(spl): genetic, developmental, and evolutionary aspects of a group of invertebrate Hes proteins with close ties to Notch signaling. *Curr. Top. Dev. Biol.* 110: 217–62. <https://doi.org/10.1016/B978-0-12-405943-6.00006-3>
- Dernburg A. F., K. W. Broman, J. C. Fung, W. F. Marshall, J. Philips, *et al.*, 1996 Perturbation of nuclear architecture by long-distance chromosome interactions. *Cell* 85: 745–759.
- Dixon J. R., S. Selvaraj, F. Yue, A. Kim, Y. Li, *et al.*, 2012 Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485: 376–380. <https://doi.org/10.1038/nature11082>
- Donohoe M. E., S. S. Silva, S. F. Pinter, N. Xu, and J. T. Lee, 2009 The pluripotency factor Oct4 interacts with Ctf and also controls X-chromosome pairing and counting. *Nature* 460: 128–132. <https://doi.org/10.1038/nature08098>
- Doren M. Van, A. M. Bailey, J. Esnayra, K. Ede, and J. W. Posakony, 1994 Negative regulation of proneural gene activity: hairy is a direct transcriptional repressor of achaete. *Genes Dev.* 8: 2729–42.
- Dunaway M., and P. Droge, 1989 Transactivation of the *Xenopus* rRNA gene promoter by its enhancer. *Nature* 341: 657–659. <https://doi.org/10.1038/341657a0>
- Duncan I. W., 2002 Transvection effects in *Drosophila*. *Annu Rev Genet* 36: 521–556. <https://doi.org/10.1146/annurev.genet.36.060402.100441>
- Duncan E. J., and P. K. Dearden, 2010 Evolution of a genomic regulatory domain: the role of gene co-option and gene duplication in the Enhancer of split complex. *Genome Res.* 20: 917–28. <https://doi.org/10.1101/gr.104794.109>
- Duttagupta R., B. Tian, C. J. Wilusz, D. T. Khounh, P. Soteropoulos, *et al.*, 2005 Global Analysis of Pub1p Targets Reveals a Coordinate Control of Gene Expression through Modulation of Binding and Stability. *Mol. Cell. Biol.* 25: 5499–5513. <https://doi.org/10.1128/MCB.25.13.5499-5513.2005>
- Duvillie B., D. Buccini, T. Tang, J. Jami, and A. Paldi, 1998 Imprinting at the mouse *Ins2* locus: evidence for cis- and trans-allelic interactions. *Genomics* 47: 52–57. <https://doi.org/10.1006/geno.1997.5070>
- Edgar B. A., and G. Schubiger, 1986 Parameters controlling transcriptional activation during early *Drosophila* development. *Cell* 44: 871–877.
- Erceg J., J. A. Abed, A. Goloborodko, B. R. Lajoie, G. Fudenberg, *et al.*, 2018 The genome-wide, multi-layered architecture of chromosome pairing in early *Drosophila* embryos. *bioRxiv* 443028. <https://doi.org/10.1101/443028>
- Erokhin M., A. Davydova, O. Kyrchanova, A. Parshikov, P. Georgiev, *et al.*, 2011 Insulators form gene loops by interacting with promoters in *Drosophila*. *Development* 138: 4097–106. <https://doi.org/10.1242/dev.062836>
- Fedorova E., N. Sadoni, I. K. Dahlsveen, J. Koch, E. Kremmer, *et al.*, 2008 The nuclear organization of Polycomb/Trithorax group response elements in larval tissues of *Drosophila melanogaster*. *Chromosom. Res* 16: 649–673. <https://doi.org/10.1007/s10577-008-1218-6>
- Feng S., S. J. Cokus, V. Schubert, J. Zhai, M. Pellegrini, *et al.*, 2014 Genome-wide Hi-C Analyses in Wild-Type and Mutants Reveal High-Resolution Chromatin Interactions in *Arabidopsis*. *Mol. Cell* 55: 694–707. <https://doi.org/10.1016/j.molcel.2014.07.008>
- Foe V. E., and B. M. Alberts, 1985 Reversible chromosome condensation induced in *Drosophila* embryos by anoxia: visualization of interphase nuclear organization. *J Cell Biol* 100: 1623–1636.
- Foe V. E., and G. M. Odell, 1993 Mitosis and morphogenesis in the *Drosophila* embryo: point and counterpoint. *Dev. Drosoph. melanogaster* 1: 149–300.
- Fritsch C., G. Ploeger, and D. J. Arndt-Jovin, 2006 *Drosophila* under the lens: imaging from chromosomes to whole embryos. *Chromosom. Res* 14: 451–464. <https://doi.org/10.1007/s10577-006-1068-z>
- Fujioka M., G. Sun, and J. B. Jaynes, 2013 The *Drosophila* eve Insulator Homie Promotes eve Expression and Protects the Adjacent Gene from Repression by Polycomb Spreading. *PLoS Genet.* 9. <https://doi.org/10.1371/journal.pgen.1003883>
- Fujioka M., H. Mistry, P. Schedl, and J. B. Jaynes, 2016 Determinants of Chromosome Architecture: Insulator Pairing in cis and in trans. *PLoS Genet.* 12: 1–25. <https://doi.org/10.1371/journal.pgen.1005889>
- Fukaya T., and M. Levine, 2017 Transvection. *Curr. Biol.* 27: R1047–R1049. <https://doi.org/10.1016/j.cub.2017.08.001>
- Fung J. C., W. F. Marshall, A. Dernburg, D. A. Agard, and J. W. Sedat, 1998 Homologous chromosome pairing in *Drosophila melanogaster* proceeds through multiple independent initiations. *J Cell Biol* 141: 5–20.
- Furriols M., and S. Bray, 2001 A model Notch response element detects Suppressor of Hairless-dependent molecular switch. *Curr. Biol.* 11: 60–4.
- Geisberg J. V., Z. Moqtaderi, X. Fan, F. Ozsolak, and K. Struhl, 2014 Global Analysis of mRNA Isoform Half-Lives Reveals Stabilizing and Destabilizing Elements in Yeast. *Cell* 156: 812–824. <https://doi.org/10.1016/j.cell.2013.12.026>
- Gelbart W. M., 1982 Synapsis-dependent allelic complementation at the decapentaplegic gene complex in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci.* 79: 2636–2640.
- Gemkow M. J., P. J. Verveer, and D. J. Arndt-Jovin, 1998 Homologous association of the Bithorax-Complex during embryogenesis: consequences for transvection in *Drosophila melanogaster*. *Development* 125: 4541–4552.
- Georgiev P., T. Tikhomirova, V. Yelagin, T. Belenkaya, E. Gracheva, *et al.*, 1997 Insertions of Hybrid P Elements in the yellow Gene of *Drosophila* Cause a Large Variety of Mutant Phenotypes. *Genetics* 146.
- Gerasimova T. I., D. A. Gdula, D. V. Gerasimov, O. Simonova, and V. G. Corces, 1995 A *drosophila* protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. *Cell* 82: 587–597. [https://doi.org/10.1016/0092-8674\(95\)90031-4](https://doi.org/10.1016/0092-8674(95)90031-4)
- Geyer P. K., M. M. Green, and V. G. Corces, 1990 Tissue-specific transcriptional enhancers may act in trans on the gene located in the homologous chromosome: the molecular basis of transvection in *Drosophila*. *EMBO J* 9: 2247–2256.
- Geyer P. K., and V. G. Corces, 1992 DNA position-specific repression of transcription by a *Drosophila* zinc finger protein. *Genes Dev* 6: 1865–1873.
- Ghysen A., C. Dambly-Chaudière, L. Y. Jan, and Y. N. Jan, 1993 Cell interactions and gene interactions in peripheral neurogenesis. *Genes Dev.* 7: 723–33.
- Giagtoglou N., P. Alifragis, K. A. Koumbanakis, and C. Delidakis, 2003 Two modes of recruitment of E(spl) repressors onto target genes. *Development* 130: 259–70. <https://doi.org/10.1242/dev.00206>
- Giagtoglou N., K. A. Koumbanakis, J. Fullard, I. Zarifi, and C. Delidakis, 2005 Role of the Sc C Terminus in Transcriptional Activation and E(spl) Repressor Recruitment. *J. Biol. Chem.* 280: 1299–1305. <https://doi.org/10.1074/jbc.M408949200>

- Giebel B., and J. A. Campos-Ortega, 1997 Functional dissection of the *Drosophila* Enhancer of split protein, a suppressor of neurogenesis. *Proc. Natl. Acad. Sci.* 94: 6250–6254. <https://doi.org/10.1073/pnas.94.12.6250>
- Gilbert S. F., 2000 *Developmental Biology*. Palgrave Macmillan.
- Gloor G. B., 2002 The role of sequence homology in the repair of DNA double-strand breaks in *Drosophila*. *Adv Genet* 46: 91–117.
- Goldsborough A. S., and T. B. Kornberg, 1996 Reduction of transcription by homologue asynapsis in *Drosophila* imaginal discs. *Nature* 381: 807–810. <https://doi.org/10.1038/381807a0>
- Golic M. M., and K. G. Golic, 1996 A quantitative measure of the mitotic pairing of alleles in *Drosophila melanogaster* and the influence of structural heterozygosity. *Genetics* 143: 385–400.
- Golovnin A., I. Biryukova, O. Romanova, M. Silicheva, A. Parshikov, *et al.*, 2003 An endogenous Su(Hw) insulator separates the yellow gene from the Achaete-scute gene complex in *Drosophila*. *Development* 130: 3249–3258. <https://doi.org/10.1242/dev.020263>
- Golovnin A., E. Melnick, A. Mazur, and P. Georgiev, 2005 *Drosophila* Su(Hw) insulator can stimulate transcription of a weakened yellow promoter over a distance. *Genetics* 170: 1133–1142. <https://doi.org/10.1534/genetics.104.034587>
- Gómez-Skarmeta J. L., S. Campuzano, and J. Modolell, 2003 Half a century of neural pre patterning: the story of a few bristles and many genes. *Nat. Rev. Neurosci.* 4: 587–98. <https://doi.org/10.1038/nrn1142>
- Grosfeld F., G. B. van Assendelft, D. R. Greaves, and G. Kollias, 1987 Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* 51: 975–85.
- Groth A. C., M. Fish, R. Nusse, and M. P. Calos, 2004 Construction of transgenic *Drosophila* by using the site-specific integrase from phage ϕ C31. *Genetics* 166: 1775–82.
- Gruzdeva N., O. Kyrchanova, A. Parshikov, A. Kullyev, and P. Georgiev, 2005 The Mcp element from the bithorax complex contains an insulator that is capable of pairwise interactions and can facilitate enhancer-promoter communication. *Mol. Cell. Biol.* 25: 3682–9. <https://doi.org/10.1128/MCB.25.9.3682-3689.2005>
- Guo Y., K. Monahan, H. Wu, J. Gertz, K. E. Varley, *et al.*, 2012 CTCF/cohesin-mediated DNA looping is required for protocadherin α promoter choice. *Proc. Natl. Acad. Sci. U. S. A.* 109: 21081–6. <https://doi.org/10.1073/pnas.1219280110>
- Guo Y., Q. Xu, D. Canzio, J. Shou, J. Li, *et al.*, 2015 CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* 162: 900–910. <https://doi.org/10.1016/j.cell.2015.07.038>
- Haaf T., and M. Schmid, 1989 Centromeric association and non-random distribution of centromeres in human tumour cells. *Hum Genet* 81: 137–143.
- Harrison D. A., D. A. Gdula, R. S. Coyne, and V. G. Corces, 1993 A leucine zipper domain of the suppressor of Hairy-wing protein mediates its repressive effect on enhancer function. *Genes Dev* 7: 1966–1978. <https://doi.org/10.1101/gad.7.10.1966>
- Hartl T. A., H. F. Smith, and G. Bosco, 2008 Chromosome alignment and transvection are antagonized by condensin II. *Science* (80-.). 322: 1384–1387. <https://doi.org/10.1126/science.1164216>
- Hartman J. L., B. Garvik, and L. Hartwell, 2001 Principles for the buffering of genetic variation. *Science* 291: 1001–4.
- Hasan A., C. Cotobal, C. D. S. Duncan, and J. Mata, 2014 Systematic Analysis of the Role of RNA-Binding Proteins in the Regulation of RNA Stability, (G. P. Copenhaver, Ed.). *PLoS Genet.* 10: e1004684. <https://doi.org/10.1371/journal.pgen.1004684>
- Heitzler P., M. Bourouis, L. Ruel, C. Carteret, and P. Simpson, 1996 Genes of the Enhancer of split and achaete-scute complexes are required for a regulatory loop between Notch and Delta during lateral signalling in *Drosophila*. *Development* 122: 161–71.
- Helfer S., J. Schott, G. Stoecklin, and K. Förstemann, 2012 AU-Rich Element-Mediated mRNA Decay Can Occur Independently of the miRNA Machinery in Mouse Embryonic Fibroblasts and *Drosophila* S2-Cells, (S. Pfeffer, Ed.). *PLoS One* 7: e28907. <https://doi.org/10.1371/journal.pone.0028907>
- Hendrickson J. E., and S. Sakonju, 1995 Cis and trans interactions between the iab regulatory regions and abdominal-A and abdominal-B in *Drosophila melanogaster*. *Genetics* 139: 835–48.
- Heride C., M. Ricoul, K. Kieu, J. von Hase, V. Guillemot, *et al.*, 2010 Distance between homologous chromosomes results from chromosome positioning constraints. *J Cell Sci* 123: 4063–4075. <https://doi.org/10.1242/jcs.066498>
- Herrera-Carrillo E., and B. Berkhout, 2017 Dicer-independent processing of small RNA duplexes: mechanistic insights and applications. *Nucleic Acids Res.* 45: 10369–10379. <https://doi.org/10.1093/nar/gkx779>
- Hinz U., B. Giebel, and J. A. Campos-Ortega, 1994 The basic-helix-loop-helix domain of *Drosophila* lethal of scute protein is sufficient for proneural function and activates neurogenic genes. *Cell* 76: 77–87.
- Hirano T., 2006 At the heart of the chromosome: SMC proteins in action. *Nat. Rev. Mol. cell Biol.* 7: 311–322.
- Hiraoka Y., D. A. Agard, and J. W. Sedat, 1990 Temporal and spatial coordination of chromosome movement, spindle formation, and nuclear envelope breakdown during prometaphase in *Drosophila melanogaster* embryos. *J Cell Biol* 111: 2815–2828.
- Hiraoka Y., A. F. Dernburg, S. J. Parmelee, M. C. Rykowski, D. A. Agard, *et al.*, 1993 The onset of homologous chromosome pairing during *Drosophila melanogaster* embryogenesis. *J Cell Biol* 120: 591–600.
- Hnisz D., D. S. Day, and R. A. Young, 2016 Insulated Neighborhoods: Structural and Functional Units of Mammalian Gene Control. *Cell* 167: 1188–1200. <https://doi.org/10.1016/j.cell.2016.10.024>
- Hogan D. J., D. P. Riordan, A. P. Gerber, D. Herschlag, and P. O. Brown, 2008 Diverse RNA-Binding Proteins Interact with Functionally Related Sets of RNAs, Suggesting an Extensive Regulatory System, (Sean R. Eddy, Ed.). *PLoS Biol.* 6: e255. <https://doi.org/10.1371/journal.pbio.0060255>
- Holdridge C., and D. Dorsett, 1991 Repression of hsp70 heat shock gene transcription by the suppressor of hairy-wing protein of *Drosophila melanogaster*. *Mol Cell Biol* 11: 1894–1900. <https://doi.org/10.1128/MCB.11.4.1894>.Updated
- Horn C., B. Jaunich, and E. A. Wimmer, 2000 Highly sensitive, fluorescent transformation marker for *Drosophila* transgenesis. *Dev. Genes Evol.* 210: 623–629. <https://doi.org/10.1007/s004270000111>
- Hou C., L. Li, Z. S. Qin, and V. G. Corces, 2012 Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol Cell* 48: 471–484. <https://doi.org/10.1016/j.molcel.2012.08.031>
- Housden B. E., A. Q. Fu, A. Krejci, F. Bernard, B. Fischer, *et al.*, 2013 Transcriptional Dynamics Elicited by a Short Pulse of Notch Activation Involves Feed-Forward Regulation by E(spl)/Hes Genes, (J. Lewis, Ed.). *PLoS Genet.* 9: e1003162. <https://doi.org/10.1371/journal.pgen.1003162>
- Hu J. F., T. H. Vu, and A. R. Hoffman, 1997 Genomic deletion of an imprint maintenance element abolishes imprinting of both insulin-like growth factor II and H19. *J Biol Chem* 272: 20715–20720.
- Huang F., C. Dambly-chaudiere, and A. Ghysen, 1991 The emergence of sense organs in the wing disc of *Drosophila*. 1095: 1087–1095.
- Inoue S., S. Sugiyama, A. A. Travers, and T. Ohyama, 2007 Self-assembly of double-stranded DNA molecules at nanomolar concentrations. *Biochemistry* 46: 164–171. <https://doi.org/10.1021/bi061539y>

- Jack J. W., and B. H. Judd, 1979 Allelic pairing and gene regulation: A model for the zeste-white interaction in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 76: 1368–72.
- Jennings B., A. Preiss, C. Delidakis, and S. Bray, 1994 The Notch signalling pathway is required for Enhancer of split bHLH protein expression during neurogenesis in the *Drosophila* embryo. *Development* 120: 3537–48. <https://doi.org/10.1109/CONGRESS.2009.33>
- Jennings V., J. de Celis, C. Delidakis, A. Preiss, and S. Bray, 1995 Role of Notch and achaete-scute complex in the expression of Enhancer of split bHLH proteins. *Development* 121.
- Jennings B. H., D. M. Tyler, and S. J. Bray, 1999 Target specificities of *Drosophila* enhancer of split basic helix-loop-helix proteins. *Mol. Cell. Biol.* 19: 4600–10. <https://doi.org/10.1128/MCB.19.7.4600>
- Jennings B. H., L. M. Pickles, S. M. Wainwright, S. M. Roe, L. H. Pearl, *et al.*, 2006 Molecular recognition of transcriptional repressor motifs by the WD domain of the Groucho/TLE corepressor. *Mol. Cell* 22: 645–55. <https://doi.org/10.1016/j.molcel.2006.04.024>
- Johnston R. J., C. Desplan, R. J. Johnston Jr., C. Desplan, R. J. Johnston, *et al.*, 2014 Interchromosomal communication coordinates intrinsically stochastic expression between alleles. *Science* 343: 661–5. <https://doi.org/10.1126/science.1243039>
- Joyce E. F., B. R. Williams, T. Xie, and C. ting Wu, 2012 Identification of genes that promote or antagonize somatic homolog pairing using a high-throughput FISH-based screen. *PLoS Genet* 8: e1002667. <https://doi.org/10.1371/journal.pgen.1002667>
- Kageyama R., T. Ohtsuka, and T. Kobayashi, 2007 The Hes gene family: repressors and oscillators that orchestrate embryogenesis. *Development* 134: 1243–51. <https://doi.org/10.1242/dev.000786>
- Kageyama R., Y. Niwa, H. Shimojo, T. Kobayashi, and T. Ohtsuka, 2010 Ultradian oscillations in Notch signaling regulate dynamic biological events. *Curr. Top. Dev. Biol.* 92: 311–31. [https://doi.org/10.1016/S0070-2153\(10\)92010-3](https://doi.org/10.1016/S0070-2153(10)92010-3)
- Kas E., and U. K. Laemmli, 1992 In vivo topoisomerase II cleavage of the *Drosophila* histone and satellite III repeats: DNA sequence and structural characteristics. *EMBO J* 11: 705–716.
- Kellum R., and P. Schedl, 1991 A position-effect assay for boundaries of higher order chromosomal domains. *Cell* 64: 941–950.
- Kim A., C. Terzian, P. Santamaria, A. Pelisson, N. Purd'homme, *et al.*, 1994 Retroviruses in invertebrates: the gypsy retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 91: 1285–1289.
- 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>
- Kiparaki M., I. Zarifi, and C. Delidakis, 2015 bHLH proteins involved in *Drosophila* neurogenesis are mutually regulated at the level of stability. *Nucleic Acids Res.* 43: 2543–2559. <https://doi.org/10.1093/nar/gkv083>
- Kitano H., 2004 Biological robustness. *Nat. Rev. Genet.* 5: 826–837. <https://doi.org/10.1038/nrg1471>
- Klämbt C., E. Knust, K. Tietze, and J. A. Campos-Ortega, 1989 Closely related transcripts encoded by the neurogenic gene complex enhancer of split of *Drosophila melanogaster*. *EMBO J.* 8: 203–10.
- Knust E., K. Tietze, and J. A. Campos-Ortega, 1987 Molecular analysis of the neurogenic locus Enhancer of split of *Drosophila melanogaster*. *EMBO J.* 6: 4113–23.
- Knust E., H. Schrons, F. Grawe, and J. A. Campos-Ortega, 1992 Seven genes of the Enhancer of split complex of *Drosophila melanogaster* encode helix-loop-helix proteins. *Genetics* 132: 505–18.
- Koelzer S., and T. Klein, 2003 A Notch-independent function of Suppressor of Hairless during the development of the bristle sensory organ precursor cell of *Drosophila*. *Development* 130: 1973–88. <https://doi.org/10.1242/dev.00426>
- Koeman J. M., R. C. Russell, M. H. Tan, D. Petillo, M. Westphal, *et al.*, 2008 Somatic pairing of chromosome 19 in renal oncocyoma is associated with deregulated EGLN2-mediated [corrected] oxygen-sensing response. *PLoS Genet* 4: e1000176. <https://doi.org/10.1371/journal.pgen.1000176>
- Kramatschek B., and J. a Campos-Ortega, 1994 Neuroectodermal transcription of the *Drosophila* neurogenic genes *E(spl)* and *HLH-m5* is regulated by proneural genes. *Development* 120: 815–826.
- Kravchenko E., E. Savitskaya, O. Kravchuk, A. Parshikov, P. Georgiev, *et al.*, 2005 Pairing between gypsy insulators facilitates the enhancer action in trans throughout the *Drosophila* genome. *Mol. Cell. Biol.* 25: 9283–91. <https://doi.org/10.1128/MCB.25.21.9283-9291.2005>
- Kravchuk O., M. Kim, P. Klepikov, A. Parshikov, P. Georgiev, *et al.*, 2016 Transvection in *Drosophila*: trans-interaction between yellow enhancers and promoter is strongly suppressed by a cis-promoter only in certain genomic regions. *Chromosoma* 1–11. <https://doi.org/10.1007/s00412-016-0605-6>
- Krejci A., and S. Bray, 2007 Notch activation stimulates transient and selective binding of Su(H)/CSL to target enhancers. *Genes Dev.* 21: 1322–7. <https://doi.org/10.1101/gad.424607>
- Krueger C., M. R. King, F. Krueger, M. R. Branco, C. S. Osborne, *et al.*, 2012 Pairing of homologous regions in the mouse genome is associated with transcription but not imprinting status. *PLoS One* 7: e38983. <https://doi.org/10.1371/journal.pone.0038983>
- Kurshakova M., O. Maksimenko, A. Golovnin, M. Pulina, S. Georgieva, *et al.*, 2007 Evolutionarily Conserved E(y)2/Sus1 Protein Is Essential for the Barrier Activity of Su(Hw)-Dependent Insulators in *Drosophila*. *Mol. Cell* 27: 332–338. <https://doi.org/10.1016/j.molcel.2007.05.035>
- Kutach a K., and J. T. Kadonaga, 2000 The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. *Mol. Cell. Biol.* 20: 4754–4764. <https://doi.org/10.1128/MCB.20.13.4754-4764.2000>
- Kvon E. Z., 2015 Using transgenic reporter assays to functionally characterize enhancers in animals. *Genomics* 106: 185–192. <https://doi.org/10.1016/j.ygeno.2015.06.007>
- Kyrchanova O., S. Toshchakov, Y. Podstreshnaya, A. Parshikov, and P. Georgiev, 2008a 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: 4188–95. <https://doi.org/10.1128/MCB.00229-08>
- Kyrchanova O., D. Chetverina, O. Maksimenko, A. Kullyev, and P. Georgiev, 2008b Orientation-dependent interaction between *Drosophila* insulators is a property of this class of regulatory elements. *Nucleic Acids Res.* 36: 7019–28. <https://doi.org/10.1093/nar/gkn781>
- Kyrchanova O., O. Maksimenko, V. Stakhov, T. Ivlieva, A. Parshikov, *et al.*, 2013 Effective blocking of the white enhancer requires cooperation between two main mechanisms suggested for the insulator function. *PLoS Genet* 9: e1003606. <https://doi.org/10.1371/journal.pgen.1003606>
- Kyrchanova O., V. Mogila, D. Wolle, G. Deshpande, A. Parshikov, *et al.*, 2016 Functional Dissection of the Blocking and Bypass Activities of the Fab-8 Boundary in the *Drosophila* Bithorax Complex. *PLoS Genet.* 12: 1–22. <https://doi.org/10.1371/journal.pgen.1006188>
- Lai E. C., and J. W. Posakony, 1997 The Bearded box, a novel 3' UTR sequence motif, mediates negative post-transcriptional regulation of Bearded and Enhancer of split Complex gene expression. *Development* 124: 4847–56.
- Lai E. C., C. Burks, and J. W. Posakony, 1998 The K box, a conserved 3' UTR sequence motif, negatively regulates accumulation of enhancer

- of split complex transcripts. *Development* 125: 4077–88.
- Lai E. C., R. Bodner, and J. W. Posakony, 2000 The enhancer of split complex of *Drosophila* includes four Notch-regulated members of the bearded gene family. *Development* 127: 3441–55. <https://doi.org/8223268>
- Lai E. C., and V. Orgogozo, 2004 A hidden program in *Drosophila* peripheral neurogenesis revealed: fundamental principles underlying sensory organ diversity. *Dev. Biol.* 269: 1–17. <https://doi.org/10.1016/j.ydbio.2004.01.032>
- 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>
- Landers M., M. A. Calciano, D. Colosi, H. Glatt-Deeley, J. Wagstaff, *et al.*, 2005 Maternal disruption of Ube3a leads to increased expression of Ube3a-ATS in trans. *Nucleic Acids Res* 33: 3976–3984. <https://doi.org/10.1093/nar/gki705>
- LaSalle J. M., and M. Lalande, 1996 Homologous association of oppositely imprinted chromosomal domains. *Science* (80-.). 272: 725–728.
- Laverty C., F. Li, E. J. Belikoff, and M. J. Scott, 2011 Abnormal dosage compensation of reporter genes driven by the *Drosophila* glass multiple reporter (GMR) enhancer-promoter. *PLoS One* 6: e20455. <https://doi.org/10.1371/journal.pone.0020455>
- Lecourtou M., and F. Schweisguth, 1995 The neurogenic suppressor of hairless DNA-binding protein mediates the transcriptional activation of the enhancer of split complex genes triggered by Notch signaling. *Genes Dev.* 9: 2598–608. <https://doi.org/10.1101/gad.9.21.2598>
- Lee Y. S., K. Nakahara, J. W. Pham, K. Kim, Z. He, *et al.*, 2004 Distinct Roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA Silencing Pathways. *Cell* 117: 69–81. [https://doi.org/10.1016/S0092-8674\(04\)00261-2](https://doi.org/10.1016/S0092-8674(04)00261-2)
- Lee A. M., and C. T. Wu, 2006 Enhancer-promoter communication at the yellow gene of *Drosophila melanogaster*: Diverse promoters participate in and regulate trans interactions. *Genetics* 174: 1867–1880. <https://doi.org/10.1534/genetics.106.064121>
- Lehmann R., F. Jiménez, U. Dietrich, and J. A. Campos-Ortega, 1983 On the phenotype and development of mutants of early neurogenesis in *Drosophila melanogaster*. *Wilhelm Roux's Arch. Dev. Biol.* 192: 62–74. <https://doi.org/10.1007/BF00848482>
- Leiserson W. M., N. M. Bonini, and S. Benzer, 1994 Transvection at the eyes absent gene of *Drosophila*. *Genetics* 138: 1171–1179.
- Levis R., T. Hazelrigg, and G. M. Rubin, 1985 Effects of genomic position on the expression of transduced copies of the white gene of *Drosophila*. *Science* 229: 558–61.
- Leviton M. W., E. C. Lai, and J. W. Posakony, 1997 The *Drosophila* gene Bearded encodes a novel small protein and shares 3' UTR sequence motifs with multiple Enhancer of split complex genes. *Development* 124: 4039–51.
- Lewis E. B., 1954 The Theory and Application of a New Method of Detecting Chromosomal Rearrangements in *Drosophila melanogaster*. *Am. Nat.* 88: 225–239. <https://doi.org/doi:10.1086/281833>
- Lewis J. P., H. J. Tanke, A. K. Raap, G. C. Beverstock, and H. C. Kluin-Nelemans, 1993 Somatic pairing of centromeres and short arms of chromosome 15 in the hematopoietic and lymphoid system. *Hum Genet* 92: 577–582.
- Li H. B., M. Muller, I. A. Bahechar, O. Kyrchanova, K. Ohno, *et al.*, 2011 Insulators, not Polycomb response elements, are required for long-range interactions between Polycomb targets in *Drosophila melanogaster*. *Mol Cell Biol* 31: 616–625. <https://doi.org/10.1128/MCB.00849-10>
- Liang J., L. Lacroix, A. Gamot, S. Cuddapah, S. Queille, *et al.*, 2014 Chromatin Immunoprecipitation Indirect Peaks Highlight Long-Range Interactions of Insulator Proteins and Pol II Pausing. *Mol. Cell* 53: 672–681. <https://doi.org/10.1016/j.molcel.2013.12.029>
- Lieberman-Aiden E., N. L. Van Berkum, L. Williams, M. Imakaev, T. Ragoczy, *et al.*, 2009 Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* (80-.). 326: 289–293. <https://doi.org/10.1126/science.1181369>
- Lim B., T. Heist, M. Levine, and T. Fukaya, 2018 Visualization of Transvection in Living *Drosophila* Embryos. *Mol. Cell* 70: 287–296.e6. <https://doi.org/10.1016/j.molcel.2018.02.029>
- Lomvardas S., G. Barnea, D. J. Pisapia, M. Mendelsohn, J. Kirkland, *et al.*, 2006 Interchromosomal interactions and olfactory receptor choice. *Cell* 126: 403–413. <https://doi.org/10.1016/j.cell.2006.06.035>
- Maeder M. L., B. J. Polansky, B. E. Robson, and D. A. Eastman, 2007 Phylogenetic footprinting analysis in the upstream regulatory regions of the *Drosophila* enhancer of split genes. *Genetics* 177: 1377–94. <https://doi.org/10.1534/genetics.107.070425>
- Mahmoudi T., K. R. Katsani, and C. P. Verrijzer, 2002 GAGA can mediate enhancer function in trans by linking two separate DNA molecules. *EMBO J* 21: 1775–1781. <https://doi.org/10.1093/emboj/21.7.1775>
- Maier D., B. M. Marte, W. Schäfer, Y. Yu, and A. Preiss, 1993 *Drosophila* evolution challenges postulated redundancy in the E(spl) gene complex. *Proc. Natl. Acad. Sci. U. S. A.* 90: 5464–8.
- Maksimenko O., M. Bartkuhn, V. Stakhov, M. Herold, N. Zolotarev, *et al.*, 2015 Two new insulator proteins, Pita and ZIPIC, target CP190 to chromatin. *Genome Res.* 25: 89–99. <https://doi.org/10.1101/gr.174169.114>
- Markstein M., C. Pitsouli, C. Villalta, S. E. Celniker, and N. Perrimon, 2008 Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nat. Genet.* 40: 476–483. <https://doi.org/10.1038/ng.101>
- Marshall W. F., A. F. Dernburg, B. Harmon, D. A. Agard, and J. W. Sedat, 1996 Specific interactions of chromatin with the nuclear envelope: positional determination within the nucleus in *Drosophila melanogaster*. *Mol Biol Cell* 7: 825–842.
- Marshall W. F., A. Straight, J. F. Marko, J. Swedlow, A. Dernburg, *et al.*, 1997 Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr Biol* 7: 930–939.
- Martínez-Laborda A., A. González-Reyes, and G. Morata, 1992 Trans regulation in the Ultrabithorax gene of *Drosophila*: alterations in the promoter enhance transvection. *EMBO J.* 11: 3645–52.
- McKee B. D., 2004 Homologous pairing and chromosome dynamics in meiosis and mitosis. *Biochim Biophys Acta* 1677: 165–180. <https://doi.org/10.1016/j.bbaexp.2003.11.017>
- Mellert D. J., and J. W. Truman, 2012 Transvection is common throughout the *Drosophila* genome. *Genetics* 191: 1129–1141. <https://doi.org/10.1534/genetics.112.140475>
- Melnikova L., F. Juge, N. Gruzdeva, A. Mazur, G. Cavalli, *et al.*, 2004 Interaction between the GAGA factor and Mod(mdg4) proteins promotes insulator bypass in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 101: 14806–14811. <https://doi.org/10.1073/pnas.0403959101>
- Metz C. W., 1916 Chromosome studies in Diptera II. The paired association of chromosomes in the Diptera and its significance. *J. Exp. Zool.* 21: 213–279.
- Miller S. W., M. Rebeiz, J. E. Atanasov, and J. W. Posakony, 2014 Neural precursor-specific expression of multiple *Drosophila* genes is driven by dual enhancer modules with overlapping function. *Proc. Natl. Acad. Sci.* 111: 17194–17199. <https://doi.org/10.1073/pnas.1415308111>
- Modolell J., and S. Campuzano, 1998 The achaete-scute complex as an integrating device. *Int. J. Dev. Biol.* 42: 275–82. <https://doi.org/10.1387/IJDB.9654009>

- Monastiriotti M., N. Giagtzoglou, K. A. Koumbanakis, E. Zacharioudaki, M. Deligiannaki, *et al.*, 2010 *Drosophila* Hey is a target of Notch in asymmetric divisions during embryonic and larval neurogenesis. *Development* 137: 191–201. <https://doi.org/10.1242/dev.043604>
- Moon H., G. Filippova, D. Loukinov, E. Pugacheva, Q. Chen, *et al.*, 2005 CTCF is conserved from *Drosophila* to humans and confers enhancer blocking of the Fab-8 insulator. *EMBO Rep.* 6: 165–70. <https://doi.org/10.1038/sj.embor.7400334>
- Morris J. R., J. Chen, S. T. Filandrinis, R. C. Dunn, R. Fisk, *et al.*, 1999a An analysis of transvection at the yellow locus of *Drosophila melanogaster*. *Genetics* 151: 633–651.
- Morris J. R., P. K. Geyer, and C. T. Wu, 1999b Core promoter elements can regulate transcription on a separate chromosome in trans. *Genes Dev* 13: 253–258.
- Morris J. R., D. A. Petrov, A. M. Lee, and C. T. Wu, 2004 Enhancer choice in cis and in trans in *Drosophila melanogaster*: Role of the promoter. *Genetics* 167: 1739–1747. <https://doi.org/10.1534/genetics.104.026955>
- Muller M., K. Hagstrom, H. Gyurkovics, V. Pirrotta, and P. Schedl, 1999 The mcp element from the *Drosophila melanogaster* bithorax complex mediates long-distance regulatory interactions. *Genetics* 153: 1333–1356.
- Müller H. P., J. M. Sogo, W. Schaffner, H. P. Mueller-Sturm, J. M. Sogo, *et al.*, 1989 An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. *Cell* 58: 767–777. [https://doi.org/10.1016/0092-8674\(89\)90110-4](https://doi.org/10.1016/0092-8674(89)90110-4)
- Muravyova E., A. Golovnin, E. Gracheva, A. Parshikov, T. Belenkaya, *et al.*, 2001 Loss of insulator activity by paired Su(Hw) chromatin insulators. *Science* 291: 495–8. <https://doi.org/10.1126/science.291.5503.495>
- Nagel A. C., D. Maier, S. Krauß, M. Mezger, and A. Preiss, 2004 Neurogenic phenotypes induced by RNA interference with bHLH genes of the Enhancer of split complex of *Drosophila melanogaster*. *Genesis* 39: 105–114. <https://doi.org/10.1002/gene.20033>
- Negre N., C. D. Brown, P. K. Shah, P. Kheradpour, C. A. Morrison, *et al.*, 2010 A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet* 6: e1000814. <https://doi.org/10.1371/journal.pgen.1000814>
- Nellesen D. T., E. C. Lai, and J. W. Posakony, 1999 Discrete enhancer elements mediate selective responsiveness of enhancer of split complex genes to common transcriptional activators. *Dev. Biol.* 213: 33–53. <https://doi.org/10.1006/dbio.1999.9324>
- Nishikawa J., and T. Ohyama, 2013 Selective association between nucleosomes with identical DNA sequences. *Nucleic Acids Res* 41: 1544–1554. <https://doi.org/10.1093/nar/gks1269>
- Nolo R., L. A. Abbott, and H. J. Bellen, 2000 Senseless, a Zn finger transcription factor, is necessary and sufficient for sensory organ development in *Drosophila*. *Cell* 102: 349–62.
- Oellers N., M. Dehio, and E. Knust, 1994 bHLH proteins encoded by the Enhancer of split complex of *Drosophila* negatively interfere with transcriptional activation mediated by proneural genes. *MGG Mol. Gen. Genet.* 244: 465–473. <https://doi.org/10.1007/BF00583897>
- Ohsako S., J. Hyer, G. Panganiban, I. Oliver, and M. Caudy, 1994 Hairy function as a DNA-binding helix-loop-helix repressor of *Drosophila* sensory organ formation. *Genes Dev.* 8: 2743–55. <https://doi.org/10.1101/gad.8.22.2743>
- Olivas W., and R. Parker, 2000 The Puf3 protein is a transcript-specific regulator of mRNA degradation in yeast. *EMBO J.* 19: 6602–6611. <https://doi.org/10.1093/emboj/19.23.6602>
- Parkhurst S. M., D. A. Harrison, M. P. Remington, C. Spana, R. L. Kelley, *et al.*, 1988 The *Drosophila* su(Hw) gene, which controls the phenotypic effect of the gypsy transposable element, encodes a putative DNA-binding protein. *Genes Dev.* 2: 1205–1215. <https://doi.org/10.1101/gad.2.10.1205>
- Parnell T. J., M. M. Vierung, A. Skjesol, C. Helou, E. J. Kuhn, *et al.*, 2003 An endogenous suppressor of hairy-wing insulator separates regulatory domains in *Drosophila*. *Proc Natl Acad Sci U S A* 100: 13436–13441. <https://doi.org/10.1073/pnas.2333111100>
- Parnell T. J., E. J. Kuhn, B. L. Gilmore, C. Helou, M. S. Wold, *et al.*, 2006 Identification of genomic sites that bind the *Drosophila* suppressor of Hairy-wing insulator protein. *Mol Cell Biol* 26: 5983–5993. <https://doi.org/10.1128/MCB.00698-06>
- Paroush Z., R. L. Finley, T. Kidd, S. M. Wainwright, P. W. Ingham, *et al.*, 1994 Groucho is required for *Drosophila* neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. *Cell* 79: 805–15.
- 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>
- Pirrotta V., H. Steller, and M. P. Bozzetti, 1985 Multiple upstream regulatory elements control the expression of the *Drosophila* white gene. *EMBO J.* 4: 3501–3508.
- Postika N., M. Metzler, M. Affolter, M. Müller, P. Schedl, *et al.*, 2018 Boundaries mediate long-distance interactions between enhancers and promoters in the *Drosophila* Bithorax complex, (M. C. Gambetta, Ed.). *PLOS Genet.* 14: e1007702. <https://doi.org/10.1371/journal.pgen.1007702>
- Preiss A., D. A. Hartley, and S. Artavanis-Tsakonas, 1988 The molecular genetics of Enhancer of split, a gene required for embryonic neural development in *Drosophila*. *EMBO J.* 7: 3917–27.
- Rabani M., M. Kertesz, and E. Segal, 2008 Computational prediction of RNA structural motifs involved in posttranscriptional regulatory processes. *Proc. Natl. Acad. Sci.* 105: 14885–14890. <https://doi.org/10.1073/pnas.0803169105>
- Ramos E., D. Ghosh, E. Baxter, and V. G. Corces, 2006 Genomic organization of gypsy chromatin insulators in *Drosophila melanogaster*. *Genetics* 172: 2337–2349. <https://doi.org/10.1534/genetics.105.054742>
- Ringrose L., 2009 Transgenesis in *Drosophila melanogaster*. *Methods Mol. Biol.* 561: 3–19. https://doi.org/10.1007/978-1-60327-019-9_1
- Rong Y. S., and K. G. Golic, 2003 The homologous chromosome is an effective template for the repair of mitotic DNA double-strand breaks in *Drosophila*. *Genetics* 165: 1831–1842.
- Ronshaugen M., and M. Levine, 2004 Visualization of trans-homolog enhancer-promoter interactions at the Abd-B Hox locus in the *Drosophila* embryo. *Dev Cell* 7: 925–932. <https://doi.org/10.1016/j.devcel.2004.11.001>
- Ruby J. G., A. Stark, W. K. Johnston, M. Kellis, D. P. Bartel, *et al.*, 2007 Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. *Genome Res.* 17: 1850–1864. <https://doi.org/10.1101/gr.6597907>
- Santamaria P., 1983 Analysis of haploid mosaics in *Drosophila*. *Dev Biol* 96: 285–295.
- Sass G. L., and S. Henikoff, 1999 Pairing-dependent mislocalization of a *Drosophila* brown gene reporter to a heterochromatic environment. *Genetics* 152: 595–604.
- Schoborg T., S. Kuruganti, R. Rickels, and M. Labrador, 2013 The *Drosophila* gypsy insulator supports transvection in the presence of the vestigial enhancer. *PLoS One* 8: e81331. <https://doi.org/10.1371/journal.pone.0081331>
- Schrons H., E. Knust, and J. A. Campos-Ortega, 1992 The Enhancer of split complex and adjacent genes in the 96F region of *Drosophila melanogaster* are required for segregation of neural and epidermal progenitor cells. *Genetics* 132: 481–503.
- Schwartz Y. B., D. Linder-Basso, P. V. Kharchenko, M. Y. Tolstorukov, M. Kim, *et al.*, 2012 Nature and function of insulator protein binding sites in the *Drosophila* genome. *Genome Res.* 22: 2188–2198. <https://doi.org/10.1101/gr.138156.112.2188>
- Schwartz Y. B., and G. Cavalli, 2017 Three-Dimensional Genome Organization and Function in *Drosophila*. *Genetics* 205: 5–24.

- <https://doi.org/10.1534/genetics.115.185132>
- Sexton T., E. Yaffe, E. Kenigsberg, F. Bantignies, B. Leblanc, *et al.*, 2012 Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* 148: 458–472. <https://doi.org/10.1016/j.cell.2012.01.010>
- Shalgi R., M. Lapidot, R. Shamir, and Y. Pilpel, 2005 A catalog of stability-associated sequence elements in 3' UTRs of yeast mRNAs. *Genome Biol.* 6: R86. <https://doi.org/10.1186/gb-2005-6-10-r86>
- Sigrist C. J. A., and V. Pirrotta, 1997 Chromatin insulator elements block the silencing of a target gene by the *Drosophila* polycomb response element (PRE) but allow trans interactions between PREs on different chromosomes. *Genetics* 147: 209–221.
- Simpson P., 1990 Lateral inhibition and the development of the sensory bristles of the adult peripheral nervous system of *Drosophila*. *Development* 109: 509–19.
- Singson A., M. W. Leviten, A. G. Bang, X. H. Hua, and J. W. Posakony, 1994 Direct downstream targets of proneural activators in the imaginal disc include genes involved in lateral inhibitory signaling. *Genes Dev.* 8: 2058–71. <https://doi.org/10.1101/gad.8.17.2058>
- Sipos L., J. Mihály, F. Karch, P. Schedl, J. Gausz, *et al.*, 1998 Transvection in the *Drosophila* Abd-B domain: extensive upstream sequences are involved in anchoring distant cis-regulatory regions to the promoter. *Genetics* 149: 1031–50.
- Smith H. F., M. A. Roberts, H. Q. Nguyen, M. Peterson, T. A. Hartl, *et al.*, 2013 Maintenance of interphase chromosome compaction and homolog pairing in *Drosophila* is regulated by the condensin cap-h2 and its partner Mrg15. *Genetics* 195: 127–146. <https://doi.org/10.1534/genetics.113.153544>
- Soshnev A. A., X. Li, M. D. Wehling, and P. K. Geyer, 2008 Context differences reveal insulator and activator functions of a Su(Hw) binding region. *PLoS Genet* 4: e1000159. <https://doi.org/10.1371/journal.pgen.1000159>
- Spana C., D. A. Harrison, and V. G. Corces, 1988 The *Drosophila* melanogaster suppressor of Hairy-wing protein binds to specific sequences of the gypsy retrotransposon. *Genes Dev* 2: 1414–1423.
- Spana C., and V. G. Corces, 1990 DNA bending is a determinant of binding specificity for a *Drosophila* zinc finger protein. *Genes Dev* 4: 1505–1515.
- Spilianakis C. G., M. D. Lalioti, T. Town, G. R. Lee, and R. A. Flavell, 2005 Interchromosomal associations between alternatively expressed loci. *Nature* 435: 637–645. <https://doi.org/10.1038/nature03574>
- Stark J. M., and M. Jasin, 2003 Extensive loss of heterozygosity is suppressed during homologous repair of chromosomal breaks. *Mol Cell Biol* 23: 733–743.
- Stevens N. M., 1908 A study of the germ cells of certain Diptera, with reference to the heterochromosomes and the phenomena of synapsis. *J. Exp. Zool.* 5: 359–374.
- Stout K., S. van der Maarel, R. R. Frants, G. W. Padberg, H. H. Ropers, *et al.*, 1999 Somatic pairing between subtelomeric chromosome regions: implications for human genetic disease? *Chromosom. Res* 7: 323–329.
- Stratigi K., M. Kapsetaki, M. Aivaliotis, T. Town, R. A. Flavell, *et al.*, 2015 Spatial proximity of homologous alleles and long noncoding RNAs regulate a switch in allelic gene expression. *Proc. Natl. Acad. Sci.* 1–10. <https://doi.org/10.1073/pnas.1502182112>
- Sun S., Y. Fukue, L. Nolen, R. Sadreyev, and J. T. Lee, 2010 Characterization of Xpr (Xpct) reveals instability but no effects on X-chromosome pairing or Xist expression. *Transcription* 1: 46–56. <https://doi.org/10.4161/trns.1.1.12401>
- 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.
- Thatcher K. N., S. Peddada, D. H. Yasui, and J. M. Lasalle, 2005 Homologous pairing of 15q11-13 imprinted domains in brain is developmentally regulated but deficient in Rett and autism samples. *Hum Mol Genet* 14: 785–797. <https://doi.org/10.1093/hmg/ddi073>
- Thibault S. T., M. A. Singer, W. Y. Miyazaki, B. Milash, N. A. Dompe, *et al.*, 2004 A complementary transposon tool kit for *Drosophila* melanogaster using P and piggyBac. *Nat. Genet.* 36: 283–287. <https://doi.org/10.1038/ng1314>
- Thorpe H. M., and M. C. Smith, 1998 In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc. Natl. Acad. Sci. U. S. A.* 95: 5505–10.
- Thurmond J., J. L. Goodman, V. B. Strelets, H. Attrill, L. S. Gramates, *et al.*, 2019 FlyBase 2.0: the next generation. *Nucleic Acids Res.* 47: D759–D765. <https://doi.org/10.1093/nar/gky1003>
- Tietze K., N. Oellers, and E. Knust, 1992 Enhancer of splitD, a dominant mutation of *Drosophila*, and its use in the study of functional domains of a helix-loop-helix protein. *Proc. Natl. Acad. Sci. U. S. A.* 89: 6152–6. <https://doi.org/10.1073/pnas.89.13.6152>
- Tsai T. F., J. Bressler, Y. H. Jiang, and A. L. Beaudet, 2003 Disruption of the genomic imprint in trans with homologous recombination at Snrpn in ES cells. *Genesis* 37: 151–161. <https://doi.org/10.1002/gene.10237>
- Vazquez J., A. S. Belmont, and J. W. Sedat, 2001 Multiple regimes of constrained chromosome motion are regulated in the interphase *Drosophila* nucleus. *Curr Biol* 11: 1227–1239.
- Vazquez J., A. S. Belmont, and J. W. Sedat, 2002 The dynamics of homologous chromosome pairing during male *Drosophila* meiosis. *Curr Biol* 12: 1473–1483.
- Vazquez J., M. Muller, V. Pirrotta, and J. W. Sedat, 2006 The Mcp element mediates stable long-range chromosome-chromosome interactions in *Drosophila*. *Mol Biol Cell* 17: 2158–2165. <https://doi.org/10.1091/mbc.E06-01-0049>
- Venken K. J. T., Y. He, R. A. Hoskins, and H. J. Bellen, 2006 P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science* 314: 1747–51. <https://doi.org/10.1126/science.1134426>
- Vieira K. F., P. P. Levings, M. A. Hill, V. J. Crusselle, S. H. Kang, *et al.*, 2004 Recruitment of transcription complexes to the beta-globin gene locus in vivo and in vitro. *J Biol Chem* 279: 50350–50357. <https://doi.org/10.1074/jbc.M408883200>
- Viets K., M. Sauria, C. Chernoff, C. Anderson, S. Tran, *et al.*, 2018 TADs pair homologous chromosomes to promote interchromosomal gene regulation. *bioRxiv* 445627. <https://doi.org/10.1101/445627>
- Vogelmann J., A. Le Gall, S. Dejardin, F. Allemand, A. Gamot, *et al.*, 2014 Chromatin insulator factors involved in long-range DNA interactions and their role in the folding of the *Drosophila* genome. *PLoS Genet* 10: e1004544. <https://doi.org/10.1371/journal.pgen.1004544>
- Wech I., S. Bray, C. Delidakis, and A. Preiss, 1999 Distinct expression patterns of different enhancer of split bHLH genes during embryogenesis of *Drosophila melanogaster*. *Dev. Genes Evol.* 209: 370–5. <https://doi.org/10.1007/s004270050266>
- Wei W., and M. D. Brennan, 2000 Polarity of transcriptional enhancement revealed by an insulator element. *Proc. Natl. Acad. Sci. U. S. A.* 97: 14518–23. <https://doi.org/10.1073/pnas.011529598>
- Wei W., and M. D. Brennan, 2001 The gypsy insulator can act as a promoter-specific transcriptional stimulator. *Mol. Cell. Biol.* 21: 7714–7720. <https://doi.org/10.1128/MCB.21.22.7714-7720.2001>
- Welshons W., 1956 Dosage sensitive interactions with split mutations in the presence of an Enhancer of split. *Drosoph. Inf Serv.* 30: 157–

- Williams B. J., E. Jones, and A. R. Brothman, 1995 Homologous centromere association of chromosomes 9 and 17 in prostate cancer. *Cancer Genet Cytogenet* 85: 143–151.
- Williams B. R., J. R. Bateman, N. D. Novikov, and C. T. Wu, 2007 Disruption of topoisomerase II perturbs pairing in drosophila cell culture. *Genetics* 177: 31–46. <https://doi.org/10.1534/genetics.107.076356>
- Williams A., C. G. Spilianakis, and R. A. Flavell, 2010 Interchromosomal association and gene regulation in trans. *Trends Genet.* 26: 188–197. <https://doi.org/10.1016/j.tig.2010.01.007>
- Wu C. T., 1993 Transvection, nuclear structure, and chromatin proteins. *J Cell Biol* 120: 587–590.
- Wu Y.-C., K.-S. Lee, Y. Song, S. Gehrke, and B. Lu, 2017 The bantam microRNA acts through Numb to exert cell growth control and feedback regulation of Notch in tumor-forming stem cells in the *Drosophila* brain, (A. W. Moore, Ed.). *PLOS Genet.* 13: e1006785. <https://doi.org/10.1371/journal.pgen.1006785>
- Wurmbach E., I. Wech, and A. Preiss, 1999 The Enhancer of split complex of *Drosophila melanogaster* harbors three classes of Notch responsive genes. *Mech. Dev.* 80: 171–80.
- Wyman C., D. Ristic, and R. Kanaar, 2004 Homologous recombination-mediated double-strand break repair. *DNA Repair* 3: 827–833. <https://doi.org/10.1016/j.dnarep.2004.03.037>
- Xu T., and G. M. Rubin, 1993 Analysis of genetic mosaics in developing and adult *Drosophila* tissues. *Development* 117: 1223–37.
- Xu N., M. E. Donohoe, S. S. Silva, and J. T. Lee, 2007 Evidence that homologous X-chromosome pairing requires transcription and Ctf protein. *Nat Genet* 39: 1390–1396. <https://doi.org/10.1038/ng.2007.5>
- Xu Z., G. Wei, I. Chepelev, K. Zhao, and G. Felsenfeld, 2011 Mapping of INS promoter interactions reveals its role in long-range regulation of SYT8 transcription. *Nat. Struct. Mol. Biol.* 18: 372–378. <https://doi.org/10.1038/nsmb.1993>
- Yaffe E., and A. Tanay, 2011 Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. *Nat Genet* 43: 1059–1065. <https://doi.org/10.1038/ng.947>
- Zarifi I., M. Kiparaki, K. A. Koumbanakis, N. Giagtzoglou, E. Zacharioudaki, *et al.*, 2012 Essential roles of Da transactivation domains in neurogenesis and in E(spl)-mediated repression. *Mol. Cell. Biol.* 32: 4534–48. <https://doi.org/10.1128/MCB.00827-12>
- Zhang F. F., D. A. Arber, T. G. Wilson, M. H. Kawachi, and M. L. Slovak, 1997 Toward the validation of aneusomy detection by fluorescence in situ hybridization in bladder cancer: comparative analysis with cytology, cytogenetics, and clinical features predicts recurrence and defines clinical testing limitations. *Clin Cancer Res* 3: 2317–2328.
- Zhou J., B. Lemos, E. B. Dopman, and D. L. Hartl, 2011 Copy-number variation: the balance between gene dosage and expression in *Drosophila melanogaster*. *Genome Biol Evol* 3: 1014–1024. <https://doi.org/10.1093/gbe/evr023>
- Zuker M., 2003 Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31: 3406–15.