HHS Public Access

Author manuscript

Dev Dyn. Author manuscript; available in PMC 2016 June 01.

Published in final edited form as:

Dev Dyn. 2015 June ; 244(6): 808–825. doi:10.1002/dvdy.24279.

Controlled expression of Drosophila homeobox loci using the *Hostile takeover* system

Naureen Javeed[*], **Nicholas J. Tardi**[*], **Maggie Maher**[*], **Swetha Singari**[*], and **Kevin A. Edwards**[*]

[*]School of Biological Sciences, Illinois State University, Normal, IL 61790, USA

Abstract

Background—*Hostile takeover* (*Hto*) is a Drosophila protein trapping system that allows the investigator to both induce a gene and tag its product. The *Hto* transposon carries a GAL4-regulated promoter expressing an exon encoding a FLAG-mCherry tag. Upon expression, the *Hto* exon can splice to a downstream genomic exon, generating a fusion transcript and tagged protein.

Results—Using rough-eye phenotypic screens, *Hto* inserts were recovered at eight homeobox or Pax loci: *cut*, *Drgx*/*CG34340*, *Pox neuro*, *araucan*, *shaven*/*D-Pax2*, *Zn finger homeodomain 2*, *Sex combs reduced* (*Scr*), and the *abdominal-A* region. The collection yields diverse misexpression phenotypes. Ectopic Drgx was found to alter the cytoskeleton and cell adhesion in ovary follicle cells. *Hto* expression of *cut*, *araucan* or *shaven* gives phenotypes similar to those of the corresponding *UAS-cDNA* constructs. The *cut* and *Pox neuro* phenotypes are suppressed by the corresponding RNAi constructs. The *Scr* and *abdominal-A* inserts do not make fusion proteins, but may act by chromatin- or RNA-based mechanisms.

Conclusions—*Hto* can effectively express tagged homeodomain proteins from their endogenous loci; the Minos vector allows inserts to be obtained even in transposon cold-spots. *Hto* screens may recover homeobox genes at high rates because they are particularly sensitive to misexpression.

Keywords

homeodomain; Pax genes; imaginal discs; misexpression screen; protein traps

Introduction

Controlled ectopic gene expression is widely used to identify and characterize developmental loci. By expressing a gene in a novel location, in the absence of its normal upstream regulators, one can determine if that gene is sufficient to trigger downstream events. This approach complements loss-of-function studies; for example, it can allow for phenotypic and genetic interaction analyses to be performed, even if the gene is not essential due to redundancy. Ectopic expression can model aspects of disease states, particularly cancers that are stimulated by expression of oncogenes at the wrong time or location. In

Corresponding author: Kevin A. Edwards, School of Biological Sciences, Campus Box 4120, Illinois State University, Normal, IL 61790, USA, Ph. (309) 438-7689, kaedwar@ilstu.edu.

Drosophila, misexpression of endogenous genes is facilitated by large collections of lines bearing *EP* and related transposons. These inserts carry *UAS* enhancers that allow downstream genes to be expressed in response to the GAL4 transcription factor. GAL4 is provided by a separate "driver" element, either a GAL4 enhancer trap or a designed GAL4 expression construct (Rørth 1996; Duffy 2002; Elliott & Brand 2008; del Valle Rodríguez *et al*. 2011). The *EP* family of transposons has some limitations: most of these constructs use the *P* element as the vector, which gives very uneven coverage of the genome (Bellen *et al*. 2011), and they do not tag the target protein. In addition, transcripts that begin in the *EP* element usually span genomic territory that is not normally part of the 5′UTR of the target gene, and this could impair expression of the protein.

We developed a system called *Hostile takeover* (*Hto*) that addresses these issues (Singari *et al*. 2014). Like *EP*, *Hto* expresses endogenous genes under GAL4 control, but it also employs a "splice-out" strategy to tag the protein product for further analyses (Fig. 1A). The addition of splicing to the GAL4/*UAS* system allows for very strong misexpression phenotypes, even when the *Hto* transposon is inserted well upstream of the target gene. To avoid the insertion site biases shown by the *P* element, *Hto* instead uses *Minos* as the vector (Bellen *et al*. 2011). Between the *Minos* inverted repeats, *Hto* carries a *UAS* enhancer with a basal promoter, followed by an artificial exon 1 and a 5′ splice site (ss). There is no suitable 3′ ss for intron 1 in the *Minos* element, and so exon 1 will be spliced to the next recognized exon in the downstream genomic DNA. Exon 1 encodes a FLAG epitope-tagged version of mCherry red fluorescent protein (RFP), ending in frame 0. Thus after splicing, the fusion transcript can encode a fusion protein, consisting of the FLAG-RFP domain at the N-terminus and the target protein at the C-terminus. The tag may be used to characterize the fusion protein by Western blot, fluorescence microscopy, chromatin immunoprecipitation, etc.

To identify inserts that cause dominant adult phenotypes, *Hto* was mobilized throughout the genome in the male germ line. In the progeny, an eye-specific GAL4 driver was used to express the new inserts, and an F1 screen for abnormal eye phenotypes was conducted. This screening method identified a diverse array of regulatory proteins, and also some regulatory RNAs, as targets (Singari *et al*. 2014). About half of the lines, though, were found to target regulatory transcription factors (TFs). Among these, genes of the homeodomain superfamily were especially well represented, perhaps because homeodomain TFs tend to function as "control switches" that can unilaterally alter cell fates when misexpressed.

The homeodomain is a compact (~60 a.a.) DNA binding domain that is employed by eleven structurally diverse classes of animal TFs, sometimes in combination with other domains (Holland *et al*. 2007). Structures and functions of various classes have been reviewed (Gehring *et al*. 1994; Cavodeassi *et al*. 2001; Robson *et al*. 2006; Maeda & Karch 2006; Hueber & Lohmann 2008; Hulea & Nepveu 2012; Ladam & Sagerström 2014). Some members have multiple homeodomains, and others have additional DNA binding elements such as the Paired (PRD) Domain/Paired Box (Pax). Homeodomain proteins tend to be expressed in tightly controlled but dynamic patterns, and they generally regulate the differentiation of those tissues by controlling suites of downstream genes.

From *Hto* screens that used *GMR-GAL4* as the driver, 23 unique target genes were characterized; five of those (22%) encode homeodomain or Pax proteins (*cut*, *CG34340*, *araucan*, *iab-8/abdominal-A*, *shaven*; Table 1). This is a strong enrichment over the fraction of homeodomain genes in the genome (~101 out of 13,955 protein coding genes, or 0.7%; Hammonds *et al*. 2013; Holland 2013). In other eye screens, we identified inserts in homeodomain genes *Scr* and *zfh2* and the Pax gene *Poxn*. Here we describe and compare phenotypes associated with these eight *Hto* inserts, discuss their molecular basis, and consider how the lines may be used for further functional studies.

Results & Discussion

Characterization of *Hto* inserts affecting Homeobox and Pax genes

We determined the insertion sites of a set of *Hto* elements that cause eye phenotypes (Singari *et al*. 2014), and eight inserts were found to target homeobox and Pax genes. The target genes all have direct orthologs (or orthologous families) in mammals, spanning five of the 11 human homeobox classes of Holland *et al*. (2007): ANTP, CUT, PRD, TALE, and ZF (Table 1). As they are recovered, each insert is given a unique 3-letter ID code in all caps; this is not a gene name, but is used as an identifier for the insert (see *Experimental Procedures*). We will refer to them by the 3-letter code here; see Table 1 for the corresponding target genes. Inserts were mapped relative to the surrounding gene structure; maps for the inserts *FAR*, *GLO*, *BRB*, *OMD*, and *BLY* are in Fig. 1C. Maps of the other inserts are in these references: *LNP*, Gummalla *et al*. (2012); *EAB*, Perea *et al*. (2013); and *BRO*, Singari *et al*. (2014).

The inserts were found to produce very diverse phenotypes upon induction. In particular, expression via *GMR-GAL4* causes a wide range of eye defects (Fig. 2A–G, N, S). *GMR-GAL4* drives *UAS* expression in the eye field as the retina differentiates. All the eyes shown in Fig. 2A–G and N–T have the same, single copy of the w^+ red-eye marker from the *GMR* construct, and so differences in color as well as texture are part of the *GMR>Hto* phenotype (genotypes are summarized in the form “Driver>Responder”; in all cases, a single copy of each element is used). The insert *OMD* yields eyes that are glazed (lacking clear separation of ommatidia) and nearly colorless, *BRB* and *BRO* are pale/mottled and glazed, *GLO* is severely rough, *EAB* is mildly glazed, *FAR* is rough (more severe in the dorsal and anterior regions of the retina), *LNP* is mild rough and mottled with some fused ommatidia near the center, and *BLY* is essentially wild type. When expressed earlier in eye development with *ey-GAL4*, the inserts often cause reduced or missing eyes (Fig. 2H–L). It has been shown that *ey-GAL4* can give an eyeless phenotype with various TF constructs, and this is not necessarily diagnostic of their functions (Jiao *et al*. 2001). However, several *Hto* lines also cause additional head defects or transformations that can be more informative (Fig. 2H, K, L and below). Expression in the wing blade via *ms1096-GAL4* causes a range of defects including small wing size, suppressed or ectopic wing veins, and a distorted wing margin (Fig. 3). Below we present each line in turn, summarizing our observations and proposing possible mechanisms of action underlying the phenotypes.

GLO targets a homolog of the mammalian neural homeobox protein DRGX/DRG11

The *GLO* insert produces severely rough eyes with *GMR-GAL4* (Fig. 2F), suppresses wing vein development (Fig. 3F), and is lethal with *ey-GAL4* (even at 18°C), producing an eyeless phenotype with a reduced head (Figs. 2J; 4B). *GLO* lies in a nested pair of genes and expresses the outer gene, *CG34340*. Of the 101 homeobox genes in *D. melanogaster*, *CG34340* is one of only 11 that remain "unnamed" (known only by its CG number), a sign that a gene is little-studied (Hammonds *et al*. 2013). The *CG34340* product has a single direct ortholog in humans, DRGX, that is 92% identical in the homeobox (Holland *et al*. 2007). The next best match in humans is ARX, the Aristaless ortholog, at 82% homeobox identity. The closest relative of *CG34340* in flies is *aristaless*, 83% identical in the homeobox. Based on this orthology, we rename *CG34340* as *Dorsal root ganglia homeobox* (*Drgx*). The mouse ortholog (called Prrxl1) is required for development of pain sensing circuits in the CNS (Chen *et al*. 2001; Rebelo *et al*. 2010; Soares-dos-Reis *et al*. 2014). Embryo in situ hybridization shows that fly Drgx is also expressed in just a subset of cells of the developing CNS (insitu.fruitfly.org; Hammonds *et al*. 2013). In the only other functional study of fly *Drgx*, Parrish *et al*. (2006) performed a large scale RNAi screen to identify TFs required for dendrite formation in the embryo. Knockdown of *Drgx* (therein called *CG2808*) was found to alter dendrite arborization and also cause muscle defects.

GLO lies in the coding region of *CG42463*, the second of four tandemly duplicated genes encoding small seminal fluid proteinase inhibitors that are nested in a *Drgx* intron (Fig. 1; Findlay *et al*. 2008). Although the *GLO* insertion disrupts *CG42463*, *GLO* homozygotes are normal and fertile. Since the closest BLAST matches to the *CG42463* protein come from its neighboring genes, *Sfp24Ba*, *b*, and *c*, and like them it shows male specific expression (Graveley *et al*. 2011; FlyBase.org), we propose to rename *CG42463* as *Sfp24Bd*. GAL4 induction of *GLO* should not affect the *Sfp24B* cluster since those genes are all transcribed in the reverse direction (Fig. 1).

RT-PCR shows that *GLO* transcripts splice to *Drgx* exon 2; the joint is …cgcaggcg/CTCGGTCCTGCAGCCG… (*Hto* exon 1 in lowercase). The resulting *GLO* protein has mCherry RFP fused to Drgx at this position: …delykrpqa/*LGPAAAGRG***MFCYQCP**…, where 'delyk' is the end of mCherry, 'rpqa' is encoded at the end of *Hto* exon 1, and the slash is the junction from *Hto* exon 1 to *Drgx* sequence (CAPS). The 5′UTR in *Drgx* exon 2 provides a 9 a.a. linker (italics) before the start of Drgx protein (bold). The *GLO* fusion is predicted to be 92.6 kDa, and migrates at ~110–120 kDa on an anti-FLAG Western blot (Fig. 1B). Mammalian DRGX also migrates more slowly than predicted (Rebelo *et al*. 2007).

As part of the basic pipeline for analysis of new *Hto* inserts, we express them in clonal patches of follicle cells (FC) in the ovary using the *FLP*-out system (Fig. 5; Pignoni & Zipursky 1997). After the FC epithelium ceases proliferation, the FCs continue to enlarge due to endoreplication, to provide enough mass to cover the growing oocyte and produce the eggshell (Horne-Badovinac & Bilder 2005). The oocyte FCs (OFCs) differentiate as a stereotypical monolayer epithelium, with an apical brush border facing the oocyte and a basal side attached to ECM. The basal cortex typically acquires parallel bundles of f-actin;

these make a corset-like structure that assists in egg elongation. Their uniform organization and large size make the FCs especially well-suited for documenting the subcellular localization of *Hto* fusion proteins. Use of *FLP*-out clones further allows us to detect alterations in the cell biology of the *Hto*-expressing FC compared to their wild type neighbors. Typical stains used for this purpose include SYBR Green for DNA, phalloidin for f-actin, and/or wheat germ agglutinin (WGA; mainly detects nuclear envelope, certain vesicles, and eggshell material).

Upon expression in the FC epithelium, the *GLO* fusion to Drgx showed strong nuclear localization as expected, with a general mottled pattern in the nucleus (Fig. 5). However, *GLO* also altered cell behaviors and tissue morphology in a very distinctive manner. In wild type, the FC epithelium evenly covers the germ cells of the egg chamber up to stage 8 (50% of the FCs lie in the posterior half). From stage 8–10, there is a relative shift such that 95% of the FCs come to lie in the posterior half, coincident with the underlying oocyte (Horne-Badovinac & Bilder 2005; Grammont 2007; Kolahi *et al*. 2009). This has been considered as either a migration of the OFCs posteriorly, or as a passive response to the growth of the oocyte toward the anterior (Kolahi *et al*. 2009). Video evidence shows the OFCs are not entirely passive, since the centrally located OFCs translocate posteriorly relative to the nurse cells during stage 9 (Poukkula *et al*. 2011). Anterior *FLP*-out clones expressing *GLO* fail to translocate posteriorly like their wild type OFC neighbors, and remain “stranded” in clusters over the nurse cells (Fig. 5A–B). The nurse cell FC (NCFC or stretch cells) of the anterior normally spread very flat to accommodate the growth of the germ cells (Kolahi *et al*. 2009; Brigaud *et al*. 2015). The *GLO* expressing cells over the nurse cells also flatten in a stage-appropriate manner, but do not spread as well (Fig. 5A–B). A stranded-FC phenotype has been noted previously for *cut*-overexpressing clones, but this is likely through a different mechanism; *cut* overexpressing cells strongly constrict their apices while the *GLO* cells do not (Levine *et al*. 2010). Anterior FC that are deficient in Notch or Dpp signaling are also impaired in spreading/posterior translocation (Grammont 2007; Brigaud *et al*. 2015).

In posterior OFC clones expressing *GLO*, the basal circumferential actin cables become highly disorganized; they are normally aligned perpendicular to the A-P axis (Fig. 5C; Delon & Brown 2009). Posterior clones tend to lack the hexagonal tiled pattern of wild type OFC and instead show smooth clone borders, suggesting increased adhesion among *GLO* cells relative to WT neighbors (compare Fig. 5D to wild type in, for example, 5B). This is also reflected in the apical-basal organization of clones: in cross sectional views, the clones’ apical surfaces usually bulge apically (pressing into the oocyte), and invaginate basally around the clone border (arrowheads in Fig. 5E–G). These irregularities of the FC layer are never observed in WT, and they are not associated with cell death, since DNA fragmentation is never observed via SYBR Green staining.

The *GLO* phenotype indicates that *Drgx* regulates specific genes involved in adhesion and/or cytoskeletal organization. Interestingly, mouse DRGX regulates the expression of the secreted signaling protein Repulsive Guidance Molecule b (RGMb)/DRAGON, which is involved in migration and adhesion in neurons and non-neural cells (Samad *et al*. 2004; Severyn *et al*. 2009; Conrad *et al*. 2010). While the RGM family is not present in the Diptera (based on BLAST analyses), fly *Drgx* might regulate functionally analogous target

genes. The *GLO* fusion displays sequence-specific DNA binding, as evidenced by its accumulation on specific polytene chromosome sites (Fig. 6, discussed below). The *Drgx* locus is a P element coldspot, and there are no other transposons reported that can drive Drgx protein expression, so the *GLO* line should prove useful in further *Drgx* functional studies (all assessments of other known inserts are based on FlyBase.org map data visualized in GBrowse; St. Pierre *et al*. 2014).

BRO expresses a fusion to the Cut protein

Cut belongs to the Cux family of genes; these have tumor suppressor properties from Drosophila to humans, regulating differentiation, apoptosis, and motility (Hulea & Nepveu 2012; Zhai *et al*. 2012; Wong *et al*. 2014). Fly Cut is required for development of numerous structures including the wing margin, external sensory organs, antenna, and egg chambers (Jack *et al*. 1991; Jackson & Blochlinger 1997; Nepveu 2001). The *BRO* transcript splices across 51,377 nt to reach exon 3 of *ct*, the longest splice so far observed among *Hto* lines (Singari *et al*. 2014). Over 25 kb of that distance is not part of any forward transcription unit, showing that long novel introns can be generated by *Hto* expression, and such introns can be compatible with strong dominant phenotypes. Regulatory elements for *ct* extend even further upstream, including a wing margin enhancer mapped to ~50 kb upstream of *BRO* (Jack *et al*. 1991). Besides *BRO*, there are no other constructs inserted in the >75 kb span between the wing margin enhancer and the *ct* promoter. *Hto* includes a site for the I-*Sce*I endonuclease, which could be used to make targeted chromosome breaks in this area.

In the RT-PCR product we recovered from *BRO* flies, *Hto* skips exon 2 (encoding 18 a.a.), and splices cleanly to exon 3 (encoding 12 a.a.) The transcript fusion joint is …cgcaggcg/GTAATACTCCCACTTC… (*Hto* exon 1 in lowercase). Exon 3 may be preferred due to its strong consensus splice sites (…cag/exon3/gtaagt…), while exon 2 has relatively poor splice sites (…tag/exon2/gttcta…). Both exon 2 and 3 are commonly skipped according to RNAseq data (Graveley *et al*. 2011), and their coding regions are not conserved across insects. Exon 1 encodes an additional 41 a.a. that is bypassed by *Hto*. The exon 1 sequence is well-conserved across insects, but not conserved beyond insects (based on BLAST analysis), and is distant from the conserved DNA binding domains. Overall, the *Hto* product is missing the first 41 (ct-PA) or 59 a.a. (ct-PC), but in either case includes ~98% of the Cut protein. The fusion occurs at the sequence …delykrpqa/**V_{60}ILPLHC**… As described above, ‘delykrpqa’ is encoded at the end of *Hto* exon 1, and the slash is the junction from *Hto* to Cut (isoform ct-PC; bold).

In the eye-antennal disc, *ct* becomes expressed in the antennal portion early in development (in the 2nd larval instar) and is absent from the eye field (Duong *et al*. 2008). Misexpression of *ct* early in the eye portion of the disc suppresses *eyeless* (*ey*) and eye formation, and can convert eye tissue to an ectopic antenna (Duong *et al*. 2008; Wang & Sun 2012). Expression of *BRO* has the same effect; the large majority of *ey>BRO* animals die as pharate adults with no eyes, and upon dissection a majority of these have some evidence of eye to antenna transformation, up to a full ectopic antenna in a minority of cases (Fig. 2L, 4C). *ct* is later expressed in the cone cells of the developing retina. Broader expression with *GMR-GAL4* leads to glazed and pale eyes, and we found the same phenotype when a *UAS-ct* transgene is

expressed with *GMR*. *UAS-ct* is more severe; its *GMR* phenotype at 19°C matches that of *GMR>BRO* at 25°C (Fig. 2M–N). In the wing disc, *ct* is expressed specifically in the strip of cells at the presumptive margin (Fig. 7B; Jack *et al*. 1991). *BRO* misexpression in the wing blade region with *ms1096-GAL4* (which does not express GAL4 at the margin) gives two consistent defects: the distal part of the wing margin is strongly contracted, and in the proximal part of the wing, the alula is very reduced (Fig. 3M).

To independently confirm that *BRO* expresses *ct* RNA, we used RNAi to knock down the fusion transcript. This approach was further generalized by testing it on the *BRB* line (below); it was effective in both cases. This method is especially useful because, if the *Hto* phenotype is suppressed, it validates both the *Hto* stock and the RNAi stock (see Conclusions). We first tested a *UAS-ct-RNAi* construct (*[TRiP.JF03304]*), and found that it gives a cut-wing phenotype when expressed at the wing margin, but has no obvious phenotype on its own when expressed with *ms1096* (not shown). We further determined *UAS-ct-RNAi* was effective based on its partial suppression of the *UAS-ct* transgene; compare *ms1096>UAS-ct* with and without *UAS-ct-RNAi* (Fig. 3J–K). *UAS-ct-RNAi* shows even stronger suppression of the *ms1096>BRO* phenotype: in *ms1096>BRO, UAS-ct-RNAi* only minor defects remain such as a weak posterior margin (Fig. 3M–N). This RNAi suppression of *BRO* was confirmed in the retina. *GMR>UAS-ct-RNAi* has a distinct glossy/sparkling eye phenotype (Fig. 2O). When *GMR* is used to drive both *BRO* and *UAS-ct-RNAi*, the opposing effects of these inserts essentially cancel each other out, and the eye is rescued to normal morphology (but slightly reduced pigmentation; Fig. 2P). Note that the flies in Fig. 2N–P are sibs from the same vial. We interpret this to mean that *UAS-ct-RNAi* brings *ct* transcript levels down to a tolerable level, while *BRO* still provides enough *ct* activity to rescue the *ct* loss-of-function (LOF) defect. Together, the results show that these *BRO* phenotypes are indeed due to legitimate *ct* activity, but in excess or in the wrong cells.

Finally, we demonstrated that *BRO* induces Ct protein production by staining tissues with anti-Ct monoclonal antibody. We used both ovary and late larval samples, which have various internal controls for Ct staining. *BRO* was expressed in clones in the ovary FC epithelium with the *FLP*-out system (Fig. 7A). The clones showed the RFP signal from the *Hto* transcripts, and strong antibody staining in the nucleus with some cell-to-cell variability; the antibody signal completely coincided with RFP. With this version of the *Hto* vector, there is some expression of unfused RFP due to a polyA signal in the Minos vector (Singari *et al*. 2014). The unfused RFP gives an unpatterned haze in the cytoplasm and nucleus, and this aspect of the RFP signal is not coincident with the antibody signal, as expected. We next drove patterned expression of *BRO* in larval imaginal tissues using *ptc-GAL4*. Whereas control wing discs had only the endogenous Ct stripe at the developing wing margin (D/V boundary; Fig. 7B), *ptc>BRO* discs had a second, perpendicular stripe with both RFP and anti-Ct signal, in the *ptc* domain along the A/P boundary (Fig. 7C). Notably, the *ptc>BRO* stripe also acquires an ectopic inward fold along its length, indicating that Ct can alter tissue morphology when placed in a new setting. The *ptc>BRO* flies were pupal lethal, so the effect on the wing was not followed further, but *ptc>BRO* mature legs could be recovered, and these showed severe morphological changes (Fig. 4M).

The *BRB* product contains the C-terminal portion of the Pax protein Poxn

BRB was recovered due to a glazed eye phenotype in the presence of *GMR-GAL4* (Fig. 2S). Ectopic *BRB* across the wing blade (*ms1096>BRB*) produces extra vein material, especially between L3–L4 and near L5, unlike the other lines (Fig. 3L). *BRB* expression in a dorsal stripe with *pannier* (*pnr*)-*GAL4* severely impairs thorax development and eliminates abdominal pigment bands (Fig. 4I–J). Although the thorax does not expand properly, the number of macrochaetae in the *pnr* zone of the thorax roughly doubles. *BRB* lies in the Pax gene *Poxn*, which regulates aspects of PNS and limb development (Vervoort *et al*. 1995; Balczarek *et al*. 1997; Awasaki & Kimura 2001; Robson *et al*. 2006; Blake *et al*. 2008). Ectopic, full-length Poxn has earlier been shown to produce rough eyes (Jiao *et al*. 2001). While most Pax members contain a complete or partial C-terminal homeobox, *Poxn* has lost the homeobox (Holland 2013). Instead it binds DNA only by the Paired domain, which is composed of two tandem helix-turn-helix subdomains, termed “PAI” and “RED” (Jun & Desplan 1996). Compared to humans, the Poxn Paired domain is most similar to that of Pax5 (74% identity), which is important in B-cell differentiation and related leukemias (O’Brien *et al*. 2011). However, the Poxn Paired domain is also 66–72% identical to all the other human Pax Paired domains except for Pax4 (57%). Poxn has about same level of conservation with the other Pax genes in the fly (70% identical in the Paired domain to *sv*; 67% to *Pox meso*; 65% to *ey*).

The strong effects of *BRB* provide an avenue to dissect structure/function relationships in this “simplified” Pax protein. *BRB* is inserted in the last intron and is predicted to splice to the last exon of *Poxn*, which contains the bulk of the coding region (Fig. 1). In the *BRB* fusion, the FLAG-RFP tag bypasses the first (PAI) helix-turn-helix subdomain and fuses to the start of the second (RED) subdomain. The fusion occurs at the sequence … delykrpqa/**Q_{79}VATPTV**… The two subdomains of a Paired Box are capable of independent and even opposing actions, as shown for PAX6 (Walcher *et al*. 2013). Thus the *BRB* fusion protein may still bind DNA, and regulate at least a subset of *Poxn* target genes, even without the PAI subdomain. We found that the *BRB* fusion to Poxn is indeed responsible for the *GMR* eye phenotype: coexpression of *BRB* with *Poxn* RNAi almost completely eliminates both the eye defect (Fig. 2S–T), and the *BRB* fusion protein (not shown). This *Poxn-RNAi* construct targets the exon to which *BRB* splices. *GMR>Poxn-RNAi* has no eye phenotype on its own (not shown).

The *GMR>BRB* retinal phenotype closely matches that of *GMR>BRO*; they are both pale and glazed, although *GMR>BRB* eyes are consistently larger (Fig. 2N, S). Poxn is known to regulate the *ct* gene during PNS development, and it binds to a *ct* enhancer even in non-neural tissue (Vervoort *et al*. 1995). Thus we used our assays for Ct misexpression to determine if *BRB* can regulate *ct. GMR>BRO, BRB* flies are qualitatively very similar to either insert alone, and only slightly more severe (Fig. 2Q), suggesting they could act in the same pathway. However, *ct* is not required for the *GMR>BRB* pale/glazed eye phenotype, nor for the *ms1096>BRB* wing phenotype, since they are not suppressed by *ct-RNAi* (Fig. 2Q–R; Fig. 3L,O). Also, when we express *BRB* using *ptc-GAL4*, we do not see ectopic activation of Ct protein in the *ptc* stripe (Fig. 7D), and the adult leg phenotype differs between *ptc>BRB* and *ptc>BRO* (Fig. 4K–M). This agrees with Vervoort *et al*. (1995), who

suggested that ectopic Poxn would likely require additional factors to activate *ct* broadly outside the PNS.

OMD expresses a fusion to full-length Shaven/D-Pax2 protein

Of all the characterized *Hto* lines, *OMD* has the most severe phenotypes across both eye and wing drivers (Singari *et al*. 2014): glazed white eyes in *GMR>OMD* (Fig. 2B), severely stunted wings in *ms1096>OMD* even at 18°C (Fig. 3B) and a mispatterned dorsal head in *ey>OMD* (Fig. 4D). This 4th chromosome insert expresses D-Pax2, the closest fly homologue to Pax2, encoded by the *sv* (AKA *sparkling*) locus. D-Pax2 is normally expressed in the sensory organ precursor cells and later in a subset of their progeny (Johnson *et al*. 2011), and in the cone and 1° pigment cells during retinal development (Fu & Noll 1997; Dziedzic *et al*. 2009).

OMD lies in an intron of the adjacent gene *Activin-beta* (Fig. 1). Zhu *et al*. (2008) found that *Activin-beta* mutants are mostly lethal with a few weak escapers; in contrast, homozygous *OMD* flies are healthy and appear normal. Thus we conclude *OMD* does not substantially compromise *Activin-beta* expression. Besides *OMD*, there are no other *UAS* insertions reported in *sv* or its upstream region. The *OMD* transcript was found by RT-PCR to splice over 13 kb into exon 2 of *sv*, which carries the D-Pax2 start codon. The transcript fusion joint is …cgcaggcg/GAACTTATACATTTTG… (*Hto* exon 1 in lowercase). The protein fusion joint is …delykrpqa/*ELIHFVESL***MLIMDIQ**… Because the *Hto* exon is in the same frame as the *sv* start codon, the 5′ UTR portion of exon 2 provides an additional 9 a.a. linker (italics) between FLAG-RFP and the full length D-Pax2 (bold). The major fusion product runs at ~120 kDa on anti-FLAG Western blots, which matches the 122 kDa size predicted for a fusion to the full-length isoform sv-PA (Fig. 1B). However, the major product could also be the Sv-PC isoform, which skips exon 11. That fusion protein would be only 101 kDa, but we have found several other cases where *Hto* tagged TFs run more slowly than predicted (Singari *et al*. 2014; *GLO* above). *GMR>OMD* yields the same, distinctive adult eye phenotype as *GMR* driving a *UAS-DPax2* construct (Kavaler *et al*. 1999; Kavaler, personal comm.), which indicates that *OMD* acts by simple misexpression of *sv*.

FAR expresses a fusion to the Iroquois/Irx family protein Ara

ara is one of a cluster of three homeobox genes (the Iro-C) that are the founding members of the Iroquois/Irx family (Gomez-Skarmeta *et al*. 1996; Kerner *et al*. 2009). The three fly genes have largely overlapping patterns of expression, and partially overlapping functions (Ikmi *et al*. 2008; Singh *et al*. 2012). Iro/Irx family members are known to play key roles as intermediates in specification and patterning of diverse tissues and cell types including the fly eye, wing, macrochaetae, and muscles, as well as the nervous system and heart in vertebrates (Cavodeassi *et al*. 2001; Gomez-Skarmeta & Modolell 2002; Carrasco-Rando *et al*. 2011). *FAR* transcripts are predicted to splice to *ara* exon 2, bypassing the first 17 aa which are encoded by exon 1 (Fig. 1). The resulting fusion joint would be … delykrpqa/**L_{18}LPPSVQ**… (bold, Ara sequence). The missing N-terminus is not conserved beyond the Diptera and Hymenoptera. There are no other transposons reported in the *ara* locus that can drive expression of full- or near-full length protein.

Phenotypically, the *FAR* insert behaves in much the same manner as a *UAS-ara* transgene (Gomez-Skarmeta *et al*. 1996; Singh *et al*. 2012), indicating it provides wild type Ara activity despite the *Hto* tag replacing exon 1. *UAS-ara* and *FAR* each caused a reduction of the ventral eye when expressed early in development with *ey-GAL4* (compare Singh *et al*. 2012 and Singari *et al*. 2014). When expressed in the wing disc, *UAS-ara* and *FAR* each led to fusion of veins L2 and L3, corresponding to the role of Ara/Caup in positioning L3; both lines also reduced overall wing size (Fig. 3D–E here vs. Fig. 6C of Gomez-Skarmeta *et al*. 1996).

EAB expresses a full-length *zfh2* fusion

At 332 kDa, Zfh2 is one of the ten largest transcription factors in *D. melanogaster*. Its 3 homeodomains are related to the LIM-class homeodomains, and are arranged in a conserved fashion, collinear with the first 3 (of 4) homeodomains in the human orthologs *ZFHX3/4* (Lundell & Hirsh 1992; Holland *et al*. 2007). The *EAB* insert lies upstream of *zfh2* transcription start, and *EAB* transcripts splice into *zfh2* at exon 3 according to RT-PCR results; the joint is …cgcaggcg/AAAACTCCTAGTCCGG…. Splicing of *Hto*-exon 3-exon 4 creates a 27 a.a. linker, encoded by the 5′ UTR, between the RFP and Zfh2 segments (details in Perea *et al*., 2013). A possible reason for skipping exon 2 is that the exon 3 splice acceptor is a much better match to the consensus than that of exon 2 (bold, polypyrimidine tract to base −25; slash, splice junction):

zfh2 Intron 1/exon 2: …aa**ttgttcacaaaaattactctt**tag/G…

zfh2 Intron 2/exon 3: …**tct**aaaa**tt**aa**ttttttttctttc**ag/A…

The *EAB* fusion was earlier shown to include a folded and functional Zfh2 segment by two criteria: anti-Zfh2 detects the fusion protein in nuclei upon GAL4 expression, and the fusion provides partial rescue of *zfh2* RNAi phenotypes (Perea *et al*. 2013). Zfh2 is critical for patterning the wing and leg (Perea *et al*. 2013; Guarner *et al*. 2014). Wing disc misexpression of *EAB* with *ms1096-GAL4* severely reduces growth of the wing blade (Fig. 3C), similar to previous results with other wing drivers; this was shown to result from Zfh2 repressing proliferative effects of the Hippo/Yorkie pathway (Perea *et al*. 2013). In the leg disc, development of the tarsal joints is disrupted by either loss (GAL4>*zfh2* RNAi) or mispatterning (GAL4>*EAB*) of Zfh2 expression (Guarner *et al*. 2014). Likewise we find that expression of *EAB* with *patched* (*ptc*)-*GAL4* prevents the proper extension of the leg, and the tarsal region is not properly patterned or segmented (Fig. 4N). Eye expression of *EAB* is also very deleterious, yielding a rough/glossy eye with *GMR-GAL4*, and a very reduced eye with *ey-GAL4* (Fig. 2E,I).

FLP-out clones expressing *EAB* were made in the ovary FC epithelium. The *EAB* product is strongly concentrated in the nucleus (Fig. 8A). *EAB FLP*-out clones grow to typical sizes, with no observed cell death, indicating that in this context *EAB* fusion protein is not sufficient to block cell proliferation or to drive apoptosis (Fig. 8). The shapes of the OFC are generally normal, unlike the case for *GLO*-expressing clones (Fig. 5). However, the nuclei are often displaced apically during mid-oogenesis when the OFC are columnar (Fig. 8B). To determine if this is due to altered cell polarity, we marked the apical membrane with a ubiquitously expressed *GFP-bazooka* (*baz*) protein trap (Buszczak *et al*. 2007). GFP-Baz is

exclusively at the apical side in both *EAB* expressing clones and their wild type neighbors (Fig. 8C). The microvilli, and the deposition of the eggshell, are also properly placed at the apical side. Together, these features show that cell polarity is maintained.

At stage 10, the staining of eggshell material by WGA becomes much more intense at the apical surface of *EAB* expressing cells than their WT neighbors (Fig. 8D). This phenotype is unique to *EAB*, and implies that the secretion of eggshell is somehow enhanced by the ectopic Zfh2 protein. To mark the lateral cell membranes, we used another ubiquitously expressed standard GFP protein trap, *GFP-ATPalpha* (GFP spliced into the main subunit of the Na pump; Morin et al. 2001). The Na pump is localized along the lateral membranes in both *EAB*-expressing and WT cells, confirming their normal polarization. However, the GFP-Na pump signal is much reduced at the membrane, and slightly increased in the interior of the cell, in the *EAB* expressing clones compared to their wild type neighbors (based on mean pixel values in confocal slices at the level of the nuclei; Fig. 8E–F). This suggests that the proper trafficking of the Na pump is reduced, even while the trafficking of eggshell material apically is increased. In future work, it may be useful to explore whether Zfh2 regulates vesicular traffic as part of its normal function.

The *BLY* insert lies in *Scr* and produces a *Deformed* (*Dfd*)-like phenotype

BLY is potentially very useful due to its location and unusual phenotype, but the molecular basis of the phenotype remains unclear. *ey>BLY* flies have a strong deformation of the head: the eyes are smaller, set back posteriorly, and bulging out laterally (Fig. 4A vs. E). The aristae are often curled instead of flat. A very similar head phenotype was documented with the original dominant allele of the *Dfd* Hox gene, reproduced in Fig. 4F (Bridges & Morgan 1923; see also Boube *et al*. 2014).

BLY is inserted in the Antennapedia Complex (ANT-C), 4.6 kb upstream of the *Scr* transcription unit (Fig. 1). The *BLY* transcript is predicted to splice to exon 2 of *Scr*, but the *Scr* coding region is in frame +1 of this exon instead of frame 0 as required for *Hto. BLY* should make RFP fused to/TPIHDNDPWL*, encoded from the 5′ UTR of *Scr*, instead of making Scr protein. This indicates *BLY* acts by a different mechanism than expressing Scr. To independently confirm this, we compared the effects of *BLY* with those of a *UAS-Scr* construct (Heuer *et al*. 1995); indeed, we found ectopic Scr bears no resemblance to *BLY*. Wings of *ms1096>BLY* are normal, but *ms1096>Scr* wings always fail to unfold after eclosion (Fig. 3H vs I). Eyes of *GMR>BLY* are normal (Fig. 2D) while *GMR>Scr* eyes are severely rough (not shown). In contrast to the bulging eyes of *ey>BLY*, *ey>Scr* flies mostly die as pupae or pharate adults, with no eyes and severely reduced head (Fig. 4E vs G).

BLY lies between *Scr*'s promoters and its more distal *cis* regulatory elements (Gindhart *et al*. 1995), and so GAL4 induction of *BLY* might be able to disrupt *Scr* transcriptional regulation. Thus, we asked if reduction of Scr levels could account for the *ey>BLY* phenotype. When *ey-GAL4* drives expression of a *UAS-Scr-RNAi* construct (*P[TRiP.JF03091]*), there is no phenotype, as expected since Scr is not expressed in the eye disc (e.g., Percival-Smith & Hayden 1998). Finally, if the *BLY* phenotype is caused by an *Scr* RNA species, then the phenotype should be suppressed by *UAS-Scr-RNAi*. When *BLY* is crossed to *ey-GAL4; [y*$^{+}$*, UAS-Scr-RNAi]*/+, the *ey>BLY* progeny have the same bulged eye

phenotype regardless of whether *UAS-Scr-RNAi* is present (Fig. 4H). Overall, neither gain nor loss of *Scr* messages appear to cause the eye/head malformation.

Longer-range impacts on the regulation of downstream genes might contribute to the *BLY* phenotype. *mir-10* lies ~45 kb downstream of *BLY* in the same orientation, and *Dfd* lies another ~7 kb further downstream, but in the opposite orientation (it cannot be transcribed by *BLY*). We tested whether misexpression of a *mir-10* construct could mimic *BLY*, but *ey>mir-10* flies have rough eyes and no head shape defect (not shown). Other local ncRNA species, or indirect effects on *Dfd* or other Hox genes, remain possible sources of the phenotype. *BLY* is the only reported transgenic insertion in the 84 kb block between *Antp* (*Mi[MIC]Antp*^{MI02272}) and the end of *Scr* (*PBac[WH]Scr*^{f05078}), including the intervening *ftz* gene. The recovery of a *UAS*- and I*Sce*I-bearing element in this position could be useful for investigating various features of ANT-C regulation.

LNP targets regulatory RNAs of the Bithorax Complex (BX-C)

LNP displays diverse homeotic phenotypes. The *ey>LNP* animals die as pupae or pharate adults, with head capsule and antennae dramatically shifted toward a limb fate (Fig. 9A–B). In some cases, the antenna is transformed strongly to leg; for example, the mounted head in Fig 9C has an antenna with tarsal segments and claws. *ms1096>LNP* flies have halteres that are shifted toward a wing fate (Fig. 9D–E). When *LNP* is expressed around the time of pupation (via heat induction of *Hsp70-GAL4*), ectopic pigmentation appears on the abdomen, indicating a conversion of anterior abdominal segments toward a posterior (A5-like) identity (Fig. 9F–G).

LNP lies in the *infraabdominal-6* (*iab-6*) domain of the *iab* regulatory region between the *abd-A* and *Abd-B* homeobox genes of the BX-C, and transcribes in the distal to proximal direction toward *abd-A* (Maeda & Karch 2006; Akbari *et al*. 2006). *LNP* is 70 kb upstream of the next coding gene *abd-A*, and so we sought to determine if transcripts from *LNP* could persist to *abd-A*. Indeed, RT-PCR showed that *LNP* transcripts can splice in to *abd-A*, but these products included a series of previously undescribed exons, now known to belong to the *iab-8* long ncRNA (Gummalla *et al*. 2012). *LNP* is in intron 2 of the longest *iab-8* isoforms, and splices to *iab-8* exon 3 with the joint …cgcaggcg/TTTGATTACACATCGA…. The *LNP* RT-PCR product exactly matches bases 1252–1484 of the iab-8-RC transcript (RefSeq NR_048413), before splicing into Abd-A (Gummalla *et al*. 2012). This *LNP-iab-8* fusion transcript cannot make any protein fusions of significant size, based on sequence analysis and Western blotting (Fig. 1B). Normally, *iab-8* is thought to repress *abd-A* in *cis* by interfering with its transcription; the wild type *iab-8* only rarely splices into *abd-A* (Gummalla *et al*. 2012).

With no Hox protein fusion produced by *LNP*, what accounts for its diverse homeotic effects? There are at least three other, indirect routes by which *LNP* could affect the intricate regulation of the BX-C: 1) misexpression of the miRNA mir-iab-8; 2) disruption of important chromatin elements in the *iab* region; and 3) disruption of *trans* pairing within BX-C or from BX-C to other loci. The miRNA genes *mir-iab-4*/*mir-iab-8* lie in the 5th intron of *iab-8; mir-iab-8* is transcribed in the same direction as *LNP* and *iab-8*

(Ronshaugen *et al*. 2005; Bender 2008; Stark *et al*. 2008; Hui *et al*. 2013). Controlled expression of mir-iab-8 results in significant loss of Ubx protein accumulation in the haltere disc and strong transformation of haltere toward wing identity as expected for loss of Ubx (Tyler *et al*. 2008). This suggests that the haltere transformation in Fig. 9E results when mir-iab-8 is expressed as a byproduct of the *LNP-iab-8* long ncRNA fusion, and proceeds to repress *Ubx* and perhaps other transcripts.

There are also several well-characterized regulatory regions that could be affected by *LNP*. The *Fab-7* boundary element/polycomb response element (PRE) spans ~0.3 to 1.8 kb upstream (distal) of *LNP*; the *LNP* insertion point is base 1050 on the *Fab-7* map in Fig 4B of Karch *et al*. 1994 (see also Li *et al*. 2011; Aoki *et al*. 2012). *LNP* may lie close enough that when it is activated by GAL4 and transcribed, this interferes with *Fab-7*. *LNP* transcription could also disrupt several downstream (proximal) boundary elements and PREs required for proper BX-C regulation, including the *Miscadastral pigmentation* (*Mcp*) and *Fab-6* regions (Karch *et al*. 1994; Holohan *et al*. 2007; Li *et al*. 2011). The *Mcp* PRE lies ~29 kb downstream of *LNP*, in intron 4 of *iab-8*. In *Mcp* mutants, segment A4 becomes fully pigmented in males like A5, and *Abd-B* expression is extended anteriorly (Celniker *et al*. 1990). Transcription through a PRE can interfere with its activity (Hogga & Karch 2002; Schmitt *et al*. 2005). Thus, disruption of *Mcp* by *LNP* transcripts could account for the ectopic pigment in Fig 9G. Finally, transcription may also disrupt the long range pairing of PRE/insulators from different loci; *LNP* could have impacts on distant sites, such as the ANT-C, by this type of mechanism (Bantignies *et al*. 2011; Li *et al*. 2011). Availability of a *UAS* element in this position should allow for further analysis of the effects of transcription through the *iab* region.

Hto-homeodomain fusions bind specific sites on polytene chromosomes

Hto fusions to Zn finger TFs (Elbow B, Kahuli, and Mamo) were previously shown to accumulate in banded patterns on the polytene chromosomes of larval salivary glands (Singari *et al*. 2014). To determine if homeodomain fusions also show polytene banding, we expressed *FAR* (*ara*), *BRO* (*ct*) and *GLO* (*Drgx*) in salivary glands and imaged SYBR-Green stained chromosome spreads by confocal microscopy. *FAR* did not display discrete bands, and the overall levels of RFP signal were low (not shown). The fusion proteins from *BRO* and *GLO* each gave clearly evident banding patterns; representative segments of the spreads are shown in Fig. 6A–E. With *BRO* we observe relatively few bands that rise moderately above the background signal; some of these bands overlap SYBR Green bands, while other do not (Fig. 6A–C). In contrast, the *GLO* fusion showed very strong banding, i.e., the intensity of the RFP-positive bands was much stronger than that of the interbands or background. Some of the difference could be due to the general low level of *BRO* fusion expression in tissues (*BRO* lies far upstream of *ct*). However, we have never observed changes in banding patterns due to different expression levels (the levels vary from cell to cell due to the driver; Singari *et al*. 2014). The *GLO* fusion protein also binds to several discrete sites in the chromocenter, a mass of repeat-rich DNA that does not form a distinct banding pattern. *GLO* is the only line that showed this pattern, of the six lines tested overall. In the spread in Fig. 6F, binding appears in three patches of dots; the two on the right have a circular arrangement. In other spreads the patches appear more banded; these circular

features may be cross sections of those bands. The density of fusion protein in the chromocenter sites must be several times that of the polytene bands, since the chromocenter signal reaches saturation when the polytene bands are still dim. The *GLO* fusion does not localize to all heterochromatin, since this pattern is unlike HP1 protein (Fanti & Pimpinelli 2008). These sites likely correspond to the bright foci near the nucleolus seen in *GLO*-expressing ovary FC clones (Fig. 5). The pattern suggests that Drgx binds, perhaps fortuitously, to a sequence that is present in high copy number in a few heterochromatic regions. If the overall *GLO* banding pattern does reflect sequence specific binding, it should be highly reproducible across spreads. This is demonstrated by Fig. 6G, in which four copies of the same region are aligned to each other. There is some ambiguity when aligning short segments of SYBR Green stained bands to the standard polytene maps, but we were able to assign this region to distal chromosome *2L* as indicated. Thus the *GLO* fusion binds reproducibly both to numerous euchromatic bands and to limited sites in heterochromatin.

Conclusions

Six of the lines presented here make tagged TF fusion proteins that can be used for several types of downstream analyses. In particular, the *Hto* system can facilitate an integrated combination of microscopy, biochemistry, and genetics to study how these TFs interact with various partners in protein/DNA complexes. Some cofactors of homeodomain proteins are known, but their full multi-subunit structures, and how these change throughout development, remain obscure (Hueber & Lohmann 2008; Mann *et al*. 2009; Ladam & Sagerström 2014). Subcellular localization and dynamics can be examined using the RFP tag, and the FLAG tag can be used to purify protein and DNA complexes containing the fusion. In addition, the phenotypes here (especially those using *GMR-GAL4*) are very distinct, consistent and amenable to genetic interaction analysis, so that genetic screens or candidate testing can be used to find interaction partners.

Interestingly, the only *Hto* inserts so far recovered in the homeotic clusters (*BLY* in ANT-C and *LNP* in BX-C) do not make TF fusion proteins, but rather act via microRNAs or possibly other disruptions at the chromatin or RNA level. These lines unfortunately do not take advantage of the protein tagging function of *Hto*, but they do provide reagents to make DNA breaks (from Minos transposon excision or I*Sce*I cleavage) and to examine the effects of controlled transcription through Hox regulatory elements.

We showed that several of the phenotypes can be suppressed by RNAi against the fusion protein, and that the fusion can even rescue the RNAi LOF phenotype in the case of *BRO*/*ct* (Fig 2O–P). This can be a useful method for confirming the specificity and wild type activity of an *Hto* line. Whether full “rescue” is achieved by an *Hto*/RNAi combination should depend on the relative expression levels of each transgene (as seen with the non-*Hto* construct *UAS-ct;* Fig 3J–K). Nonetheless, the *BRB* fusion is expressed quite strongly (as evaluated by microscopy), and it was still almost completely suppressed (Fig 2S–T). RNAi will knock down the *Hto* transcript regardless of whether the *Hto* product has gain of wild type activity, or is a dominant negative/antimorph, and so suppression of the *Hto* phenotype does not distinguish between these two mechanisms. However, we expect that if an RNAi

construct gives a specific LOF phenotype on its own, then *Hto* should tend to be a phenotypic suppressor only if the *Hto* product has wild type activity (as in Fig 2O–P).

The target genes described here are normally expressed in complex and highly restricted patterns (Hammonds *et al*. 2013). Since they produce such strong misexpression phenotypes across a variety of tissues, it appears that these genes are primarily regulated at the transcriptional level; the organism does not have a good "backup plan" to restrict their activity posttranscriptionally (discussed in Ladam & Sagerström 2014). MicroRNAs may help to tailor their protein expression levels, but any such mechanism is clearly overwhelmed by the GAL4/*UAS* system that is commonly used for controlled misexpression. This reliance on transcriptional controls presents an interesting tradeoff for the animal. On one hand, it facilitates evolution of new adaptations: mutations that arise in individual regulatory elements of homeobox genes can immediately cause dominant, selectable phenotypes, such as changes in limb morphogenesis, PNS patterning, CNS organization, etc. (Heffer & Pick 2013). On the other hand, this makes the animal vulnerable to deleterious mutations; for example, homeobox gene misexpression is implicated in various cancers, where these genes can drive cancer phenotypes such as proliferation, motility, and resistance to apoptosis (Shah & Sukumar 2010; Wong *et al*. 2014).

In the screens we have performed so far, no DNA-binding TF genes were recovered more than once. Thus we have not come close to saturating the potential targets for *Hto*, even though we have only used one of the three reading frames to make the protein fusions. In future work, the current frame zero version of *Hto* should be screened further until duplicates begin to appear more frequently. Then frame +1 and +2 versions of the *Hto* vector could be deployed to recover entirely different sets of target proteins.

Experimental Procedures

Genetics

The genetic screens and molecular biology were performed as described (Singari *et al*. 2014), except *BLY* was recovered in a screen using both *ey-GAL4* and *GMR-GAL4* as drivers to express inserts in the F1 flies. There are two versions of the *Hto* vector: the original version is *Mi*[*Hto-WP*] (for "Wild type PolyA"), and the upgraded version is *Mi*[*Hto*]. All the *Hto* inserts here except for *BRB* are the *Mi*[*Hto-WP*] version, and were derived from transposition of the *Mi*[*Hto-WP*]*Starter2* element. *BRB* was derived from the upgraded starter element *Mi*[*Hto*]*PM2a*. The *UAS-Hto* exon 1 segment is identical in both versions of *Hto*, but the upgraded version carries the mutations $T_{1792}C$ and $T_{1858}C$ (based on GenBank # JN049642 sequence) in order to disrupt the natural polyA signals in the Minos right inverted repeat. The upgraded version also carries minimal *attP* sites flanking the *UAS-exon 1* region. The full name of an insert is, for example, *Mi*[*Hto-WP*]*BRO*; but only the unique 3-letter code is used in the text. FlyBase has adopted the 3-letter codes to further use as allele names, for cross-referencing purposes. This does not imply that the target gene has always been disrupted; in many cases the insert lies well upstream of the target gene.

The lines *BRO*, *FAR*, *GLO*, *OMD*, and *EAB* are deposited at the Bloomington Drosophila Stock Center (BDSC, Indiana University). Additional stocks were obtained from BDSC:

'FLP-out' experiments employed *P[hsFLP]12* and *P[GAL4-Act5C(FRT.CD2).P]*; GAL4 drivers were *GMR-GAL4* on *2; ey-GAL4; pnr-GAL4; ptc-GAL4* (*P[GawB]ptc*$^{559.1}$); and heat-shock and salivary gland driver *P[GAL4-Hsp70.PB]89-2-1* (BDSC #1799, herein called *Hsp70-GAL4*). The *UAS-ct* chromosome was *P[w*$^{+}$*,UAS-ct.P]2*, *P[w*$^{+}$*,UAS-mCD8-GFP.L]LL5*. Scr and mir-10 expression lines: *P[w*$^{+}$*, UAS-Scr.M]EE2* and *P[y*$^{+}$*, w*$^{+}$*, UAS-LUC-mir-10.T]attP2*. RNAi lines: *ct* RNAi, *P[y*$^{+}$*, v*$^{+}$*, TRiP.JF03304]attP2; Poxn*-RNAi, *P[y*$^{+}$*, v*$^{+}$*, TRiP.JF02136]attP2; Scr* RNAi, *P[y*$^{+}$*, v*$^{+}$*, TRiP.JF03091]attP2* (Transgenic RNAi Project/Harvard Medical School). Note that the *P[w*$^{+}$*, UAS-Scr.M]EE2* and *Scr* RNAi constructs do not overlap, and so we could not test the efficiency of RNAi against Scr overexpression with these lines. Standard protein trap lines were *P[PTT-GC]Atpα*^{G00109} (BDSC), and *P[PTT-GC]baz*^{CC01941} (gift of the Spradling Lab/FlyTrap project). The *ms1096w-GAL4* line was described (Park & Edwards 2004). To mobilize *Hto* in the screens, the transposase line *w*^{1118}; *sna*^{Sco}*/SM6a*, *P[hsILMiT]2.4* was obtained from H. Bellen/Gene Disruption Project. All flies shown are female and raised at 25°C unless noted.

Immunostaining and microscopy

Sample preparation and microscopy were done as described (Singari *et al*. 2014). Ct was detected by 2B10 monoclonal antibody concentrate used 1:90 (Developmental Studies Hybridoma Bank.) Fluorescence images were collected on a Leica SP2 confocal; the mCherry RFP channel and the anti-Ct channel are shown with the original contrast (no sigma curve) unless noted. For polytene chromosomes, salivary glands from *Hsp70-GAL4; Hto* larvae (without heat shock) were fixed 8 min in 2% formaldeyde/PBS, washed in PBS and then in PBS with 0.2% Triton X-100, stained with SYBR Green, and spread on a slide in a drop of Vectashield (Vector Labs, Burlingame, CA) mounting medium (procedure modified from DiMario *et al*. 2006).

Acknowledgments

Thanks to Bloomington Drosophila Stock Center at Indiana U., Transgenic RNAi Project (TRiP) at Harvard Medical School, FlyTrap project, and H. Bellen for fly stocks. Thanks to the reviewers and to S. Horne-Badovinac for helpful comments on the manuscript. Supported by National Institutes of Health GM62185; Illinois State University; and a Weigel grant from the Beta Lambda chapter of Phi Sigma Biological Sciences Honor Society, Illinois State University, to N.J.

References

Akbari OS, Bousum A, Bae E, Drewell RA. Unraveling cis-regulatory mechanisms at the *abdominal-A* and *Abdominal-B* genes in the Drosophila bithorax complex. Dev Biol. 2006; 293:294–304. [PubMed: 16545794]

Aoki T, Sarkeshik A, Yates J, Schedl P. Elba, a novel developmentally regulated chromatin boundary factor is a hetero-tripartite DNA binding complex. Elife. 2012; 1:e00171. [PubMed: 23240086]

Awasaki T, Kimura K. Multiple function of *poxn* gene in larval PNS development and in adult appendage formation of Drosophila. Dev Genes Evol. 2001; 211:20–29. [PubMed: 11277402]

Balczarek KA, Lai ZC, Kumar S. Evolution of functional diversification of the paired box (Pax) DNA-binding domains. Mol Biol Evol. 1997; 14:829–842. [PubMed: 9254921]

Bantignies F, Roure V, Comet I, Leblanc B, Schuettengruber B, Bonnet J, Tixier V, Mas A, Cavalli G. Polycomb-dependent regulatory contacts between distant Hox loci in Drosophila. Cell. 2011; 144:214–226. [PubMed: 21241892]

Bellen HJ, Levis RW, He Y, Carlson JW, Evans-Holm M, Bae E, Kim J, Metaxakis A, Savakis C, Schulze KL, Hoskins RA, Spradling AC. The Drosophila gene disruption project: progress using transposons with distinctive site specificities. Genetics. 2011; 188:731–743. [PubMed: 21515576]

Bender W. MicroRNAs in the *Drosophila* bithorax complex. Genes Dev. 2008; 22:14–19. [PubMed: 18172161]

Blake JA, Thomas M, Thompson JA, White R, Ziman M. Perplexing Pax: from puzzle to paradigm. Dev Dyn. 2008; 237:2791–2803. [PubMed: 18816860]

Boube M, Hudry B, Immarigeon C, Carrier Y, Bernat-Fabre S, Merabet S, Graba Y, Bourbon HM, Cribbs DL. *Drosophila melanogaster* Hox transcription factors access the RNA polymerase II machinery through direct homeodomain binding to a conserved motif of mediator subunit Med19. PLoS Genet. 2014; 10(5):e1004303. [PubMed: 24786462]

Bridges CB, Morgan TH. The third-chromosome group of mutant characters of *Drosophila melanogaster*. Publs Carnegie Instn. 1923; 327:1–251.

Brigaud I, Duteyrat JL, Chlasta J, Le Bail S, Couderc JL, Grammont M. Transforming Growth Factor β/activin signalling induces epithelial cell flattening during Drosophila oogenesis. Biol Open. 2015 pii: BIO201410785.

Buszczak M, Paterno S, Lighthouse D, Bachman J, Planck J, Owen S, Skora AD, Nystul TG, Ohlstein B, Allen A, Wilhelm JE, Murphy TD, Levis RW, Matunis E, Srivali N, Hoskins RA, Spradling AC. The Carnegie protein trap library: a versatile tool for Drosophila developmental studies. Genetics. 2007; 175:1505–1531. [PubMed: 17194782]

Carrasco-Rando M, Tutor AS, Prieto-Sánchez S, González-Pérez E, Barrios N, Letizia A, Martín P, Campuzano S, Ruiz-Gómez M. Drosophila *araucan* and *caupolican* integrate intrinsic and signalling inputs for the acquisition by muscle progenitors of the lateral transverse fate. PLoS Genet. 2011; 7(7):e1002186. [PubMed: 21811416]

Cavodeassi F, Modolell J, Gómez-Skarmeta JL. The Iroquois family of genes: from body building to neural patterning. Development. 2001; 128:2847–2855. [PubMed: 11532909]

Celniker SE, Sharma S, Keelan DJ, Lewis EB. The molecular genetics of the bithorax complex of Drosophila: cis-regulation in the *Abdominal-B* domain. EMBO J. 1990; 9:4277–4286. [PubMed: 2265608]

Chen ZF, Rebelo S, White F, Malmberg AB, Baba H, Lima D, Woolf CJ, Basbaum AI, Anderson DJ. The paired homeodomain protein DRG11 is required for the projection of cutaneous sensory afferent fibers to the dorsal spinal cord. Neuron. 2001; 31:59–73. [PubMed: 11498051]

Conrad S, Stimpfle F, Montazeri S, Oldekamp J, Seid K, Alvarez-Bolado G, Skutella T. RGMb controls aggregation and migration of Neogenin-positive cells in vitro and in vivo. Mol Cell Neurosci. 2010; 43:222–231. [PubMed: 19944164]

del Valle Rodríguez A, Didiano D, Desplan C. Power tools for gene expression and clonal analysis in Drosophila. Nat Methods. 2011; 9:47–55. [PubMed: 22205518]

Delon I, Brown NH. The integrin adhesion complex changes its composition and function during morphogenesis of an epithelium. J Cell Sci. 2009; 122:4363–4374. [PubMed: 19903692]

DiMario P, Rosby R, Cui Z. Direct visualization of GFP-fusion proteins on polytene chromosomes. Dros Inf Serv. 2006; 89:115–118.

Duffy JB. GAL4 system in Drosophila: a fly geneticist's Swiss army knife. Genesis. 2002; 34:1–15. [PubMed: 12324939]

Duong HA, Wang CW, Sun YH, Courey AJ. Transformation of eye to antenna by misexpression of a single gene. Mech Dev. 2008; 125:130–141. [PubMed: 18037276]

Dziedzic K, Heaphy J, Prescott H, Kavaler J. The transcription factor D-Pax2 regulates crystallin production during eye development in *Drosophila melanogaster*. Dev Dyn. 2009; 238:2530–2539. [PubMed: 19718746]

Elliott DA, Brand AH. The GAL4 system: a versatile system for the expression of genes. Methods Mol Biol. 2008; 420:79–95. [PubMed: 18641942]

Fanti L, Pimpinelli S. HP1: a functionally multifaceted protein. Curr Opin Genet Dev. 2008; 18:169–174. [PubMed: 18329871]

Findlay GD, Yi X, Maccoss MJ, Swanson WJ. Proteomics reveals novel Drosophila seminal fluid proteins transferred at mating. PLoS Biol. 2008; 6(7):e178. [PubMed: 18666829]

Fu W, Noll M. The Pax2 homolog sparkling is required for development of cone and pigment cells in the Drosophila eye. Genes Dev. 1997; 11:2066–2078. [PubMed: 9284046]

Gehring WJ, Affolter M, Bürglin T. Homeodomain proteins. Annu Rev Biochem. 1994; 63:487–526. [PubMed: 7979246]

Gindhart JG Jr, King AN, Kaufman TC. Characterization of the cis-regulatory region of the Drosophila homeotic gene *Sex combs reduced*. Genetics. 1995; 139:781–795. [PubMed: 7713432]

Gomez-Skarmeta JL, Diez del Corral R, de la Calle-Mustienes E, Ferre-Marco D, Modolell J. Araucan and caupolican, two members of the novel iroquois complex, encode homeoproteins that control proneural and vein-forming genes. Cell. 1996; 85:95–105. [PubMed: 8620542]

Gomez-Skarmeta JL, Modolell J. *Iroquois* genes: genomic organization and function in vertebrate neural development. Curr Opin Genet Dev. 2002; 12:403–408. [PubMed: 12100884]

Grammont M. Adherens junction remodeling by the Notch pathway in *Drosophila melanogaster* oogenesis. J Cell Biol. 2007; 177:139–150. [PubMed: 17420294]

Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, Brown JB, Cherbas L, Davis CA, Dobin A, Li R, Lin W, Malone JH, Mattiuzzo NR, Miller D, Sturgill D, Tuch BB, Zaleski C, Zhang D, Blanchette M, Dudoit S, Eads B, Green RE, Hammonds A, Jiang L, Kapranov P, Langton L, Perrimon N, Sandler JE, Wan KH, Willingham A, Zhang Y, Zou Y, Andrews J, Bickel PJ, Brenner SE, Brent MR, Cherbas P, Gingeras TR, Hoskins RA, Kaufman TC, Oliver B, Celniker SE. The developmental transcriptome of *Drosophila melanogaster*. Nature. 2011; 471:473–479. [PubMed: 21179090]

Guarner A, Manjón C, Edwards K, Steller H, Suzanne M, Sanchez-Herrero E. The *zinc finger homeodomain-2* gene of *Drosophila* controls Notch targets and regulates apoptosis in the tarsal segments. Dev Biol. 2014; 385:350–365. [PubMed: 24144920]

Gummalla M, Maeda RK, Castro Alvarez JJ, Gyurkovics H, Singari S, Edwards KA, Karch F, Bender W. *abd-A* regulation by the iab-8 noncoding RNA. PLoS Genetics. 2012; 8(5):e1002720. [PubMed: 22654672]

Hammonds AS, Bristow CA, Fisher WW, Weiszmann R, Wu S, Hartenstein V, Kellis M, Yu B, Frise E, Celniker SE. Spatial expression of transcription factors in Drosophila embryonic organ development. Genome Biol. 2013; 14:R140. [PubMed: 24359758]

Heffer A, Pick L. Conservation and variation in Hox genes: how insect models pioneered the evo-devo field. Annu Rev Entomol. 2013; 58:161–179. [PubMed: 23317041]

Heuer JG, Li K, Kaufman TC. The Drosophila homeotic target gene *centrosomin* (*cnn*) encodes a novel centrosomal protein with leucine zippers and maps to a genomic region required for midgut morphogenesis. Development. 1995; 121:3861–3876. [PubMed: 8582295]

Hogga I, Karch F. Transcription through the iab-7 cis-regulatory domain of the bithorax complex interferes with maintenance of Polycomb-mediated silencing. Development. 2002; 129:4915–4922. [PubMed: 12397100]

Holland PW. Evolution of homeobox genes. Wiley Interdiscip Rev Dev Biol. 2013; 2:31–45. [PubMed: 23799629]

Holland PW, Booth HA, Bruford EA. Classification and nomenclature of all human homeobox genes. BMC Biol. 2007; 5:47. [PubMed: 17963489]

Holohan EE, Kwong C, Adryan B, Bartkuhn M, Herold M, Renkawitz R, Russell S, White R. CTCF genomic binding sites in Drosophila and the organisation of the bithorax complex. PLoS Genet. 2007; 3(7):e112. [PubMed: 17616980]

Horne-Badovinac S, Bilder D. Mass transit: epithelial morphogenesis in the Drosophila egg chamber. Dev Dyn. 2005; 232:559–574. [PubMed: 15704134]

Hueber SD, Lohmann I. Shaping segments: Hox gene function in the genomic age. Bioessays. 2008; 30:965–979. [PubMed: 18798525]

Hui JH, Marco A, Hunt S, Melling J, Griffiths-Jones S, Ronshaugen M. Structure, evolution and function of the bi-directionally transcribed iab-4/iab-8 microRNA locus in arthropods. Nucleic Acids Res. 2013; 41:3352–3361. [PubMed: 23335784]

Hulea L, Nepveu A. CUX1 transcription factors: from biochemical activities and cell-based assays to mouse models and human diseases. Gene. 2012; 497:18–26. [PubMed: 22306263]

Ikmi A, Netter S, Coen D. Prepatterning the Drosophila notum: the three genes of the *iroquois* complex play intrinsically distinct roles. Dev Biol. 2008; 317:634–648. [PubMed: 18394597]

Jack J, Dorsett D, Delotto Y, Liu S. Expression of the *cut* locus in the Drosophila wing margin is required for cell type specification and is regulated by a distant enhancer. Development. 1991; 113:735–747. [PubMed: 1821846]

Jackson SM, Blochlinger K. *cut* interacts with Notch and protein kinase A to regulate egg chamber formation and to maintain germline cyst integrity during Drosophila oogenesis. Development. 1997; 124:3663–3672. [PubMed: 9342058]

Jiao R, Daube M, Duan H, Zou Y, Frei E, Noll M. Headless flies generated by developmental pathway interference. Development. 2001; 128:3307–3319. [PubMed: 11546747]

Johnson SA, Harmon KJ, Smiley SG, Still FM, Kavaler J. Discrete regulatory regions control early and late expression of D-Pax2 during external sensory organ development. Dev Dyn. 2011; 240:1769–1778. [PubMed: 21644243]

Jun S, Desplan C. Cooperative interactions between paired domain and homeodomain. Development. 1996; 122:2639–2650. [PubMed: 8787739]

Karch F, Galloni M, Sipos L, Gausz J, Gyurkovics H, Schedl P. Mcp and Fab-7: molecular analysis of putative boundaries of cis-regulatory domains in the bithorax complex of Drosophila melanogaster. Nucleic Acids Res. 1994; 22:3138–3146. [PubMed: 7915032]

Kavaler J, Fu W, Duan H, Noll M, Posakony JW. An essential role for the Drosophila Pax2 homolog in the differentiation of adult sensory organs. Development. 1999; 126:2261–2272. [PubMed: 10207150]

Kerner P, Ikmi A, Coen D, Vervoort M. Evolutionary history of the iroquois/Irx genes in metazoans. BMC Evol Biol. 2009; 9:74. [PubMed: 19368711]

Kolahi KS, White PF, Shreter DM, Classen AK, Bilder D, Mofrad MR. Quantitative analysis of epithelial morphogenesis in Drosophila oogenesis: New insights based on morphometric analysis and mechanical modeling. Dev Biol. 2009; 331:129–139. [PubMed: 19409378]

Ladam F, Sagerström CG. Hox regulation of transcription: more complex(es). Dev Dyn. 2014; 243:4–15. [PubMed: 23765878]

Levine B, Hackney JF, Bergen A, Dobens L 3rd, Truesdale A, Dobens L. Opposing interactions between Drosophila cut and the C/EBP encoded by slow border cells direct apical constriction and epithelial invagination. Dev Biol. 2010; 344:196–209. [PubMed: 20450903]

Li HB, Müller M, Bahechar IA, Kyrchanova O, Ohno K, Georgiev P, Pirrotta V. Insulators, not Polycomb response elements, are required for long-range interactions between Polycomb targets in *Drosophila melanogaster*. Mol Cell Biol. 2011; 31:616–625. [PubMed: 21135119]

Lundell MJ, Hirsh J. The *zfh-2* gene product is a potential regulator of neuron-specific DOPA decarboxylase gene expression in Drosophila. Dev Biol. 1992; 154:84–94. [PubMed: 1426635]

Maeda RK, Karch F. The ABC of the BX-C: the bithorax complex explained. Development. 2006; 133:1413–1422. [PubMed: 16556913]

Mann RS, Lelli KM, Joshi R. Hox specificity: unique roles for cofactors and collaborators. Curr Top Dev Biol. 2009; 88:63–101. [PubMed: 19651302]

Morin X, Daneman R, Zavortink M, Chia W. A protein trap strategy to detect GFP-tagged proteins expressed from their endogenous loci in Drosophila. Proc Natl Acad Sci USA. 2001; 98:15050–15055. [PubMed: 11742088]

Nepveu A. Role of the multifunctional CDP/Cut/Cux homeodomain transcription factor in regulating differentiation, cell growth and development. Gene. 2001; 270:1–15. [PubMed: 11403998]

O'Brien P, Morin P Jr, Ouellette RJ, Robichaud GA. The Pax-5 gene: a pluripotent regulator of B-cell differentiation and cancer disease. Cancer Res. 2011; 71:7345–7350. [PubMed: 22127921]

Park H, Edwards K. "Marker removal" screen to generate an improved wing disc GAL4 driver. Drosophila Inf Serv. 2004; 87:96–99.

Parrish JZ, Kim MD, Jan LY, Jan YN. Genome-wide analyses identify transcription factors required for proper morphogenesis of Drosophila sensory neuron dendrites. Genes Dev. 2006; 20:820–835. [PubMed: 16547170]

Percival-Smith A, Hayden DJ. Analysis in *Drosophila melanogaster* of the interaction between *sex combs reduced* and *extradenticle* activity in the determination of tarsus and arista identity. Genetics. 1998; 150:189–198. [PubMed: 9725838]

Perea D, Molohon K, Edwards K, Díaz-Benjumea FJ. Multiple roles of the gene *zinc finger homeodomain-2* in the development of the Drosophila wing. Mech Dev. 2013; 130:467–481. [PubMed: 23811114]

Pignoni F, Zipursky SL. Induction of Drosophila eye development by decapentaplegic. Development. 1997; 124:271–8. [PubMed: 9053304]

Poukkula M, Cliffe A, Changede R, Rørth P. Cell behaviors regulated by guidance cues in collective migration of border cells. J Cell Biol. 2011; 192:513–524. [PubMed: 21300853]

Rebelo S, Reguenga C, Lopes C, Lima D. Prrxl1 is required for the generation of a subset of nociceptive glutamatergic superficial spinal dorsal horn neurons. Dev Dyn. 2010; 239:1684–1694. [PubMed: 20503365]

Rebelo S, Reguenga C, Osorio L, Pereira C, Lopes C, Lima D. DRG11 immunohistochemical expression during embryonic development in the mouse. Dev Dyn. 2007; 236:2653–2660. [PubMed: 17676639]

Robson EJ, He SJ, Eccles MR. A PANorama of PAX genes in cancer and development. Nat Rev Cancer. 2006; 6:52–62. [PubMed: 16397527]

Ronshaugen M, Biemar F, Piel J, Levine M, Lai EC. The Drosophila microRNA iab-4 causes a dominant homeotic transformation of halteres to wings. Genes Dev. 2005; 19:2947–2952. [PubMed: 16357215]

Rørth P. A modular misexpression screen in Drosophila detecting tissue-specific phenotypes. Proc Natl Acad Sci USA. 1996; 93:12418–12422. [PubMed: 8901596]

Samad TA, Srinivasan A, Karchewski LA, Jeong SJ, Campagna JA, Ji RR, Fabrizio DA, Zhang Y, Lin HY, Bell E, Woolf CJ. DRAGON: a member of the repulsive guidance molecule-related family of neuronal- and muscle-expressed membrane proteins is regulated by DRG11 and has neuronal adhesive properties. J Neurosci. 2004; 24:2027–2036. [PubMed: 14985445]

Schmitt S, Prestel M, Paro R. Intergenic transcription through a polycomb group response element counteracts silencing. Genes Dev. 2005; 19:697–708. [PubMed: 15741315]

Severyn CJ, Shinde U, Rotwein P. Molecular biology, genetics and biochemistry of the repulsive guidance molecule family. Biochem J. 2009; 422:393–403. [PubMed: 19698085]

Shah N, Sukumar S. The Hox genes and their roles in oncogenesis. Nat Rev Cancer. 2010; 10:361–371. [PubMed: 20357775]

Singari S, Javeed N, Tardi NJ, Marada SK, Carlson JC, Kirk S, Thorn JM, Edwards KA. Inducible protein traps with dominant phenotypes for functional analysis of the Drosophila genome. Genetics. 2014; 196:91–105. [PubMed: 24172131]

Singh A, Tare M, Puli OR, Kango-Singh M. A glimpse into dorso-ventral patterning of the Drosophila eye. Dev Dyn. 2012; 241:69–84. [PubMed: 22034010]

Soares-dos-Reis R, Pessoa AS, Matos MR, Falcão M, Mendes VM, Manadas B, Monteiro FA, Lima D, Reguenga C. Ser^{119} phosphorylation modulates the activity and conformation of PRRXL1, a homeodomain transcription factor. Biochem J. 2014; 459:441–453. [PubMed: 24564673]

St Pierre SE, Ponting L, Stefancsik R, McQuilton P. FlyBase Consortium. FlyBase 102 - advanced approaches to interrogating FlyBase. Nucleic Acids Res. 2014; 42:D780–788. [PubMed: 24234449]

Stark A, Bushati N, Jan CH, Kheradpour P, Hodges E, Brennecke J, Bartel DP, Cohen SM, Kellis M. A single Hox locus in Drosophila produces functional microRNAs from opposite DNA strands. Genes Dev. 2008; 22:8–13. [PubMed: 18172160]

Tyler DM, Okamura K, Chung WJ, Hagen JW, Berezikov E, Hannon GJ, Lai EC. Functionally distinct regulatory RNAs generated by bidirectional transcription and processing of microRNA loci. Genes Dev. 2008; 22:26–36. [PubMed: 18172163]

Vervoort M, Zink D, Pujol N, Victoir K, Dumont N, Ghysen A, Dambly-Chaudière C. Genetic determinants of sense organ identity in Drosophila: regulatory interactions between *cut* and *poxn*. Development. 1995; 121:3111–3120. [PubMed: 7555735]

Walcher T, Xie Q, Sun J, Irmler M, Beckers J, Öztürk T, Niessing D, Stoykova A, Cvekl A, Ninkovic J, Götz M. Functional dissection of the paired domain of Pax6 reveals molecular mechanisms of coordinating neurogenesis and proliferation. Development. 2013; 140:1123–1136. [PubMed: 23404109]

Wang CW, Sun YH. Segregation of eye and antenna fates maintained by mutual antagonism in Drosophila. Development. 2012; 139:3413–3421. [PubMed: 22912416]

Wong CC, Martincorena I, Rust AG, Rashid M, Alifrangis C, Alexandrov LB, Tiffen JC, Kober C, Green AR, Massie CE, Nangalia J, Lempidaki S, Döhner H, Döhner K, Bray SJ, McDermott U, Papaemmanuil E, Campbell PJ, Adams DJ. Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium. Inactivating *CUX1* mutations promote tumorigenesis. Nat Genet. 2014; 46:33–38. [PubMed: 24316979]

Zhai Z, Ha N, Papagiannouli F, Hamacher-Brady A, Brady N, Sorge S, Bezdan D, Lohmann I. Antagonistic regulation of apoptosis and differentiation by the Cut transcription factor represents a tumor-suppressing mechanism in Drosophila. PLoS Genet. 2012; 8(3):e1002582. [PubMed: 22438831]

Zhu CC, Boone JQ, Jensen PA, Hanna S, Podemski L, Locke J, Doe CQ, O'Connor MB. Drosophila Activin- and the Activin-like product Dawdle function redundantly to regulate proliferation in the larval brain. Development. 2008; 135:513–521. [PubMed: 18171686]

Key Findings

- The *Hostile takeover* (*Hto*) transposon system was used to generate fluorescent protein traps affecting eight Drosophila loci of the Homeobox and Pax superfamilies
- Each *Hto* insert generates a unique set of dominant, GAL4-dependent phenotypes suitable for genetic interaction studies
- The little-studied *CG34340* gene is renamed *Dorsal root ganglia homeobox* (*Drgx*) after its human ortholog; it alters cell adhesion when ectopically expressed
- An *Hto* insert between Hox genes *abdominal-A* and *Abdominal-B* causes diverse homeotic phenotypes, probably via noncoding RNAs and effects on Hox transcriptional regulation

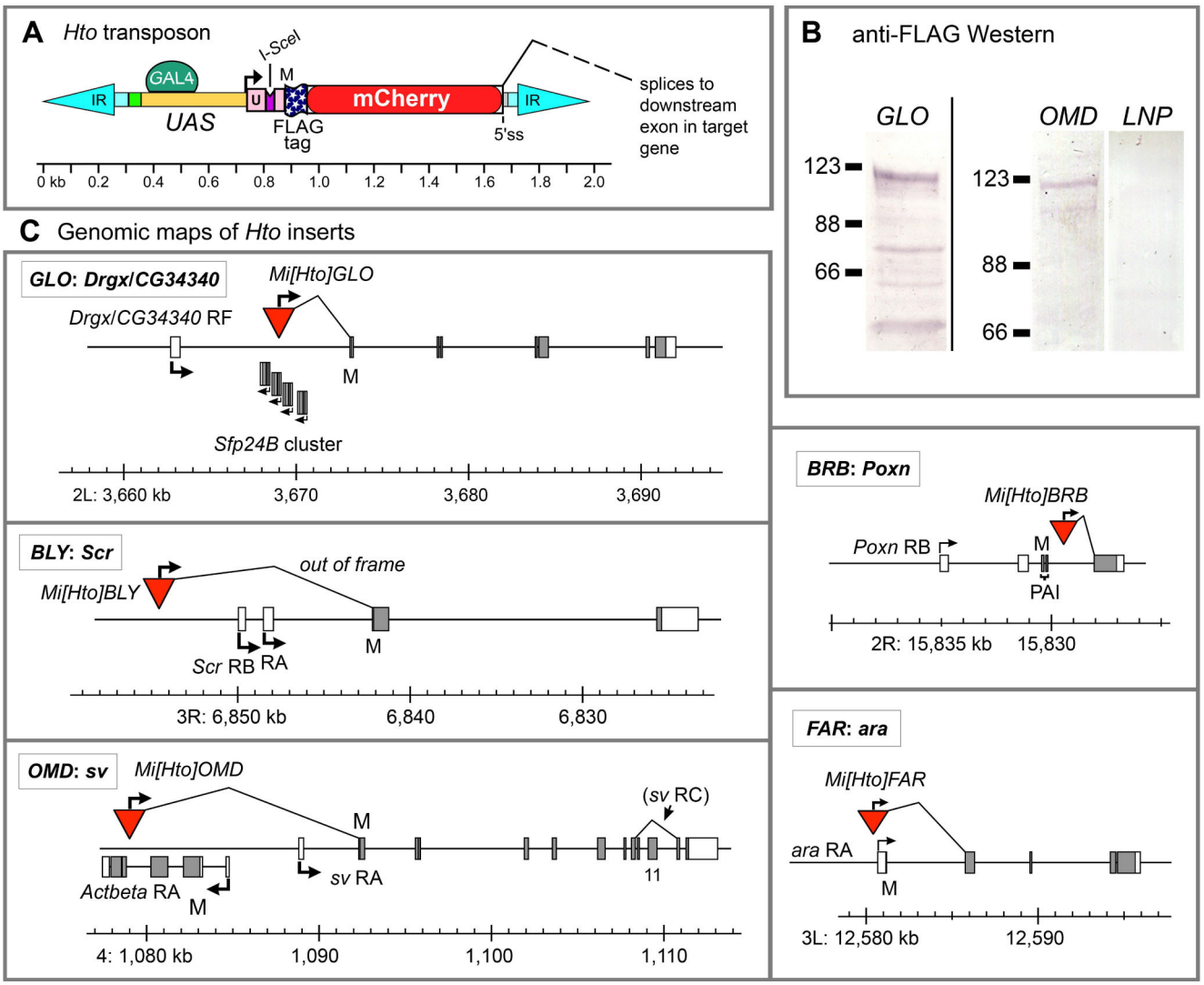

Figure 1. *Hto* schematic, Western analysis, and insert maps

A. Schematic diagram of the *Hto* element. IR, *Minos* inverted repeats; black arrow, transcription start; U, 5′ UTR; M, start codon; 3xFLAG epitope tag and mCherry RFP coding regions are indicated. See Singari *et al*. (2014) for sequence and complete annotation. **B.** SDS/PAGE-Western blot. Each lane is whole adult protein from heat shock-induced *Hsp70-GAL4>Hto* flies as indicated, probed with anti-3xFLAG; marker bands given on the left (kDa). The lines also produce some unfused FLAG-RFP at ~32 kDa (Singari *et al*. 2014). *LNP* shows no fusion protein, as expected since the *LNP-iab-8* fusion has no extended open reading frame. **C.** Genomic maps of *Hto* insertions *GLO*, *BLY*, *OMD*, *BRB*, and *FAR* (red triangles). Maps show the relevant transcripts; for complete transcript maps see FlyBase.org (St. Pierre *et al*. 2014). Lower scale bars, genome coordinates based on *D. melanogaster* Genome Release 6. Arrows, transcription starts; boxes, exons; gray boxes, coding regions; M, endogenous start codon; angled lines show *Hto* splices or endogenous alternative splices.

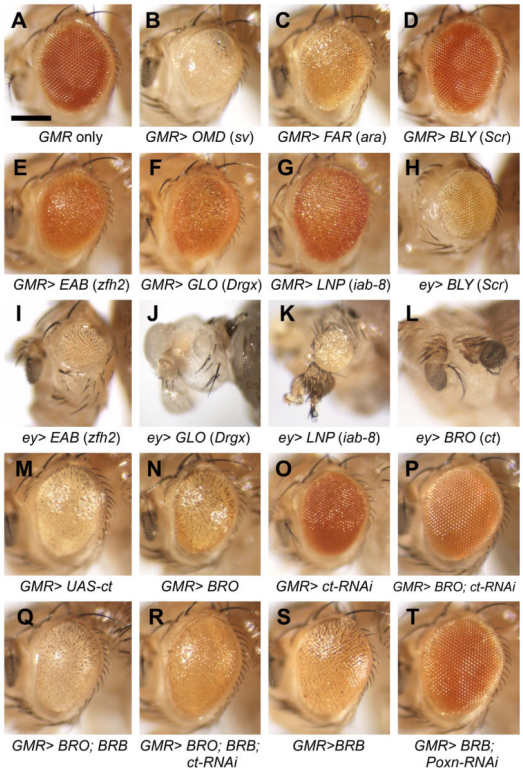

Figure 2. *Hto* inserts driven by eye-specific GAL4 constructs cause diverse phenotypes

All lateral views; bar, 200 μm for all panels. All 25°C except (**M**), 19°C. **A**. Control: eye with *GMR-GAL4* but no responder appears wild type. **B**–**G**, **N**, **S**, characteristic phenotypes of each of the eight inserts driven by *GMR-GAL4* as indicated. The *BLY* insert, which does not make a fusion protein, does not disrupt retina development (**D**); the remaining lines each cause strong rough and/or glazed eyes, also with reduced pigmentation in *OMD*, *FAR*, *BRO*, and *BRB*. **H**–**L**. Five inserts expressed during early eye disc development using *eyeless* (*ey*)-*GAL4*. The eyes are eliminated by *ey>GLO* (**J**). For *ey>BRO*, most have partial to full conversion of eye to antenna; (**L**) shows an extreme case where a posterior antenna has replaced the eye. **M**–**N,** A *UAS-ct* construct and *BRO* produce similar phenotypes with *GMR-GAL4*. **N**–**T**. RNAi analysis of *BRO* and *BRB* lines. Presence of an RNAi construct is scored by its segregation from *TM6,Tb*. **N**–**P**, sibs from a cross that combines *BRO* with *UAS-ct-RNAi*. *GMR>BRO* with no RNAi is severe pale, rough, and glazed (**N**). *GMR>ct-RNAi* alone has a glossy/sparkling phenotype (**O**). With both inserts present, the eye is restored to wild type morphology, but slightly paler (**P**). **Q**–**R**, sibs with both *BRO* and *BRB* inserts, either without (**Q**) or with (**R**) *UAS-ct-RNAi*. Coexpression of *BRO* and *BRB* gives a phenotype similar to each on its own. The *ct-RNAi* construct does not suppress *BRB*'s contribution to the phenotype. **S**–**T**, sibs with *GMR>BRB*, either without (**S**) or with (**T**) *UAS-Poxn-RNAi*. The *BRB* phenotype is almost completely suppressed by RNAi against *Poxn*.

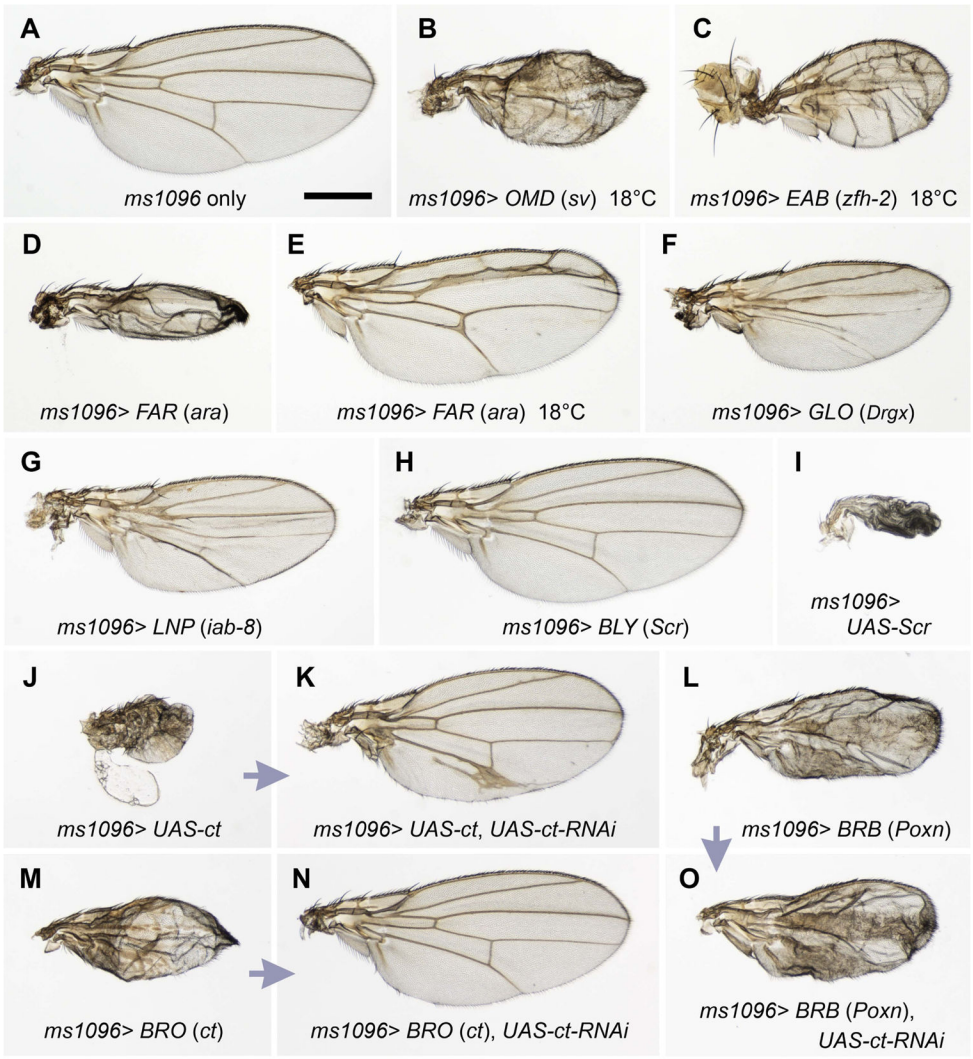

Figure 3. Phenotypes from *Hto* expression in the wing blade

Bar, 500 μm for all panels. **B**, **C**, and **E** reared at 18°C, the rest at 25°C. **A**. Control: with only the wing driver *ms1096-GAL4*, the wing appears wild type. **B**–**H**, **L**. Induction of each of the eight inserts with *ms1096* causes a unique wing phenotype; only *BLY* has little or no effect on the wing. **D**–**E**. The GAL4 system is temperature sensitive; at 18°C, the *FAR* small-wing phenotype is suppressed, revealing specific vein defects. **H**–**I**. *BLY* does not match the *UAS-Scr* phenotype; *ms1096>UAS-Scr* wings never unfold. **J**–**O**. RNAi analysis of *BRO* and *BRB* inserts. The pairs (indicated by arrows) are *ms1096* siblings without or with *ct-RNAi*. **J**–**K**. A *UAS-ct* construct has a small wing phenotype similar to *BRO*, but more severe. It is substantially suppressed by *ct-RNAi*, though the posterior region remains mispatterned. **M**–**N**. *BRO* has a small wing phenotype that is almost fully suppressed by *ct-RNAi*. **L**,**O**. *BRB* leads to extra vein material around the A/P boundary and L5. Poxn is known to activate *ct* under some circumstances, but this phenotype does not require *ct* since it is not suppressed by *ct-RNAi*.

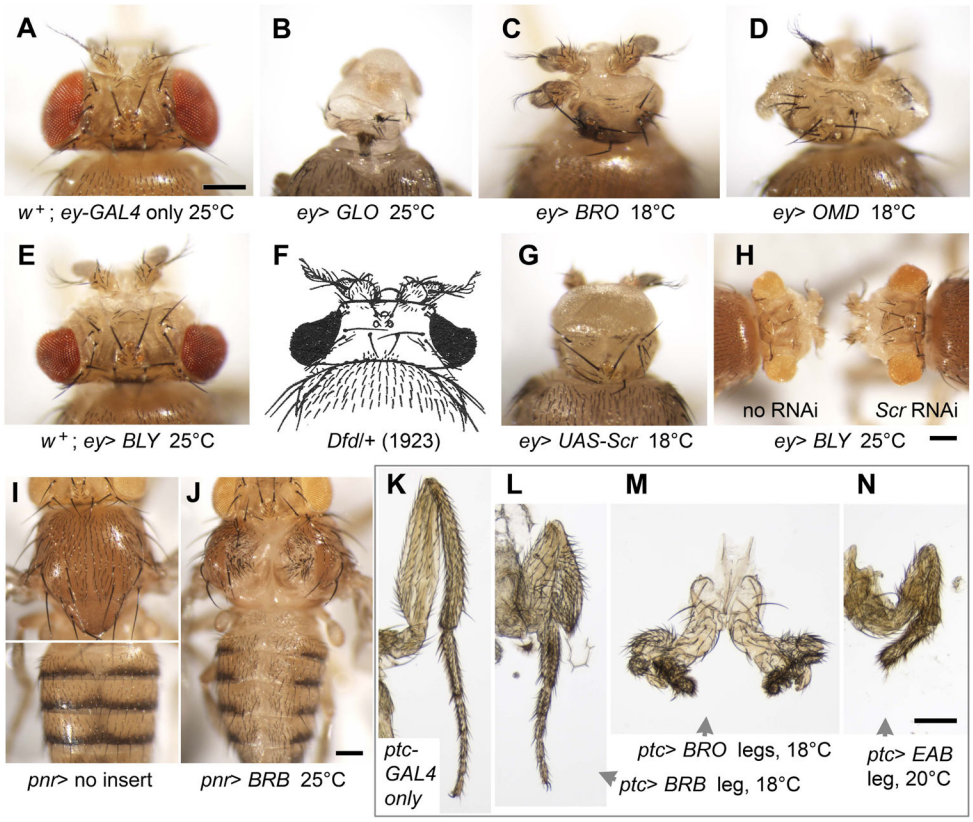

Figure 4. Head, limb, and body phenotypes from *Hto* expression

Bars, 200 μm. The head and leg panels in **A**–**E**, **G,** and **K**–**N** are all the same magnification. **A–H**, dorsal views of the head. **A.** Control: wild type head with *ey-GAL4* only. **B**. *ey>GLO* results in severe reduction of the head (as in Fig. 2J). **C**. *ey>BRO* with one eye missing and one eye transformed to full antenna (lower temperature and expression level than Fig. 2L). **D.** Severe derangement of the dorsal head pattern caused by *ey>OMD*. **E.** *ey>BLY* gives a severely misshapen head with the eyes protruding. **F**. Edith Wallace drawing of the original *Dfd* phenotype (Bridges & Morgan 1923) resembles *ey>BLY* (**E**). **G.** In contrast to *BLY*, *ey>UAS-Scr* construct gives a 100% eyeless phenotype; this sample is the mildest phenotype observed. **H**. Two *y*/*Y; ey-GAL4*/+ sib males: *BLY*/+ (left) and *BLY*/*[y+, UAS-Scr-RNAi]* (right). An *Scr* RNAi construct does not suppress *BLY*, as both have the typical *ey>BLY* phenotype. **I–J**. *pnr-GAL4* expression of *BRB* in a dorsal stripe along the body causes a cleaved thorax with excess macrochaetae, no scutellum, and loss of abdominal pigment bands (**J**), compared to sib without *BRB* (**I**). **K–N**, Leg disc expression via *ptc-GAL4* causes diverse malformations in the rear legs, compared to *ptc-GAL4*/+ only (**K**). **L.** Pharate *ptc>BRB* animal; the femur and tibia failed to elongate. **M.** Pharate *ptc>BRO* legs are severely mispatterned. **N.** In *ptc>EAB*, the tarsal segments (distal end) fail to form properly.

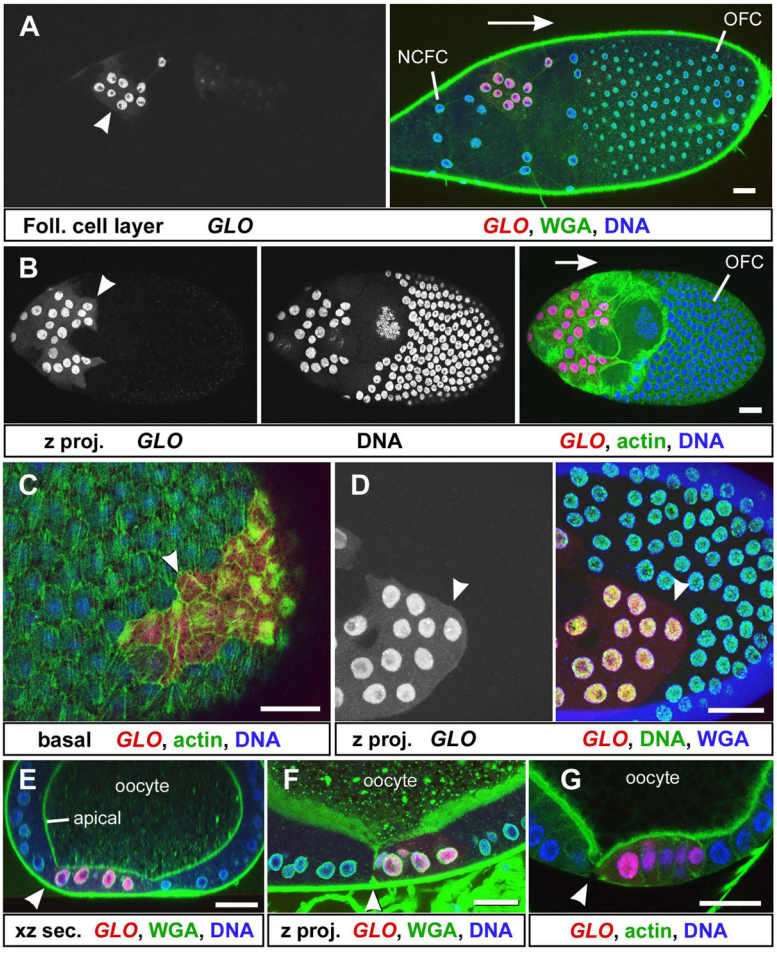

Figure 5. Localization and phenotypes of ectopic Drgx fusion protein from the *GLO* insert

Confocal analyses of egg chambers expressing the *GLO* product (red in each overlay image), in clones of FC via the *Act5C-GAL4* FLP-out system. Arrowheads indicate clone borders. The original-contrast RFP signal is shown in **A**, **B**, and **F**. The RFP channel was brightened to highlight the shape of the clone in **C**, **D**, **E**, and **G**. Bars, 20 μm in each set. **A**. FC layer of early stage 10 chamber; the stretched nurse cell FC (NCFC) lie at the anterior end (left) and oocyte FC (OFC) lie at the posterior end as indicated. The *GLO*-expressing cells have failed to either separate like normal NCFC or reach the oocyte with the other OFC. Green, WGA; blue, SYBR Green. Arrows in **A** and **B** indicate the normal direction of OFC translocation relative to the underlying germ cells. **B**. Z-projection through the entire FC layer on the upper side of a stage 9 chamber. Large *GLO* clone fails to translocate posteriorly as the WT cells do. Green, f-actin; blue, SYBR Green. All three channels are shown in original contrast. **C.** Basal circumferential actin cables are disrupted only in the clone (arrowhead) of a stage 10 chamber. Green, f-actin; blue, SYBR Green. **D.** Clone in stage 9 OFCs, Z-projection. The edge of the clone is rounded (arrowhead), unlike a typical group of cells, which would show a hexagonal pattern (as in Fig. 8C,E). Green, SYBR Green; blue, WGA. **E–G.** *GLO* also alters cell shapes in the apical-basal direction. Cross-sections of OFC clones; the apical (inner) margin of the epithelium, adjacent to the oocyte

(indicated), is invaginated toward the basal (outer) edge just at the border of each clone (arrowheads). **E.** XZ section through posterior end of a stage 8 chamber, bottom edge of the chamber is pressed against the coverslip. Green, WGA; blue, SYBR Green. **F.** Z-projection along OFC, ~stage 10. WGA-stained material is usually a marker for the apical surface, and never seen laterally in wild type; here is it being deposited laterally at the left edge of the clone (arrowhead). Green, WGA; blue, SYBR Green. **G.** Single cross section of a clone, end of stage 9; green, f-actin; blue, SYBR Green.

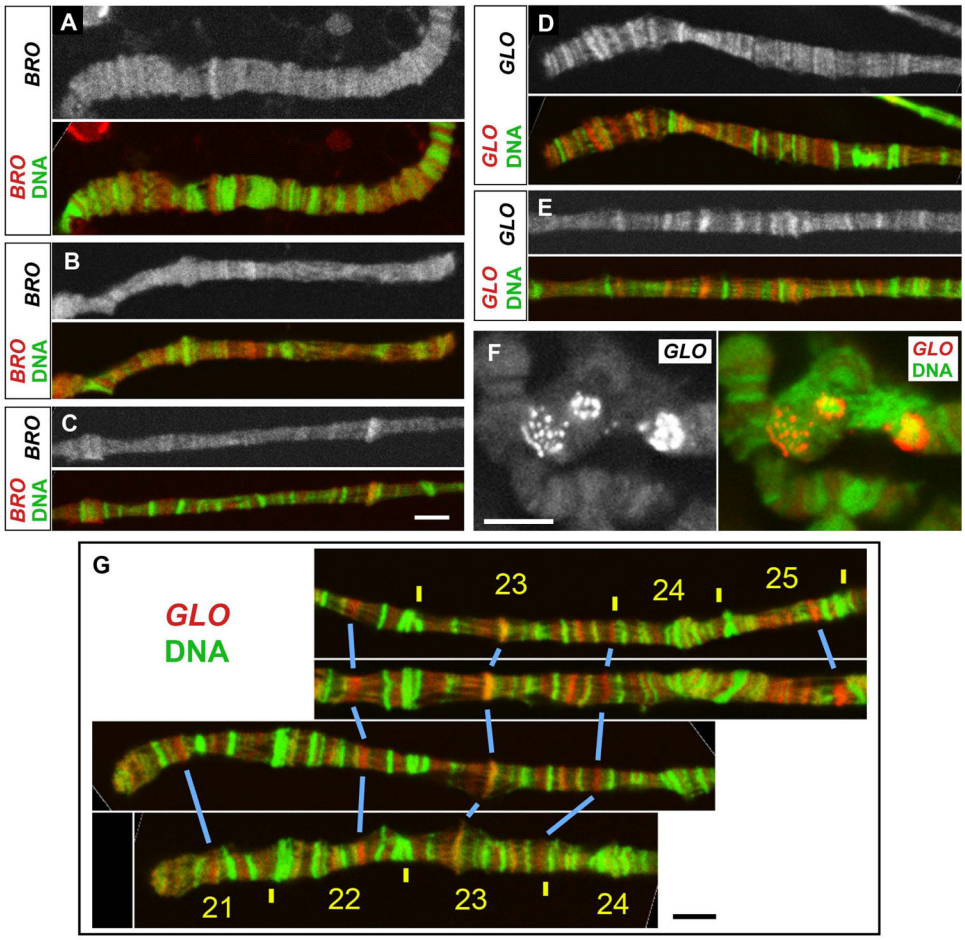

Figure 6. Banding patterns of Cut (*BRO*) and Drgx (*GLO*) fusion proteins on polytene chromosomes

In each pair of images, as indicated, the RFP fusion protein is shown in grayscale and in the red channel of the overlay, and SYBR Green staining of DNA is shown in the green channel of the overlay. Bars, 5 μm (bar in **C** is for panels **A**–**E**). **A**–**C**. *BRO* fusion shows weak banding, with most bands only moderately brighter than the background. **D**–**E**. *GLO* fusion accumulates in numerous strong bands. **F**. *GLO* fusion binds to discrete sites in the chromocenter. **G**. Reproducibility of the *GLO* banding pattern seen in four chromosome segments corresponding to distal *2L*; numbered polytene divisions are in yellow; homologous RFP bands are indicated by blue lines.

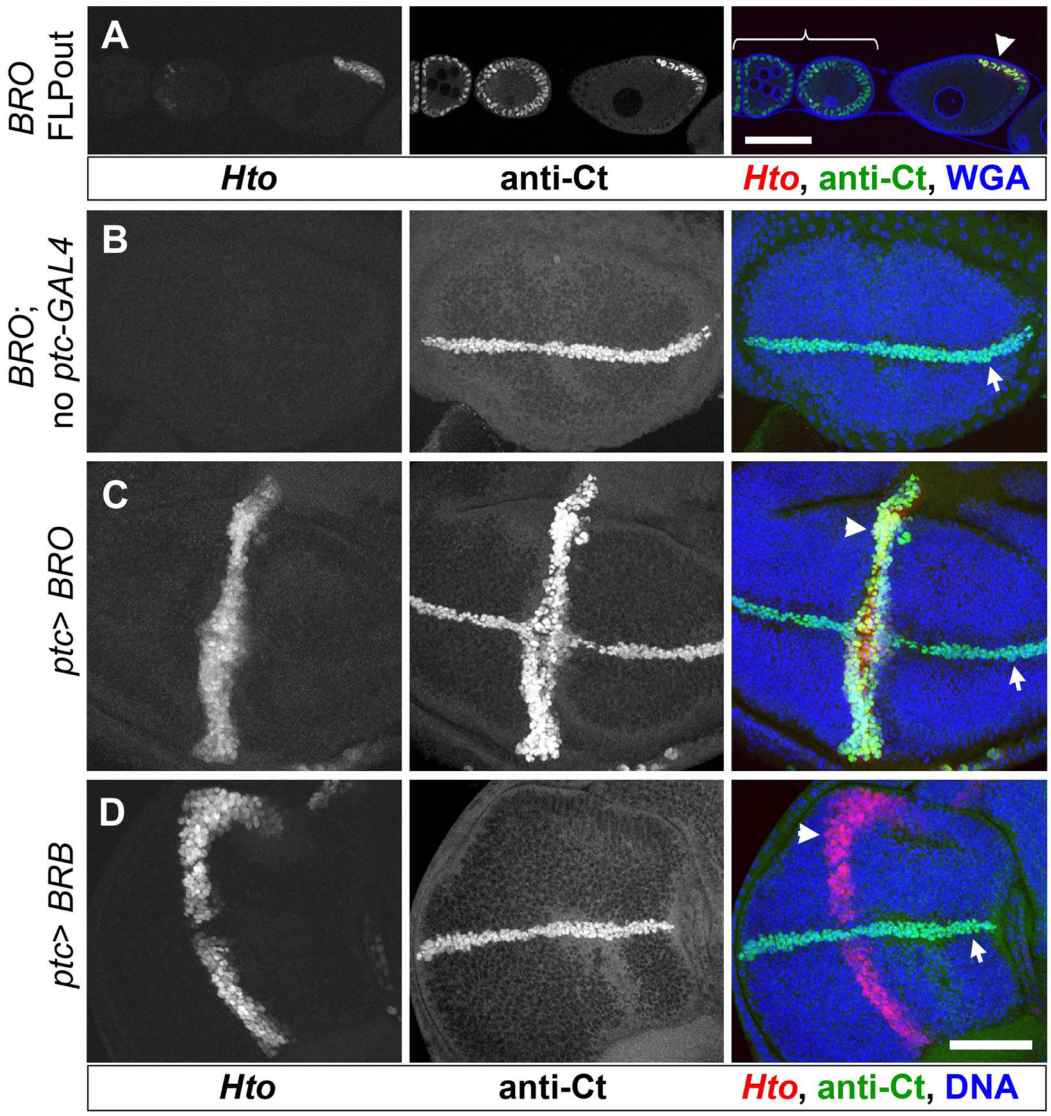

Figure 7. Anti-Ct immunostaining

Ct antibody 2B10 (middle column and green in merge) detects *BRO* expression in egg chambers and imaginal discs. **A**. *FLP*-out clone (arrowhead) expressing *BRO* in a stage 7 egg chamber is detectable by anti-Ct antibody. Ct is also expressed endogenously by FC up to stage 6 (bracket; Jackson & Blochlinger 1997), but the clone expresses Ct at a higher level. Bar, 50 μm. **B**–**D**. Ectopic *BRO* and *BRB* in 3rd instar wing discs stained with Ct antibody (Z-stack projections). Each disc has a stripe of Ct-positive cells at the presumptive wing margin (arrows). *Hto* inserts *BRO* or *BRB* were expressed in a perpendicular stripe (arrowheads) using *ptc-GAL4*, visible by RFP expression (left column). **B**. Control with *BRO* insert but lacking *ptc-GAL4* has only the endogenous Ct pattern. **C**. *ptc>BRO* also has strong Ct expression in the *ptc* stripe, generating a cross pattern. The epithelium is folded inward along the *ptc* stripe, visible here as the vertical gap in the center. **D**. *ptc>BRB*, expressing Poxn fusion protein, does not induce Ct expression in the *ptc* stripe. In later discs

(not shown), the *ptc>BRB* stripe grows to over double the width shown here, but does not fold inward like the *ptc>BRO* stripe. Bar in **D**, 50 μm for **B**–**D**.

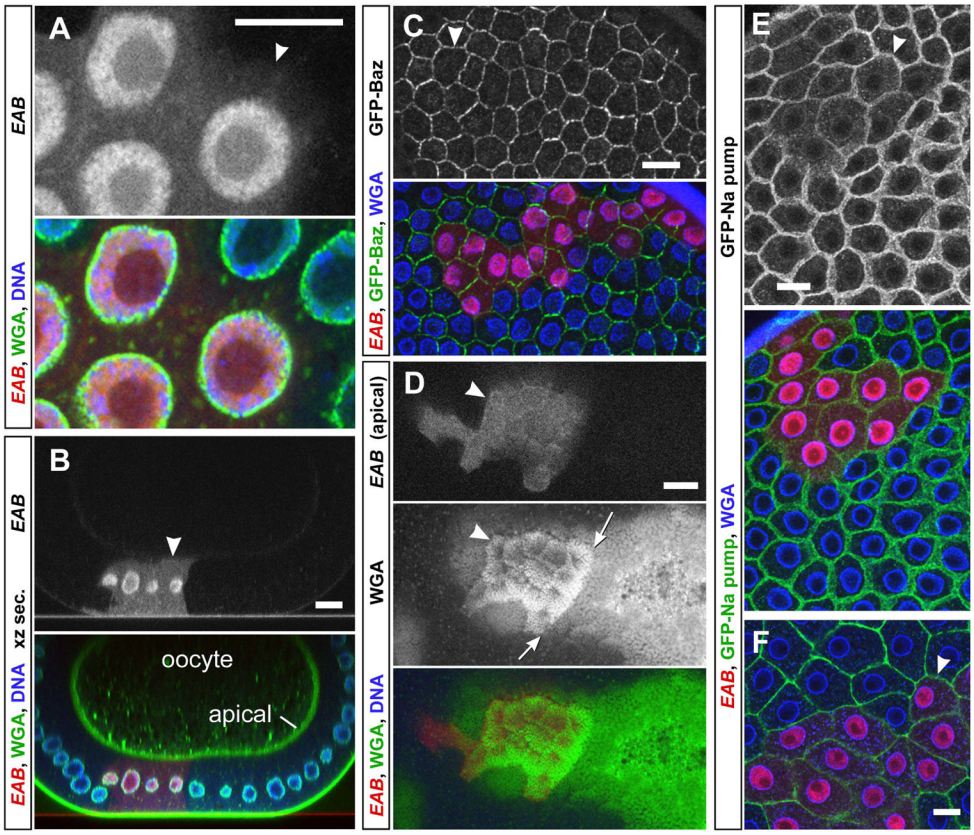

Figure 8. Effects of ectopic Zfh2 fusion protein produced by the *EAB* insert

Clones of ovary FCs expressing the *EAB* product (arrowheads) are shown as in Fig 5. *EAB* fusion protein is red in each overlay image; other markers are as indicated. Bars, 10 μm in each set. **A**. Closeup of stage 10 *EAB* clone shows mottled distribution of *EAB* fusion in the nucleus. **B**. XZ section through stage 9 chamber, bottom edge of the chamber is pressed against the coverslip. The apical surface (indicated) is uniform, unlike *GLO* clones (Fig. 5), but the nuclei are shifted apically. **C.** Z-projection of an *EAB* clone in OFC layer of stage 8 chamber; GFP-Baz (protein trap $baz^{CC01941}$, green) is expressed in all cells. **D.** Section tangential to the apical edge of the OFC layer where it contacts the oocyte, stage 10. The speckled texture of the WGA signal (green, shown in original contrast) is due to the presence of OFC microvilli interspersed with WGA-positive material at this position. Part of the clone border is indicated by arrows; note the large enrichment of WGA-positive eggshell material produced by the clone. **E**–**F.** OFC clones expressing *EAB*; the chambers also carry a GFP-Sodium pump protein trap ($Atp\alpha^{G00109}$, green). The GFP-Sodium pump should be evenly expressed by all cells, but is greatly reduced on the plasma membrane within each clone. **E.** Z-projection through stage 10 OFC layer. **F**. Plane through stage 12 chamber, when the OFC layer becomes squamous; GFP-Sodium pump plasma membrane signal remains very low in the clone.

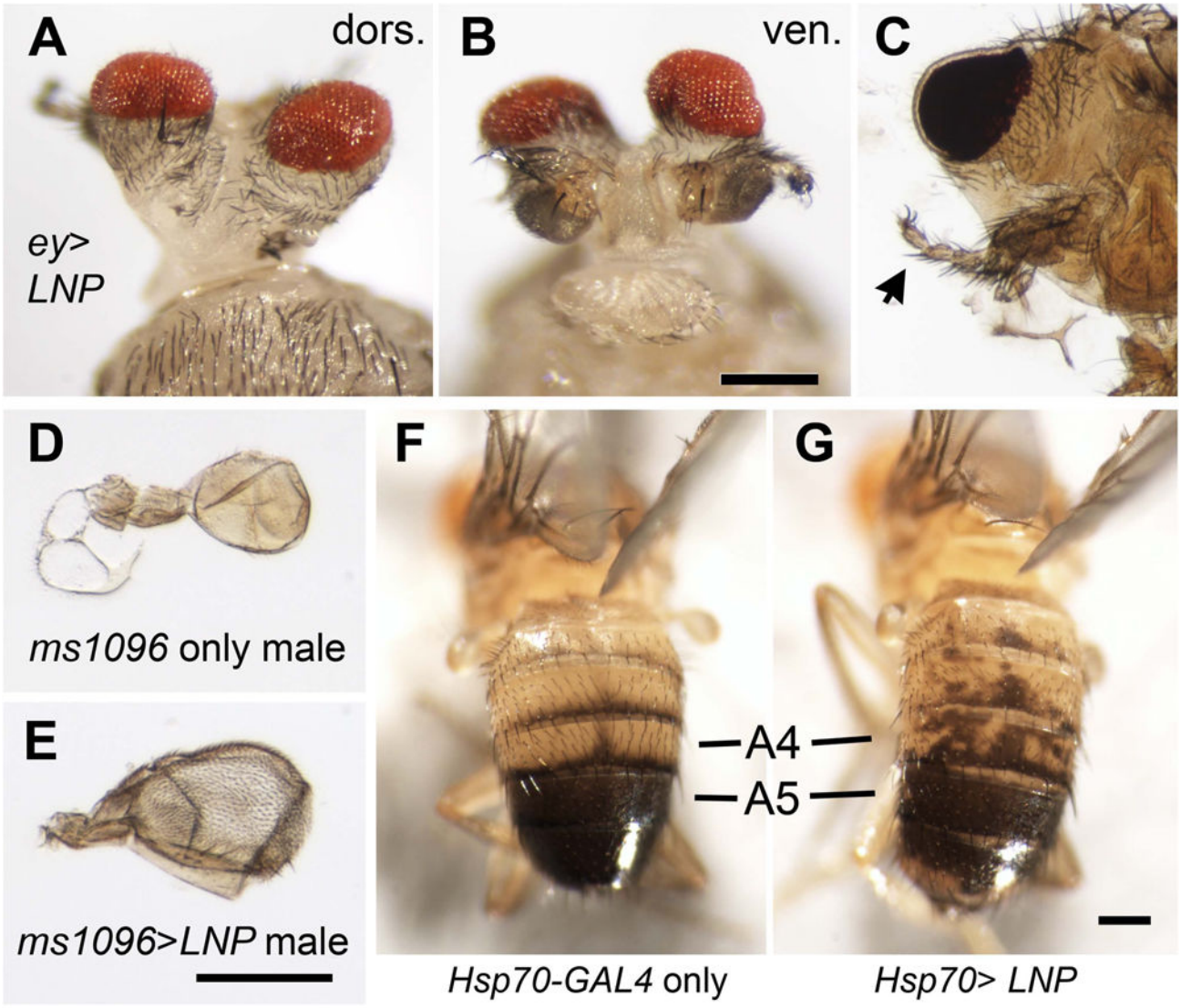

Figure 9. Homeotic effects of *LNP* expression

Bars, 200 μm for each set. **A**–**B**. Pharate adult of *ey>LNP* at 18°C; head pattern is converted to appendage-like stalks; **A**, dorsal view; **B**, ventral view. **C**. Pharate adult head of *ey>LNP* at 18°C was mounted flat to view the antenna more clearly (arrow); note antenna conversion toward leg identity. **D**–**E.** *ms1096>LNP* causes haltere transformation toward wing; note wing margin formed on anterior (upper) edge in (**E**). **F**–**G**. Ectopic expression of *LNP* under heat shock control leads to ectopic dark pigment in abdominal segments A3–4; A5 is normally dark (**F**).

Table 1

Hto inserts affecting Homeobox and Pax genes

***Hto* insert name**	**Target gene**	**Homeobox Class and other notes on the target gene/protein**	**Chromosome: Insertion site (orientation)[1]**
Inserts that make FLAG-mCherry-transcription factor fusion proteins			
BRO	*cut* (*ct*)	CUT Class. 2383 aa protein with 3 Cut domains and one homeodomain	***X***: 7,581,688 (+)
GLO	*Dorsal root ganglia homeobox* (*Drgx*)/*CG34340*	PRD Class. Ortholog of Human Dorsal Root Ganglia Homeobox	***2L***: 3,669,091 (+)
BRB	*Pox neuro* (*Poxn*)	Similar to Human Pax 2/5/8; Paired box transcription factor lacking a homeodomain	***2R***: 15,829,437 (−)
FAR	*araucan* (*ara*) [in Iro-C]	TALE Class; IRO family homeodomain	***3L***: 12,580,466 (+)
EAB	*zn finger homeodomain 2* (*zfh2*)	ZF Class. 3005 aa protein with 3 homeodomains and multiple Zn fingers	***4***: 494,420 (+)
OMD	*shaven* (*sv*)	PRD Class. Also called D-Pax2; Paired box transcription factor with partial homeodomain	***4***: 1,078,986 (+)
Inserts that affect Hox loci without making transcription factor fusion proteins			
BLY	*Sex combs reduced* (*Scr*); downstream genes? [in ANT-C]	ANTP Class. *Hto* exon 1 is out of frame with *Scr* coding region	***3R***: 6,854,635 (−)
LNP	*iab-8* and *abdominal-A* (*abd-A*) [in BX-C]	*iab-8* is a long ncRNA; *LNP* transcript can further splice in to *abd-A* and may express mir-iab-8	***3R***: 16,898,163 (−)

[1] Genome coordinates are from Release 6.