Investigating the use of enhancer trapping and Cas9-induced rearrangements to characterize *cis*-regulatory elements in the *Drosophila* gene *hindsight*

by

Brittney Lato

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Science
in
Biology

Waterloo, Ontario, Canada, 2022

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis including any required final revisions, as accepted by my examiners. I understand that my thesis may be electronically available to the public.

Abstract

The gene *hindsight (hnt)* encodes a transcription factor that is essential for all stages of *Drosophila* development. Expression of *hnt* has been observed to promote cell specialization and differentiation of several cell lineages at different stages. Consequently, the temporal and spatial expression of *hnt* is both dynamic and complex. A tremendous amount of research in this gene has been dedicated to its protein activity; however, little is known about the regulatory mechanisms that modulate its expression. Two genetic approaches have been applied to the 5' region of *hnt* to investigate its role in *hnt* regulation, including enhancer-based techniques and Cas9 multiplexing. Additionally, the use of Cas9 multiplexing was investigated for its feasibility in inducing chromosomal rearrangements (CRs). This was performed in the upstream region of *hnt,* as a means to study potential regulatory sequences associated with the gene.

In investigating this region with an enhancer trap-based method, putative elements that directly influence *hnt* expression have been characterized in terms of the cells and stages in which they influence *hnt* expression. In investigating the upstream region of *hnt* using Cas9-induced CRs, two lines with putative deletions in the *ruby (rb)* region have been recovered. Both lines display misexpression of *hnt*, indicating the presence of *hnt*-associated elements within the deleted region. While deletions using this system was successful, no translocations or inversions were recovered from Cas9 multiplexing. Results indicate that the use of Cas9 multiplexing without the use of homology-based techniques does not provide a feasible method for inducing CRs.

Acknowledgments

Thank you to Dr. Bruce Reed for the opportunity to work in your laboratory. I am truly appreciative for the skills I have acquired and count myself lucky to have learned “the awesome power of fly genetics” under your supervision. I would also like to show gratitude to my committee members. Thank you to Dr. Susan Lolle for your endless support and guidance. Your input has been more valuable than I can express. Thank you to Dr. David Rose for your accommodation and encouragement; I am truly grateful.

To all the “Department of Biology” friends that I have had the privilege of meeting throughout my degree, thank you for your encouragement and friendship. A special thank you to present and past members of the Reed lab, including Emily Baker, Alex Ng, Richard Do, and Mackenzie Valley. I am so thankful for all your kind words and support. As well, thank you April Wetting for your being a kind face and a helpful hand throughout my degree. Finally, thank you to all my friends and family who were there for me throughout this process. Special thanks to my cat Einstein, for doing this degree for me. I appreciate you all more that words can describe.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Bacteria-one hybrid (B1H)

Bolwig's Organ (BO)

Bone Morphogenic Protein (BMP)

Central Nervous System (CNS)

Chromatin Immunoprecipitation (ChIP)

Chromosomal Rearrangements (CRs)

Chromosome Confirmation Capture (3C)

Clustered Regulatory Interspaced Short Palindromic Repeats (CRISPR)

Chordotonal Organ Precursors (COPs)

Dead Cas9 (dCas9)

Dendritic Arborisation (DA)

Double Strand Break (DSB)

Epidermal Growth Factor Receptor (EGFR)

Gene of interest (GOI)

Germ Band Retraction (GBR)

Glass Multiple Reporter (GMR)

Green Florescent Protein (GFP)

Guide RNA (gRNA)

Hindsight (Hnt)

Homology Directed Repair (HDR)

Homologous Recombination (HR),

Jitterbug (jbug),

Jun kinase (JNK)

Leading Edge (LE).

Loss-Of-Function (LOF)

Mesenchyme to Epithelial (M/E)

Microhomology-Mediated End Joining (MMEJ).

Mother of Dpp (Mad)

Multi-Dendritic (MD)

Nanos (nos)

Natural Goat Serum (N g/s)

Non-Coding RNAs (ncRNAs)

Non-Homologous End Joining (NHEJ)

Olfactory Sensory Neurons (OSN)

Photoreceptor (PR)

Peripheral Nervous System (PNS)

Pre-Initiation Complex (PIC)

Protospacer Adjacent Motif (PAM)

Red fluorescent protein (RFP)

ruby (rb)

Scalloped (sd)

Senseless (sens)

Sensory Organ Precursor (SOP)

Systematic Evolution of Ligands by EXponential (SELEX)

Transcription Factor Binding Site (TFBS)

Transcription Start Site (TSS)

Untranslated Region (UTR)

Upstream Activation Sequence (UAS)

Ventral Nerve Cord (VNC)

Vienna tile (VT)

White Eraser (WE)

Chapter 1: General Background

1.1. The role of intergenic DNA in gene regulation

1.1.1 Function and activity of a gene's regulatory regions

It has been long appreciated that a remarkable portion of a genome's nonprotein-coding DNA is functionally important for the regulation of gene expression. This is particularly true for eukaryotes, where differential gene expression is brought about by mechanisms of gene regulation. DNA regions outside of genes, which make up most of the eukaryotic genome, is known as intergenic DNA. When considering humans, for example, less than 15% of the genome is transcribed into RNA. Originally, the remaining intergenic DNA was thought of as "junk DNA". Further research has recognized that "junk DNA" is required for the regulation of gene expression during embryogenesis and homeostasis. Regulatory mechanisms controlling differential gene expression are a major driving force in the evolution of eukaryotes. This is exemplified by observations where two organisms contain nearly identical gene, but vary greatly in terms of complexity and phenotype (Rodriguez-Trelles, Tarríol, & Ayala, 2003). This change in phenotype can often be attributed to subtle alterations in the regulation of genes that direct embryogenesis.

In addition to development and evolution, differential gene expression may also be responsible for differences between healthy and disease conditions. Gene expression can be regulated at many different levels. Mainly, regulation can be through RNA processing, transport, and localization; moreover, at the protein level there may be translational and post-translational modifications. Intergenic DNA often houses *cis*-regulatory elements involved in gene expression regulation at the level of transcription. The process of transcription, carried out by RNA

polymerase, is finely controlled by transcription factors. Interestingly, close to 8% of genes in the human genome encode transcription factors (Messina, Glasscock, Gish, & Lovett, 2004). These proteins bind to DNA in a sequence-specific fashion and interact with RNA polymerase to regulate its activity, but the intricacies of transcription factor interactions and function remain topics of active research. Transcription factors can be expressed ubiquitously or specifically in response to their cellular environment. Since the type and concentration of transcription factors can be selectively altered depending on cell properties or environmental signals, transcription can significantly vary across the spatial and temporal conditions of cells. Transcription factors bind to transcription factor binding sites (TFBS). These sites, referred to as *cis*-regulatory elements, are typically housed in intergenic DNA.

Generally, transcription starts at a gene's core promoter, which houses the transcriptional start site (TSS) and a series of consensus sequences that specify protein binding interactions to a site. The TSS is the site of pre-initiation complex (PIC) formation and activation; the PIC includes basal transcription factors and RNA polymerase II (Carlberg & Molnár, 2014). The activity of the core promoter itself does not contribute any significant function for a gene's regulation, as gene activity is largely modulated by the transcription factors that bind to a complex organization of *cis*-regulatory elements. These elements are typically six to ten nucleotides long, and can be distal or proximal to the core promoter (Rodriguez-Trelles et al., 2003). They can interact or come in proximity with the PIC using looping mechanisms, ultimately permitting transcription factors to regulate RNA polymerase activity. A *cis*-regulatory element can be located virtually anywhere in the genome relative to its associated gene, however the probability of finding these sites decrease with increasing distance from the gene's core promoter (Carlberg & Molnár, 2014).

Regulatory elements outside of the core promoter include proximal/distal promoter elements, enhancers, silencers, and insulators. The proximal/distal promoter elements are located close to the core promoter. In addition to acting as a traditional activator of transcription, it is hypothesized that distant regulatory elements interact with these proximal/distal promoter elements through the association of various protein-protein interactions. This is important since *cis*-regulatory elements can be very distant from their target core promoter (up to 1 Mb away from a gene) (Mora, Sandve, Gabrielsen, & Eskeland, 2016). Transcription factors can interact with proteins in the PIC or proximal/distal promoters to influence the formation of the PIC, or alter PIC-promoter affinity. While mechanisms of enhancer activity are not universally defined, it is generally thought that several transcription factors and different *cis*-regulatory elements are often required in regulating the expression of a given gene. While *cis*-regulatory elements are generally known as enhancers, they do not always work to upregulate expression. *Cis*-regulatory elements, referred to as silencers, can down-regulate transcription. Accordingly, some transcription factors bind *cis*-regulatory elements to block the formation of the PIC. These are known as repressors.

A single enhancer or silencer can alter the transcription of multiple genes. Such interactions, however, may be deleterious. To prevent these interactions, intergenic DNA may contain sequences that function as insulators. Insulators are usually situated between regulatory elements and are thought to block the interaction between *cis*-regulatory elements and promoters. One mechanism of this involves interfering with the DNA looping that is required for these elements to be brought within proximity of each other (Carlberg & Molnár, 2014; Riethoven, 2010). Small changes, even by a single base pair, within the sequence of enhancers, silencers and insulators (which can be collectively referred to as *cis*-regulatory elements) can greatly alter gene

expression. Differential gene expression conferred by *cis*-regulatory elements demonstrates the significance of intragenic DNA with respect to embryogenesis, as well as the survival and evolution of an organism (Rodriguez-Trelles et al., 2003). In order to fully explore and understand the intricacy of the genome, it is imperative to develop a better understanding of *cis*-regulatory elements.

1.1.2 Methods of studying cis-regulatory elements

Bioinformatic approaches alone are quite limited when it comes to identifying *cis*-regulatory elements. Techniques to compare multiple genomes of related species can identify functionally important DNA, as such sequences may show evolutionary conservation. However, this method is more useful for identifying coding sequences, as the complexity of *cis*-regulatory DNA often make alignment analyses difficult to implement for several reasons. First, unlike protein coding DNA, regulatory regions lack comparable properties. There is no consistent hallmark that indicates the presence of *cis*-regulatory elements solely through DNA sequence analysis. Additionally, *cis*-regulatory elements are not binary systems that act to simply increase or decrease transcription. The complexities of these regions make the molecular interactions of these regulatory regions tremendously context-dependent. Examples include competitive or cooperative binding of proteins at the site of regulatory regions, as well as chromatin bending that can alter transcription. Such features complicate the identification and analysis of regions of genome that may harbour *cis*-regulatory elements (Rodriguez-Trelles et al., 2003).

Biochemical approaches offer a more reliable method of predicting elements using the binding sequences of isolated transcription factors. This was originally accomplished using electrophoretic mobility shift assays. This technique is based on the notion that protein-bound probes will travel slower than unbound probes when mixtures are separated by gel

electrophoresis. Random DNA fragments can be run with a purified transcription factor to isolate sequences with a TFBS that can bind to the protein (Garner & Revzin, 1981). Systematic Evolution of Ligands by EXponential (SELEX) treatment cross-links mixtures of DNA fragments to a purified transcription factor. Following this, DNA-protein complexes are immunoprecipitated and sequences are separated from the transcription factor. This process is repeated several times before the DNA fragments are purified, and a consensus binding site is made based on sequences that are preferentially recovered. Alternatively, a method known as chromatin immunoprecipitation (ChIP) permits the analysis of TFBS *in vivo.* Protein interactions within the genome are stabilized by cross-linking, and DNA is sheared to into 100-500bp fragments. A transcription factor of interest can be immunoprecipitated and purified to identify any interactions with DNA (Geertz & Maerkl, 2010).

The complexities of regulatory regions also make this method difficult to implement, partly due to the nucleotide sequences not completely representing what makes a genomic location a regulatory region. More importantly, the short sequences of DNA that make up regulatory regions are likely to appear in the genome by chance. (Rodriguez-Trelles et al., 2003). More recent methods to study gene regulation are based on the detection of inter-and-intra chromosomal interactions. Chromosome confirmation capture (3C)-based technologies can map interactions between *cis*-regulatory elements using the DNA loop that is formed from protein-protein and protein-DNA interactions. Such interactions are stabilized by cross-linking them and isolated using immunoprecipitation. Chimeric DNA molecules from the two non-adjacent sequences can be made from these interactions, and used to identify regulatory elements influencing a given gene (Carlberg & Molnár, 2014). The bacterial one-hybrid (B1H) system can identify target sites of binding domain without the use of antibodies. DNA fragments that

represent putative *cis*-elements are cloned into individual vectors with positive and negative selection markers into *E. coli,* as well as a vector encoding a fusion protein of a desired transcription factor and RNA polymerase. If the transcription factor binds to the fragment, the promoter of the selectable markers will recruit RNA polymerase and activate gene expression.

Genetic approaches are often used to truly confirm the regulatory function of DNA sequences. One approach inserts reporter gene constructs in the genome. Since genes are influenced by proximal regulatory elements, the expression of this gene will reflect the activity of endogenous elements. Larger constructs that can be inserted randomly in the genome include reporter genes under the influence of candidate regulatory elements. This approach often comes with the caveat that results may not reflect the normal pattern of endogenous gene expression. Unfortunately, there is a limited size of constructs that can be recovered and tested. Additionally, position effects observed by regulatory elements can result in altered expression of the reporter gene. One possible explanation for this is steric hindrance of the regulatory elements. Finally, the promoter of the reporter gene may respond to regulatory factors endogenous to the integration site of the construct.

Reporter constructs can be coupled with deletion mutagenesis. This is followed by detection for altered reporter expression. Deletion mutagenesis of sequences upstream or downstream of a gene can also be followed by protein or RNA expression analysis; this may provide exact locations of regulatory elements (Daniel, Nagy, & Nagy, 2014; Geertz & Maerkl, 2010) Outside of TFBSs, other regulation forces can alter gene regulation. For instance, genomic location may also influence gene regulation. DNA methylation has been associated with the repression of gene expression. (Carlberg & Molnár, 2014; Daniel et al., 2014). Such mechanisms would be refractory to many of the methods discussed above.

1.2. *Hindsight*

1.2.1. Characterized expression of *hindsight* within the *Drosophila* life cycle

Immunostaining analysis has demonstrated Hnt expression is highly regulated in the embryo. In the developing embryo, *hnt* expression is observed in the extraembryonic amnioserosa, peripheral nervous system (PNS), midgut, oenocytes and developing tracheae. In larval stages, expression is present in the lymph gland, differentiated crystal cells, imaginal myoblasts, and salivary glands. During adult stages, it is present in the follicle and border cells, midgut, spermathecae, and wing neurons (Kim et al., 2020; Reed, Wilk, & Lipshitz, 2001).

1.2.2. The characterization and function *Hindsight*

The gene *hindsight* (*hnt*), a gene located on the X chromosome of *Drosophila* ("FlyBase," 2020), encodes a C_2H_2 type zinc finger transcription factor. Several genes have been identified as likely being direct transcriptional targets of Hnt. This includes *hnt* itself through negative autoregulation, and the genes *nervy* and *jitterbug (jbug)* (Kim et al., 2020; Reed et al., 2001). SELEX identified two consensus binding sequences of Hnt as "CAGCATCC" and "YGGWCCA" (Ming, Wilk, Reed, & Lipshitz, 2013). An independent study using the B1H system also identified "AGCATCY" as a consensus binding sequence for Hnt (Zhu et al., 2011). The human homolog of *hnt*, *Ras Responsive Element Binding Protein-1* (*RREB-1*), also encodes a C_2H_2 type zinc finger transcription factor that is linked to the offset of several cancers due to its role in epidermal growth factor receptor (EGFR) signalling (Ming et al., 2013). The distinctive U-shaped embryonic lethal phenotype of *hnt* mutants was first characterized in genetic screens performed in the 1980s. This U-shaped phenotype is due to premature degradation and death of the amnioserosa and a failure in the embryonic morphogenic process known as germ band retraction (GBR) during embryogenesis. Interestingly, expression of various genes that activate

the EGFR/Ras/MAPK pathway, as well as expression of baculovirus caspase inhibitor p35, can rescue this phenotype (Kim et al., 2020). Overall, the expression of *hnt* is dynamic and complex. This expression is required for proper embryo, larva, and adult development (Ming et al., 2013).

In addition to its role in embryogenesis, *hnt* has been implicated in the regulation of several other developmental processes. These include maintenance of the integrity of the developing tracheal system (Ming et al., 2013), axon guidance in the developing adult visual system (Oliva et al., 2015), enteroblast to enterocyte differentiation (Kim et al., 2020), the differentiation and homeostasis of the adult midgut epithelium (Baechler, McKnight, Pruchnicki, Biro, & Reed, 2016), stages of follicle cells during development and egg rupture during adult fly oogenesis (Deady, Li, & Sun, 2017). In addition, *hnt* has been identified as a direct target of the Notch signalling pathway in the context of the myoblast differentiation during the larval/pupal transition, (Krejci, A., Bernard, Housden, & Bray, 2009), as well as differentiation of crystal cells in the larval/pupal lymph system (Baonza & Garcia-Bellido, 1999)

Overall, the function of *hnt* appears to be highly context-specific, as it acts to transcriptionally activate and repress genes that are involved in several different types of cell differentiation. With respect to its role in the regulation of downstream target genes, the regulation of the *hnt* itself is of interest. One region that is a probable candidate for cis-regulatory elements of *hnt* includes the large 5' region of the gene, in which there is ~40 Kb of non-protein coding sequence. Some characterization of this region has been reported, including three Notch-responsive elements (Terriente-Felix et al., 2013), as well as a region implicated in *hnt* negative autoregulation (Ming et al., 2013). It is clear, however, that many other aspects of the complex pattern of *hnt* expression remain to be elucidated

1.3 Research Rational

The regulation of *hnt* remains relatively understudied, which can be attributed to two factors. First, the gene contains a large upstream region (>50Kb), which provides a likely location for the gene's regulators; and second, the importance of gene regulation has only recently become of interest, as functions of the protein-coding aspect of genes have been the primary focus of research within the Reed lab as well as other research groups. This project uses two approaches to characterize segments of DNA within the 5' *hnt* region. The first method used is based on a technique called enhancer trapping. The second method investigates the use of Cas9 to induce CRs targeted to the 5' regulatory region of *hnt*. While any CRs recovered may be useful in characterizing the large 5' regulatory region of *hnt*, these experiments will also serve as a proof-of-principle for the feasibility of using CRISPR/Cas9 genome editing to design and recover CRs.

Chapter 2: Characterizing the 5' regulatory region of *hnt* using enhancer trapping

2.1 Introduction

2.1.1 Enhancer traps

The enhancer detection method, referred to as enhancer trapping, offers a convenient approach to observing transcription without performing whole mount *in situ* hybridization. The first instance of this method used *LacZ* under the control of a minimal promoter, cloned into a defective P element transposon (Bellen et al., 1989). This transgene was inserted randomly in different locations in the genome, and many different single insertion transgenic lines were recovered. It came as a surprise that many of these lines displayed unique patterns the *LacZ* reporter gene expression. The explanation for this was based on position effects, whereby the expression of *LacZ* is influenced by enhancer elements within proximity to the P element insertion site (Bellen et al., 1989). Since these initial experiments, the enhancer trap technique has evolved to become a versatile "toolbox" for studying regulatory elements. This includes using different reporter genes and bipartite systems for inducible gene expression, as discussed below. Such methods are now standard approaches in *Drosophila,* as well as other model organisms, for characterizing *cis*-regulatory elements.

2.1.2 The Gal4/UAS system

The most commonly used and standard system for inducible reporter gene expression in *Drosophila* is based on a bipartite system adapted from *Saccharomyces*. The yeast's Gal4 transcription factor binds and activates an element known as the Upstream Activation Sequence (UAS). A gene's expression can be made inducible if it is placed under the control of a UAS

element. Transcription of this gene, frequently referred to as the "responder", will become dependent on Gal4 activation of UAS. Hence, there will only be expression of the reporter gene when Gal4 is present. Expressing Gal4 under the activity of heterologous enhancers, which normally function on their endogenous gene, can result in spatially or temporally restricted expression of the UAS transgene. The Gal4/UAS system has been adapted in several ways, including using a heat-shock promoters and temperature sensitive repressors of Gal4 ($Gal80^{ts}$). A gene placed under UAS control, the so-called responder gene, can be anything, including coding sequence for GFP, a dsRNA construct to facilitate RNAi knock-down, a guide RNA (gRNA) for genome editing, or any gene of interest for ectopic or overexpression experiments. Thousands of Gal4 driver lines have been either recovered or designed by *Drosophila* researchers. As Figure 1 shows, the offspring of GAL4xUAS crosses allow targeted expression of a transgene (McGuire, Roman, & Davis, 2004).

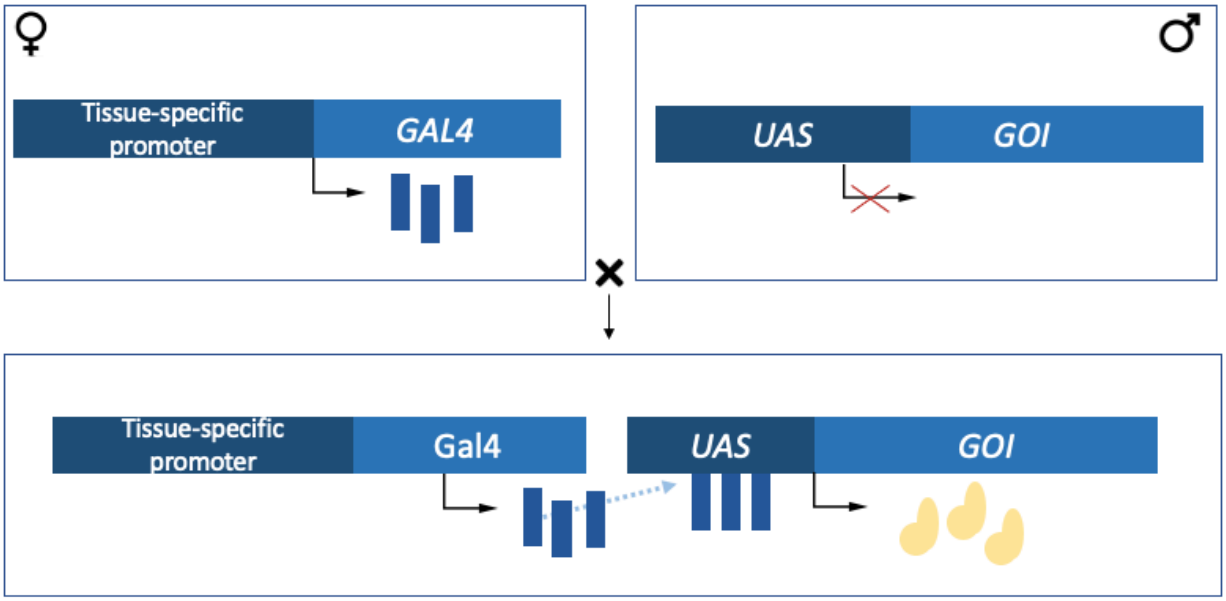


Figure 1. Targeted expression of a GOI using the GAL4/UAS system. In the parental cross, one line is a "GAL4 driver" that expresses of GAL4 under the control of a promoter. The promoter permits inducible expression of GAL4, promoter activity will be specific to tissue or time points during the life cycle. The other parent contains the GOI under the control of UAS. Offspring that inherit both chromosomes will have inducible expression of a GOI, as dictated by the GAL promoter. This is because UAS activation is restricted to regions in which GAL4 is present. In general, the expression of GAL4 alone does not disrupt endogenous genes expression during any stage of *Drosophila* development

2.1.3. Vienna Tile lines

The Vienna tile (VT) library was developed by the Dickson and Stark group at the Research Institute of Molecular Pathology (Kvon et al., 2014). The lines contain inserts of the Gal4 transgene under the control of short sequences, which have been identified as candidate regulatory elements throughout the genome. This massive project included candidate sequences located in the 5' region of *hnt*. The corresponding genomic locations of these sequences are shown in Figure 2 and Table 1 ("Vienna BioCentre Core Facilies," n.d.). The VT library uses the concept of enhancer trapping to generate a powerful way to study regulatory regions. Like traditional enhancer traps, the VT lines uses Gal4 as a reporter gene to identify transcriptional activity of endogenous *cis*-regulatory elements. The library was made using plasmid vectors that contain an attP2 site for phic31 mediated integration as a transgene, in addition to the *Gal4* gene under the expression of a minimal promoter and candidate regulatory sequences of ~2Kb in size (Kvon et al., 2014). For the purpose of this study, the VT lines are crossed with flies containing inserts of UAS-green florescent protein (GFP). This serves as a method to visualize Gal4 expression in live embryos.

The VT Gal4-lines are available at minimal cost to the *Drosophila* research community. Like most of the lines in the library, the *hnt* VT-Gal4 lines have not been characterized outside of their expression in the adult brain. Expression of UAS-GFP was visualized following crosses to VT-Gal4 lines to determine if these sequences had influence on the transcriptional activity of the *Gal4* reporter. The line with GFP expression has an insert of ubiquitously expressed histone-tagged red fluorescent protein (RFP), so that GFP can be visualized within a background where all tissues are labeled with RFP. This enhancer trap-based method permitted the visualization of UAS-GFP expression, as regulated by *hnt* regulatory elements driving Gal4 expression.

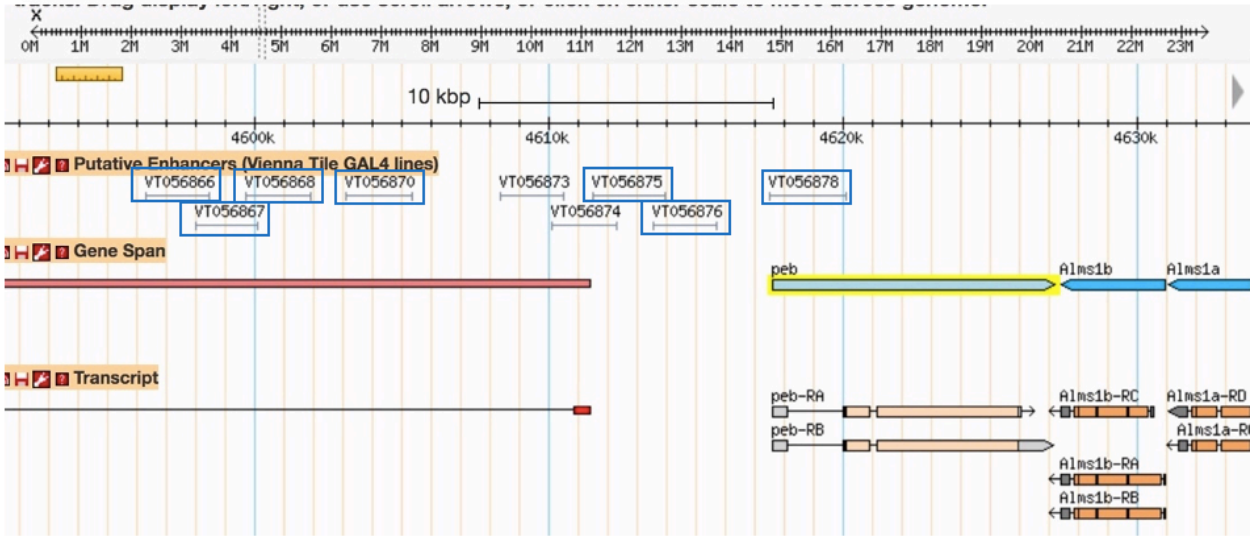


Figure 2. Inserts of VT constructs in the upstream region of *hnt*. The VT library was made using constructs of candidate enhancers, fused to a reporter gene with a minimal promoter. The gene *hnt* is labelled as "*peb*", the name used for *hnt* in Flybase. The VT inserts used for analysis upstream of *hnt* includes VT056966, VT056868, VT056867, VT056870, VT056875, VT056876, and VT056878. The genomic locations are shown above, their corresponding coordinates are present in Table 1 below. This picture is from Flybase (Flybase.Com). The studied lines are highlighted in blue boxes. Note that the VT056878 line houses a sequence that overlaps the 5' UTR and the first intron of *hnt*.

Construct Name	Coordinates
VT056866	4,596,278 - 4,598,442
VT056867	4,597,994 - 4,600,095
VT056868	4,599,681 - 4,601,865
VT056870	4,603,111 - 4,605,324
VT056875	4,611,478 - 6,613,920
VT056876	4,613,550 - 4,615,681
VT056878	4,617,474 - 4,620,070

Table 1. List of VT lines and associated genomic coordinates. The sequence coordinates of VT line inserts, labelled in figure 2. Information is extracted from Flybase (version FB2022_01, released February 8, 2022). Search for coordinate information was performed on February 22, 2022. At this time, the given coordinates for *hnt* is 4,617,564 - 4,627,206. Note that all coordinates provided are on the X chromosome.

2.2 Methods

2.2.1 Mounting Vienna Tile lines samples

Embryos were collected from the surface of standard *Drosophia* media 12 hours after crossing. Embryos were washed in 50% commercial bleach for 3-4 minutes, rinsed in water, and spread into a monolayer on a coverslip within a suspension of halocarbon oil. The coverslip was inverted onto a depression slide and embryos visualized by confocal microscopy. Larvae (2^{nd} – 3^{rd} intra) were dissected in halocarbon oil on a microscope slide by removing the anterior portion by mouth hooks so that the salivary glands, trachea, midgut, ganglian structures and gonads were separated body. Finally, 3–5-day old flies were dissected in halocarbon oil by removing heads and pulling out the tip of the abdomen. Larvae and adult dissections performed in halocarbon oil were covered with a cover slip. Dissected tissues were imaged by confocal microscopy.

2.2.2 Confocal imaging of Vienna Tile line expression

Z stacks of mounted embryos, larvae and adult flies were collected and processed using a Nikon Eclipse 90i microscope equipped with Nikon EZ-C1 software. Z stacks used a 2.00 μm step size intervals, and typically 30 steps were collected for each Z stack. Specimens were visualized using a 20X objective.

2.2.3 JASPAR analysis of putative enhancer sequences

Putative enhancer sequences of the VT-Gal4 constructs that reflect known expression patterns of *hnt* were compared to the conserved binding site sequences of known transcription factors. Using the publicly available JASAR program (http://jaspar.genereg.net). These analyses identified candidate sequences of intreat including *sensless (sens)* and elements that interact with the Mother of DPP (Mad) transcription factor.

2.3 Results

The VT lines resulted in GFP expression within the embryo, larvae and adult fly. Embryonic expression is mainly discussed, as this provides insight on the role of *hnt* during developmental stages. Studies of this nature are of interest due to the discussed role of *hnt* in cell differentiation. Given the diverse role of *hnt* in the embryo, immunostaining using anti-Hnt was done to compare confocal microscopy to a wild type model of *hnt* expression.

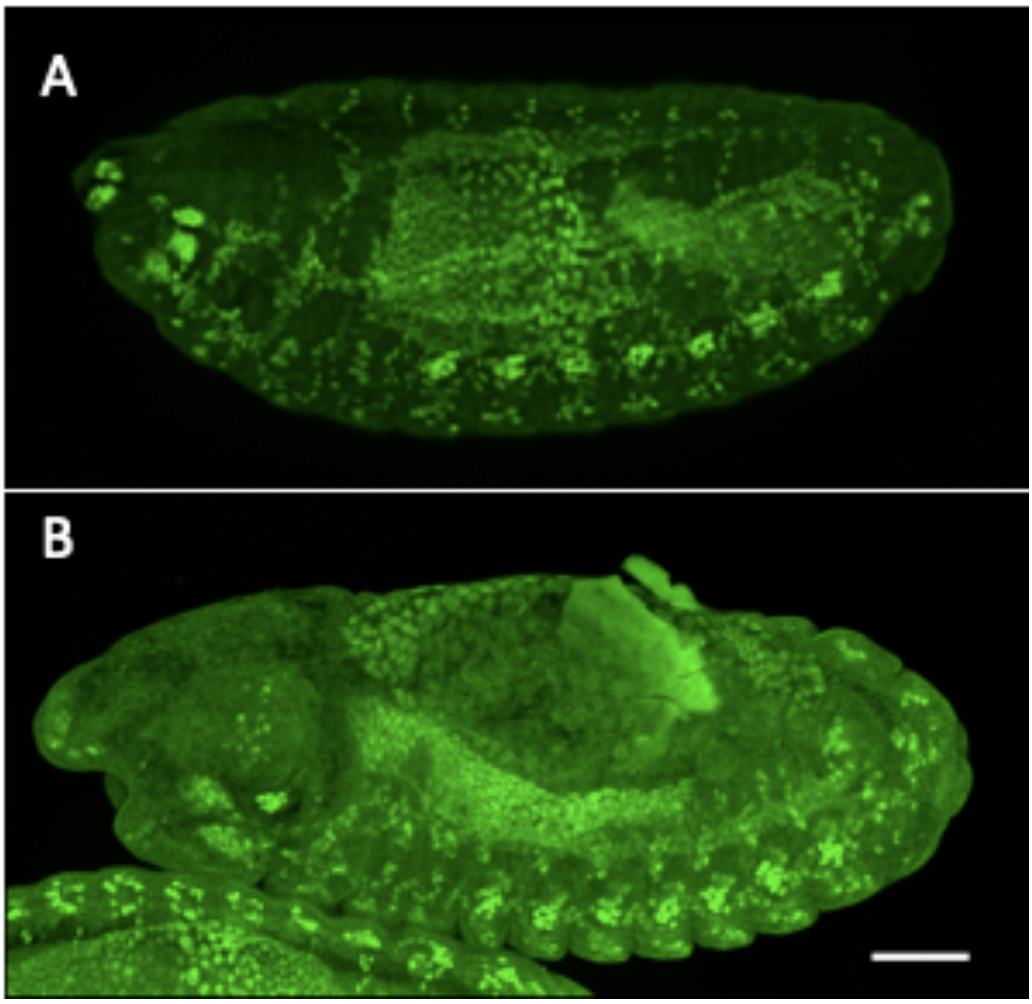


Figure 3. Expression of wild type Hnt using anti-Hnt immunostaining. Expression of the Hnt protein is mainly observed in 7 clusters along the dorsal side of the embryo. This represents sites of the developing oenocytes. There is expression within the amnioserosa throughout its development. There is also expression in four district clusters within the embryonic head. This is thought to represent the BO and developing olfactory organs. Scale bar represents 50 μm.

2.3.1 GFP expression in the *VT056866* Gal4-line

The *VT056866* Gal4-line is the most distal from the TSS of *hnt* of the lines examined. In this line, GFP expression begins in two small clusters of cells within the anterior embryo prior to dorsal closure (during mid embryogenesis). This appears to be presumptive neurons associated with the Bolwig's organ (BO), the larval light sensing organ (Figure 4, A). Expression of *hnt* within the BO is expected, as it has been observed within the Reed Lab from whole mount embryo immunostaining (B. Reed, personal communication). This is accompanied by expression throughout various components of the PNS, where cells are known to express *hnt* during embryogenesis. Within the larval stages, there is GFP expression within the axons of the brain's ventral nerve cord (VNC), as well as within brain hemispheres (Figure 4, B). The larval brain is not known to express *hnt*. However, *hnt* has been recognized to have a role in axonal development within the visual system (Oliva et al., 2015). The larval crystal cells have intense GFP expression (Figure 4, C), a tissue characterized to expresses *hnt*. This expression is a product of upregulation by the Notch signalling pathway, as the cell's differentiation into crystal cells is Notch-dependent (Terriente-Felix et al., 2013). Additionally, there is GFP expression within tracheal cells (Figure 3, D); this tissue is also known to express *hnt* during its development (Kim et al., 2020) . Like the larval stages, the line also activates GFP expression within the eye (Figure 4, E). While there is characterized expression of *hnt* within photoreceptor (PR) cells of the developing adult retina, a required step for eye development within pupae (Wilk, Pickup, Hamilton, Reed, & Lip, 2004), this expression does not normally persist in cells of the adult eye.

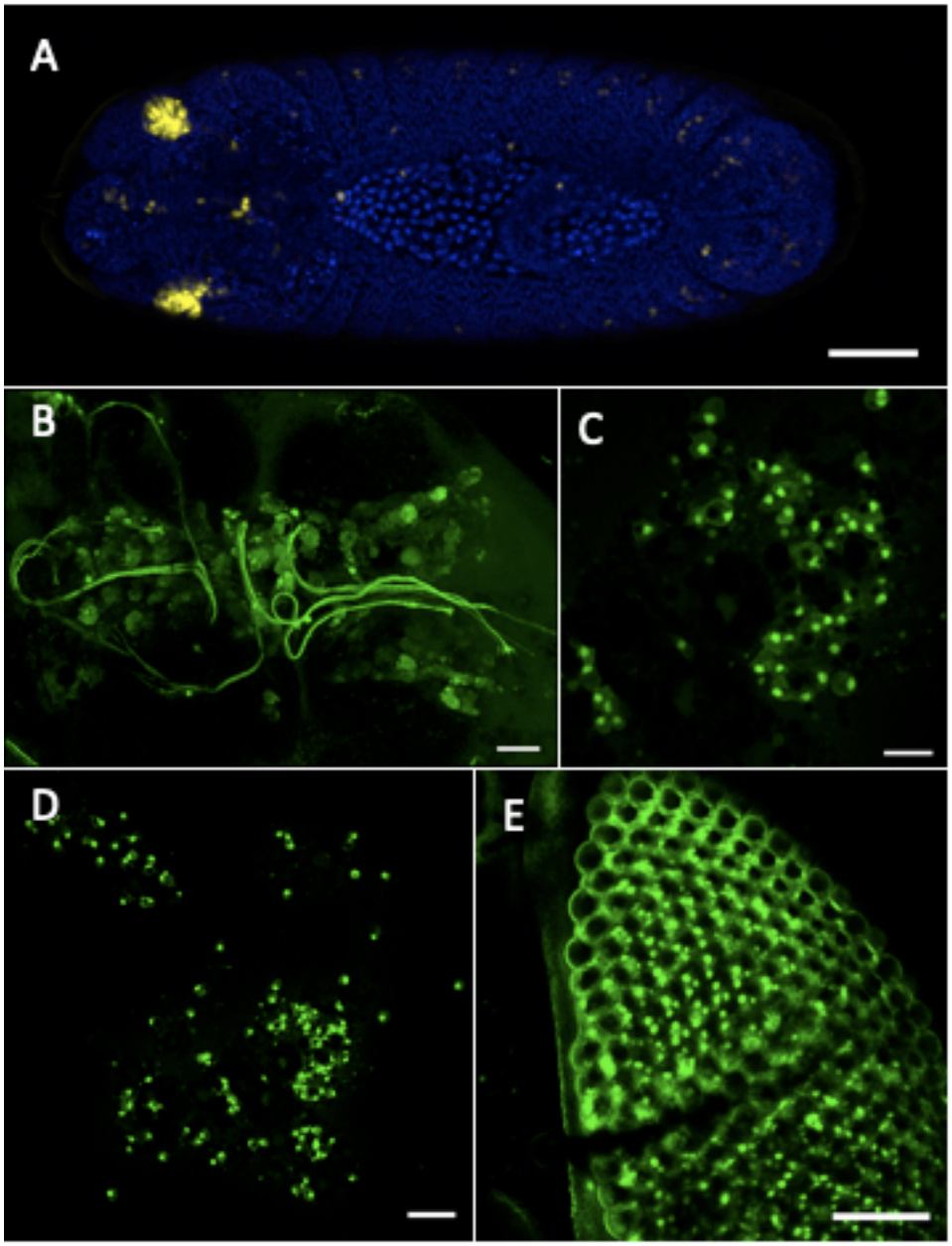


Figure 4. Expression of GFP within the VT056866Gal4-line. GFP expression due to of GAL4-enhancer trap within the 5' region of *hnt* (A-F). Strong expression was observed within the developing BO during embryonic stages (A). Axons of the ventral nerve cord expressed GFP within in the larval brain (B). Within adults, expression is observed within crystal cells (C), trachea cells (D) and compound eye (E). Scale bar represents 50 μm.

2.3.5 GFP expression in the *VT056867* Gal4-line

The *VT056867* Gal4-line drives GFP expression throughout the PNS during early embryonic stages. Unlike the PNS expression of the previous *VT056866* Gal4-line, where the reporter is expressed within clusters in the embryonic head, expression is specific to the neurons of the lateral epidermis. The location and organization of these neurons are consistent with the formation sites of oenocyes, indicating that expression is associated with chordotonal neurons (Figure 5, A). Expression of *hnt* within the PNS, as well as oenocytes, has been characterized both in embryos and larvae (Kim et al., 2020). Later in embryogenesis, expression increases in these neurons of the PNS. Expression also accumulates in two clusters within the embryo head, representing neurons of the PNS in the site of the developing BO (Figure 5, B). Embryos also display strong expression in the CNS, within the ventral cord (Figure 5, B). To date, there has been no clear characterization of *hnt* expression within the embryonic CNS. However, embryonic expression of *hnt* within glial cells of the CNS has been reported (Yip, Lamka, & Lipshitz, 1997). Like other lines, the regulatory elements drive expression in the larval brain (Figure 5, C). During larval stages, GFP expression is present within the multi-dendritic (md) neurons of the PNS (Figure 5, D). While PNS expression of *hnt* has been characterized in larvae, expression in these specific neurons has not previously been observed. Expression of the *hnt* continues into the dendritic arborisation (da) cells, which extend from md nerurons.

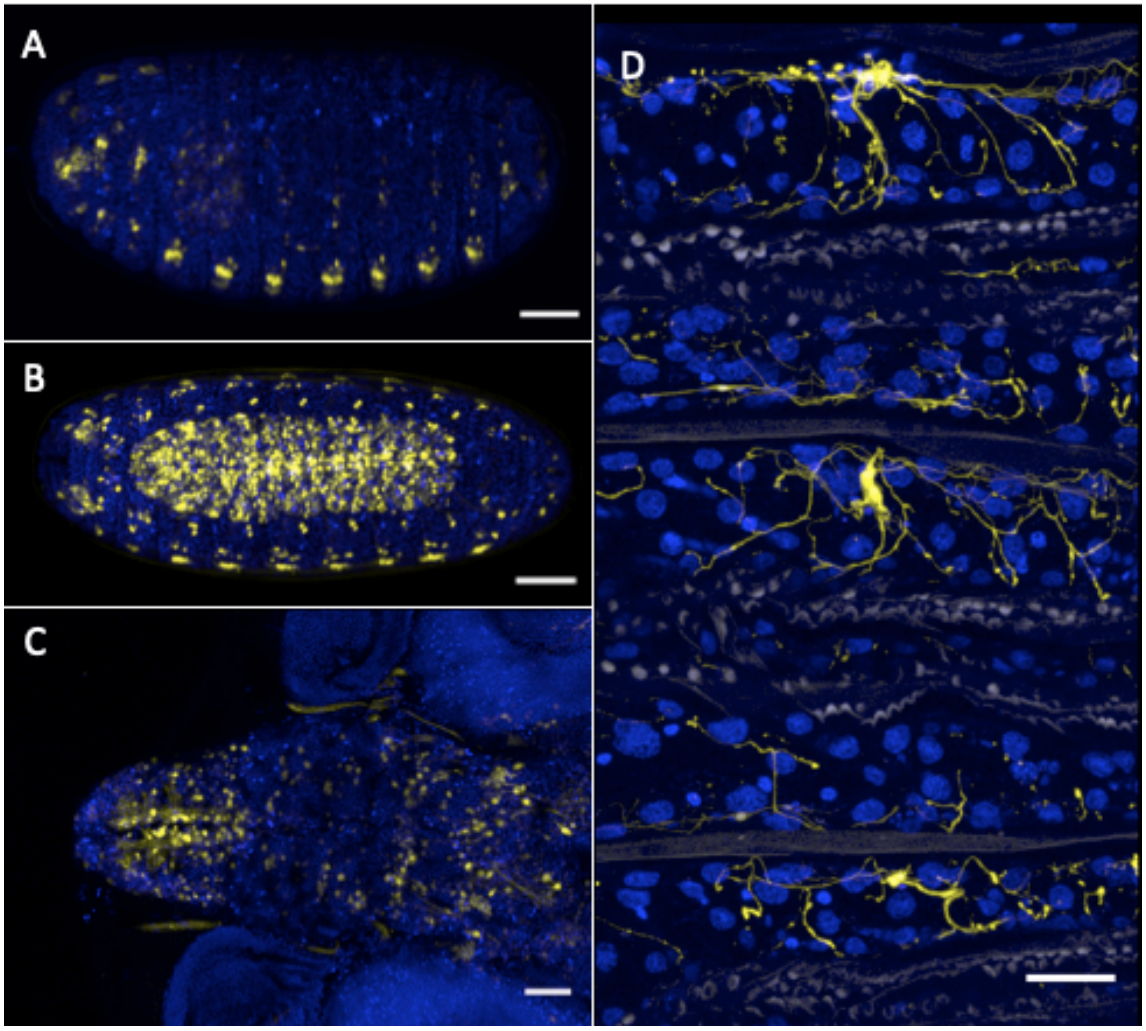


Figure 5. Expression of GFP within the VT056867 Gal4-line. GFP expression due to of GAL4-enhancer trap within the 5' region of *hnt* (A-E). Within embryos, expression in peripheral neurons begins in oenocytes-containing clusters (A). Later in development, expression is additionally in CNS, as well as PNS clusters in the site of the BO (B), and within the larval brain (C). During the larval stage, GFP expression is present in the md sensory neurons (D). Scale bar represents 50 μm.

2.3.3 GFP expression in the *VT056868* Gal4-line

Expression of GFP in the *VT056868* Gal4-line initiates within two locations on either side of the early embryo's developing head (Figure 6, A). These expressing cells disperse throughout the embryo as embryogenesis proceeds (Figure 6, B). The organization and migration of these cells strongly suggests these cells are macrophages, a blood cell lineage within *Drosophila.* While this expression has not been documented in the literature, *hnt* expression in these cells has been observed in the Reed lab during embryonic stages using immunolocalization (unpublished work). Expression is also evident in distinct clusters across the lateral epidermis. It is highly likely that this expression is a result of gene activation by the regulatory sequences of the *VT056867* Gal4-line, which overlaps with the sequences of this line. As discussed in the results of the *VT056867* enhancer sequence, the reporter gene expresses in the location of oenocytes (Figure 5, A-B). During larval stages, the regulatory sequences drive GFP expression within the axons of the VNC (Figure 6, C), like the expression of GFP within larva of the *VT056866* Gal4-line. As discussed earlier, this *cis* element may be involved within axon targeting that takes place within developing neurons. Additionally, there is expression within the nuclei of the salivary gland during the late stages of salivary gland development (Figure 6, D). Expression of *hnt* within the salivary glands at this stage has been previously noted (Kim et al., 2020; Ming et al., 2013). The *hnt*-encoded transcription factor targets *nevy* and *hnt* itself within the salivary glands of third instar larvae. Expression begins to the anterior segment of the gland and spreads to the posterior end. Expression is then discontinued in the anterior end in all parts except within the intersection between the gland itself and the imaginal discs (Ming et al., 2013).

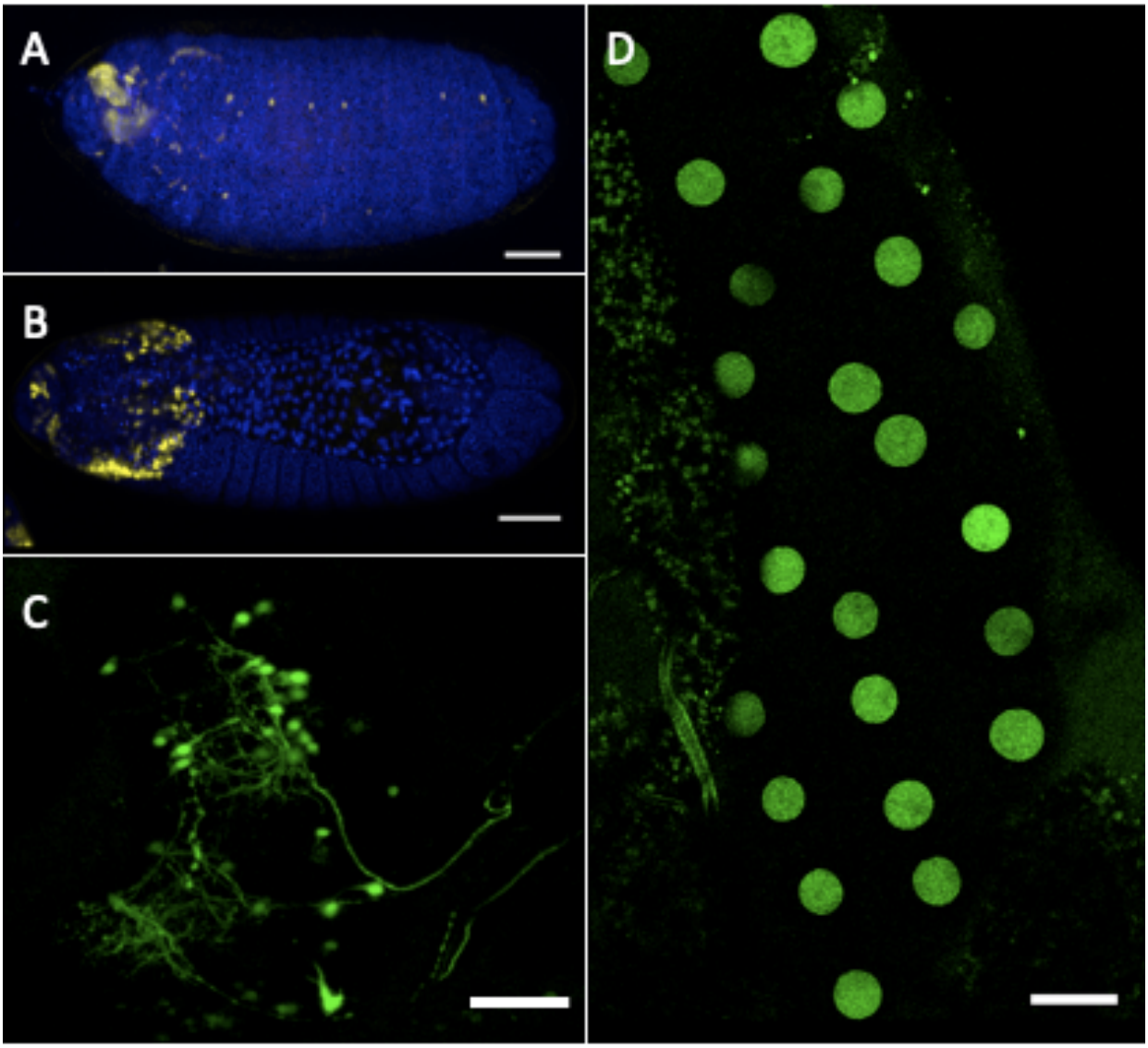


Figure 6. Expression of GFP within the VT056868 Gal4-line. GFP expression due to of GAL4-enhancer trap within the 5' region of *hnt* (A-D). Within embryos, macrophage precursors express GFP (A-B). Expression originates from the head (A), expressing cells spread into in developing embryo (B). During the larval stage, axons of the ventral nerve cord expressed GFP within in the larval brain (C). There is also expression within the larval salivary gland (D). Scale bar represents 50 μm.

2.3.4 GFP expression in the *VT056870* Gal4-line

During gastrulation of the *VT056870* Gal4-line, GFP is expressed in two clusters of cells on either side of the embryo (Figure 7, A). During progression to dorsal closure, these cells eventually disperse and migrate throughout the embryo. The cell organization and migration pattern strongly suggest expression in macrophage cells. The pattern displayed by this line is similar to the GFP-expressing cells observed in the *VT056868* Gal4-line embryos (Figure 7, B). The regulatory sequence of this line also influences GFP expression in the larval brain, within a junction that connects the two brain hemispheres (Figure 7, C). There is also GFP expression within the cyst cells of the larval testis, following the initiation of spermatogenesis of 3rd larval instar males (Figure 7, D). Interestingly, there was no GFP expression observed within the developing reproductive structures of female 3rd instar larvae. In female adults, expression from this line was observed in ovaries within the anterior stretch follicle cells (Figure 7, E). Additionally, expression was evident within border cells of the developing egg chamber (Figure 7, F). This expression of *hnt* within the ovaries have has been characterized (Sun & Deng, 2007). In males, expression is present within the cyst cells at the anterior of the testis (Figure 7, G). While expression of *hnt* within cyst cells is not documented, *hnt* plays a Notch-dependent role in the differentiation of follicle cells (Sun & Deng, 2007). Since follicle and cyst cells are both characterized as somatic cells within the male and female germline respectively, it is possible that cyst cell differentiation using this putative element is also involved with Notch signalling. Finally, expression is evident within some enterocytes of the adult midgut (Figure 7, H). Immunolocalization has shown *hnt* is present within all enterocytes of the midgut, indicating that this putative enhancer element may only be associated with expression in a subset of these cells.

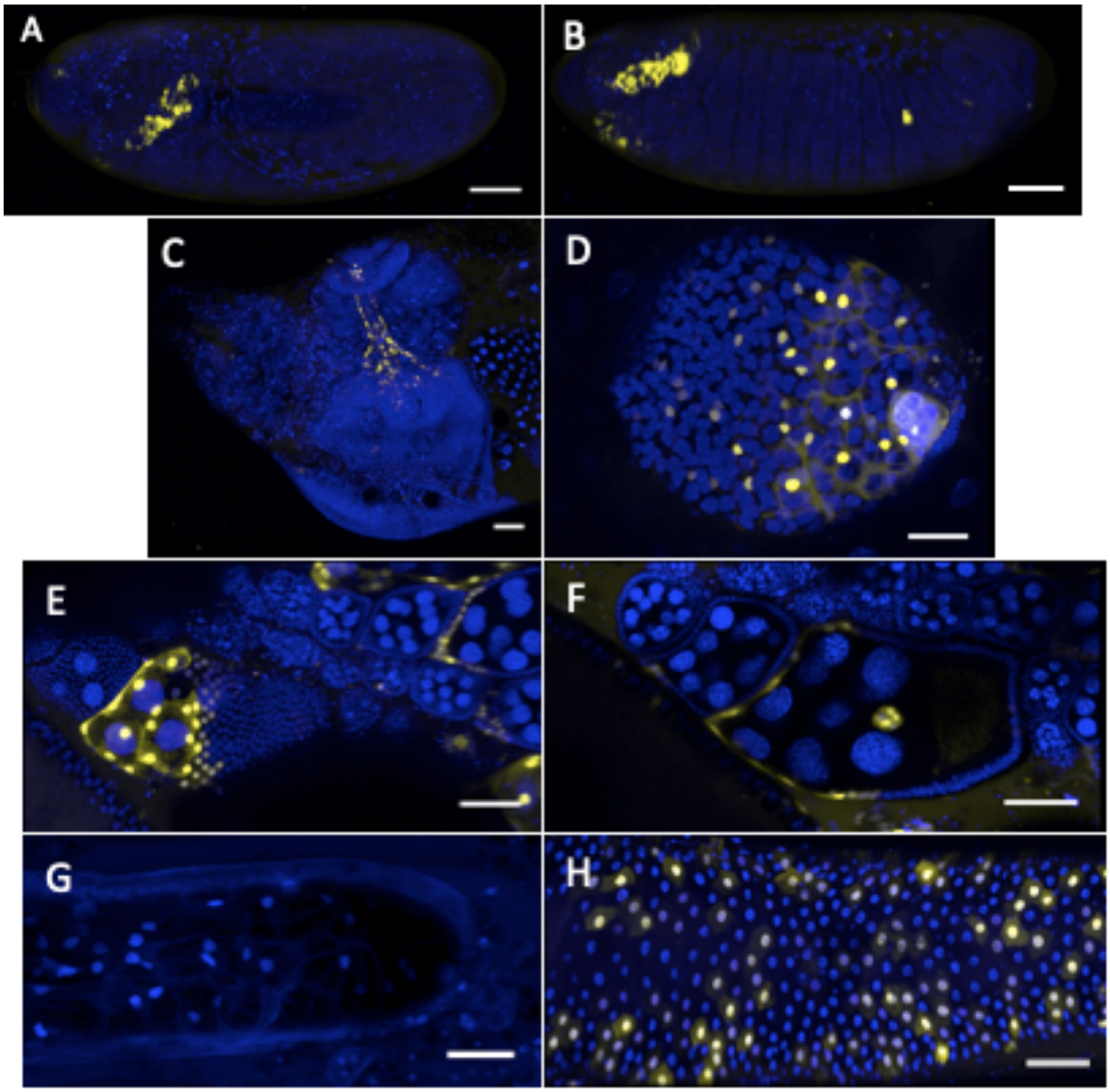


Figure 7. Expression of GFP within the VT056870 Gal4-line. GFP expression due to of GAL4-enhancer trap within the 5' region of *hnt* (A-D). Within embryos, phagocyte precursors express GFP (A, B). Expression originates from the head (A), expressing cells spread intro in developing embryo (B). During the larval stage, there is expression of GFP within in the brain (C), as well as in testis cyst cells (D). In the adult, expression is in present in anterior stretch follicle cells (E) and within some of the border cells of the egg chamber (F) There is also expression within cyst cells of the testis (G) and enterocytes of the midgut (H). Scale bar represents 50 μm.

2.3.6 GFP expression in the *VT056875* Gal4-line

The *VT056875* Gal4-line displays expression within the amnioserosa following GBR (Figure 8, A). Expression continues in these cells until dorsal closure (Figure 8, B). The expression and role of *hnt* in the amnioserosa is well characterized (Kim et al., 2020). It is required for GBR (Lamka & Lipshitz, 1999), and dorsal closure (Kim et al., 2020). There is also expression in epidermal cells where the amnioserosa and lateral epidermis meet, known as the leading edge (LE). Immunostaining of embryonic *hnt* expression does not include the LE in the lateral epidermis. The pattern of expressing cells along the LE is consistent in cells that display strong Jun kinase (JNK) signaling. The JNK pathway is down regulated in the amnioserosa and upregulated in cells of the epidermis along the LE boundary (Reed et al., 2001). This expression is also consistent with expression of the bone morphogenic protein (BMP) signalling pattern during dorsal closure, where the BMP ligand is high at the LE and forms a gradient along the epidermis dorsally (Hoppe et al., 2020). Within larvae, certain cells within various regions of the larval brain have *VT056875* expression (Figure 8, C). As mentioned previously, the larval brain does not typically express *hnt*. The putative regulatory sequences of this *VT056875* also influences GFP expression within the malpighian tubules of larvae (Figure 8, D). Expression in the malpighian tubes persists in adults (Figure 8, F). This expression of *hnt* has been characterized documented (Yip et al., 1997). Adults also display GFP expression within fat bodies (Figure 8, E), which is not reflective of true *hnt* expression.

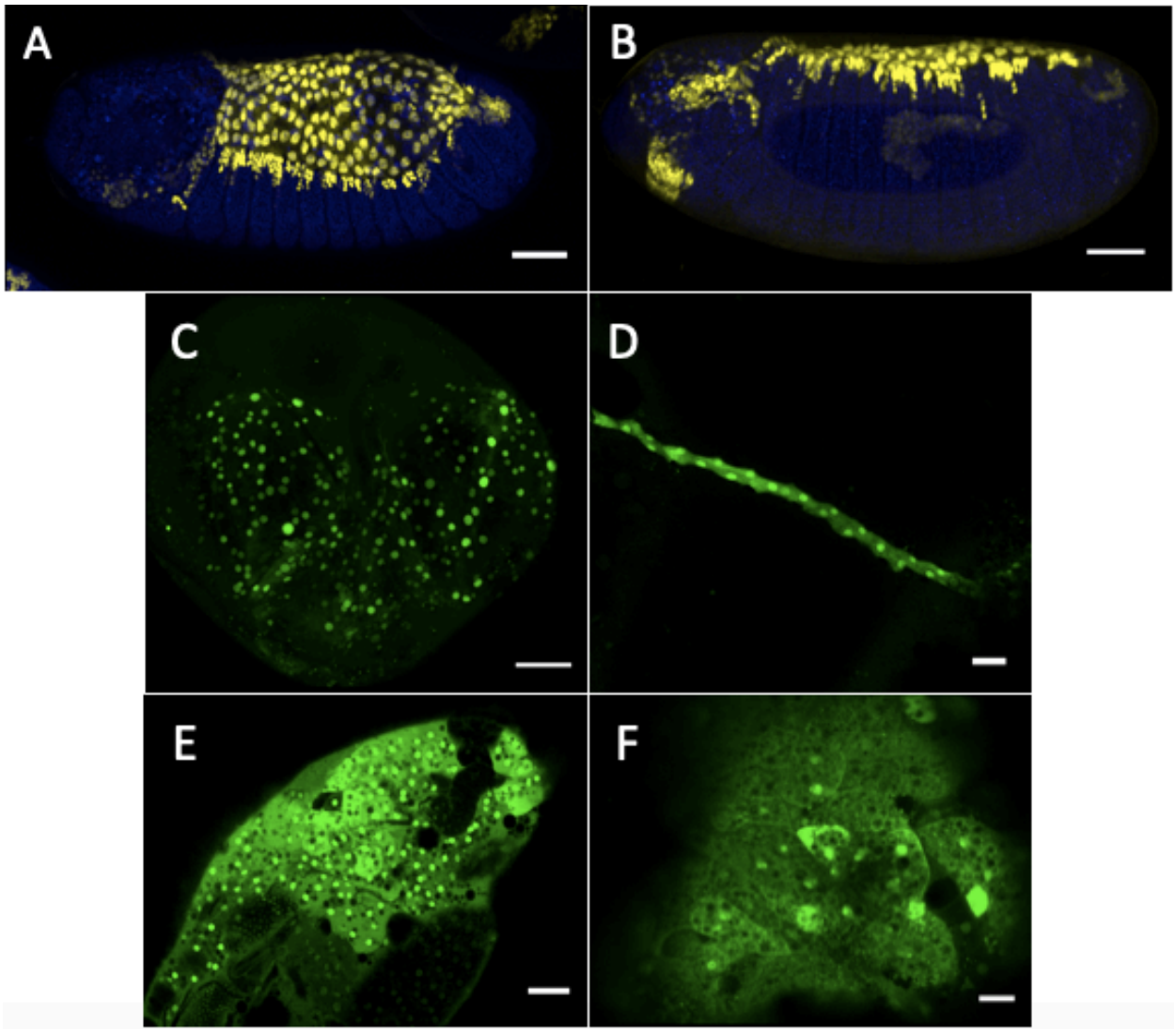


Figure 8. Expression of GFP within the VT056875 Gal4-line. GFP expression due to of GAL4-enhancer trap within the 5' region of *hnt* (A-F). Expression within the amnioserosa is present during dorsal closure, as well as in cells in groups along the border of the amnioserosa and the dorsal epidermis (A). This expression continues throughout dorsal closure (B). During larval stages, expression is seen in the brain (C), larval body fat (D) and malpighian tubules (E). In the adult, GFP is expressed within the malpighian tubules (F). Scale bar represents 50 μm.

2.3.7 GFP expression in the *VT056876* Gal4-line trap

The *VT056876* Gal4-line results in robust expression of GFP within the amnioserosa (Figure 9, A), a characterized site of *hnt* expression required for GBR and dorsal closure (Kim et al., 2020; Lamka & Lipshitz, 1999). This expression diminishes as dorsal closure reaches completion (Figure 9, B). This is the expected expression pattern of *hnt*; contrary to the expression pattern of the *VT056875* Gal4-line where there is reporter expression in the lateral epidermis of the LE in addition to the amnioserosa (Figure 8, A & B). The insert of this line overlaps the insert of the *VT056875*-Gal4 line, which is also associated with amnioserosa expression (Figure 9, A). Like the reporter expression of the *VT056868* Gal4-line, this line influences GFP expression in the 3rd instar larval salivary glands. This is a known region for *hnt* expression (Kim et al., 2020). Within the adult, expression of the reporter gene is evident within the follicle cells encapsulating the egg chambers in the ovariole (Figure 9, D). In late-stage egg chambers, GFP expression is present in follicle cells surrounding the entire chamber. This is a feature observed during transition from stage 13 to stage 14 (Deady et al., 2017). During the egg maturation, expression of GFP is restricted to the posterior follicle cells over the developing oocyte (Figure 9, E). Expression of *hnt* has been characterized in follicle cells as a result of the Notch-dependent mesenchyme to epithelial (M/E) transition of cells during stage 7-11 (Sun & Deng, 2007). The *VT056876* Gal4-line used also displayed membrane localized GFP expression. This GFP was due to the Ubi-De-Cadherin-GFP construct carried by this particular stock. It is not associated with *VT056876* expression (Figure 9, C-D).

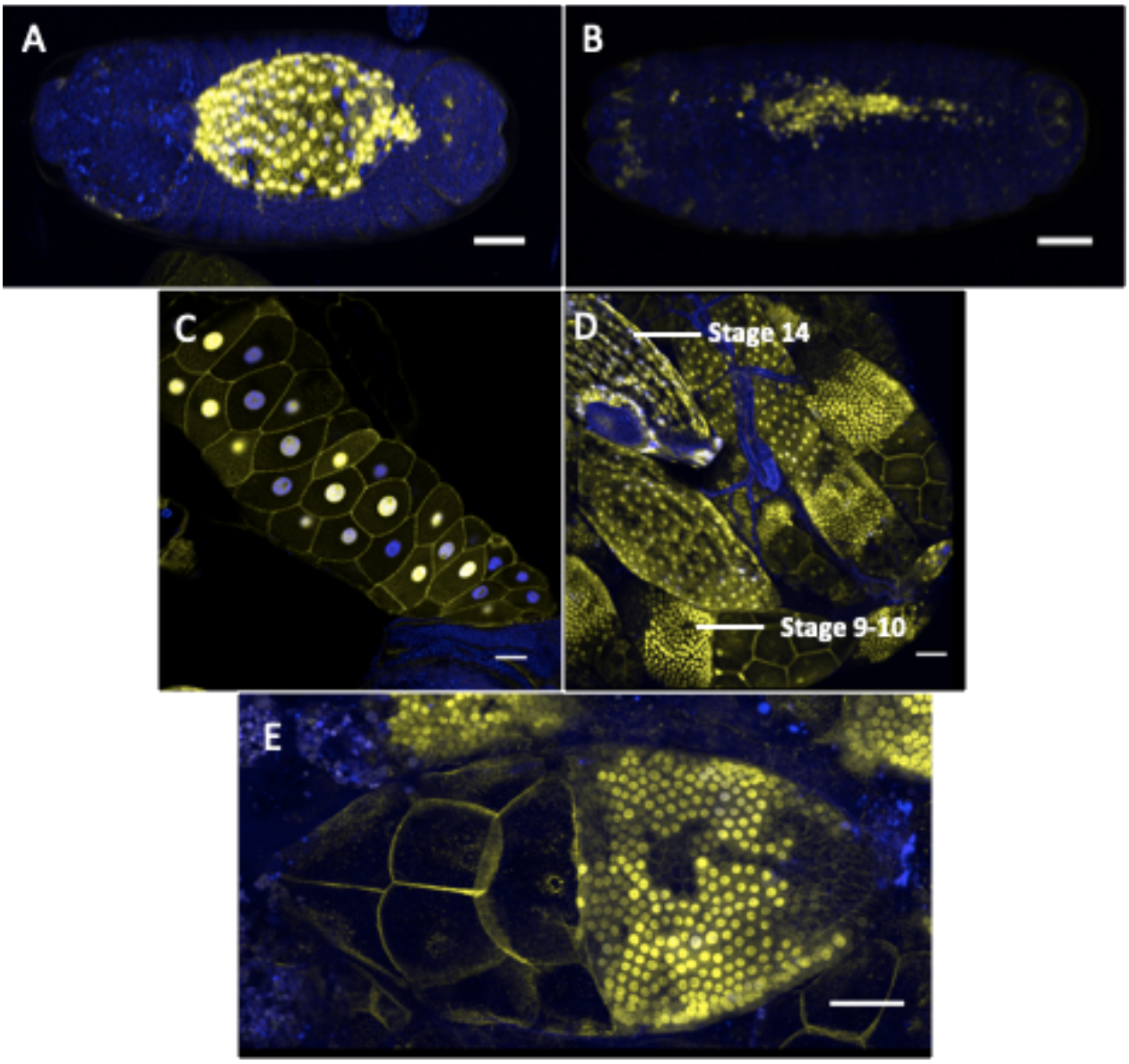


Figure 9. Expression of GAL4 within the VT056876-line. GFP expression due to of GAL4-enhancer trap within the 5' region of *hnt* (A-E). Expression within the amnioserosa is present during germ band extension to dorsal closure (A) and ceases after closure (B). During larval stages, expression is driven in the salivary glands (C) There is expression within the ovary in stage 14 egg chambers (D) The expression of cells in the posterior egg chamber (stage 9-10) indicates expression is due in follicle cells (E). Scale bar represents 50 μm.

2.3.8 GFP expression in the *VT056878* Gal4 line

Expression of the *VT056878* Gal4-line, which includes sequences from the 5' untranslated region (UTR) and first intron of *hnt*, results in strong expression of GFP in mid-late stages of embryogenesis in cells present in the embryo head (Figure 10, A). After dorsal closure, these cells become distinct clusters and migrate to the outer peripheral of the embryo (Figure 10, B). The pattern and migration of these clusters suggests this line could be associated with neuronal cells. The location of these cells is consistent with the developing BO, visualized as the two outer clusters on the embryos head, and antennal olfactory sense organ, present in clusters closer to the midline. Both of these components are "organs" of the PNS, a characterized site for *hnt* expression (Kim et al., 2020). During the larval stages, *VT0556878* is associated with expression within the developing brain (Figure 10, C). To date, there is no recognized expression of *hnt* within the central nervous system (CNS) during larval stages. The regulatory sequence of this line does not demonstrate activity during adult stages. It is likely that the activity of this element controls the expression of *hnt* during the earliest developmental stages of the PNS. Specifically, this regulatory sequence is likely reflecting a role *hnt* has in the differentiation cells in the PNS.

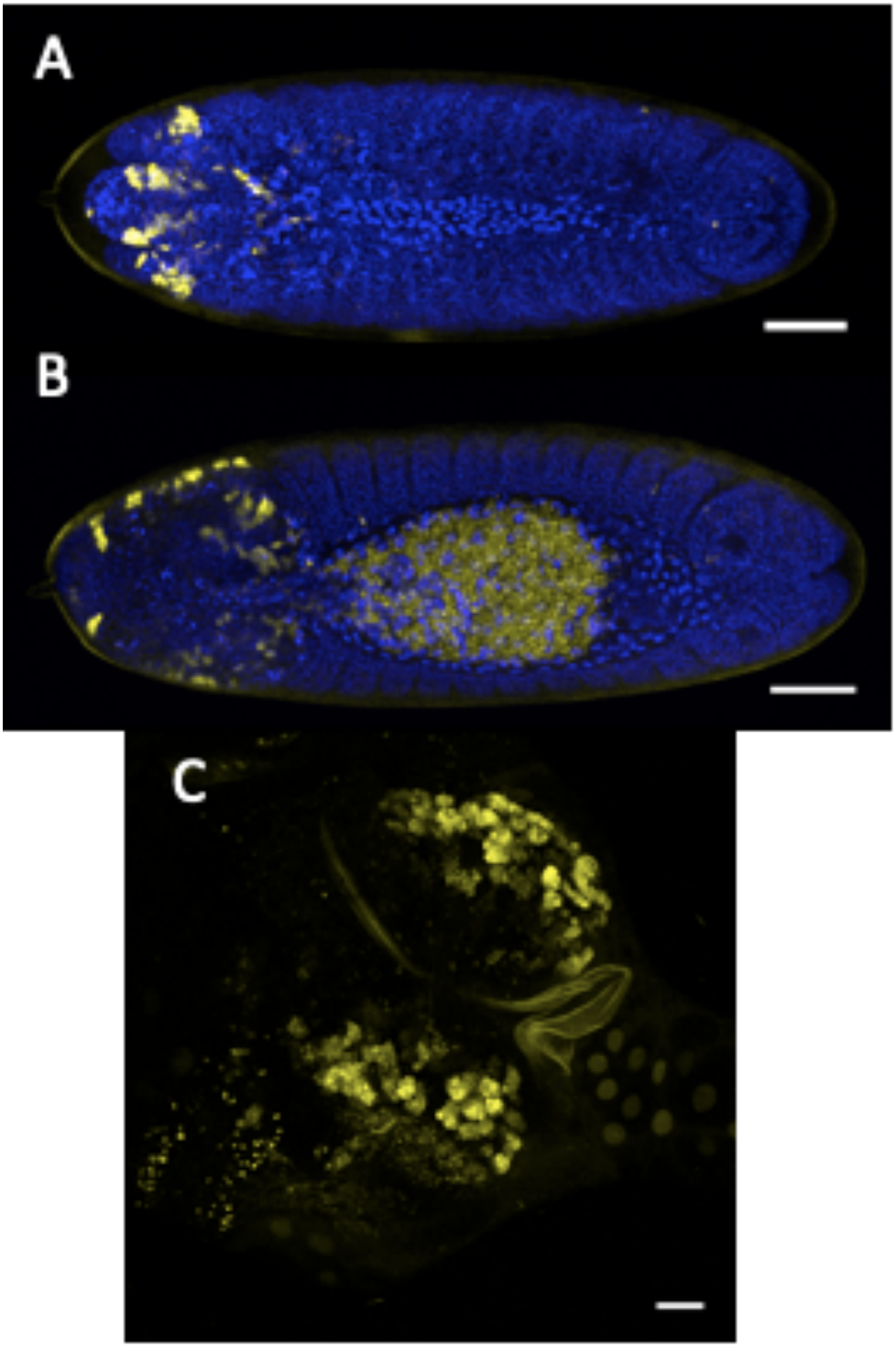


Figure 10. Expression of GFP within the VT056878 Gal4-line. GFP expression as a result of GAL4-enhancer trap within the 5' region of *hnt* (A-C). Expression is higher during early development (A), and decreases as it is localized to the outer regions of the embryo (B).GFP expression within distinct clusters hints t a potential role of the enhancer in neuronal fate (A, B). GFP expression within the larval brain indicates activation of *hnt* (C). Scale bar represents 50 μm.

2.4 Discussion and future directions

2.4.1 The role of *hnt* during PNS development

The *hnt* VT-Gal4 lines show a variety of expression patterns and many correspond to the known expression of *hnt*, characterized by immunostaining. For reference of the wild type embryo with immunostaining of Hnt, see figure 1. *VT056866, VT056867* and *VT056878* constructs were associated with expression within the embryonic neurons, or associated neurons, of the BO. Moreover, the *VT056878* constructs influences GFP expression in the antennal olfactory sense neurons (OSN). There is also expression in the *VT056867* construct within the chordotonal neurons of embryos, which are associated with the oenocytes. Collectively, these lines may be associated with the *cis*-regulatory elements involved in the expression of *hnt* within the PNS.

While expression of *hnt* in the BO is not characterized in published work, embryonic expression within this organ has been observed by immunostaining in the Reed lab (unpublished data). Cells that form the BO, as well as cells associated with the BO, differentiate during the embryonic stages. The BO bundles within larva are organized into PR cells, in groups ranging between 8-16 cells. These cells are precursors of the "larval eye", which is developed exclusively for larval stages (Singh & Kango-Singh, 2013). Given the known association of *hnt* in the axon targeting of PR cells in adults and 3rd instar larvae (Oliva et al., 2015), it is possible that the putative enhancers of these lines have a role in the specification of PR axons within the developing BO. In support of this, a downstream target of *hnt* during this process is *jbug*, where *hnt* likely plays a role in upregulating *jbug* during development of the adult visual system. This was identified by microarray analysis in the third instar larval eye imaginal disc. Overall, the

transcriptomes of wild type and HNT over-expressing larvae identified nine target genes of *hnt*, *jbug* being the only one to be downregulated in *hnt* LOF mutants (Oliva et al., 2015).

It would of interest to track how the expression of *jbug* activity is altered if *hnt* activity is not regulated by the putative elements identified by *VT056866, VT056867* and *VT056878*. This can confirm the involvement of these construct's putative enhancers in axon targeting through *hnt* expression. To further investigate these putative enhancers, mutagenesis of the element could give insight on the element's role in *hnt* expression. In addition, the activity of *jbug* transcription can be monitored to identify any changes in these *hnt* mutants. If these enhancers do have a significant role in axon targeting, this will also likely be observable as an irregular phenotype of stained axons in larvae.

Interestingly, whole embryo ChiP analysis has identified the *sens* TFBS within the region corresponding to the constructs of the *VT056866*, *VT056867*, and *VT056878* lines (Flybase, 2022). Given the role of the *sens* in simulating proneural gene expression and PNS differentiation, this may be the activating transcription factor causing reporter gene expression in the developing BO. This would explain expression associated with PR cells in the larval eye. Interestingly, the *VT056867* construct only showed faint expression in BO. It is likely that any expression in these neurons is the result of enhancers housed within the *VT056866* sequence. In these embryos, the intensity of reporter gene expression within the BO indicates significant activity of this enhancer for the organ's development. Interestingly, expression was also observed in the PR cells of adult flies.

It is important to note that the compound eye does not originate from the larval eye, as the BO is a separate structure that is considered a "primitive form" of the adult fly's eye (Singh & Kango-Singh, 2013). The compound eye of the adult fly is made up of 750+ ommatidia, each

group consisting of ~10 PR cells (Singh & Kango-Singh, 2013). Expression in the adult eye is likely linked to axon targeting of the PR cells. Overall, *Drosophila* is a useful model for studying the development of the eye. The BO offers a simplified model of the eye since it is considered the relict of the compound eyes. It contains PR cells; however, it lacks accessory cells that make up the specialized parts of the compound eye, such as the lens and pigments (Singh & Kango-Singh, 2013). Studying this "simpler" model of the eye offers a less complex model for studying the formation of PR cells. For instance, the role of *hnt* in axon targeting within the compound eye can be studied in the larval eye, as it offers a much simpler system.

Given the important role of *hnt* in cell differentiation into sensory organs, it is probable that the expression of the *VT056878* Gal4-line reflects *hnt* expression in specification of embryonic sensory neurons into the antennal olfactory sense neurons (OSN). *Hnt* expression has been used as a molecular marker for olfactory sensory neurons (OSN) (Arguello et al., 2021). *Hnt* is also expressed in the sensory organ precursors (SOPs) within pupae (Kim et al., 2020), which are the developmental origin sites of OSNs. Overall, little is known about the development OSNs or SOPs (Chai, Cruchet, Wigger, & Benton, 2019). Further characterization of the putative enhancer within the *VT056878* construct could lead to a better understanding of the mechanisms involved in embryonic specification of the antennal OSN, and the role of *hnt* in the differentiation of cells into sensory neurons.

The putative enhancer of the *VT056867* Gal4-line results in expression of GFP within clusters of chordotonal neurons associated with the embryonic oenocytes. The development of oenocytes begin in individual clusters of the first seven abdominal segments of the embryo. Delamination of these cells from the dorsal ectoderm is partly induced by the EGFR signalling pathway. The ligand for this pathway, Spitz, is secreted by the chordotonal organ precursors; this

ligand is necessary for the formation of oenocytes (Gould, Elstob, & Veronique, 2001). Null *hnt* embryos lack chordotonal organs, and fail to differentiate oenocytes. These are both phenotypes associated with the downregulation of the EGFR pathway. While *hnt* has not been recognized as part of the Spitz group, which includes genes that are components of the EGFR pathway, a myriad of evidence suggests that in certain contexts, *hnt* expression is required for maintenance or upregulation of the EGFR/Ras/MAPK signalling pathway (Kim et al., 2020). Given the suggestive role of this putative enhancer in the developing oenocytes, the analysis of this element could elucidate *hnt's* potential role in oenocyte differentiation through EGFR signalling.

Hnt expression has been characterized within type 1 neurons, both of which have been discussed above in the context of *hnt* expression. Type 2 neurons include md and da neurons (Crozatier & Vincent, 2008). Typically, there is no expression of *hnt* in these neurons. It is peculiar that reporter gene expression was visualized in md and da neurons in the *VT056865* Gal4-lines. It is possible this enhancer works with a repressor to differentially express *hnt* in presumptive type 1 and 2 neurons. Further investigating the role of a repressor in this expression may clarify the role *hnt* has in cell differentiation within the PNS.

2.4.2 The role of *hnt* in blood lineages

The putative enhancer identified by the *VT056866* Gal4-line, which influences expression in larval crystal cells, may contain an element that mediates *hnt* expression in a Notch-dependent manner. The implicated role of *hnt* during crystal cell differentiation is Notch-dependent (Terriente-Felix et al., 2013). Blood lineage differentiation occurs in two phases: a primitive phase originating within the embryonic head, and a definitive stage within larvae. Evidence suggests that there are clusters of progenitors that are the origin of *de novo* crystal cells in adult

flies. In addition to crystal cells, precursors are also capable of generating macrophages (Ghosh, Singh, Mandal, & Mandal, 2015; Meister & Lagueux, 2003).

To confirm if the expression of the *VT056866* enhancer is downstream of Notch signalling, expression of Gal4 could be monitored in the background of a null allele of Notch. Given Notch's reported role in the upregulation of *hnt* within these cells (Terriente-Felix et al., 2013), a decrease in *hnt*-expression in crystal cells would be an expected result in a Notch null genetic background. CHiP analysis can also identify transcription nfactors that bind to particular sequences. This may provide insight on how *hnt* responds to the Notch signalling pathway. The JASPAR core database was used to probe the *VT056866* sequence for elements that can associate with the central transcription factors in Notch signalling. One putative TFBS for Suppressor of Hairless (Su(H)), a regulatory element for Notch signalling, was identified within this construct. The details of this binding site is noted in appendix B, table B.2.

The expression of the reporter gene influenced by the *VT056868* and *VT056870* constructs in macrophages of the embryo head, where *hnt* is not expressed, suggests the putative enhancers of these lines may work in concert with a repressor in the context of their native genomic location. In the case of cells destined to be macrophages or crystal cells, activation or repression of *hnt* expression inferred by such elements may reflect the differential expression patterns that result in cell differentiation into these two cell types. Given the interest of "molecular switches" that lead to the development of different cell types, this pathway for blood cell lineage differentiation can be further investigated to better understand how *hnt* acts to specify the fate of crystal blood cells.

2.4.3 The involvement of *hnt* in reproductive organs and development

Within the adult ovary of *Drosophila*, *hnt* is expressed within follicle cells starting in stage 7-10 egg chambers, and this expression is known to be dependent on Notch signalling. The

upregulation of Delta, a ligand for Notch signalling, within germline cells actives Notch signalling within the overlying follicles cells. This leads to the upregulation of several downstream targets including *hnt*. This activity of Notch ultimately promotes the M/E transition of follicle cells (Sun & Deng, 2007).*Hnt* is also expressed within follicle cells of the anterior portion of egg chambers during stage 12 oogenesis, which may be due to Notch signalling. *Hnt* is finally upregulated within follicle cells of the mid chamber in stage 14 egg chambers, a required process in follicle rupture.

Since Notch is not upregulated in stage 14 egg chambers, it is unlikely that this expression is associated with the Notch pathway (Deady et al., 2017). *Hnt* is also required for the migration of border cells, a small group which derive from anterior follicle cells. These cells will undergo the epithelial to mesenchyme (E/M) transition and migrate through nurse cells. Border cells are eventually situated at the anterior pole of the developing oocyte, and form the structure known as the micropyle. Border cells mutant for *hnt* remain at the anterior tip of the egg chamber due to failed initiation of migration (Melani, Simpson, Brugge, & Montell, 2008). It is suspected that this role of *hnt* for border cell migration may be Notch-dependent (Sun & Deng, 2007). Interestingly, the V*T056870* Gal4-line influences GFP expression within the anterior stretch follicle cells during later stages of the egg chamber, as well as within boarder cells. This putative enhancer may therefore be one of the regulatory elements that mediate *hnt* expression during these events.

The genomic sequences of *VT056870* line were analyzed in JASPAR for complementation to central Notch transcription factors. No putative sites with a relative profile score above an 80% threshold was identified. This could be due to the restricted number of TFBS documented within the JASPAR database. Since the *VT056870* regulatory element influences expression in two

known sites of Notch-dependent *hnt* expression, the involvement of this enhancer in Notch signalling should be further investigated. The *VT056870* Gal4-line also shows expression of GFP within cyst cells of the larval and adult testes. Cyst cells within the male testis are analogous to the follicle cells in females. Like follicle cells, cyst cells enclose the developing germline cells. The expression of *Notch* has been shown to be activated within cyst cells in the apical region of the testis, and this is required for germ cell differentiation (C. Ng, Qian, & Schulz, 2019). Given the role *hnt* has in the Notch pathway for follicle cell differentiation, the role of *hnt* within cyst cells presents a topic of interest. Expression of *hnt* within cyst cells has, at this time, has yet to be confirmed by immunostaining.

The *VT056876* sequences were also analyzed for TFBS motifs involved in Notch signalling, as it influenced reporter expression within follicle cells in stage 7-10 chamber follicle. The Su(H) TFBS motif was identified in the sequence. Details of this putative binding site is available in appendix B, table B.3. During stage 14, when follicle cells completely encapsulate the oocyte, expression of the VT056876 reporter gene is present in all follicle cells. This enhancer could be involved in expression of *hnt* required follicle rupture during stage 14 egg chambers, however it is unlikely related to the Notch signalling pathway. The transcriptional activation of *hnt* for this stage is still unknown (Deady et al., 2017). If this enhancer is involved in *hnt* expression at this stage, it could provide insight on how *hnt* upregulation from stage 13 to stage 14 is mediated.

2.4.4 The role of *hnt* in the amnioserosa

The reporter gene expression observed in the amnioserosa of the *VT056875* Gal4-line is typical of *hnt* expression. Expression within the lateral epidermis along the LE of the *VT056875* Gal4-line embryos, as observed in Figure 7a, however, is distinctly atypical. Given the pattern of this ectopic expression, the reporter gene may be responding to BMP signalling. A gradient of

Decapentaplegic (Dpp) signalling follows the pattern of the BMP ligand gradient established during dorsal closure. This includes a high concentration of expression in the epidermal cells along the LE, and a decrease of expression in the epidermis as cells get further away from the LE.

Dpp is a secreted morphogen that is part of the T6F-B super family of signalling molecules and is similar to the vertebrate BMP2 and BMP4 morphogen. Upon activation, the BMP receptor will phosphorylate Mad (pMad). The protein Medea interacts with pMad to activate and repress downstream genes, including *hnt* (Hoppe et al., 2020). Since the gene *hnt* is a known downstream target for Dpp during the early cellular blastoderm stage, it is possible that the putative enhancer of the *VT056875* is responding to BMP signalling. The atypical expression of *hnt* within the lateral epidermis suggests that the activity observed by the *VT056875* putative enhancer is typically repressed in LE cells, given that this element drives expression of the reporter within the amnioserosa reflective of *hnt*. However, in the context of the reporter gene construct used in this experiment, expression of the reporter gene may be activated by BMP signalling. The expression of this reporter can be tested in an altered BMP background, which can confirm any genetic interaction with the putative enhancer and this pathway.

To determine if the sequence of this construct consists of putative sites responsive to the BMP pathway, it was searched for similarity to known TFBS in the JASPER database. There are 11 putative sites which in which binding of the Mad transcription factor is predicted. This information is available in appendix B, table 4. This suggests that the putative *VT056875* enhancer is responding to the BMP/pMad signalling pathway, but this element is repressed during normal development. The embryonic amnioserosa, a region that displays high *hnt* expression, had high expression of the reporter gene within the *VT056875*-line (discussed above)

and the VT056876-line. Notably, these two inserts overlap each other. This 500bp overlap is of interest for identifying enhancers involved with amnioserosa expression. Interestingly, this overlap contains more than half of the Mad TFBS found in the *VT056875* Gal4-line.

Amniserosa expression of *hnt* is required for proper GBR, as well as dorsal closure. It is suggested that HNT regulates these two morphogenic processes through the JNK/Dpp signalling pathways. It has been shown that *hnt* mutants have persistent expression of JNK signalling within the amnioserosa, which has been suggested to inhibit the formation of focal complexes between the amnioserosa and LE. These complexes are thought to form because of the boundary between high and low JNK signalling within the epidermis and amnioserosa, respectively. The boundary of JNK signalling fails to form within *hnt* mutants (Melani et al., 2008; Reed et al., 2001). Given that cells in the epidermis of the LE display high JNK signalling, a necessary step for the formation of focal complexes, *hnt* is normally expected to be downregulated within these cells. Interestingly, the *VT056875* Gal4-line does not display this downregulation, whereas the *VT056876* Gal4-line does. This suggests the possibility of defining a silencer element within the sequences of the latter. Introducing small deletions within the regions of both putative enhancer sequences may confirm and identify elements involved in GBR and dorsal closure. With the advert of CRISPR/Cas9 genome editing, it is now possible to target and modify the sequences of the VT-Gal4 lines directly within transgenic lines.

2.4.5 Experimental considerations and conclusion

The expression patterns of the VT-Gal4 lines do not necessarily define the genomic sequences that act as *cis*-regulatory elements, as sequences influencing GFP expression are not in their native genomic locations. This is the largest disadvantage of using enhancer trap-based techniques. However, this method does provide possible insights into the general nature of the

regulatory sequences, in this case sequences corresponding to ~2Kb fragments that are in the region ~20Kb upstream to the *hnt* TSS. Based on immunostaining results (published and unpublished work by the Reed Lab), many of the expression patterns displayed by the VT-Gal4 lines are consistent with expected expression of *hnt*. Furthermore, these observations also support preliminary reports indicating *hnt* expression is influenced by major signalling pathways mediated by Notch and BMP.

As mentioned, the changed genomic location of Gal4-driver inserts of the VT lines may result in unexpected GFP expression, which is not reflective of *hnt* expression observed in immunostaining experiments. In the context of the endogenous positions of these candidate enhancers, another element within proximity may interact with this sequence to repress *hnt* expression. Another outcome of changing the genomic position of enhancers includes the inability for enhancers to interact with proximal elements due to distance constraints. In this case, expression may appear incomplete. An observed example of this includes the putative enhancer of the *VT056870* Gal4 line, where Gal4 expression was only observed in a subset of enterocytes within the midgut. While *hnt* is expressed in all enterocytes, the putative enhancer may only be involved in a subset of these cells. Alternatively, GFP expression influenced by this candidate enhancer may also be dependent on external factors outside of genomic locations, such as cell stage. For example, *hnt* expression may be present in newly differentiated cells – but decline as these cells age. These factors should be considered when interpreting unexpected results in the VT-Gal4 lines.

Chapter 3: Using Cas9 multiplexing to induce targeted rearrangements in the 5' region of *hnt*

3.1 Investigating the use of Cas9 to induce chromosomal rearrangements

3.1.1 Introduction

3.1.1.1 Inducing chromosomal rearrangements

When a double strand break (DSB) is induced in DNA, the cell's endogenous repair mechanisms can repair the broken DNA strands using one of two pathways: homology directed repair (HDR), or non-homologous end joining (NHEJ). The HDR pathway repairs DNA with high fidelity, as it uses a template with an identical sequence for the repair process. This type of repair is usually accomplished using homologous recombination (HR), where the sister chromatid or a homologous chromosome is used as a template. Alternatively, the NHEJ pathway ligates broken ends of DNA non-specifically. This repair involves first making blunt ends of DNA strands, such that they can be ligated without depending on homologous sequences.

When mutagenesis uses a multiplexing approach, where two or more DSBs are induced simultaneously, the sequence between the induced DSBs may be lost from the genome. This creates a deletion (also known as a deficiency). At a certain frequency, if the NHEJ pathway is adopted for repair, DSBs are generally expected to be rejoined in a different position or orientation in the genome - or excluded from the genome entirely. These events may yield chromosomal rearrangements (CRs), which include duplications, transpositions, deletions, inversions, and translocations.

While it is generally true that the mutations caused by CRs will have negative implications on an organism's health, there are contexts where they are favourable. For instance, CRs can form the basis of genome reorganization that occurs between species. These CRs, which

are the outcome of spontaneous mutations, can drive evolution (Sankoff & Nadeau, 2003). Moreover, CRs can serve as valuable genetic tools for several applications. A prime example is the use of inversions and translocations to make balancer chromosomes; these are routinely used by researchers using model organisms such as *Drosophila* and *Caenorhabditis*. CRs in balancer chromosomes suppress meiotic recombination between homologous chromosomes, which permits a mutation of interest to be stable and maintained as a heterozygote in diploid organisms. These balancer chromosomes are essential for maintaining and studying these mutations over generations (Robert, 1986).

This project focuses on the feasibility of designing and inducing inversions, deletions, and translocations in *Drosophila* using Cas9 multiplexing to make DSBs. The creation of these CRs using the described method is elaborated in appendix C, Figure C.1. The proposed CRs in this project are generated in the upstream regulatory region of *hnt* as a means to investigate the regulatory properties encoded in this region. CRs will be the basis by which putative regulatory elements are dissected from the protein coding region of *hnt* and identified as being necessary for wild type *hnt* expression. As discussed below, this could provide a rapid and generally applicable method for investigating large sequences of regulatory DNA. The generation of CRs following two or more DSBs have been mostly attributed to the NHEJ pathway, as the random ligation of DNA is more likely to create CRs in the genome. However, CRs have also been created using the HR pathway. The various methods to generate rearrangements using both pathways will be discussed below.

The first instance of lab generated CRs used x-ray mutagenesis (Kaufmann, 1946). While this can recover CRs with a reliable frequency, the targeting of DSBs is entirely random (Kaufmann, 1946). More recently, defined DSBs using the 18 bp yeast mitochondrial

endonuclease I-SceI, site housed in P elements has been used. Upon I-SceI expression, inserts of the I-SceI recognition sites are cleaved. Following cleavage of multiple genomic sites, translocations were recovered with an efficiency of 1-4% while inversions varied from <1%-4%. These were a by-product of HR between the identical sequences of the *P*-elements vectors in which the I-SceI sites are housed. Deletions were also recovered as a product of NHEJ (Egli, Hafen, & Schaffner, 2004). This method is limited in that *P*-element integration is itself random, but does show preferential transposition within "hotspots". Systems for site-specific recombination (SSR), such as the FLP/FRT and Cre/loxP, have also been used to induce CRs in *Drosophila* using HR. These can be used to recover inversions, translocations and deletions between two recognition sites (Branda & Dymecki, 2004). CRs recovered using this method occur at very low efficiency, ranging from 0.04%-0.08% (Brunet et al., 2009; Egli et al., 2004).

Engineered systems that have also been used to induce CRs use site-specific nucleases, such as zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs). These use DNA/protein recognition to target endonucleases to a particular DNA site. Simultaneous DSBs using this system followed by the NHEJ pathway yields CRs with an efficiency ranging from >0.1%-5% (Brunet et al., 2009). In addition to having an undependable efficiency, this method for inducing CRs can be labour intensive and costly. For example, one DSB requires the creation and expression of two DNA-specific proteins.

These techniques have been superseded by the advent of CRISPR/Cas9 genome editing methods, which offer a more effective approach for inducing targeted CRs. CRISPR/Cas9 is also broadly applicable to any model organism, provided that transformation methods are available. This system is discussed below. In the context of this study, the design and recovery of potential CRs was targeted to the upstream region of *hnt*. The dissection of this region using CRs can give

insight on the gene's regulation beyond the direct disruption caused by individual mutagenesis events, or enhancer trap-based methods. Rearranging regulatory elements commonly causes disruption in regulation due to position effects. Inversions and translocations have shown there is an inverse correlation between the interaction frequency and distance of a cis-regulatory element and promoter. Position effects from these CRs can also be caused by the displacement of an element into a different chromatin environment. Finally, *trans*-effects in expression can be disrupted using these CRs. The activities of these *trans*-elements, referred to as the transvection effect, often induce the cis-regulatory elements of homologous genes. Consequently, this regulation is pairing dependent; hence genomic position of an element is of high importance for regulation. Cytogenetically visible deletions that do not include protein-coding sequences can affect gene regulation, due to loss of one or more regulatory elements (Harewood & Fraser, 2014). Translocations, inversions, and deletions generally offer a diverse method for studying gene regulation. Since large portions of the genome can be rearranged, CRs can provide an efficient approach for studying large-scale stretches of sequences at once, which can be useful for genes with large upstream regulatory regions.

3.1.1.2 The CRISPR/Cas9 system

Targeted DSBs can be conveniently induced using the CRISPR/Cas9 system which was adapted from a bacterial immune system. This is the basis by which designed CRs will be targeted within the 5' region of *hnt*. The system relies on the Cas endonuclease, which can be directed to a particular DNA sequence to induce a targeted DSBs. The Cas protein, conserved amongst bacteria, is located within proximity of the repetitive sequences known as CRISPR elements, which are separated by spacers. Spacers are transcribed into RNA and interact with the Cas endonuclease to guide it to a complementary sequence, where it will induce a DSB. If

phages infect a bacteria cell that contains a spacer with its genetic information, the spacer is transcribed and incorporated into the Cas9/RNA complex. This results in the specific cleavage of the attacking phage DNA. In this system, it is necessary to distinguish between bacterial and phage DNA and Cas proteins require a recognition site known as the protospacer adjacent motif (PAM) within proximity of the targeted sequence to properly bind and cleave the DNA. This motif is not included within the spacers sequence, so that bacteria do not target their own spacer sequences for a DSB.

The gene encoding *Cas9* was isolated from *Streptococcus pyogenes*. It is now routinely used in many genetic systems, including *Drosophila*, to induce targeted DSBs. The system guides Cas9 to a sequence determined by gRNAs. This is a chimeric transcript comprising the two RNA molecules that normally make up the guiding molecule for Cas9 within the endogenous bacterial system. The complementary sequence of the Cas-associated gRNA molecule must contain the PAM sequence "NGG". The PAM sequence is required 3bp upstream of the target site for successful Cas9 mutagenesis (Adli, 2018; Ran et al., 2013). A customizable "target sequence" can be encoded within a synthetic gRNA to guide Cas9 to any desired sequence within the genome, provided it includes the PAM motif. Following a Cas9-induced DSB, the cell can repair breaks using HDR or NHEJ. The active Cas9 will continuously bind and cleave target DNA until the recognition sequence has been altered. This can happen following NHEJ, which will result in deletions or insertions that will alter the original sequence. The HDR pathway can also be used, which is advantageous for genetic knock-ins. If a desired sequence with flanking regions of homology is introduced with the Cas9 + gRNA to induce a site specific DSB, the HDR pathway can be used to replace this fragment of DNA into a desired genomic position complementary to the gRNA target sequence (Ran et al., 2013).

3.1.2 Methods

3.1.2.1 Cas9-induced mutations within the germline using the Nos promoter

The Cas9 system can be used to induce inheritable mutations by expressing the endonuclease within the developing germline cells of an organism. As a general resource to the *Drosophila* research community, the *Cas9* gene has been placed under the control of the *Nanos* promoter *(Nos-Cas9)*. Nanos is a protein required in the germline of both males and females. *Nos* is also a maternal effect gene and is required in establishing polarity in the early oocyte. As a result, females of the *Nos-Cas9* line express Cas9 in their eggs during oogenesis. When flies with a transgene of *Nos-Cas9* are crossed to flies with a ubiquitously expressed sgRNA transgene, offspring that inherit both transgenes will have the active sgRNA-Cas9 complex formed within their primordial germ cells. The mutated chromosomes of the germ cells are subsequently inherited by the next generation, and progeny heterozygous for the Cas9-induced mutation can be recovered (Kondo & Ueda, 2013). This process is illustrated in figure 10 below.

The *nos-Cas9* construct provides an interesting method for inducing CRs due to nuclear organization during Cas9 expression. While pairing of homologous chromosomes is exceptionally prevalent within both germ and somatic cells of *Drosophila*, unpairing events of chromosomes during embryogenesis occurs almost immediately after germline differentiation. Chromosomes remain unpaired until the formation of the adult gonad (Joyce, Apostolopoulos, Beliveau, & Wu, 2013) This lack of pairing may enhance the NHEJ pathway of repair following Cas9 mutagenesis, as HDR often uses the homologues chromosome for repair via HR.

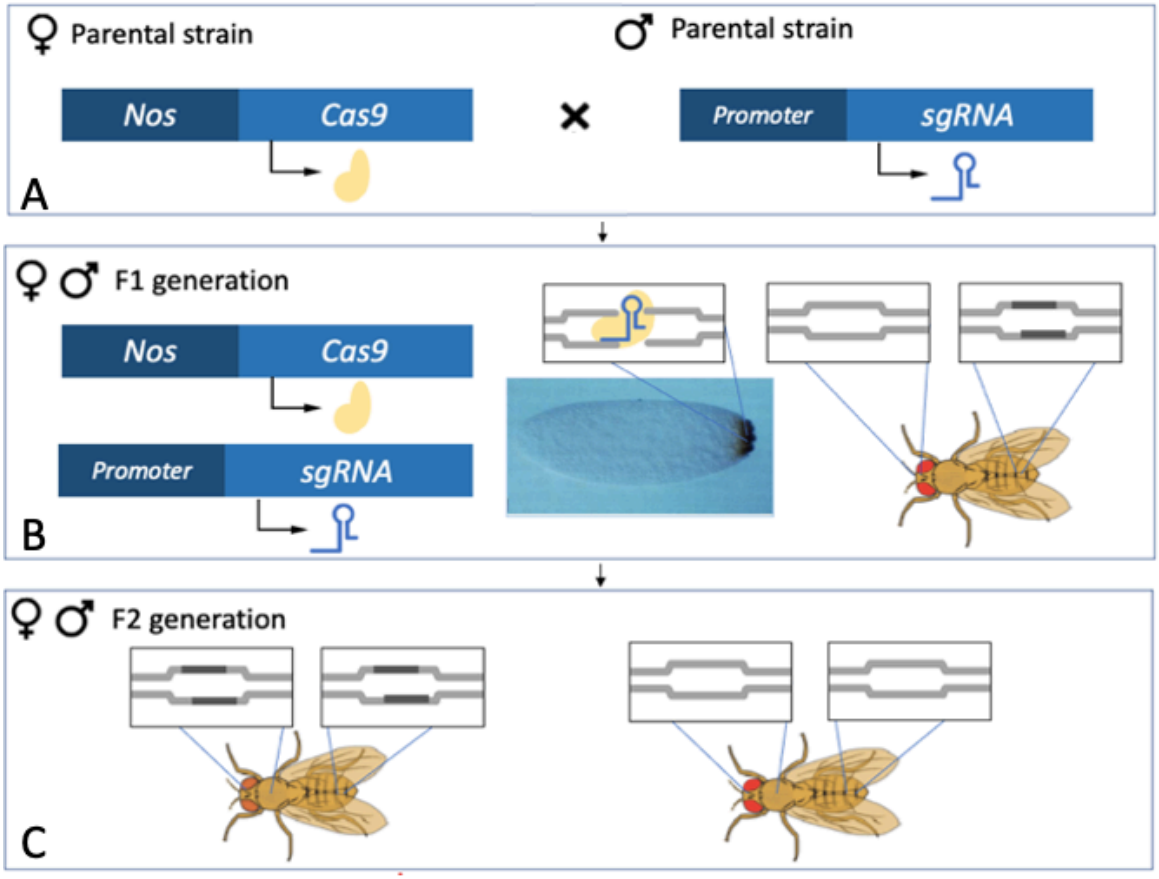


Figure 11. Inducing germline mutations with Cas9 using the *Nos* promoter. One parental strain expresses the Cas9 endonuclease under *Nanos*, which will only result in Cas9 expression in the developing germline. The other parent ubiquitously expresses a sgRNA that will target Cas9 to a desired genomic location (A). In the F1 generation, organisms have expression Cas9 within developing germ cells (B). The fly itself is still wild type, but its germ cells that will be used for embryogenesis of F2 offspring will contain the desired mutation is Cas9 is successful. Flies in the F2 generation that inherit a mutated germ cell will be heterozygous for the mutation within somatic cells. 50% of germ cells will also contain a mutated chromosome (C, left). Alternatively, if flies in the F2 generation inherit a wild type germ cell from the F1 generation, it will have a wild type genotype (C, right)

3.1.2.2 Screening the efficiency of the nos-cas9 system

As a pilot project to determine the efficiency of the *Nos-Cas9* system, the gene *scalloped (sd)* was selected as a target. This was in part due to its well characterized LOF recessive visible phenotype, but also because the gRNA expressing stock were readily available from the Reed lab stock collection. The F1 offspring of the designed cross will have transgenes for both *nos-cas9* and gRNA-sd transgenes. Due to the recessive nature of *sd,* putative mutants were crossed to flies homozygous for *sd*. Using this method, both female and male offspring can be used to determine the efficiency of *nos-Cas9* mutagenesis by measuring the frequency of homozygous *sd* progeny.

3.1.3 Results

3.1.3.1 Testing and validating the use of nos-Cas9 to recover mutations in the germline

F1 offspring from the Cas9 mutagenesis of *sd* within the female and male germline was successfully induced with 62.6% and 58.2% efficiency respectively. Offspring counts is shown in table 2. Interestingly, the *sd* mutant offspring from both F2 sexes had varying degrees of the *sd* phenotype, ranging from thickening of the most distal wing margin to the complete absence of wing development. This is illustrated in figure 12.

Phenotype	C1	C2
sd	374	238
***sd*$^{+}$**	580	332
Total	971	570
Estimate of Cas9 mutagenesis (%)	**62.6**	**58.2**

Table 2. F2 generation females with the *sd* phenotype. F2 flies of the *sd* mutagenesis cross were used to determine the frequency of which a cas9-induced mutation was made in the germline of F1 flies. C1 refers to offspring of Cas9 mutagenesis within the female germline. C2 refers to offspring of Cas9 mutagenesis within the male germline.

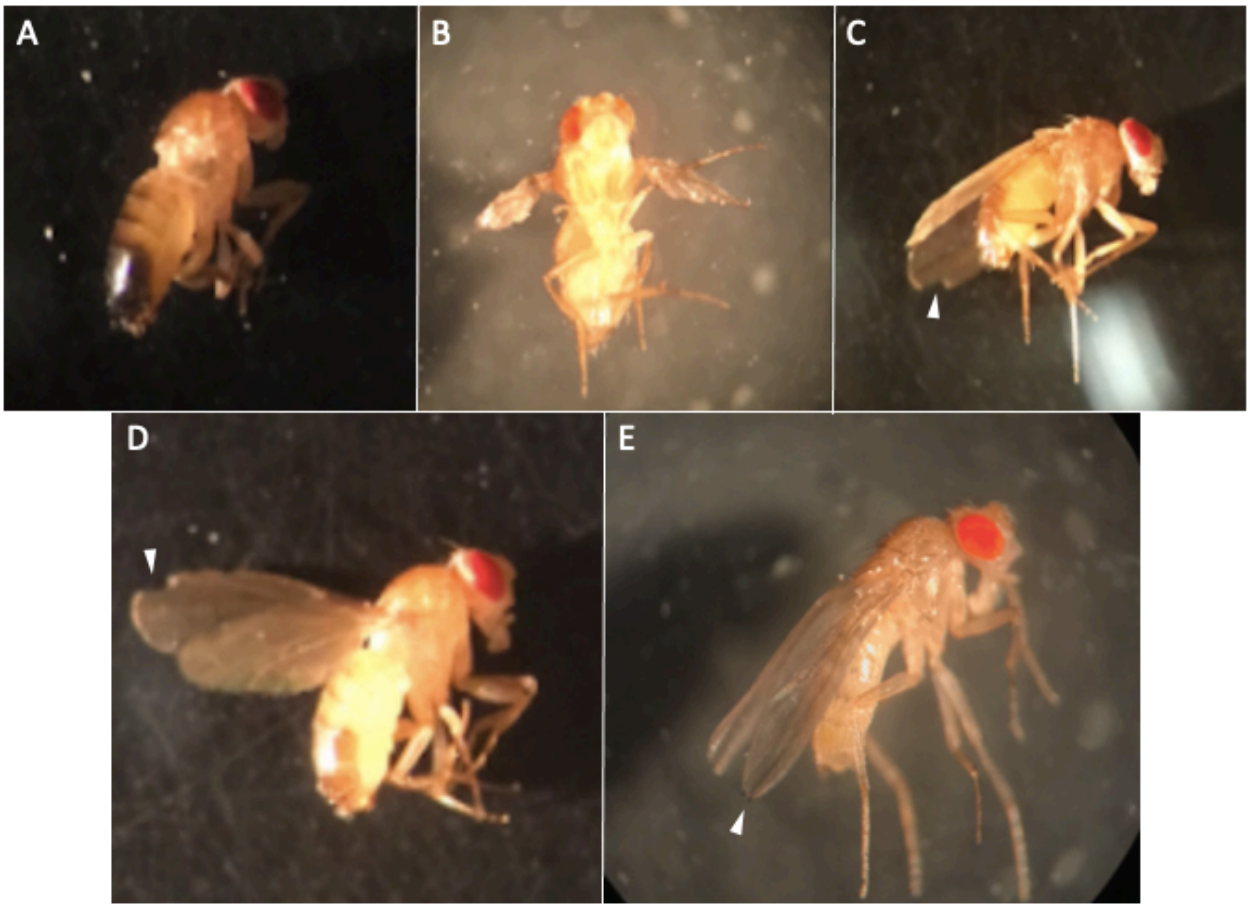


Figure 12. F2 mutants of the *sd* mutagenesis screen. The degree of the *sd* phenotype varied across both sexes (A-E). White arrows highlight the wing phenotype that arises from *sd* germline mutagenesis. Given the recessive nature of the *sd* allele, flies with the mutation should be homozygous for the mutation. One mutant allele would have come from the germline mutagenesis of *sd* in one parent, while the other comes from homozygous *sd* mutant that flies with Cas9 mutagenesis were crossed with.

3.1.4 Discussion

The efficiency of *nos-Cas9* targeted to the *sd* locus was similar in the female and male germlines. The scoring of offspring from the screen, shown in Table 2, shows that transgenic females demonstrated a mutation efficiency of ~63%, while transgenic males induced mutations with ~58% efficiency. Kondo and Ueda's study (2013) with the *nos-Cas9* system show the male germline had a consistently higher efficiency of Cas9 activity. The system in the Kondo and Ueda study (2013) used six gRNAs targeted to different regions within *white.* In most cases, males had a higher rate of mutagenesis by than more than a 2-fold, with a difference as large as 98.7% and 37.2%. In some cases, the efficiency was about equal. However, mutagenesis was still favoured towards the male germline. (Kondo & Ueda, 2013).

Given some targets of *white* were favoured for mutagenesis in the male germline, while other targets had roughly equal mutation efficiencies in both germlines, it is possible that *nos-Cas9* mutagenesis is favoured by the male germline. However, it is unclear if this is dependent on the genomic coordinates of the target. The selected *white* targets were overlapping throughout a ~2Kb span, however displayed favoured germline mutagenesis in only some of these targets.

The *sd* locus is ~13 Kb away from *white* and displayed a different pattern to that observed when *white* was used as a target. The mutation efficiencies within the two germlines were roughly equal but was favoured in females. Given the powerful method the *nos-Cas9* system offers to the *Drosophila* research community, more tests should be carried out to confirm the deviation of efficiency within female and male germ cells, as well as genomic locations.

The frequency of *sd* mutants within the female germline screen, described here, could have been inflated due to occurrences of somatic mutations. Due to the role of *Nos* as a maternal effect gene, Cas9 that is maternally deposited in the oocyte can induce mutations within early somatic

cells (Kondo & Ueda, 2013). This would most likely take the form of a weak *sd* phenotype. Mutations that result from maternal effect false positives for screening mutant offspring within the female germline. This was not considered at the time of the screen. For the majority of flies, the weak *sd* phenotype was observed when female transgenic flies were used to induce germline mutagenesis, as shown in figure 12, D and E.

While somatic mutations may be the reason *sd* mutagenesis appeared to favour the female germline, it isn't expected that this inflated results significantly. Somatic expression was only evident in ~5% of offspring in the *white* mutagenesis screen (Kondo & Ueda, 2013). However, it is possible that the frequency of somatic mutations is dependent on the target site of mutations. Given the requirement of a multiple DSBs to make such large-scale mutations, a method with high mutagenesis efficiency should be used to ensure a detectable frequency of CRs is achieved. The high efficiency of the *nos-Cas9* construct provides a reasonable method to investigate the potential of Cas9 to induce chromosomal rearrangements.

3.2 Screening for inversions using Cas9/gRNA and double phenotypes

3.2.1 Introduction

Having determined that CRISPR/Cas9 was highly efficient in inducing germline mutations for the gene *sd*, the next step was to use the Cas9 system to induce CRs targeted to the 5' regulatory region of *hnt.* The first strategy was to use two gRNAs to simultaneously target the 5' regulatory region of *hnt* and a site elsewhere on the X chromosome. Taking advantage of the rich resources of *Drosophila* genetics allowed for the development of a screen in which double mutants could be selected in a visible F1 screen. This method relies on inducing two LOF mutations, each of which is associated with a dominant phenotype. In particular, this involved the haplo-insufficiency of the *Notch* locus and a previously developed method for over-expressing *hnt* in the developing pupal eye using CRISPR/Cas9.

Since both the *Notch* gene and the CRISPR/Cas9 mediated over-expression of *hnt* are dose sensitive, it was possible to identify double mutants by screening for visible phenotypes in the F1 progeny following mutagenesis. The basis of this screen is described in further detail below. Over the years, many inversions on the X chromosome have been isolated and described. Many of these have been used to make X balancers (*FM1, FM3,* and *FM7).* In addition, inversions with breakpoints in the vicinity of the two gRNA targets (*Notch* at 3C7; *hnt* at 4C10) have been described. These include In(1)Mud (3C3-4;5A6-B1).

An important distinction to make regarding the historical recovery of inversions is that these were for the most-part recovered by random X-ray mutagenesis. X linked mutations, being recessive visible, recessive lethal or dominant mutations were often found to be associated with CRs (Sturtevant & Beadle, 1936). Usually, these CRS were described by cytological analysis of the salivary gland polytene chromosomes.

The screen presented in this chapter uses CRISPR/Cas9 as the method of mutagenesis and simultaneously targets the gene *Notch* and the 5' regulatory region of *hnt* for Cas9 multiplexing. The possibility of inducing CRs through multiplexing is often cited as an undesired event, but there is little data available on the frequency with which CRs are induced by multiplexing. The screen presented, to the best of my knowledge, represents the first attempt to specifically induce a CR using CRISPR/Cas9 multiplexing in *Drosophila*. While the recovery of a desired inversion may be useful for further analysis of the regulatory region of *hnt*, this screen also serves to address whether or not Cas9 multiplexing can induce CRs at any appreciable frequency.

3.2.2 Methods

3.2.2.1 Identifying Notch mutants

Notch signalling is a cell-to-cell communication pathway conserved in practically all metazoans through mammals. The transmembrane receptor encoded by *Notch* is required for the maintenance of stem cells and cell determination. The two ligands of the Notch protein, also transmembrane proteins encoded by *Delta* and *Serrate,* activate the receptor to trigger the release of the Notch intracellular domain, which will ultimately alter gene expression. The dominant wing phenotype observed in *Notch* mutants is the result of failed signal transduction during the morphogenic process that establishes the wing margin's dorsal/ventral boundary during larval stages.

Mutations of *Notch* were first identified in *Drosophila,* when the deletion of the gene resulted in large indentations, referred to notches, in the wing margin. The phenotype was characterized as a dominant X-linked mutation, which was associated with haploinsufficiency (Mohr, 1919). Haploinsufficiency is uncommon and is associated with few genes in diploids that result in a mutant phenotype when heterozygous. Additional phenotypes associated with overexpression of such genes indicate that these genes are generally sensitive to gene dosage effects. Generally, dosage effects are associated with genes that encode proteins that form structural and regulatory complexes, as well as proteins that act in signalling pathways (Morrill & Amon, 2019). Viable alleles that display haploinsufficiency, such as *Notch*, offer a convenient and simple screening strategy that can be used to detect NHEJ following a DSB event.

3.2.2.2 Identifying hindsight mutants

If a mutant phenotype cannot be achieved for a given gene by haploinsufficiency, there are several methods available in *Drosophila* to confer a dominant phenotype to what is normally

recessive. One method used to detect mutations within *hnt* relates to such a technique. The system, developed within the Reed lab, permits the screening of a LOF *hnt* mutation within the F1 progeny of a cross between putative mutants and a line developed in the Reed lab, the "RGV" line. The RGV line takes advantage of the UAS-GAL4 system and a dead Cas9-VPR (dCas9-VPR) fusion protein to overexpress *hnt* within the eye disk of *Drosophila.* This overexpression of wild type *hnt* results in a "rough eye" phenotype, which is dosage sensitive such that heterozygous *hnt* LOF alleles could be selected as a visible phenotype. LOF mutants result in the wild type eye phenotype, as they are refractory to Hnt overexpression by Cas9-VPR (W. A. Ng, Ma, Chen, & Reed, 2019). This system is further explained in section 3.2.2.6.

3.2.2.3 Screening inversions within double mutants

Double mutants that result from this screen, identified by a notched wing and a wild type eye phenotype, are not conclusive of an inversion between the respective genes. These putative inversions can be subjected to a complementation test with a *hnt* null to confirm a lethal *hnt* mutation. The proposed inversion will fail to complement a null *hnt* allele*,* as the entire regulatory region of *hnt* will be misplaced from its native position. A true CR, however, cannot be confirmed without cytological analysis and/or PCR diagnostics.

3.2.2.4 Crossing scheme for recovery of putative inversions between *Notch* and *hindsight*

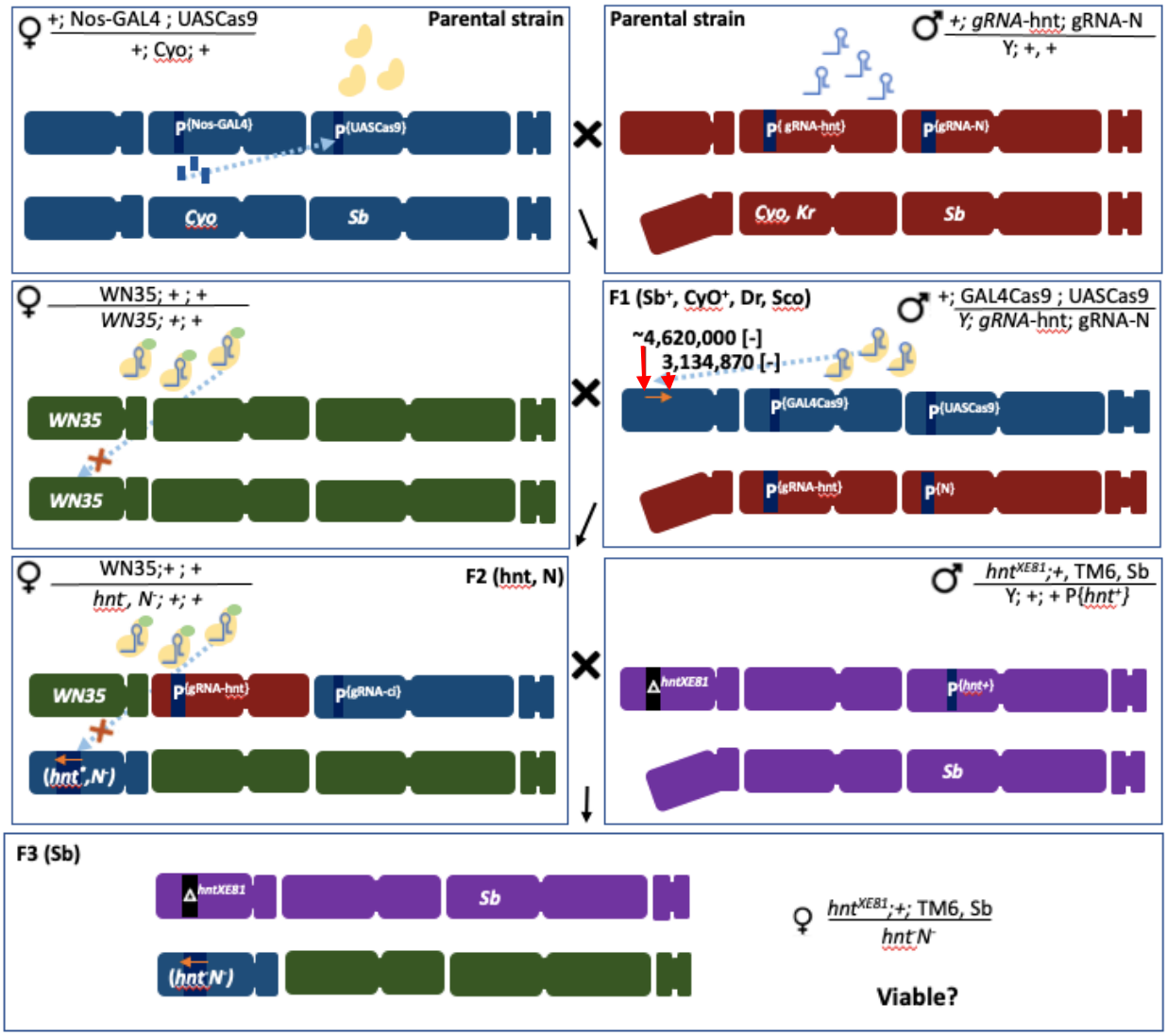


Figure 13. **Crossing of transgenic lines to induce and screen for recovery of an inversion.** The parental cross will produce F1 offspring with the Cas9 and gRNA expression within germ cells. Two screens were done in parallel using different multiplexed mutations. Genomic locations of targeted mutagenesis are indicated using red arrows. The upstream DSB is targeted to *N* (3,134,870), while the downstream mutation is targeted to a site upstream to the TSS of *hnt* (4,617,564). The first screen targets Cas9RNAs 109-131 bp upstream of the TSS, while the other screen targets Cas9 181-203bp upstream of the TSS. F1 double mutants with germline mutagenesis are crossed to *WN35* flies. While this stock expresses dCas9-VPR targeted to overexpress *hnt*, the stock is refractory to overexpression owing to a previous mutation in the designated *hnt* target sequence. F2 offspring that inherit the transgenes required for *hnt* overexpression and a double mutated chromosome will have the notch wings and wild type eyes. If an inversion is induced between these two sites, *hnt* will be lethal. This would not be viable over a *hnt* null, hence the F3 generation illustrated would not be produced.

3.2.2.5 Expression and targeting of the Cas9 endonuclease

This screen uses the *Nos* promoter to drive expression of the GAL4 transcription factor. The gRNAs were cloned into the pCED4 which expresses gRNAs ubiquitously under control of gRNA:U6:96Aa and gRNA:U6:96A).

3.2.2.6 The RGV line

The RGV line includes a dCas9-VPR fusion protein expressed under the UAS element. The line also contains a transgene that ubiquitously expresses gRNAs designed to target the dCas9 complex to two sites upstream to the TSS of *hnt*. Overexpression of *hnt* using this complex is restricted to the pupal eye by using the glass multiple reporter (GMR) element as a GAL4 driver for dCas9. This promoter shows restricted activity in the developing eye of pupae (W. Li, Li, Zheng, Zhang, & Xue, 2012). The over-expression of *hnt* using the GMR-GAL4 driver line results in the rough eye phenotype. The RGV line is made refractory to overexpression, as the target sites of dCas9 are mutated. The inability of dCas9 to bind to these targets prevents overexpression, resulting in a wild type eye phenotype. When this line is crossed with flies with X chromosomes that are not refractory to overexpression, the dCas9 system is able to overexpress *hnt* of the diploid's homologous chromosome resulting in the rough eye phenotype. In cases where the line is crossed to *hnt* mutants, overexpression is abrogated and flies present a wild type eye (W. A. Ng et al., 2019). This system is illustrated in figure 14 below.

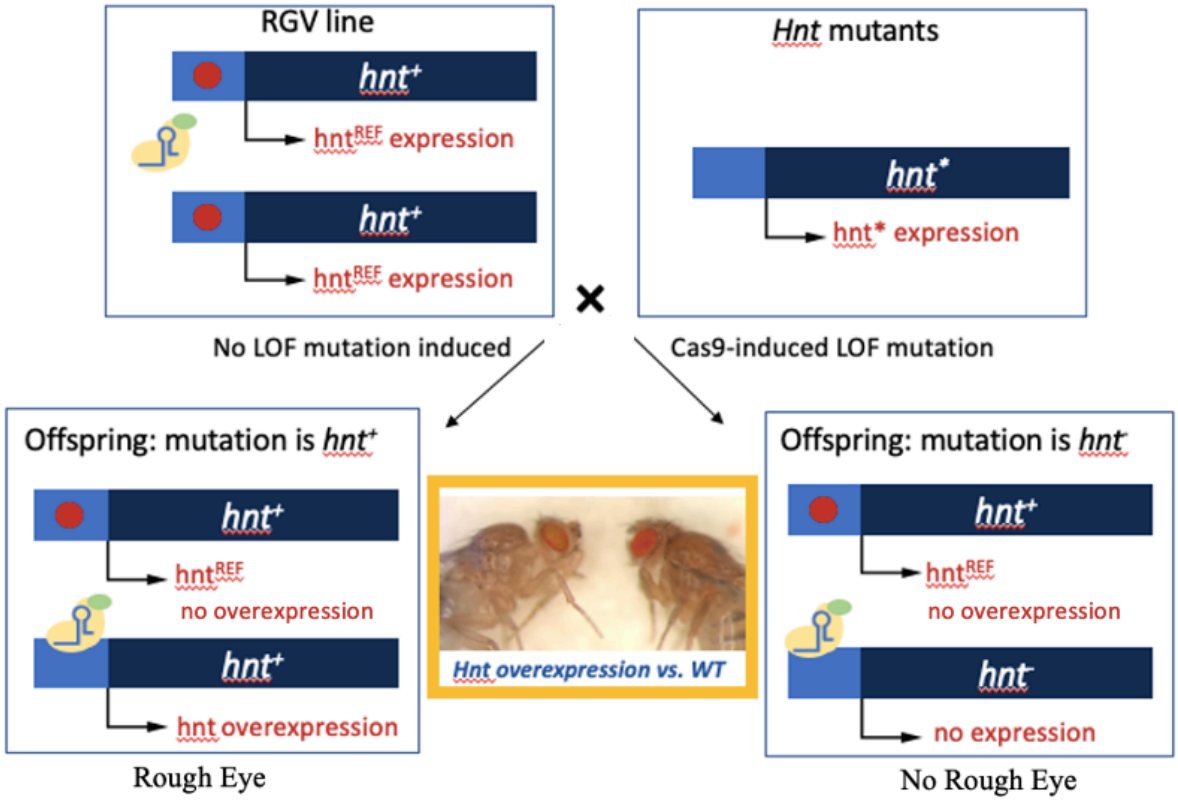


Figure 14. Using the RGV line to identify LOF *hnt* mutants. One parent is from the RGV line and contains the components necessary for *hnt* overexpression in cells encompassed in the eye disc. This includes GAL4 under a pupal-specific promoter, dCas9-VPR under the control of the UAS element, a sgRNA targeted upstream to *hnt*. However, target sites of Cas9 are mutated to prevent overexpression. This is crossed to a putative *hnt* mutant. Offspring in this cross inherit the components necessary for pupal overexpression, as well as a chromosome with the *hnt* allele in question. If the mutated *hnt* allele is wild type, overexpression is possible leading to a rough eye phenotype. If *hnt* is a LOF allele, overexpression is not possible, hence there is a wild type phenotype.

3.2.2.8 Complementation testing

Double mutant females from the inversion screen were crossed with male flies harbouring a null hnt^{XE81} mutation. The cross used is illustrated in Figure 15. While the mutant line used for the complementation test is deficient for *hnt* in its native position, gene activity is rescued by an insert of hnt^{+} on the third chromosome. Female offspring was screened for double *hnt* mutants (hnt^{XE81}/hnt^{REF}). Deviation from the expected 25% of female offspring with this genotype indicates failure, or partial failure, of null *hnt* complementation by the double mutated chromosome. Double mutant stocks that produced unexpected numbers of hnt^{XE81}/hnt^{REF} female offspring were also used in a complementation test with flies harbouring the hnt^{308} mutation. This mutation is a *hnt* hypomorph, which only has partial loss of *hnt* activity.

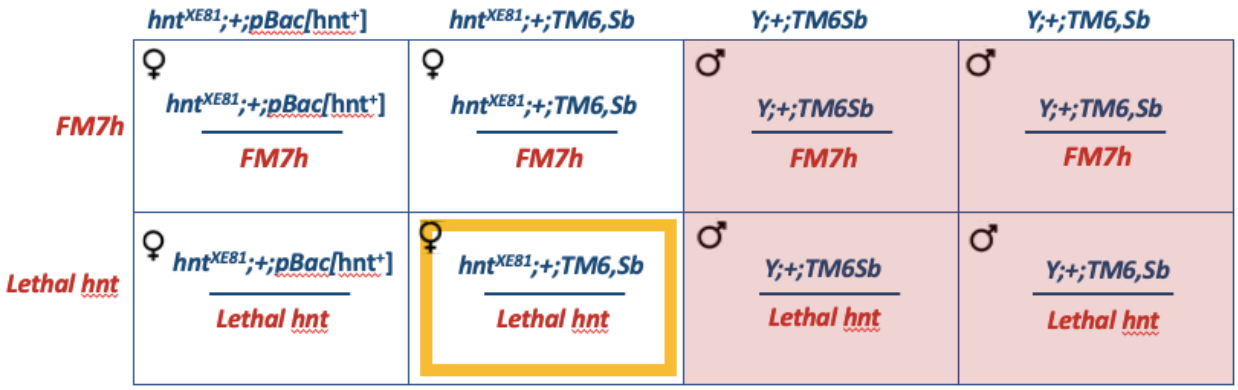


Figure 15. Complementation test of flies with double mutations. Red and blue chromosomes represent maternal and paternal chromosomes respectively. Mothers with a potential lethal *hnt* alleles are crossed fathers deficient for *hnt*. Activity of *hnt* is rescued by an insert of the wild type gene on the third chromosomes. Female offspring that have not inherited the rescuing gene (they have inherited the balancer with a dominant marker of the stubbled phenotype) are screened. If the *hnt* allele is lethal, the only expected female offspring will have FM7h balancer (with a dominant marker for the bar phenotype). If the *hnt* is not lethal, there will also be female progeny without the bar phenotype. The females being screened are indicated in the yellow box; if the mutated *hnt* allele is truly lethal, these females will not exist amongst offspring.

3.2.2.7. Female sterility assay

To identify female fertility effects, females from the complementation cross (discussed above) were used to recover *N hnt* mutants balanced over a null *hnt* allele (hnt^{XE81}). Females were crossed to their siblings. Of the N *hnt* stocks, 12 lines with the hnt^{4344} allele and 19 lines with the hnt^{4849} allele were selected.

3.2.2.9 Polytene squash analysis

Polytene chromosomes were visualized using phase contrast microscopy. Salivary glands were dissected from 3rd stage larva in 45% acetic acid. The glands were stained with orcein (1% orcein dye in 45% acidic acid solution) and squished to spread the polytene chromosomes.

3.2.2.10 Embryo immunostaining

Embryos were washed with 50% bleach 18 hours following a cross and transferred into a heptane/aqueous fixation solution (8mL dH_2O, 1 mL 10x PBS, 1 mL 3.7% formaldehyde). After fixation the aqueous phase was removed, and 95% methanol was used to separate embryos from their vitelline envelope. Embryos were transferred through a methanol gradient (75%, 50% and 25%), washed with 1X PBS, and rinsed with 1X PBS. Embryos were incubated in blocking solution (95% PBS-TX with 2.5% BSA, 5% natural goat serum (N g/s)), then in mouse monoclonal anti-*hindsight* (425 uL PBT-BSA, 25 uL Ng/s, 50 uL d-*hnt*). After washing with PBT-BSA, they were incubated in blocking agent (475 uL PBS, 25 uL Ng/s). Samples were incubated with the secondary antibody (2.5 uL anti-mouse antibody, 473 uL PBT-BSA), washed with PBT and PBS, then transferred through a glycerol gradient (25%, 50% and 75%). Visualization using DABCO containing moutant (70% glycerol/PBS) was done using 20X objective and confocal microscopy.

3.2.3 Results

3.2.3.1 Characterization of offspring from the inversion screen

Table 3 shows the efficiency of male germline mutagenesis using Cas9 targeted to *N* and ~3Kbp upstream of *hnt* (the target site further from the *hnt* TSS will be referred to as hnt^{4344}). An impressive 98% of offspring had hnt^{4344} mutations, while 66% had *Notch* mutation. Almost all *N* mutations manifested as double mutants (65%). Given 1% of offspring showed a single mutation of *Notch,* occurrences of this mutation are favoured when hnt^{4344} is also mutated. When a single mutation was induced, the hnt^{4344} locus was favoured. This pattern was observed when mutagenesis was carried out in the female germline.

Table 4 shows the efficiency of female germline mutagenesis using the inversion screen. Offspring showed a mutagenesis efficiency of hnt^{4344} and *Notch* to be 93% and 66% respectively. As observed in the male germline, most of these mutations were manifested as *hnt, N* double mutants (62%). When Cas9 was used to multiplex *Notch* and a site closer to the *hnt* TSS (and further from the *Notch* locus), the efficiency of *hnt* mutagenesis decreased (93%). *Notch* mutations maintained a similar efficiency. Despite this lower frequency of *hnt* mutations in this screen, the same trends described for Table 3 were observed. Most *Notch* mutations occurred when *hnt* was also mutated, and the *hnt* locus was favoured for mutagenesis when a single mutation resulted from Cas9 multiplexing.

Table 5 shows the efficiency of male germline mutagenesis when Cas9 was targeted to *Notch* and ~2.5 Kbp upstream of *hnt* (the target closer to the *hnt* TSS will be referred to as hnt^{4849}). The efficiency of *hnt* and *Notch* mutagenesis was 74% and 58% respectively. Only 37% of mutations were manifested as double mutants, which still make up most offspring with *Notch* mutations. Table 6 shows the efficiency of female germline mutagenesis for the same targets

used in table 4. The efficiency of *hnt*^{4849} and *Notch* mutagenesis was 62% and 55% respectively. Only 33% of offspring were double mutants, as the event of a single mutation was much higher in the female germline. While >5% of offspring had a wild type phenotype in all other experiments, 15% of flies from this experiment remained unmutated. Overall, this may suggest that when the distance between two multiplexed sites is increased, the probability of a double mutation decreases. Additionally, the favouring of one site for single mutagenesis during Cas9 multiplexing may decrease with this increased distance. Although there are discrepancies between the rate of single and multiplexed mutations based on the germline and genomic location, all screens produced double mutant offspring with a reliable efficiency.

Cross	hnt+, N+	hnt-,N+	hnt+,N-	hnt-,N-
1	0	33	2	45
2	0	38	0	64
3	1	10	0	42
4	0	33	1	72
5	0	12	0	52
6	0	21	0	31
7	0	6	0	18
8	0	41	3	79
9	0	0	0	6
10	0	26	0	49
11	0	1	0	7
12	0	1	1	2
13	0	7	0	14
14	1	18	0	35
15	0	22	0	43
16	0	24	2	43
17	0	44	3	70
18	0	31	0	76
19	0	10	0	3
20	0	8	0	13
21	0	7	0	18
22	0	10	0	11
23	0	2	0	2
24	0	22	0	25
25	0	10	0	15
26	0	5	0	32
27	0	13	1	41
28	0	3	0	5
Total	**2**	**458**	**13**	**913**

Table 3. Frequency of sgRNA-hnt^{4344} and sgRNA-N mutagenesis in offspring of the inversion screen using Cas9 in male germline. Offspring of the *N, hnt^{4344}* mutagenesis screen using Cas9 mutagenesis within the male germline. Successful mutagenesis will be observed as flies with notches in the wing margin, due to mutagenesis of the *N* locus, and wild type eyes, due to aborted overexpression from dCas9 due to *hnt* mutagenesis. Double mutants were then subjected to a complementation test to identify inversions within the *hnt* regulatory region.

Cross	hnt+, N+	hnt-,N+	hnt+,N-	hnt-,N-
1	0	12	0	21
2	0	8	0	20
3	1	24	4	45
4	0	9	0	24
5	0	15	0	61
6	0	3	0	28
7	0	1	7	5
8	0	3	0	14
9	0	13	0	33
10	0	9	0	38
11	2	17	2	54
12	0	11	0	30
13	0	8	0	3
14	1	31	4	93
15	0	7	0	18
16	2	13	1	33
17	0	5	0	26
18	0	10	0	8
19	0	6	0	35
20	0	23	0	20
21	0	7	0	3
22	0	12	0	38
23	1	17	1	8
24	0	9	0	25
25	6	21	0	55
26	1	18	6	6
27	3	21	8	6
28	12	21	5	10
29	5	30	21	18
30	0	0	0	4
Total	**34**	**384**	**59**	**782**

Table 4. Frequency of sgRNA-hnt^{4344} and sgRNA-N mutagenesis in offspring of the inversion screen using Cas9 in female germline. Offspring of the *N, hnt^{4344}* mutagenesis screen using Cas9 mutagenesis within the female germline. Successful mutagenesis will be observed as flies with notches in the wing margin, due to mutagenesis of the *N* locus, and wild type eyes, due to aborted overexpression from dCas9 due to *hnt* mutagenesis. Double mutants were then subjected to a complementation test to identify inversions within the *hnt* regulatory region.

Cross	hnt+, N+	hnt-,N+	hnt+,N-	hnt-,N-
1	7	48	22	38
2	2	32	18	31
3	0	16	9	13
4	3	23	7	31
5	3	9	6	12
6	4	7	1	0
7	2	2	7	3
8	4	27	14	26
9	4	25	12	23
10	2	17	6	8
11	0	12	7	13
12	6	22	17	28
13	3	34	19	42
14	1	15	8	23
15	0	24	12	23
16	6	42	15	35
17	4	28	14	32
18	2	10	10	9
19	5	43	28	36
20	2	18	22	39
22	1	15	12	20
23	3	32	12	22
Total	**64**	**501**	**278**	**507**

Table 5. Frequency of sgRNA-hnt^{4849} and sgRNA-N mutagenesis in offspring of the inversion screen using Cas9 in male germline. Offspring of the *N, hnt^{4849}* mutagenesis screen using Cas9 mutagenesis within the male germline. Successful mutagenesis will be observed as flies with notches in the wing margin, due to mutagenesis of the *N* locus, and wild type eyes, due to aborted overexpression from dCas9 due to *hnt* mutagenesis. Double mutants were then subjected to a complementation test to identify inversions within the *hnt* regulatory region.

Cross	hnt+, N+	hnt-,N+	hnt+,N-	hnt-,N-
1	6	34	13	28
2	8	29	12	35
3	5	35	16	48
4	7	6	10	5
5	14	11	23	14
6	23	22	18	18
7	10	12	22	14
8	49	12	31	31
9	23	65	37	92
10	10	22	6	18
11	3	17	9	17
12	2	8	2	5
13	0	9	3	2
14	5	12	8	20
15	1	9	2	10
16	4	13	7	18
17	3	22	12	26
18	6	8	8	22
19	3	18	10	24
20	7	6	10	5
22	3	0	9	1
23	9	1	6	3
24	22	1	27	2
25	9	3	5	2
26	2	14	17	22
27	0	40	2	14
Total	**234**	**429**	**325**	**494**

Table 6. Frequency of sgRNA-hnt^{4849} and sgRNA-N mutagenesis in offspring of the inversion screen using Cas9 in female germline. Offspring of the *N, hnt^{4849}* mutagenesis screen using Cas9 mutagenesis within the female germline. Successful mutagenesis will be observed as flies with notches in the wing margin, due to mutagenesis of the *N* locus, and wild type eyes, due to aborted overexpression from dCas9 due to *hnt* mutagenesis. Double mutants were then subjected to a complementation test to identify inversions within the *hnt* regulatory region.

3.2.3.3 Complementation testing of the Notch and *hindsight* double mutants *and female fertility assays*

Single mutant females with either *Notch, hnt*^{4344}, or *hnt*^{4849} mutations were able to complement a *hnt*^{XE81} and *hnt*^{308} allele in a complementation test (demonstrated in Figure 14). Additionally, offspring from these complementation crosses demonstrated no fertility effects. Of the double mutants recovered from the screen, 54 flies were selected to establish balanced stocks and subjected to a complementation test (28 lines with the *Notch hnt*^{4344} genotype, and 25 lines with the *N hnt*^{4849} genotype). Partial failure of complementation was observed by three *Notch* and *hnt*^{4344} double mutants when crossed with flies expressing the null *hnt*^{XE81} allele. This was observed by a lower-than-expected recovery of *Notch hnt*^{4344} */ hnt*^{XE81} flies (>25%). Partial failure of complementation was also observed by two *Notch hnt*^{4849} double mutant lines. Progeny from 12 of the *hnt*^{XE81} complementation tests were crossed to their siblings to observe fertility effects of *Notch hnt*^{4344}/hnt^{XE81} females. No offspring was produced by three of the lines displaying partial complementation, or two other lines that complemented the null allele. When a complementation test using the hypomorphic *hnt*^{308} allele was performed on these lines, all flies displayed complete complementation, including the three lines displaying partial complementation to the *hnt*^{XE81} allele. Two of these lines displayed infertility when the *N hnt*^{4344} mutation was coupled with the hypomorphic allele. All *Notch hnt*^{4849}*/ hnt*^{XE81} flies demonstrated complete complementation of both the *hnt*^{XE81} and *hnt*^{308} alleles. When female offspring from the complementation test were crossed to their siblings, no effects in fertility was observed. However, six of the 19 lines tested were sterile when offspring from the *hnt*^{308} complementation test was crossed to siblings. Both complementation tests were also carried out in parallel by Dr. Reed. This screen recovered two *N hnt*^{4849} lines that failed to complement a null *hnt*^{XE81} allele, as well as three *N hnt*^{4849} lines that partially failed to complement. Several *Notch hnt*^{4344} also failed

or partially failed to complement this allele. Three *N hnt*^{4849} lines and one *N hnt*^{4344} line were selected for *hnt* expression analysis using immunostaining.

3.2.3.4 Immunostaining of recovered Notch and hindsight mutants

Homozygous *N hnt* double mutants were embryonic lethal with both *hnt*^{4344} and *hnt*^{4849} alleles, as seen in Figure 17 and 18 respectively. The established line of *Notch hnt*^{4344} mutants will be referred to as the X8 line. Some homozygous mutants of this line displayed partial absence of *hnt* expressing cells. Expression was normal in the amnioserosa; expression was absent in all other normally expressed regions, except for faint expression in the oenocytes (Figure 17, B). In other cases, *hnt* expression was completely absent, and embryos failed in GBR (Figure 17, C and D). Homozygotes from the three established lines of *N hnt*^{4849} mutants, referred to as D2, G3 and M1, demonstrated similar phenotypes to *N hnt*^{4344} homozygotes. D2 embryos had normal expression of *hnt* in the PNS, but irregular clustering of *hnt* expressing cells in oencoytes (Figure 18, B). Some D2 mutants also had no expression of *hnt* outside of faint expression within some cells in the site of oenocytes. These embryos failed in GBR (Figure 18, C). G3 *N hnt*^{4849} homozygotes displayed no *hnt* expression, followed by failure of GBR (Figure 18, D and E). M1 *N hnt*^{4849} homozygotes also displayed weak expression of *hnt* in oenocytes in some cases (Figure 18, G). In other cases, *hnt* expression was absent and embryos failed in GBR (Figure 18, F).

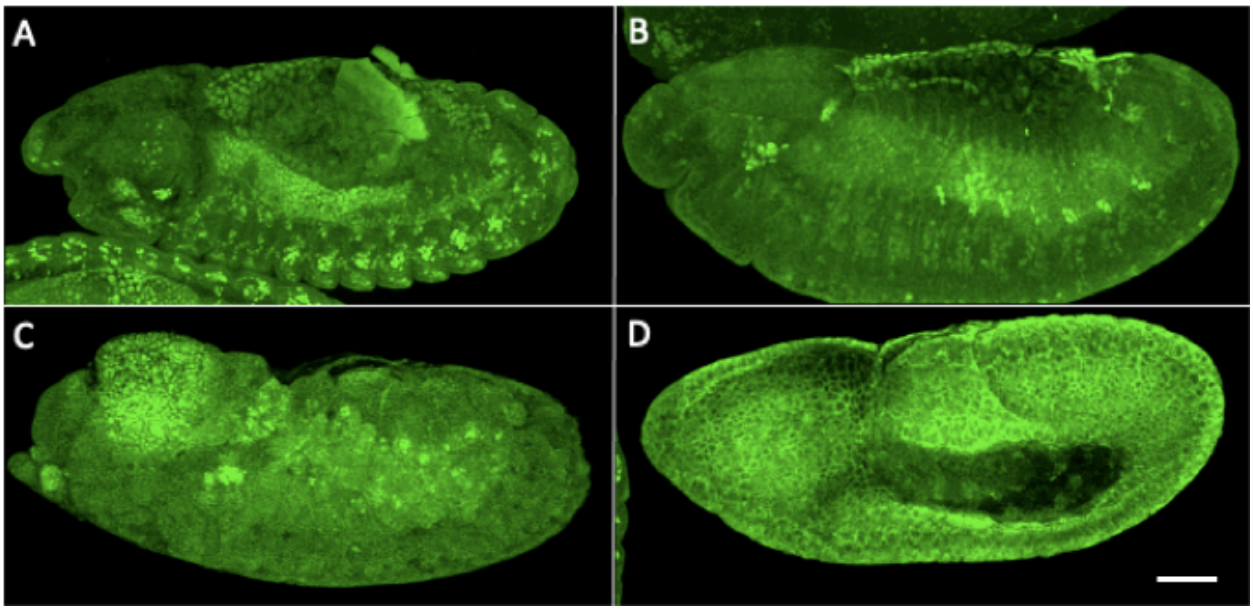


Figure 17. Immunostaining of anti *hnt* within double *N hnt*^{4344} mutants. Wild type *hnt* has distinctive patterning in the embryo, including expression within seven clusters corresponding to oenocytes and three clusters in the embryo head (A). Flies of the X8 line displayed irregular clustering of *hnt* expressing cells in oenocytes (B) and failure of GBR, where expression of *hnt* was almost absent with the exception of light expression in oenocytes (C-D). Scale bar represents 50 μm.

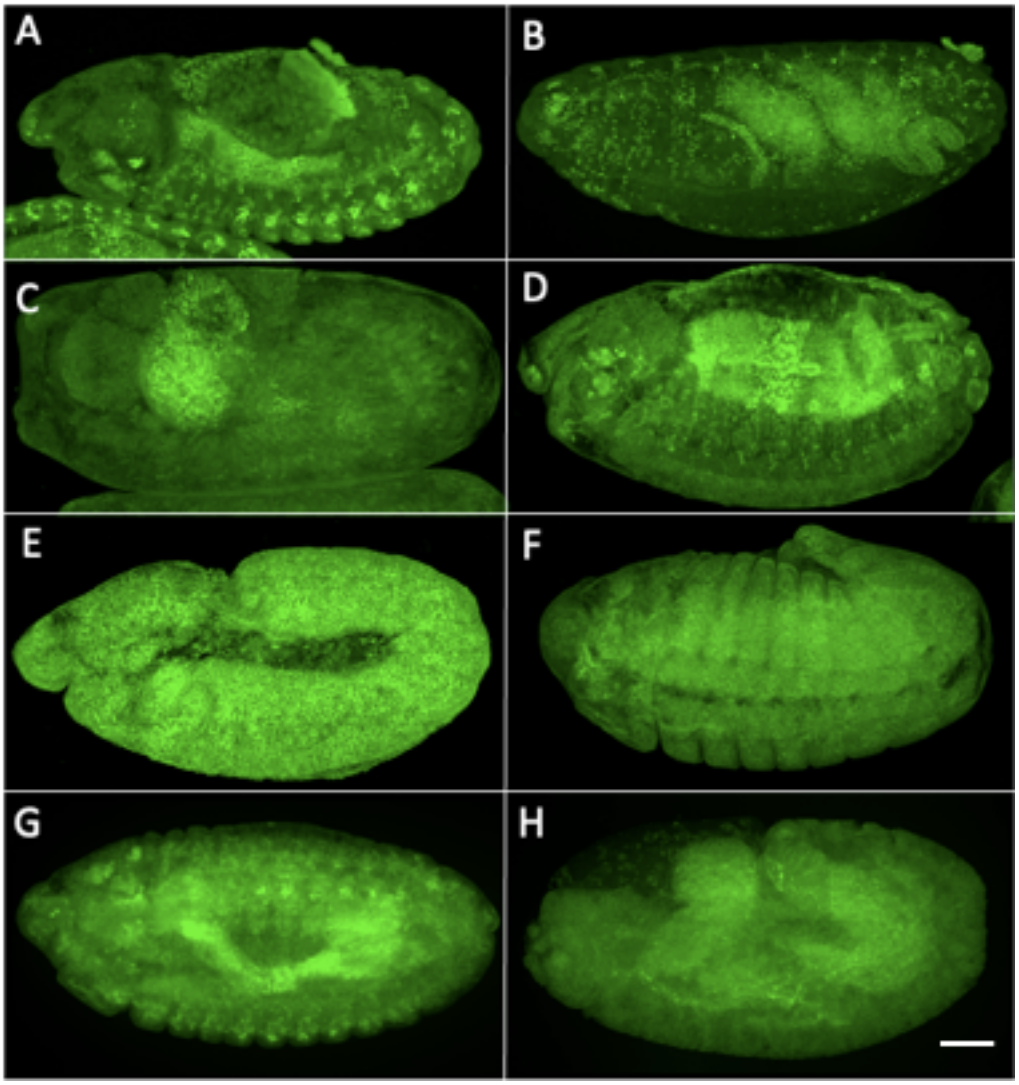


Figure 18. Immunostaining of anti *hnt* within double *N hnt*^{4849} mutants. Wild type *hnt* displays a distinctive patterning in embryos, including expression within seven clusters corresponding to oenocytes and three clusters in the embryo head (A). D2 mutants displayed irregular clustering of *hnt* expressing cells in oenocytes (B) and failure of GBR, where expression of *hnt* was almost absent except for light expression in oenocytes (C). G3 mutants also showed irregular clustering of oenocyte cells. Additionally, *hnt* expression in these cells was low (D). G3 mutants also displayed absence of *hnt* expression, coupled with failure of GBR. M1 mutants also displayed irregular clustering and light expression of *hnt* in oenocytes, as well as GBR coupled with faint but irregular *hnt* expressing cells. Scale bar represents 50 μm.

3.2.4 Discussion

The multi-phenotype screen indicates that *nos-Cas9* provides an efficient method for inducing two DSBs in the genome simultaneously. In both screens, mutagenesis for both single and double loci was more efficient in the male germline. This is consistent with the literature and the *sd* screen described above, which suggests that the male germline has a higher rate of induced mutations compared to the female germline when using *nos*-Cas9 (Kondo & Ueda, 2013). Data of germline mutagenesis efficiencies using *nos-Cas9* only includes *white* and *sd* for single mutations, and *Notch/hnt* for multiplexed mutations. These are all present within close proximity of each other on the X chromosome. Since genomic location may influence which germline induces mutations more frequently, this trend should be verified in other chromosomes. Males also lack a homologous chromosome to the X chromosome. Since the homologous chromosome is the prime template used for HDR repair, higher mutagenesis may only be favoured in the male germline when targets are on the X chromosome.

The mutation frequency of hnt^{4344} was >90% in both germlines, however this decreased by ~15% in both germlines when Cas9 was targeted 0.5Kbp upstream to hnt^{4849}. The rate of double mutations decreased over 30%. This can indicate that the frequency of locus mutagenesis when multiplexing may be dependent on the distance between the two Cas9 target sites. Furthermore, results suggest that multiplexed mutations that have target sites within proximity of each other may bias one site for mutagenesis. This can be attributed to several factors. For example, one feature exhibited by RNAs is their ability to titrate proteins away from a particular genomic location. This can decrease transcription and translation (Carlberg & Molnár, 2014). Similar experiments with this screen's target sites can be repeated using different locations for chromosomal expression of Cas9 and gRNAs. Furthermore, different distances between the two

target sites can be compared. The location for *hnt* mutagenesis can be changed while the *Notch* locus remains fixed. If a change in distance does affect the favouring of *hnt* mutagenesis, results indicate that this bias is less prominent with increasing distance between the two target sites. It is also possible this bias becomes negligible as the distance between the two targets sites increases.

Complementation tests offspring indicated that no inversions were recovered in the screen. Flies that failed to complement the null *hnt* allele were subjected to polytene squash analysis. The polytene chromosome showed no visible inversions in the X chromosome, an event that is quite distinct in cytological analysis. This could indicate that Cas9 multiplexing is not a reliable method for inducing inversions in *Drosophila*, as recovery of these CRs was not detectable in >3,200 mutagenesis events. However, certain factors can result in the designed inversion having a lethal or sterile phenotype. This would mean that inversions induced would not be passed down to offspring, hence would be undetectable.

Given the role of both *hnt* and *Notch* in fertility, it was of interest to determine if the mutations induced by this screen affected fly fertility. Findings indicate that the two locations selected for mutagenesis may result in female sterility when both mutations are present. A type of combined female fertility effect is observed by double mutants in complementation test offspring, where the mutation of *hnt* is likely a point mutation resulting from a single Cas9 mutagenesis event. This mutation can result a hypomorphic *hnt* allele that displays some types of fertility effects when combined with the haploinsufficient *Notch* mutation. In the case of an inversion, the mutation of *hnt* would displace the *hnt* protein coding gene away from a large portion of its 5' regulation region. This large-scale mutation will likely recover a null *hnt.* It is possible that while a hypomorphic *hnt* mutation paired with the haploinsufficient *Notch* mutation

can cause some fertility effects, a null *hnt* allele paired with this *Notch* mutation causes complete sterility. This would make inversions unrecoverable in the screen.

The combined haploinsufficiency *hnt* and *Notch* is likely due to effects during oogenesis, as the expression of *hnt* and *Notch* work together to regulate follicle cell proliferation and differentiation. During this process, the Notch signalling pathway upregulates *hnt* in follicle cells. Additionally, *hnt* mediates downregulation of *Cut*, *string* and *Hedgehog* signalling in a Notch-dependent manner (Sun & Deng, 2007). While *hnt* has been recognized to regulate egg competency, Notch signalling likely does not have a role in this process since it is not present in stage 14 follicle cells (Deady et al., 2017).

A change in *hnt* gene dosage resulted in a change in the severity of the sterile phenotype. In *Notch* hnt^{4344}/hnt^{XE81} and *Notch* hnt^{4849}/hnt^{XE81} females no eggs produced. In *Notch* hnt^{4344}/hnt^{308} and *Notch* hnt^{4849}/hnt^{308} females, a different fertility effect was observed. Eggs were produced but not released, as evident by the accumulation of embryos within the female abdomen. In cases where eggs where successfully laid, embryos died after a few days. This supports the idea that there is relationship between the severity of *hnt* mutagenesis and sterility effects; in the case of hypomorphic *hnt* alleles paired with *Notch* there is some sterility effects, while a null *hnt* allele paired with *Notch* results in complete sterility. This assumes that the hnt^{4344} and hnt^{4849} alleles do not result in fertility effects when paired with a null *hnt* allele. While these alleles have displayed fertility effects in the control screen, it is possible that fertility effects were not detected due to small test sizes.

While inversions were not recovered in this screen, double *Notch hnt* mutations were recovered that display abnormal activity of *hnt*. Double mutants that failed to complement a null *hnt* allele were analyzed for *hnt* expression using immunostaining. The wild type expression of

hnt was completely missing from some *Notch hnt* homozygous embryos. The absence of *hnt* eventually results in late-stage embryos failing to complete embryogenesis. In cases where *hnt* expression was present, it was faint. Given that two gene targets can result in combined lethality or sterility, the complementation test can be used to screen any combined effects that result from a multiplexed mutation. The activity of genes directly adjacent to the break points of a multiplexed mutation will likely be affected by mutagenesis, hence screening a double mutation using complementation tests and fertility assays should be done to establish feasible targets for inducing CRs.

While the combined haploinsufficiency of *Notch* and *hnt* may be one explanation as to why no inversions were recovered, it is also possible that inversions are being induced at a very low frequency. Based on this screen, Cas9 would be an impractical method for inducing inversions. One method to increase the probability of recovering inversions is to induce DSBs in regions of homology. Similar cut sites may support DNA strand ligation into the genome in a non-specific manner. Sequences of homology may also support HR between the two sites, which may recover inversions more readily. This method was tested to induce and screen deletions using the strategy discussed in the next chapter.

3.3 Screening Deletions using Allelic Combinations

3.3.1 Introduction

Drosophila eye colour is a polygenic trait that is affected by genes associated with synthesis or transportation of pigments into lysosome-like organelles in the eye. These structures are known as pigment granules. The *white* gene encodes a transporter protein required for the entry of substrates to the pigment granules in the developing pigment cells of the pupal eye; homozygous *white* mutants display a white eye phenotype due to the lack of transportation for both major eye pigments (Grant et al., 2016). It has been demonstrated that *white* gene dosage can be rescued with insertions of the gene within ectopic sites of the X chromosome.

The *white gene* is routinely used as a genetic marker for many different transgene constructs. The modified *white* transgene is known as *mini-white* (Qian' & Pirrotta, 1995). The allelic combinations of *mini-white* and the gene *rb,* also involved with the pigmentation observed in the *Drosophila* eye, creates the basis of the deletion screen described below.

A transgenic line carrying two *PBac{Rb}[mini-white]-FRT* insertions was previously described and is available in stock centres. These two inserts, which are present in a null *white* background, are arranged in *cis* on the same chromosome and map ~5Kb and ~180Kbp upstream of the *hnt* TSS. Additionally, a CRISPR/Cas9 stock that is designed to target the *mini-white* gene was recently described and is also available in stock centres (Liu et al., 2020). This line, called the 'white eraser' (WE) line, expresses two gRNAs targeted to different regions within the *mini-white* gene, as well as active Cas9. Mutation of *mini-white* using this line has been observed with a 57% efficiency (Liu et al., 2020). Using these two stocks, it is possible to test if the WE line could result in a ~180Kb deletion between the two *Pbac{RB}[mini-white]-FRT* inserts, which are known as $PBac\{rb\}bi^{e02388a}$ and $PBac\{rb\}^{e02388b}$.

In addition to using the WE approach, the FRT sites contained in each *PBac{RB}* insert allows the recovery of the same ~180Kb deletion using FLP/FRT-mediated SSR. In this way, the comparison of CRISPR/Cas9 and FLP/FRT recombination in the generation of the ~180Kb deletion can be made. Since the full extent in which genomic location and fragment size affects the probability of a CR event is not fully understood, inducing a deletion using the same target sites avoids deviations in results associated with these factors. Since the two *PBac{RB}* sites a flank the gene *r*b, this offers a mechanism to screen for potential deletions, as flies with both *mini-white* mutations can be crossed to flies with the visible recessive rb^1 allele. If the deletion did occur, the allelic combination will give rise to offspring that display ruby coloured eyes. Furthermore, the complementation test highlighted in chapter 2 can be used to identify deletions, as discussed below.

3.3.2 Materials and Methods

3.3.2.1 Screening deletions using ruby

The recessive *rb* gene is involved with *Drosophila* eye pigmentation. Prior to the formation and budding of transport vesicles from a cell, a protein coat is assembled. The specificity of this coat is dictated in part by the adaptor protein complex. One subunit of this complex is encoded by *rb*. Certain allelic combinations of *rb* result in the ruby eye phenotype. (Grant et al, 2016). Since heteozygous $rb^1/Df(rb)$ display the ruby phenotype, any deletion of *rb* can be detected as visible recoveries over rb^1. This is the basis of the proposed screen for deletions, as the WE and FLP/FRT-based techniques for inducing a deletion are targeted to sites flanking *rb*.

3.3.2.2 Confirming putative deletions of ruby

Flies with putative deletions, distinguished by the ruby phenotype, were subjected to a complementation test to confirm the deletion of *rb*. Deficiencies of *rb* can be tested for its complementation of null *hnt* alleles in a similar way null *hnt* was screened. While this provides evidence of a deletion, a true rearrangement cannot be confirmed without molecular or cytological analysis of the chromosomal region.

3.3.2.3 Crossing scheme for inducing and screening putative deletions

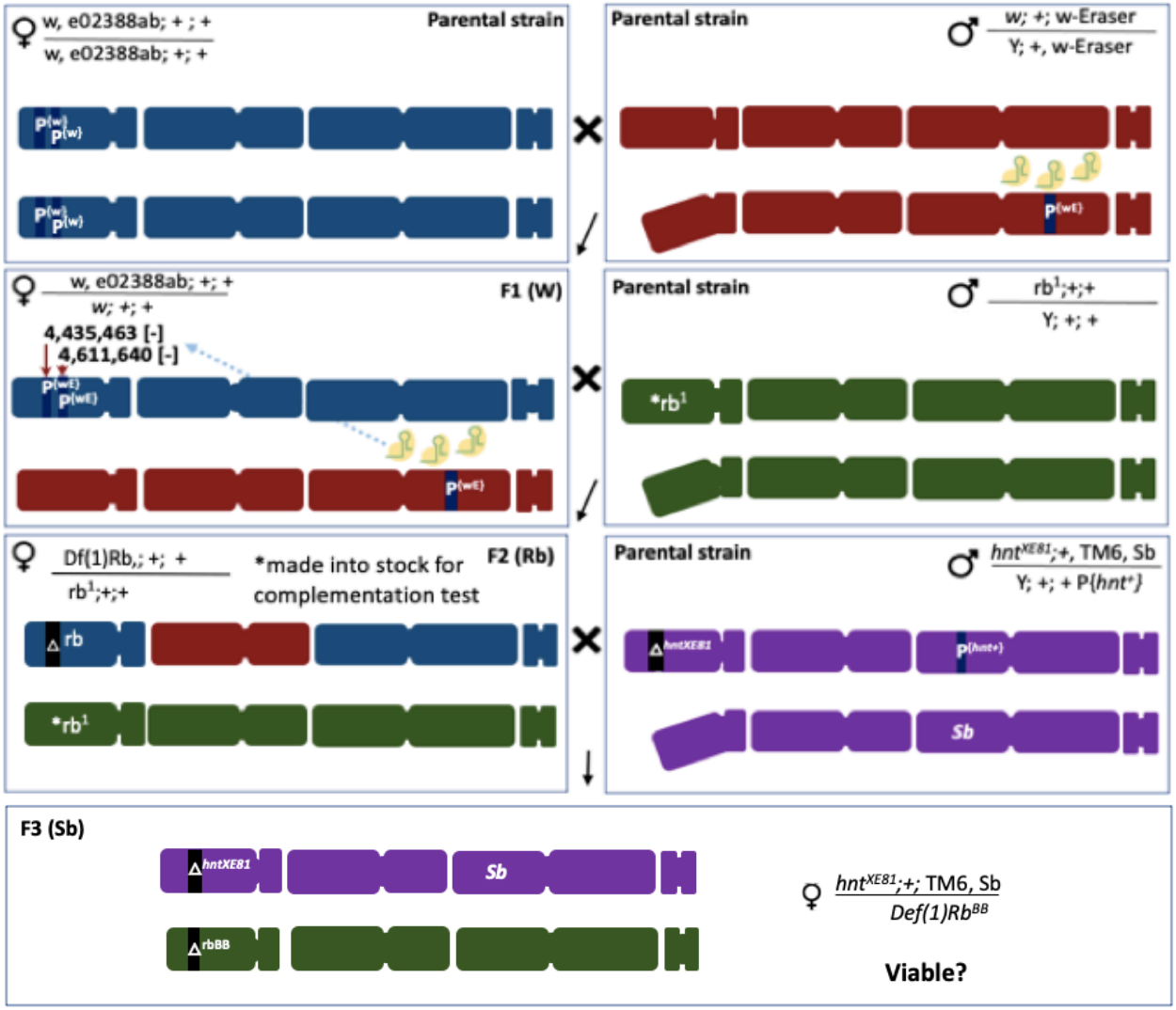


Figure 19. Crossing of transgenic lines to induce and screen a deletion. The F1 offspring will have Cas9-sgRNA activity within germ cells. All *mini-white* genes, which are present within P element inserts, are targeted by the Cas9. This is indicated by the red arrows (genomic location also show). There are two targets within the *mini-white* gene, expressed by the "white eraser" chromosome. F2 offspring that have both *mini-white* genes mutated will have white eyes. These are crossed to a stain which is homozygous for the rb^1 deficiency. If F2 offspring inherit a chromosome with the rb^1 deficiency and a chromosome with a deletion event, it will display the ruby eye phenotype. If a deletion was made in ruby, F3 progeny will not be viable because the *rb* deletion is lethal when paired with a *hnt* null.

3.3.2.4 Expression and targets of the Cas9 endonuclease encoded by the white eraser line

The Cas9 protein is expressed under the Act5C regulatory sequence. The gRNAs are expressed ubiquitously under the U6 promoter and targeted to *mini-white*. The two insert sites of the *PBac* element that houses *mini-white* is indicted in the crossing scheme above (Figure 19).

3.3.2.5 Heat shock to induce the FRT/FLP system

FRT sites are available on the same *PBac{RB}* element that houses the *mini-white* genes. Larvae with insertions heatshocked at 37°C for 40 minutes during the 2nd instar stage to activate the heat sensitive promoter of FLPase.

3.3.2.6 Complementation test

A complementation test was carried out as described in figure 15 of section 3.2.2.8.

3.3.2.7 Polytene squash analysis

The protocol for a polytene squash analysis was done as described in section 3.2.2.9. When visualizing chromosomes, the banding pattern of the X chromosome was mapped to determine missing or distorted patterning within the *ruby* region

3.3.2.8 Immunostaining of mutant embryos

Immunostaining of embryos was carried out as indicated in section 2.2.

3.3.2.9 Singly fly genomic extraction and PCR

50uL of extraction buffer (10 mM Tris-HCl pH8 1 mM EDTA 25 mM NaCl 200 g/ml Proteinase K) was used to crush flies. The solution was incubated at 30°C for 30 minutes, then incubated at 80°C for 5 minutes. PCR analysis was used to determine if break point were present at each targeted site for Cas9-mutagenisis.

3.3.3. Results

3.3.3.1. Characterization of deletion screen offspring using the CRISPR-Cas9 and FRT/FLP system

Of the five established stocks made from putative *rb* mutants, identified as flies displaying the ruby eye phenotype, two failed to complement the null *hnt* allele. This suggests that the expected deletion was successfully induced using Cas9 multiplexing. The two mutant lines recovered are referred to as rb^{e027} *and* rb^{e038}. Given only 0.18% of flies from the screen carry putative deletions, this technique did not represent an efficient method for the recovery of deletions.

Using the FRT/FLP system, 0.91% of offspring displayed the ruby eye phenotype (Table 8). All lines recovered were subjected to a complementation test to confirm deletion of the *rb* gene. Unfortunately, all lines complemented the null *hnt* allele. The results of the screen carried out using WE and FLP/FRT SSR are summarized below.

Cross	phenotype of offspring eyes		
	Wild type	Mosaic	Ruby
1	17	42	0
2	0	32	1
3	0	71	1
4	2	63	0
5	45	21	0
6	12	79	0
7	28	0	0
8	1	63	1
9	0	62	0
10	23	46	0
11	0	51	1
12	5	92	0
13	1	73	1
14	24	58	0
15	2	62	0
16	11	20	0
17	28	58	0
Total	**199**	**893**	**5**

Table 7. Offspring of the Cas9-induced deletion screen. Offspring from the cross will have wild type or mosaic eyes if no deletion was induced by WE Cas9 multiplexing. If a deletion of the *rb* region was made, flies will have the ruby eye phenotype due to the deletion of *rb* from Cas9 multiplexing in combination with the rb^{1} allele. Offspring with ruby eyes were subjected to a complementation test to confirm the deletion of *rb*.

Cross	phenotype of offspring eyes		
	Wild type	White	Ruby
1	26	4	0
2	6	12	1
3	18	7	2
4	18	4	0
5	14	2	0
6	10	1	0
7	11	4	2
8	7	3	1
9	12	4	2
10	9	1	0
11	20	8	0
12	7	2	0
13	18	9	1
14	9	7	0
15	23	5	0
16	10	3	0
17	11	6	0
18	17	3	0
19	4	0	0
20	28	8	0
21	11	3	1
22	24	5	0
23	7	4	0
24	12	6	0
25	32	9	0
26	7	0	0
Total	**371**	**120**	**10**

Table 8. Offspring of the FRT/FLP-induced deletion screen. Offspring from the cross will have wild type of white eyes if no deletion was induced by FRT/FLP system. If a deletion of the *rb* region was made, flies will have the ruby eye phenotype due to the deletion of *rb* from Cas9 multiplexing in combination with the rb^{1} allele. Offspring with ruby eyes were subjected to a complementation test to confirm the deletion of *rb*.

3.3.3.2. Confirmation of deletions in chromosomes using polytene squashes and PCR

Polytene squash analysis indicated no cytologically visible deletions in the polytene chromosome in the rb^{027} and rb^{038} lines, both of which failed to complement a null *hnt* allele. Polytene analysis of flies with known deletions in the ruby region can be cytologically visualized (Figure 20 A and B). When visualizing the polytene chromosomes of the lines with a putative deletion, no irregular banding patterns were observed in rb^{027} (Figure 20, C) and rb^{038} (Figure 20, D). This is possibly due to the small size of the deletion (~180Kb), which many be insufficient for observation using this method. While a diagnostic PCR was attempted to confirm deletions, this method could not be carried out as the primers failed to successfully amplify two of the four diagnostic regions in the control fly stock (*w, e02388ab / w)*. While the inner bounds of the two P elements inserted in the control stock genome did amplify (including ~600 bp of endogenous sequence adjacent to the insert sequence), the primers used to confirm the region of the outer bounds of P elements did not support amplification. A diagnostic PCR using solely the two diagnostic regions cannot be used to determine if a deletion was induced, as PCR analysis of the outer bounds of P elements only provide a positive control.

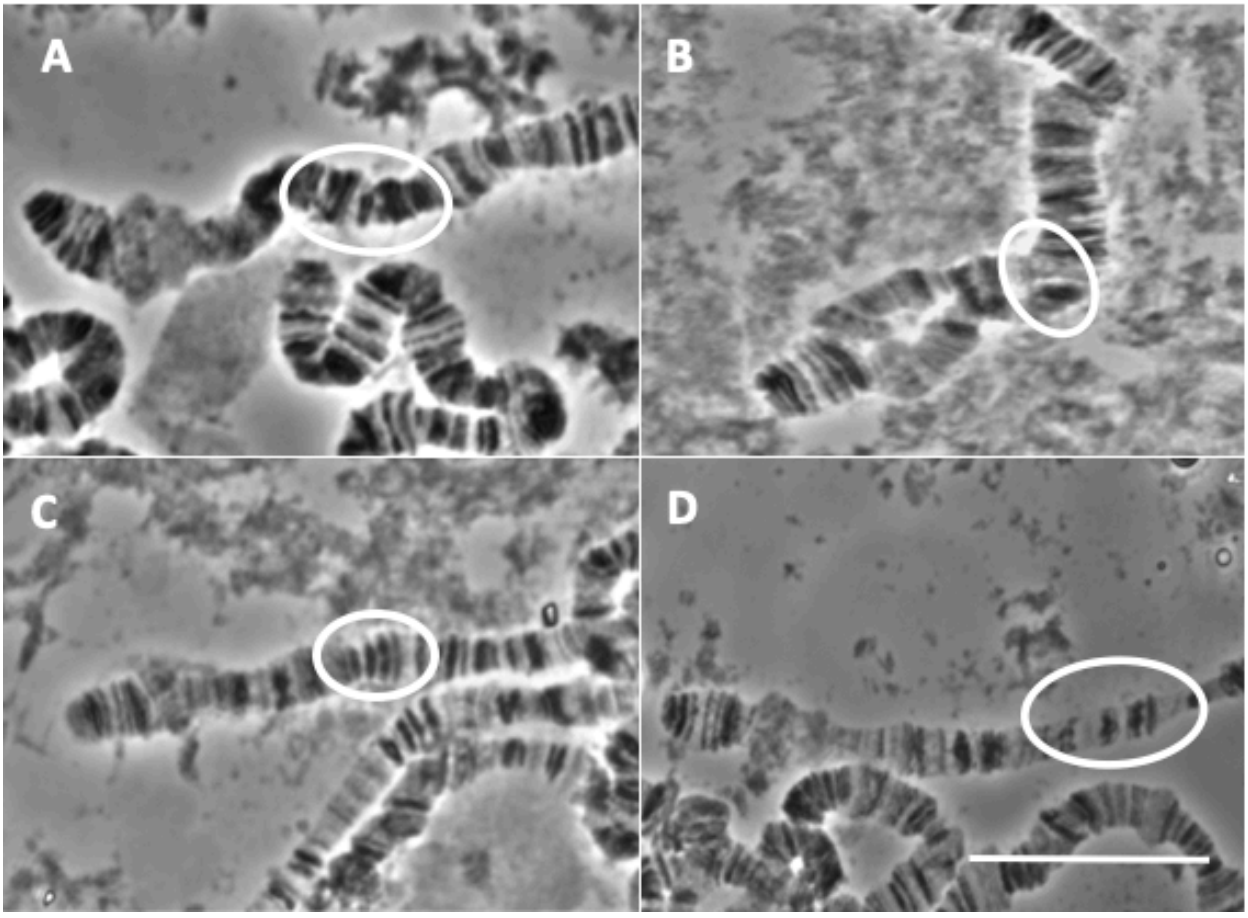


Figure 20. Polytene squash analysis of the *rb*^{027} and *rb*^{038} lines. The wild type chromosome has all bands in symmetrical lines between paired chromosomes (A). Established fly lines with deletions of this region, such as *rb*^{46}, demonstrate a cytological difference in their paired chromosomes of the polyene chromosome (B). Lines with a Cas9 induced deletion do not display cytological visible deletions, including *rb*^{027} (C) and *rb*^{038}(D). The banding patterns within the region corresponding to *rb* (circled) are symmetrical in the paired chromosome, however demonstrate slight curving of the parallel lines. Scale bar represents 20 μm.

3.3.3.3. Immunostaining of recovered Notch and hindsight mutants

The *hnt* expression in the rb^{027}*and* rb^{038} allele background demonstrated the Cas9-induced *rb* deletions were embryonic lethal. Compared to wild type embryos, mutant embryos displayed malformations. The embryos of the rb^{027} line showed incomplete formation of the midgut (Figure 21, B), as well as disorganized formation of oenocyte and oenocyte-associated cells (Figure 21, D). Expression of *hnt* in the wild type embryo is readily observed in the amnioserosa, BO, and midgut, only show weak expression of *hnt* in these embryos (Figure 21, B and D). During stages 7-9, expression of *hnt* in the rb^{e038} embryos showed faint expression in the early stage amnioserosa (Figure 20, B), which usually displays high expression of the protein (Figure 22, A). As development progresses, the mutant embryos display disorganized and asymmetrical formation of the embryonic segmentation during stage 11-13. During these stages, expression of *hnt* in oenocytes and oenocyte-associated cells, which is usually strong in wild type embryos, was very faint. No expression was observed in the BO, or the PNS of mutants (Figure 22, B and C). Mutated embryos also fail to undergo GBR during stage 12-14, resulting in the "tail up" phenotype (Figure 22, C and D). Expression of *hnt* in tracheal cells was more scattered than expected, as compared to the wild type embryo (Figure 22, B and D).

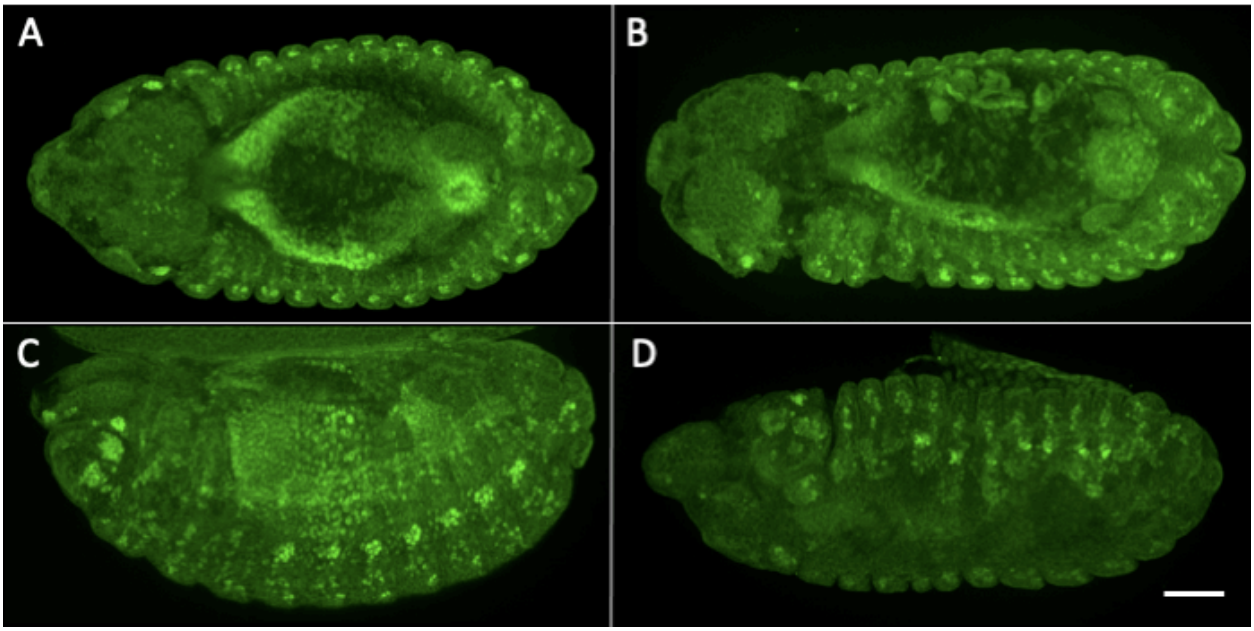


Figure 21. Immunostaining using anti-Hnt antibodies of the rb^{027} line. Embryos on the left display the wild type expression of *hnt* (A and C). Mutant embryos with the *rb^{027}* deletion fail to complete embryogenesis. Embryos are morphologically abnormal, and display irregular expression of Hnt (B and D). Scale bar represents 50 μm.

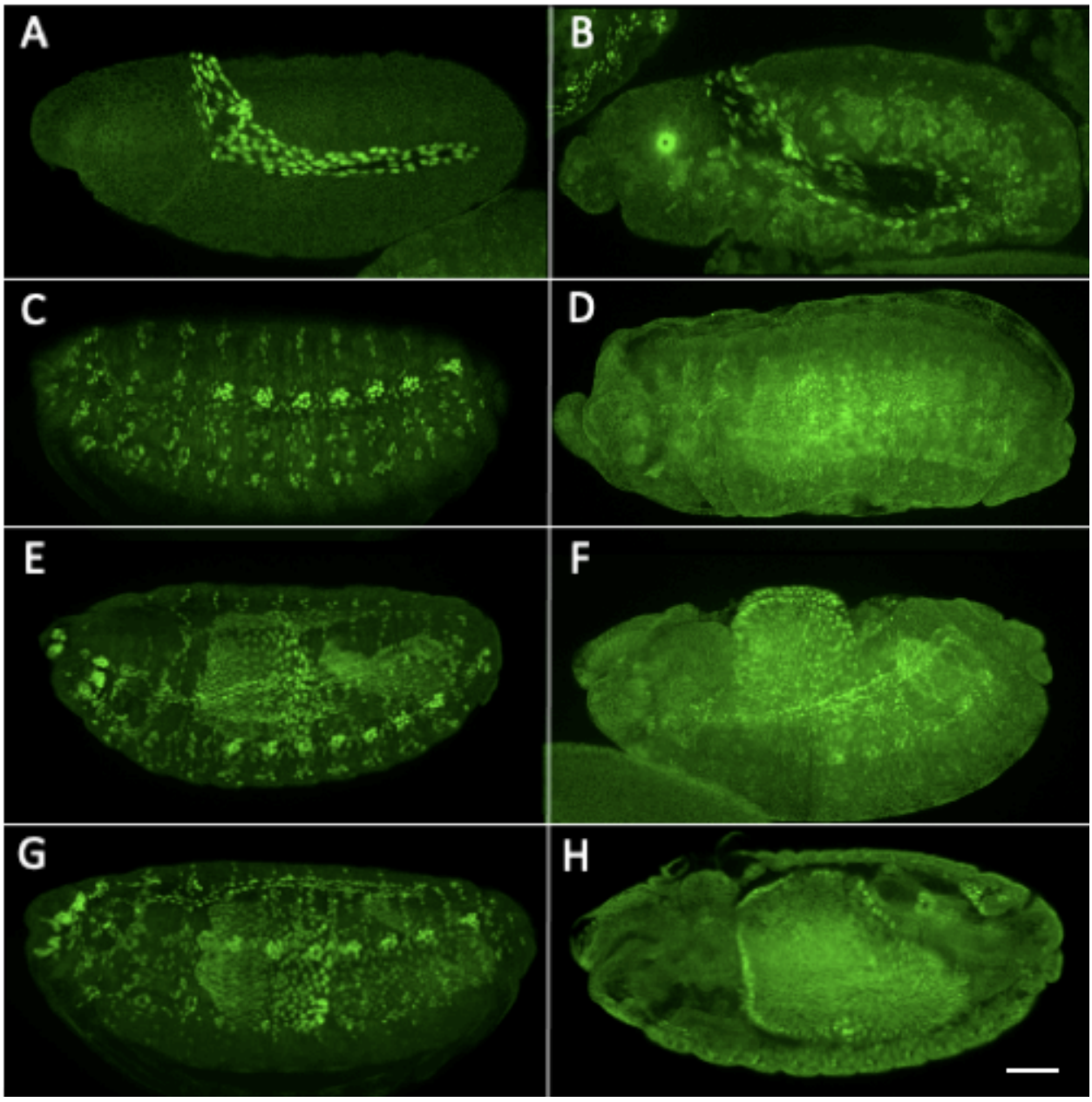


Figure 22. Immunostaining of rb^{038} line. Embryos on the left display the wild type expression of *hnt* (A, C, E, and G). Mutant embryos with the *rb^{038}* deletion fail to complete embryogenesis. Embryos are morphologically abnormal, and display irregular expression of Hnt (B, D, F and H). Scale bar represents 50 μm.

3.3.4 Discussion

Using the WE system, putative deletions were only recovered in 0.19% of offspring. It is important to note that while the complementation tests and immunostaining results are consistent with a deletion of *rb*, deletions will need to be subjected to diagnostic PCR to truly confirm the presence of the CR. This was not successfully carried out in the lab due to limitations in time, as the outer bounds of the WE inserts could not be successfully amplified. One possibility is that the sites for primer annealing have been changed due to knock-in of the WE insert in the lines used to perform the screen.

When investigating the use of SecI restriction enzyme as a technique to induce translocations and inversions, as described in the introduction, Egli et al (2004) recovered a high frequency of deletions as a product of NHEJ following DSBs. Additionally, it was found that translocations and inversions were recovered exclusively as a product of HR between the two homologous P-element constructs that housed the SecI recognition site (Egli, Hafen, & Schaffner, 2004). The SSR systems discussed also recovers inversions and translocations using HR between FRT recognition sites, however the SecI approach recovered inversions and translocations at a much higher frequency (Egli, Hafen, & Schaffner, 2004). When this method was investigated with the *rb* deletion screen using the FRT/FLP system, no offspring with a putative deletion was recovered. Since this screen was only carried out in a sample of ~500 flies, further tests using this method should be conducted to determine an accurate frequency of deletion mutagenesis with the proposed screen.

Inversions and translocations induced using the *nos-Cas9* system were designed based on the notion that CRs will be recovered due to the non-specific strand repair of the NHEJ pathway. Inversions and translocation screens can be repeated using the WE technique, which will induce

DSBs in regions of homology (the *mini-white* gene). In a study using Cas9 multiplexing to induce CRs in mammalian cells, inversions and deletions were recovered using microhomology-mediated end joining (MMEJ). This alternative end joining repair pathway uses repeats 4-25bp near the DSBs induced during multiplexing. Given the abundance of repetitive sequence in the eukaryotic genome, strategically targeting Cas9 to identical sequence repeats may provide a feasible method for inducing CRs without the need of transgenes (Li et al., 2015). However, repair of broken fragments may differ between various organisms. For example, no inversions were recovered as a result of NHEJ in *Drosophila* (Egli et al., 2004), while Cas9 efficiently induced inversions using NHEJ in mammalian cells (Li et al., 2015).

Several improvements to this screen can be considered for further experiments. Firstly, the screened ruby phenotype was difficult to select, largely due to the "deep pseudopupil" within the eye disc. This structure produced a colour in the eye disc makes the ruby phenotype difficult to select. Given only two of the five flies selected to have the ruby phenotype had a putative deletion in ruby, it is highly likely that the flies are almost indistinguishable – particularly for the "untrained" eye for fly phenotypes. To circumvent this issue, the screen can be redesigned so that the parental stain of the F1 cross harbours a rb^1 deficiency in the background of the Cherry-like eye phenotype. This phenotype results from a particular mutation of *white*, known as w^{CL}, which results in a pigment similar to the wild type phenotype. Under certain allelic combinations with mutations of *white* and w^{CL}, different pigments can be made. When testing this mutation in the background of *hnt* and *rb* deficiencies, offspring have a distinctive apricot eye phenotype. This provides a feasible method for screening offspring with putative deletions using phenotype.

Additionally, induced inversions resulting from the deletion screen could have produced false-negative results, since eye colour may be unaffected by the resulting inversion. A method

to detect inversions should be used in conjunction with the deletion screen to distinguish flies with an inversion from the wild type without a CR. This was also a consideration with the inversion screen, as a deletion may cause lethality resulting in false negative results. Creating such a method will complicate the screening strategies proposed to identify CRs, however it will provide a more accurate method for determining the frequency of deletions and inversions. In practice this may require designing an inversion that will result in a functional gene following the inversion event (Egli et al., 2004). This would require the insertion of the two selection gene halves on either side of two targeted breakpoints. If an inversion is induced, the marker gene is expressed. This results in an observable phenotype, which can confirm an inversion was made.

As alluded to, HR may be the preferred pathway to recover translocations and inversions. The WE screen offers one convenient approach for this, as an abundance of fly stocks with *mini-white* genes randomly inserted in the genome are readily available in collections. If more precise CRs are desired, the CRISPR/Cas9 system can be used to knock-in FRT or *mini-white* loci in desired locations. As a means to induce CRs more efficiently, the MMEJ pathway may supersede this approach. Sequences of homology that span 4-25 bp would be easily located in the *Drosophila* genome, and hence can provide a feasible method for inducing CRs depending on the recovered efficiency from this technique. Alternatively, templates can be designed when inducing Cas9 mutagenesis to promoter HDR with the template. Templates to induce a deletion would include nonprotein-coding fragments that can replace the targeted fragment, while templates to induce inversions would include the sequence of the targeted fragment inverted.

References

Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-04252-2

Arguello, J. R., Abuin, L., Armida, J., Mika, K., Chai, P. C., Benton, R., & Ramaswami, M. (2021). Targeted molecular profiling of rare olfactory sensory neurons identifies fate, wiring, and functional determinants. *ELife*, *10*, 1–33. https://doi.org/10.7554/eLife.63036

Aza-Blanc, P., & Kornberg, T. B. (1999). Ci: A complex transducer of the Hedgehog signal. *Trends in Genetics*, *15*(11), 458–462. https://doi.org/10.1016/S0168-9525(99)01869-7

Baechler, B. L., McKnight, C., Pruchnicki, P. C., Biro, N. A., & Reed, B. H. (2016). Hindsight/RREB-1 functions in both the specification and differentiation of stem cells in the adult midgut of Drosophila. *Biology Open*, *5*(1), 1–10. https://doi.org/10.1242/bio.015636

Baonza, A., & Garcia-Bellido, A. (1999). Notch signaling directly controls cell proliferation in the Drosophila wing disc. Retrieved from www.pnas.org

Bellen, H. J., O'Kane, C. J., Wilson, C., Grossniklaus, U., Pearson, R. K., & Gehring, W. J. (1989). P-element-mediated enhancer detection: a versatile method to study development in Drosophila. *Genes & Development*, *3*(9), 1288–1300. https://doi.org/10.1101/gad.3.9.1288

Carlberg, C., & Molnár, F. (2014). *Mechanisms of gene regulation*. *Mechanisms of Gene Regulation*. Springer Netherlands. https://doi.org/10.1007/978-94-007-7905-1

Chai, P. C., Cruchet, S., Wigger, L., & Benton, R. (2019). Sensory neuron lineage mapping and manipulation in the Drosophila olfactory system. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-08345-4

Crozatier, M., & Vincent, A. (2008). Control of multidendritic neuron differentiation in Drosophila: The role of Collier. *Developmental Biology*, *315*(1), 232–242. https://doi.org/10.1016/j.ydbio.2007.12.030

Daniel, B., Nagy, G., & Nagy, L. (2014). The intriguing complexities of mammalian gene regulation: How to link enhancers to regulated genes. Are we there yet? *EEBS Letters*. https://doi.org/10.1016/j.febslet.2014.05.041

Deady, L. D., Li, W., & Sun, J. (2017). The zinc-finger transcription factor hindsight regulates ovulation competency of Drosophila follicles. *ELife*, *6*, 1–22. https://doi.org/10.7554/eLife.29887

Egli, D., Hafen, E., & Schaffner, W. (2004). An efficient method to generate chromosal rearrangements by targeted DNA double-strand breaks in Drosophila melanogaster. *Genome Research*, *14*(7), 1382–1393. https://doi.org/10.1101/gr.2279804

FlyBase. (2020). Retrieved April 8, 2020, from https://flybase.org/reports/FBgn0004647

Garner, M. M., & Revzin, A. (1981). A gel electrophoresis mothod for quantifying the binding of proteins to specific DNA regions. *Nucleic Acids Research*, *9*(13), 3047–3060.

Geertz, M., & Maerkl, S. J. (2010). Experimental strategies for studying transcription factor^DNA binding specificities, *9*, 362–373. https://doi.org/10.1093/bfgp/elq023

Ghosh, S., Singh, A., Mandal, S., & Mandal, L. (2015). Active Hematopoietic Hubs in Drosophila Adults Generate Hemocytes and Contribute to Immune Response. *Developmental Cell*, *33*(4), 478–488. https://doi.org/10.1016/j.devcel.2015.03.014

Gould, A., Elstob, P., & Veronique, B. (2001). Insect oenocytes: a model system for studying cell-fate spesification by Hox genes. *Journal of Anatomy*, *199*, 25–33.

Grant, P., Maga, T., Loshakov, A., Singhal, R., Wali, A., Nwankwo, J., … Johnson, D. (2016). An eye on trafficking genes: Identification of four eye color mutations in drosophila. *G3: Genes, Genomes, Genetics*, *6*(10), 3185–3196. https://doi.org/10.1534/G3.116.032508

Harvard Medical School. (2022). Retrieved from https://perrimon.med.harvard.edu/stock-lists
Hoppe, C., Bowles, J. R., Minchington, T. G., Sutcliffe, C., Upadhyai, P., Rattray, M., & Ashe, H. L. (2020). Modulation of the Promoter Activation Rate Dictates the Transcriptional Response to Graded BMP Signaling Levels in the Drosophila Embryo. *Developmental Cell*, *54*(6), 727-741.e7. https://doi.org/10.1016/j.devcel.2020.07.007
Joyce, E. F., Apostolopoulos, N., Beliveau, B. J., & Wu, C. -tin. (2013). Germline Progenitors Escape the Widespread Phenomenon of Homolog Pairing during Drosophila Development. *PLoS Genetics*, *9*(12), 1004013. https://doi.org/10.1371/JOURNAL.PGEN.1004013
Kim, M., Du, O. Y., Whitney, R. J., Wilk, R., Hu, J., Krause, H. M., … Reed, B. H. (2020). A Functional Analysis of the Drosophila Gene hindsight: Evidence for Positive Regulation of EGFR Signaling. *Genes, Genomes, Genetics*, *10*(1), 117–127. https://doi.org/10.1534/g3.119.400829
Kondo, S., & Ueda, R. (2013). Highly Improved Gene Targeting by Germline-Specific Cas9 Expression in Drosophila. *Genetics*, *195*(3), 715. https://doi.org/10.1534/GENETICS.113.156737
Krejci, A., F., Bernard, B. E., Housden, S. C., & Bray, S. . (2009). Direct response to Notch activation: signaling crosstalk and incoherent logic. *Science Signalling*, *2*(55). https://doi.org/Https://doi.org/10.1128/MCB.24.2.573-583.2004, 24: 573–583.
Kretzschmar, D., Poeck, B., Roth, H., Ernst, R., Keller, A., Porsch, M., … Pflugfelder, G. O. (2000). *Defective Pigment Granule Biogenesis and Aberrant Behavior Caused by Mutations in the Drosophila AP-3 Adaptin Gene ruby*. Retrieved from https://academic.oup.com/genetics/article/155/1/213/6047931
Kvon, E. Z., Kazmar, T., Stampfel, G., Yáñez-Cuna, J. O., Pagani, M., Schernhuber, K., … Stark, A. (2014). Genome-scale functional characterization of Drosophila developmental enhancers in vivo. *Nature*, *512*(1), 91–95. https://doi.org/10.1038/nature13395
Lamka, M. L., & Lipshitz, H. D. (1999). Role of the amnioserosa in germ band retraction of the Drosophila melanogaster embryo. *Developmental Biology*, *214*(1), 102–112. https://doi.org/10.1006/dbio.1999.9409
Li, J., Shou, J., Guo, Y., Tang, Y., Wu, Y., Jia, Z., … Wu, Q. (2015). Efficient inversions and duplications of mammalian regulatory DNA elements and gene clusters by CRISPR/Cas9. *Journal of Molecular Cell Biology*, *7*(4), 284–298. https://doi.org/10.1093/JMCB/MJV016
Li, W., Li, S., Zheng, H., Zhang, S., & Xue, L. (2012). A broad expression profile of the GMR-GAL4 driver in Drosophila melanogaster. *Genetics and Molecular Research*, *11*(3), 1997–2002. https://doi.org/10.4238/2012.August.6.4
Liu, C.-H., Chen, Z., Oliva, M. K., Luo, J., Collier, S., Montell, C., & Hardie, R. C. (2020). Rapid Release of Ca 21 from Endoplasmic Reticulum Mediated by Na 1 /Ca 21 Exchange. https://doi.org/10.1523/JNEUROSCI.2675-19.2020
Locke, J., & Tartof, K. D. (1994). Molecular analysis of cubitus interruptus (ci) mutations suggests an explanation for the unusual ci position effects. *MGG Molecular & General Genetics*, *243*(2), 234–243. https://doi.org/10.1007/BF00280321
McGuire, S. E., Roman, G., & Davis, R. L. (2004). Gene expression systems in Drosophila: a synthesis of time and space. *Trends in Genetics*, *20*(8), 384–391. https://doi.org/10.1016/J.TIG.2004.06.012
Meister, M., & Lagueux, M. (2003). Drosophila blood cells. *Cellular Microbiology*, *5*(9), 573–580. https://doi.org/10.1046/j.1462-5822.2003.00302.x
Melani, M., Simpson, K. J., Brugge, J. S., & Montell, D. (2008). Regulation of Cell Adhesion

and Collective Cell Migration by Hindsight and Its Human Homolog RREB1. *Current Biology*, *18*(7), 532–537. https://doi.org/10.1016/j.cub.2008.03.024
Messina, D. N., Glasscock, J., Gish, W., & Lovett, M. (2004). An ORFeome-based analysis of human transcription factor genes and the construction of a microarray to interrogate their expression. *Genome Research*, *14*(10 B), 2041–2047. https://doi.org/10.1101/gr.2584104
Ming, L., Wilk, R., Reed, B. H., & Lipshitz, H. D. (2013). Drosophila Hindsight and mammalian RREB-1 are evolutionarily conserved DNA-binding transcriptional attenuators. *Differentiation*, *86*(4–5), 159–170. https://doi.org/10.1016/j.diff.2013.12.001
Mohr, O. L. (1919). *Character Changes Caused By Mutation of an Entire Region of a Chromosome in Drosophila*. *Genetics* (Vol. 4). https://doi.org/10.1093/genetics/4.3.275
Mora, A., Sandve, G. K., Gabrielsen, O. S., & Eskeland, R. (2016). In the loop: promoter-enhancer interactions and bioinformatics. *Briefings in Bioinformatics*, *17*(6), 980–995. https://doi.org/10.1093/bib/bbv097
Morrill, S. A., & Amon, A. (2019). Why haploinsufficiency persists, *116*(24), 11866–11871. https://doi.org/10.1073/pnas.1900437116
Ng, C., Qian, Y., & Schulz, C. (2019). Notch and Delta are required for survival of the germline stem cell lineage in testes of Drosophila melanogaster. *PLoS ONE*, *14*(9). https://doi.org/10.1371/journal.pone.0222471
Ng, W. A., Ma, A., Chen, M., & Reed, B. H. (2019). A method for rapid selection of randomly induced mutations in a gene of interest using CRISPR/Cas9 mediated activation of gene expression. *BioRxiv*, 788372. https://doi.org/10.1101/788372
Oliva, C., Molina-Fernandez, C., Maureira, M., Candia, N., ıa opez, E. L., Hassan, B., … Sierralta, J. (2015). Hindsight Regulates Photoreceptor Axon Targeting Through Transcriptional Control of jitterbug/Filamin and Multiple Genes involved in Axon Guidance in Drosophila. *Wiley Periodicals, Inc*, *75*, 1018–1032. https://doi.org/10.1002/dneu.22271
Patterson, J., Stone, W., Bedichek, S., & Suchte, M. (1934). The production of translocations in Drosophila. *American Society of Naturalists* , 361. Retrieved from https://www.jstor.org/stable/2456935
Qian', S., & Pirrotta, V. (1995). *Dosage Compensation of the Drosophila white Gene Requires Both the X Chromosome Environment and Multiple Intragenic Elements*. Retrieved from https://academic.oup.com/genetics/article/139/2/733/6013211
Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), 2281–2308. https://doi.org/10.1038/nprot.2013.143
Reed, B., Wilk, R., & Lipshitz, H. D. (2001). Downregulation of Jun kinase signaling in the amnioserosa is essential for dorsal closure of the Drosophila embryo. *Current Biology*, *11*(14), 1098–1108. https://doi.org/10.1016/S0960-9822(01)00318-9
Riethoven, J.-J. M. (2010). Regulatory Regions in DNA: Promoters, Enhancers, Silencers, and Insulators. In *Regulatory Regions in DNA: Promoters, Enhancers,Silencers, and Insulators* (pp. 33–40). https://doi.org/10.1007/978-1-60761-854-6_3
Rodriguez-Trelles, F., Tarríol, R., & Ayala, F. J. (2003). Evolution of cis-regulatory regions versus codifying regions. *International Journal of Developmental Biology*, *47*(7–8), 665–673. https://doi.org/10.1387/ijdb.14756342
Singh, A., & Kango-Singh, M. (2013). *Molecular Genetics of Axial Patterning, Growth and Disease in the Drosophila Eye*. Springer. Retrieved from https://link-springer-com.proxy.lib.uwaterloo.ca/content/pdf/10.1007%2F978-1-4614-8232-1.pdf

Sturtevant, A. H., & Beadle, G. W. (1936). the Relations of Inversions in the X Chromosome of Drosophila Melanogaster To Crossing Over and Disjunction. *Genetics*, *21*(5), 554–604. https://doi.org/10.1093/genetics/21.5.554

Sun, J., & Deng, W. M. (2007). Hindsight Mediates the Role of Notch in Suppressing Hedgehog Signaling and Cell Proliferation. *Developmental Cell*, *12*(3), 431–442. https://doi.org/10.1016/j.devcel.2007.02.003

Terriente-Felix, A., Li, J., Collins, S., Mulligan, A., Reekie, I., Bernard, F., … Bray, S. (2013). Notch cooperates with Lozenge/Runx to lock haemocytesinto a differentiation programme. https://doi.org/10.1242/dev.086785

Vienna BioCentre Core Facilies. (n.d.). Retrieved January 17, 2022, from https://www.viennabiocenter.org/vbcf/vienna-drosophila-resource-center/enhancer-gal4-driver-library/

Wilk, R., Pickup, A., Hamilton, J., Reed, B., & Lip, H. (2004). Dose-Sensitive Autosomal Modifiers Identify Candidate Genes for TissueAutonomous and Tissue Nonautonomous Regulation by the DrosophilaNuclear Zinc-Finger Protein, Hindsight. *Genetics Aociety of America*, 281–300.

Wilk, R., Reed, B. H., Tepass, U., & Lipshitz, H. D. (2000). The hindsight gene is required for epithelial maintenance and differentiation of the tracheal system in Drosophila. *Developmental Biology*, *219*(2), 183–196. https://doi.org/10.1006/dbio.2000.9619

Yip, M. L. R., Lamka, M. L., & Lipshitz, H. D. (1997). Control of germ-band retraction in Drosophila by the zinc-finger protein Hindsight. *Development*, *124*(11), 2129–2141. https://doi.org/10.1242/dev.124.11.2129

Zhu, L. J., Christensen, R. G., Kazemian, M., Hull, C. J., Enuameh, M. S., Basciotta, M. D., … Brodsky, M. H. (2011). FlyFactorSurvey: A database of Drosophila transcription factor binding specificities determined using the bacterial one-hybrid system. *Nucleic Acids Research*, *39*, 111–117. https://doi.org/10.1093/nar/gkq858

Appendices

Appendix A: JASPAR results

ID	Genomic Location	Strand orientation	Sequence
MA0535.1	X:4611500..4613999	+	ccgacgccgctgtat
MA0535.1	X:4611500..4613999	-	caatcgccgccactg
MA0535.1	X:4611500..4613999	-	gtggcgctgaggggc
MA0535.1	X:4611500..4613999	+	gcggcgcgtgcgcta
MA0535.1	X:4611500..4613999	+	gtgtcgtcgcagtgc
MA0535.1	X:4611500..4613999	+	ccggagccaaggcca
MA0535.1	X:4611500..4613999	+	cagactccgatgatg
MA0535.1	X:4611500..4613999	+	tcagcgccaccgcgc
MA0535.1	X:4611500..4613999	-	cacgcgccgcgtttc
MA0535.1	X:4611500..4613999	-	gcgacgacacccgaa
MA0535.1	X:4611500..4613999	+	ttcccgacgccgctg

Table A.1. Putative TFBS of Mad to the VT056875 construct region. Putative Mad TFBS within the VT056875 construct sequence. Analysis searched on JASPAR core. Sequence fastas obtained on Flybase using VT line.

ID	Genomic Location	Strand orientation	Sequence
MA0085.1	X:4596000..4598699	-	tcgttagaaaccaatt

Table A2. Putative TFBS of Su(H) to the VT056866 construct region Putative Su(H) TFBS within the VT056866 construct sequence. Analysis searched on JASPAR core. Sequence fastas obtained on Flybase using VT line.

ID	Genomic Location	Strand orientation	Sequence
MA0085.1	X:4613600..4615699	-	tcgttagaaaccaatt

Table A3. Putative TFBS of Su(H) to the VT056876 construct region. Putative Su(H) TFBS within the VT056876 construct sequence. Analysis searched on JASPAR core. Sequence fastas obtained on Flybase using VT line.

Appendix B: Screening Translocations using Position Effects

B.1 Introduction

In addition to the incredible number of phenotypes associated disrupting a gene's sequence, *Drosophila* has a range of unique phenomena that can be exploited for phenotypic screens. For example, changing gene location can also change gene activity. One such effect includes a process known as transvection. Generally, this process refers to a homolog-pairing event in which a gene's transcription can be altered *in trans* by the regulatory region of its homologous gene. The gene *cubitus interruptis* (*ci*) displays transvection, a phenomenon first dubbed the *Dubinin Effect*. When translocations were associated with ci^{+} (T(*ci*)/ci^{+}), a wild type phenotype is observed. However, if the translocation is paired with a mutated copy of *ci,* such as ci^{1} (T(*ci*)/ci^{1}), the *ci* mutation acts as a dominant due to lack of regulation for the mutant's homolog (ci^{1}). The position effects of *ci* associated with transveciton will be used to identify translocations in the following screen. Ci expression is vital for the proper formation of the anterior and posterior developmental compartments during embryonic stages, in which the wing imaginal disc develops from. This is due to the important role it has in *hh* repression and Hh signalling activation. The *ci* mutants display an interrupted or absent wing vein phenotype due to the misexpression of Ci within the posterior compartment (Aza-Blanc & Kornberg, 1999). The inappropriate expression of Ci observed by the ci^{1} allele is recessive, but the phenotype is observed as a dominant when a CR displaces the genes homolog (Locke & Tartof, 1994). The basis of the proposed screen identifies the translocation event by recreating the R(ci^{+})/ci^{1} genotype, in which the ci^{1} displays a dominant phenotype. The regions targeted for DSBs to induce this translocation are in the proximal region of *ci,* and the regulatory region of *hindsight.* The screen is described below.

B.2. Methods

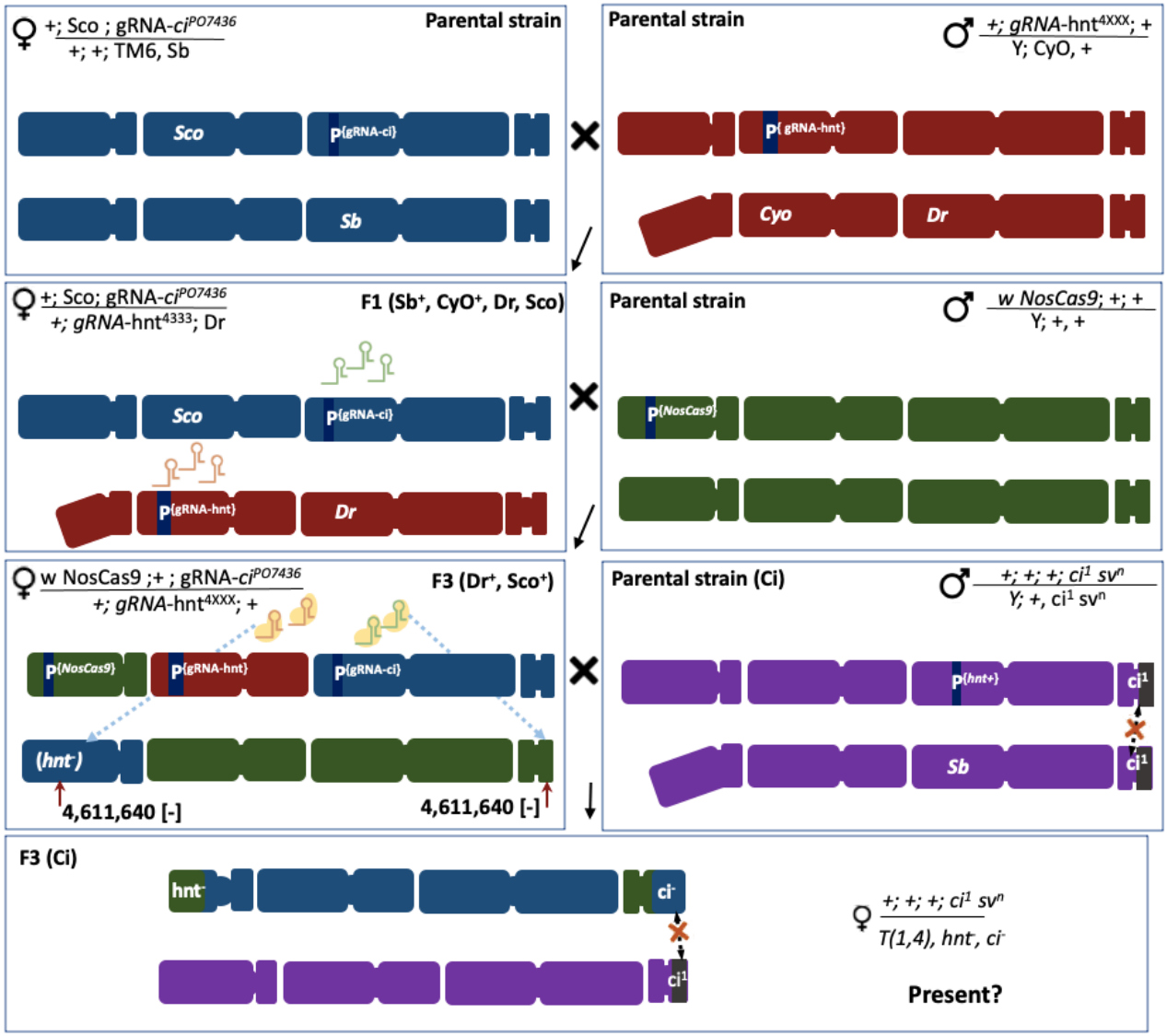


Figure B.1. Crossing of transgenic lines to induce and screen a translocation. The F2 offspring will have Cas9-sgRNA activity within germ cells. The first DSB is targeted to the regulatory region of *hnt,* the second DSB is targeted within *ci*. The arms of chromosomes 1 and 4 will break off proximal to these target sites, indicated by the red arrows (with genomic location). These are crossed to a ci^{1} homozygous stock, where transvection cannot occur between the two *ci* alleles. Due to lack of transvection, the flies display the Ci phenotype. The F3 that inherit the ci^{1} allele and a 4th chromosome with a translocation event (ci^{-}) cannot have transvection between it's 4th chromosome pair. The result is a F3 fly that displays the Ci phenotype, which indicated a translocation event has successfully occurred.

B.3 Results

Using both the female and male germline for mutagenesis, no offspring from the proposed translocation screen displayed the Ci^{l} phenotype. This indicates that no translocations were induced by the multiplexed mutation. Table C.1 and C.2 shows the counts for offspring from the proposed cross. Additionally, no translocations were visible from eight random flies selected for polytene squash analysis.

Cross	Germline of *NosCas9*	$C1^{+}$ phenotype	Ci phenotype
1	male	216	0
2	male	98	0
3	male	258	0
4	male	126	0
5	male	127	0
6	male	105	0
7	male	95	0
8	male	173	0
9	male	240	0
10	male	217	0
11	male	220	0
12	male	196	0
13	male	190	0
14	male	142	0
15	male	155	0
16	male	187	0
17	male	223	0
18	male	22	0
19	male	236	0
20	male	126	0

Table B.2. Ci phenotype in offspring of the translocation screen using Cas9 mutagenesis in the male germline
Each offspring will inherit a fourth chromosome that has been targeted for Cas9 mutagenesis within the parent's germline. A successful translocation event will result in the Ci^{1} phenotype, as the inherited chromosome will not be able to able to regulate it's homolgous *ci* gene *in cis* due to the position effects. If no translocation event occurs, *ci* will have wild type regulation and no

phenotype will occur. The Ci phenotype is characterized as interrupted or absent vein growth on the fly wing.

Cross	Germline of *NosCas9*	C1+ phenotype	Ci- phenotype
1	female	254	0
2	female	17	0
3	female	184	0
4	female	142	0
5	female	242	0
6	female	153	0
7	female	134	0
8	female	201	0
9	female	202	0
10	female	61	0
11	female	119	0
12	female	96	0
13	female	151	0
14	female	163	0
15	female	217	0
16	female	232	0
17	female	59	0
18	female	173	0
19	female	211	0
20	female	212	0
21	female	67	0
22	female	139	0
23	female	82	0

Table B.3. Ci phenotype in offspring of the translocation screen using Cas9 mutagenesis in the female germline. Each offspring will inherit a fourth chromosome that has been targeted for Cas9 mutagenesis within the parent's germline. A successful translocation event will result in the Ci[1] phenotype, as the inherited chromosome will not be able to able to regulate it's homologous *ci* gene *in cis* due to the position effects. If no translocation event occurs, *ci* will have wild type regulation and no phenotype will occur. The Ci phenotype is characterized as interrupted or absent vein growth on the fly wing.

B.4 Discussion

Of all crosses, a single offspring displayed the ci phenotype indicating a lack of *ci* regulation that may result from the proposed translocation event. Analysis of the polytene chromosome concluded that the fly had wild type chromosome rearrangement. In the case of this screen, where the targeted breakpoint disrupts the regulation activity of *ci*, no observable translocation event means the phenotype is a result of two *ci* mutations *in trans*. If the parental cross was not properly cleared from vials when conducting this screen, this may have been a source of this false positive. Alternatively, it is possible that the *hnt* allele generated by the *GPO7436* sgRNA is not sufficient for *in trans* regulation of the ci^1 allele, resulting in lack of gene regulation. This would result in the ci phenotype, despite no translocation events.

Given the results, it can be concluded that multiplexing with Cas9 does not offer an efficient method to induce translocation rearrangements when compared to current methods. However, the choice of targets to induce rearrangements may have resulted in translocations being undetectable in the proposed screen. Gametes harbouring the translocation could be "chromosomally unbalanced", resulting in no embryonic development following fertilization. Even if the gametes are balanced with the translocation, the degree of the disruption may not support embryonic development. In either case, events of a translocation produce observable offspring. Divisions of *ci* from its 5' region as the result of a translocation has shown to have a wild type phenotype ($R(ci)/ci^+$) (Locke & Tartof, 1994). The sgRNA-6PO7436 is targeted over 1kbp upstream of the *ci* locus, within a long ncRNA region. If the proposed translocation was inviable, it would be likely due to the targeting of the distal end of the X chromosome. Translocations with the fourth and X chromosome have been recovered in which the X chromosome breakpoint was close to previously observed translocations. In these studies, the

break point was situated between *rb,* ~80kbp upstream of the sgRNA-hnt^{4344} target site, and *crossveinless (cv),* ~1Mbp upstream of the target site (Patterson, Stone, Bedichek, & Suchte, 1934). The sgRNA-hnt^{4344} should be replaced with an exact target site known to result in a viable translocation of the X chromosome. Especially since translocations involving the upstream region of the *ci* locus only include distal part of other autosomes (Locke & Tartof, 1994).

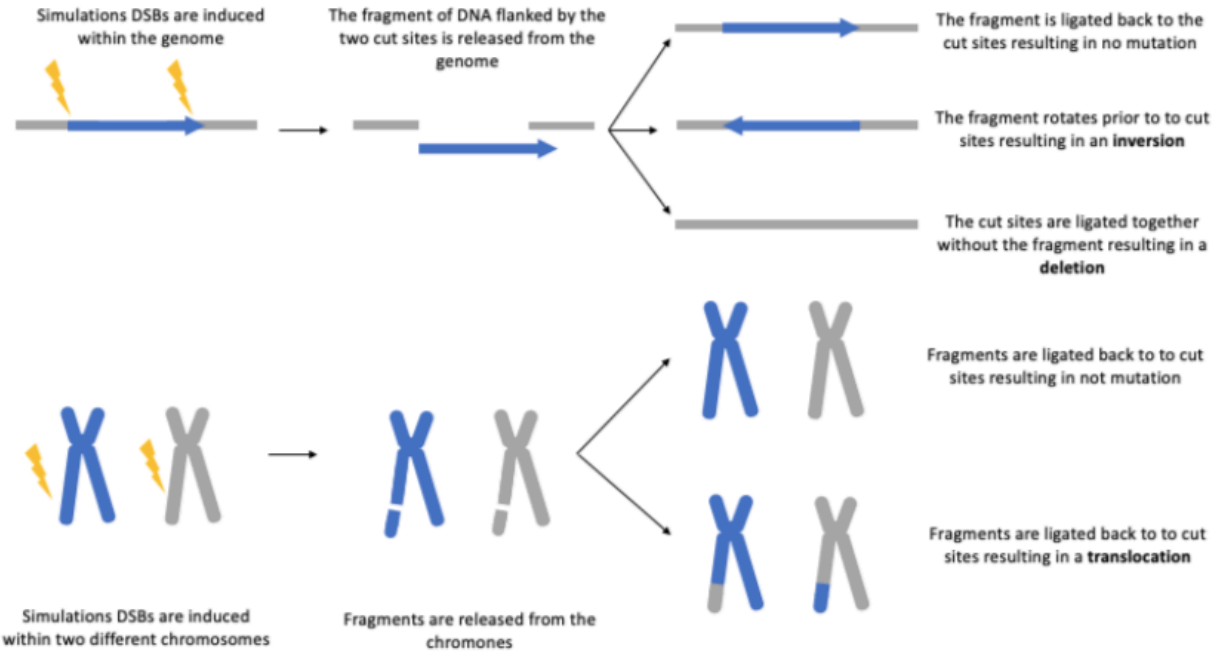
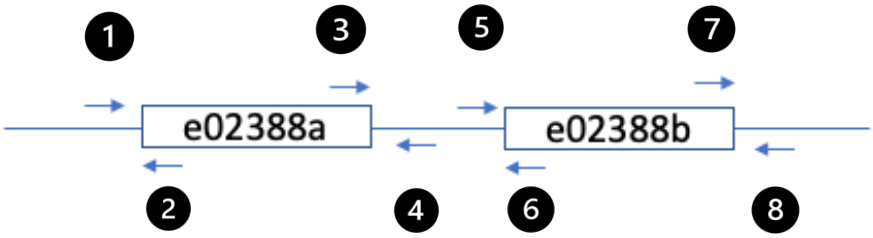


Figure C.1: Inducing chromosomal rearrangements using multiplexed mutations.



Primer	Sequence
1	CGACAATGCCTGACGAATGTG
2	AGCCTGACGTCATCGTTTATGC
3	TGGACGTCATCTTCACTTACGTG
4	AGATCTGAAATCACAGGATGCTG
5	CCATCTTCGCAAGAGAGGAATC
6	AGCCTGACGTCATCGTTTATGC
7	TGGACGTCATCTTCACTTACGTG
8	GGATCGTCATCATCATCGGAG

	PCR Sample			
	1	2	3	4
Wild type	✓	✓	✓	✓
Deletion	✓	✗	✗	✓

Figure C.2. Diagnostic PCR used to confirm deletion. Primers (demonstrated by the blue arrows) flank the *mini-white* inserts. The sequence of these inserts are targeted for a Cas9-induced deletion. The sequence of these primers is shown in the table. The wild type stock, with the e02388a and e02388b inserts, will produce four PCR products. If a deletion is present, the inner bounds of the two inserts will not anneal to primers. Therefore, a genome with a deletion will produce 2 PCR products.