Investigating the Use of CRISPR/Cas9 for Genetic Screens in Drosophila

by

William Ng

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

Waterloo, Ontario, Canada, 2021

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 made electronically available to the public.

Abstract

The CRISPR/Cas9 system has been adapted into a powerful tool for genetic research. The ability to target regions of a genome in a highly specific manner with a reprogrammable guide RNA has been used in a variety of applications. These applications include targeted double stranded breaks for inducing mutations or gene knock-ins, as well as hijacking the targeting ability of CRISPR/Cas9 to guide proteins to a defined sequence within the genome. These methods have been established in multiple model organisms including Drosophila. Since the establishment of the Drosophila model organism in 1909, methods have been discovered and refined to allow researchers to lay the groundwork for our understanding of eukaryotic genetics. This thesis evaluates two methods that combine previously established techniques found in the Drosophila toolbox with the CRISPR/Cas9 system and its modifications. The first is an F1 visible screen for loss-of-function mutations utilizing overexpression methods adapted from CRISPR/Cas9 technology. The second method attempts to evaluate the efficiency of creating translocations with CRISPR/Cas9 using a phenomenon called the “Dubinin Effect” which was first characterized in 1934.

Acknowledgements

I'd like to thank Andrew Ma, your establishment of the initial refractory stock made my project possible. I'd also like to think Molly Chen, whose work on creating the mutant allelic series (Figure 2-5) and immunostaining of embryos (Figure 2-6) contributed key data to the project.

Genetic lines that house the guides for the *ci* gene (Table3-1) were obtained in collaboration with the Perrimon Lab at Harvard Medical School.

To my supervisor Dr. Bruce Reed, thank you for the opportunity to work in the lab. I have gained a tremendous amount of experience through your guidance throughout my apprenticeship with you and I hope to take the skills that I've learned to my future endeavours. I'd like to thank the masters students that I've had the pleasure of spending time with, Donya Saghattchi and Brittney Lato. Sharing lab space with the both of you has made this journey fun and exciting. To all the undergraduate thesis students who I've spent time with in the lab; Molly Chen, Tristan Woo, Karolina Zarichna and Emily Baker, thanks for being awesome.

I owe a big thank you to Dr. Bernard Dunker and Dr. Barbara Moffatt as committee members your support is greatly appreciated. I also have to thank Dr. Barb Katzenback, who took a gamble on an undergraduate thesis student ultimately resulting in me gaining the initial skills to get as far as I have.

Finally, I want to thank my family for their unconditional love and support.

Table of Contents

List of Figures

List of Tables

Chapter 1
Background development of CRISPR/Cas9

1.1 Drosophila as a Model Genetic Organism

The species *Drosophila melanogaster* has been established as a key model organism for researching eukaryotic genetics. Since 1909 Drosophila have been used for their high fecundity and low cost. This organism has become a cornerstone of modern genetics (Hales et al., 2015). It takes roughly 10 days for an egg to develop into a mature adult fly (Hales et al., 2015). The developing embryo is an excellent model for developmental biology. Many of the processes typical of generalized body development; segmentation, specification of body axes are readily observed in Drosophila (Duffy, 2002). Moreover, using the genetic toolbox available for Drosophila, mutations disrupting these processes can be isolated and maintained as balanced stocks (Hales et al., 2015). In the late 1980's it was the ability to isolate mutations, and ultimately map, clone and sequence the genes associated with particular mutants that led to a major breakthrough in the field of animal development (Hales et al., 2015).

After the Drosophila genome was sequenced in 1999 and compared the sequenced human genome in 2000, it was estimated that 75% of known human diseases genes have orthologues in Drosophila (Pandey & Nichols, 2011). This makes Drosophila a great model organism for modeling the function of these human disease genes. At this point the power of Drosophila as a model organism is due in part to the history of Drosophila genetics. Drosophila has a storied history of contributions to modern genetics. In addition to the proof of the chromosomal theory of inheritance (Bridges, 1919), it was through Drosophila that we learnt that X-rays are mutagenic (H. Muller, 1928). More recently HOX genes which are the main genes that control body segment, in all metazoans were discovered in Drosophila (McGinnis et al., 1984).

Throughout the years different technologies have been implemented in Drosophila to further our understanding of eukaryotes. Earliest technologies included random mutagenesis to determine gene function based on phenotype (H. Muller, 1928). Drosophila geneticists also benefitted greatly from a high-

resolution physical map of the genome in the form of the salivary gland polytene chromosomes (Zhimulev & Koryakov, 2009). These chromosomes are formed through endoreduplication cycles of repeated DNA replication without mitosis, and up to 1000 copies of interphase chromatids remain aligned forming a highly reproducible banding pattern (Zhimulev & Koryakov, 2009). In addition, these polytene chromosomes are paired with their homologues, which enables easy identification of large scale chromosomal rearrangements (Zhimulev & Koryakov, 2009).

In addition to these polytene chromosomes, one of most powerful genetic tools used in Drosophila are balancer chromosomes. Balancer chromosomes are ultimately useful for the maintenance of mutations as heterozygotes (Hentges & Justice, 2004). Balancers achieve this by containing multiple inversions and these inversions have two effects (Hentges & Justice, 2004). First, they disrupt pairing of homologues during meiosis and supress meiotic recombination (Hentges & Justice, 2004). Second, if there is successful pairing and a meiotic recombination event, which invariably will occur within an inverted region of a paracentric or pericentric inversion, the products will either result in dicentric or acentric chromosome fragments and/or be excessively aneuploid (Hentges & Justice, 2004). Through the suppression of meiotic recombination, a chromosome carrying a recessive lethal is not lost through recombination with the homologue in the case where the balancer chromosome carries a second unlinked recessive lethal mutation (Hentges & Justice, 2004). These chromosomes also carry a dominant visible marker that allows the chromosome to be followed within a population. By using balancer chromosomes, any mutation, including recessive lethal mutations, can be maintained as a heterozygous stock. This facilitates further genetic and phenotypic analysis. Indeed, it is noteworthy that the famous saturation mutagenesis screens for embryonic lethal mutations performed in the 1980's by Christiana Nusslein-Volhard and Eric Weischaus would not have been possible without balancer chromosomes (Nusslein-Volhard & Eric, 1980).

With the advent of molecular biology Drosophila genetics continued to acquire new and powerful tools. These include single P element mutagenesis (transposon tagging), and the development of modified transposons for transgenesis as well as enhancer trap analysis of gene expression (Bier et al., 2018). In addition, systems of inducible gene expression as well as somatic recombination have been developed, permitting analysis of ectopic gene expression and genetic mosaic analysis, respectively (Bier et al., 2018). The Drosophila research community has a long history of collaboration and the database FlyBase was created in 1992 as a hub for consolidating information on the fly genome (Drysdale et al., 2005).

The most used system of inducible gene expression in Drosophila is known as the GAL4/UAS system (Duffy, 2002). This tool, which was "imported" to Drosophila from yeast, consists of the GAL4 transcription factor and the upstream activation sequence (UAS) (Duffy, 2002). GAL4 can be expressed in various cell types and areas of *Drosophila* which can then be used to upregulate any sequence that has a UAS sequence attached up stream of the transcription start site (Duffy, 2002).

Unlike mammalian systems and yeast, *Drosophila* does not readily undergo homologous recombination and as such the advantages of specific gene knock-out or specific gene replacement strategies has been a shortcoming of *Drosophila* as a model genetic system (Hales et al., 2015). Only recently, targeting specific parts of the genome became a reality with the introduction of zinc finger nucleases (ZFNs) and TALENs (Ma & Liu, 2015). However, these technologies need to be specifically manufactured for each use and are labour intensive. All of this has recently changed with the introduction of CRISPR/Cas9 genome editing, which allows for quick and specific targeting of any point within the *Drosophila* genome (Bassett et al., 2013). CRISPR/Cas9 is now routinely used and is the method of choice for generating mutations or facilitating gene replacement in *Drosophila* and many other model genetic organisms (Anzalone et al., 2020).

In the short time since the CRISPR/Cas9 system was introduced, there have been numerous modifications to the system for use in *Drosophila*. The first of these was implemented in 2013 showed that it was possible to adapt the CRISPR/Cas9 system for *Drosophila*. This was first done by using plasmids containing Cas9 and a guide RNA (gRNA) for the yellow gene. Mutagenesis occurred in 5.9% of progeny (Bassett et al., 2013). Although rates were low it showed promise for the technique. The next major advancement happened when researchers transformed embryos to produce in vivo transcribed Cas9 and gRNAs with mutagenesis rates as high as 42.9% (Lin et al., 2015).

The purpose of this thesis was to explore the use and limitations of CRISPR/Cas9 genome editing for two genetic applications in *Drosophila*. By combining CRISPR/Cas9 technology with other *Drosophila* based tools that have been developed over multiple decades, we have been able to test the usage and limitations of CRISPR/Cas9 as a genetic tool.

1.2 Development of CRISPR/Cas9 as a Genetic Tool

Clustered regularly interspersed short palindromic repeats (CRISPR) are repeat regions of DNA that were first discovered in *Streptococcus thermophiles* (Bolotin et al., 2005). Their true function would not be deduced until 2005 when multiple research groups made the realization that sections in this CRISPR locus are associated with various viruses and archaea (Pourcel et al., 2005). Within what are called the spacer regions of the CRISPR locus, DNA sequences with direct homology to plasmids, bacteriophages and prophages were identified. Furthermore, hosts that contained sequences homologous to a particular phage or plasmid were immune to infection. By integrating part of invading DNA into the CRISPR locus, the organism can recognize the same viral threat. With this stored information, CRISPR-associated genes (Cas) complexed with RNA form an endonuclease that creates double-stranded breaks (DSBs) in the foreign DNA, effectively neutralizing it (Bolotin et al., 2005). In order to recognize the viral DNA threat, various sections of the CRISPR locus are transcribed into RNA that is recognizable by various Cas proteins.

This Cas RNA complex binds using the RNA, which is transcribed from the stored sequences within the spacer regions of the CRISPR locus, as a template to target foreign DNA degradation (Bolotin et al., 2005).

In 2007 it was discovered that the CRISPR locus functions as a form of adaptive immunity in bacteria (Barrangou et al., 2007). A strain of *Steptococcus thermophillus* had sections of a phage sequence integrated into the CRISPR locus as a template to target foreign DNA for degradation (Barrangou et al., 2007). The modified strain was then shown to be immune to infection by phage, proving the hypothesis of acquired immunity. Subsequently, in 2008 sections of the CRISPR locus were found to be directly derived from phages. These sequences, usually about 20 bps in length, were found to be transcribed into RNAs and were termed CRISPR RNAs (crRNAs). These crRNAs were found to act as guides when combined with Cas proteins, leading to endonuclease activity and the creation of DSBs at sites homologous to the crRNA sequence (Brouns et al., 2008). While several different CRISPR/Cas systems have been discovered, it has been predominantly the Cas9 protein from *Steptococcus pyogenes* that has been developed as a tool for genome editing. This is due, in part, to the discovery 2010 that Cas9 was sufficient for targeted endonuclease activity (Garneau et al., 2010). The last major component in the system was found to be a trans-activating CRISPR RNA (tracrRNA) which complexes with crRNA in order to guide Cas9 to a target (Deltcheva et al., 2011).

In 2012 the full mechanism for how the CRISPR system works was elucidated. In order for target sequences to be cleaved by the Cas RNA complex, it was found that targets had to be associated with protospacer adjacent motif (PAM) which is found in the target sequences but not in the gRNA itself (Gasiunas et al., 2012). In this way, it is the presence of the PAM that permits the bacterial adaptive immune system to distinguish self from non-self. For the Cas9 system, the PAM is the sequence “NGG” although multiple different Cas proteins use different PAM sequences (Gasiunas et al., 2012). For Cas9, this is a region is found 3-4 nucleotides downstream from the targeted cut site (Gasiunas et al., 2012). Additionally,

the crRNA could also be shortened to 20 nucleotides in order for cleavage to take place and more importantly this sequence can be changed to allow Cas9 to cut at different sites (Gasiunas et al., 2012). Also, it was found that the tracrRNA and crRNA could be fused and that could be transcribed as a single RNA molecule known as guide-RNA (gRNA). Subsequently this system was refined to the engineered minimal 2-part system that is the most widely used genome editing (gRNA +Cas9) (Jinek et al., 2012).

1.3 Generalized Uses of Engineered CRISPR/Cas9

As an engineered genome editing technique, the initial common application for this tool was to create DSBs. What's more revolutionary is that as long as Cas9 and gRNAs can be introduced into a model system the corresponding DSBs can be created in a highly specific manner. Off-site targeting is also rare due to the long and very specific 20bp guide (Jinek et al., 2012). These targeted DSBs are repaired using imprecise cellular machinery. This has a wide variety of applications, including creating mutations in the given target sequence through non-homologous end joining (NHEJ) and the incorporation of template strands into genomic DNA through homology-directed repair (HDR) (Bier et al., 2018).

One of the more obvious applications of the CRISPR/Cas9 system is a gene knock-in. This can be done easily using Cas9, a gRNA that targets a sequence at which the knock-in should take place, and an oligonucleotide that has areas of homology to the target sequence flanking the desired knock-in sequence (Ma & Liu, 2015). In this approach repair of the targeted DSB by HDR using the provided oligonucleotide as a template is an efficient method for gene replacement. Methods of delivery can vary depending on the model organism used and might include lipid-mediated transfection, microinjection or electroporation among others.

The CRISPR/Cas9 system has been further engineered using a number of different modifications. The basic idea of some of these modifications is to utilize the ability of Cas9 to be directed to a sequence of interest. If the endonuclease activity is removed, this can be a good method of delivering proteins to a

specified sequence. Such a modified Cas9 does not have any nuclease activity and is referred to as a dead Cas9 (dCas9) (Qi et al., 2013). One such application has been the design of DNA base editors. In this case dCas9 is augmented with a ribonuclease protein complex that creates a single base substitution when targeted to a sequence of interest.

Another application of dCas9 is used when investigating the composition of chromatin (Fujita & Fujii, 2013). Using DNA-binding molecule-mediated Chromatin immunoprecipitation (enChIP) it is possible to tag a region of interest by having a FLAG-tagged dCas9 targeted to a genomic area of interest. These fragments can then be isolated via FLAG-tag specific antibodies and put through mass spectrometry (Fujita & Fujii, 2013). This allows for the identification of proteins that are associated with specific loci. A similar process to this is known as CasID, where dCas9 is fused with promiscuous biotin ligase BirA* (Schmidtmann et al., 2016). Biotin is then added to the system which results in proteins near (~10nm) the loci to which dCas9 was targeted to being tagged with biotin (Schmidtmann et al., 2016). These tagged proteins can also be isolated and identified by mass spectrometry.

The CRISPR/Cas9 system has also been adapted to repress gene transcription. This was first done in *E.coli* with a system called CRISPR interference (CRISPRi). In this case, a dCas9 physically blocks transcription from initiating (Bier et al., 2018). It can also be used to bind the cis element targets of specific a transcription factor, thus blocking its activity. This was later refined by adding a repressor domain to dCas9 called Kruppel-associated box (KRAB) (Bassett & Liu, 2014).

Another application of dCas9 is to overexpress targeted genes of interest. This is done by fusing it to a transcriptional activation domain consisting of VP64, p65 and Rta domains (VPR) (Lin et al., 2015). Altogether, this $dCas9^{VPR}$ has no nuclease activity but has the ability to upregulate the transcription of a gene of interest if the gRNAs are targeted to a sequence in the proximity and usually upstream of any given gene's transcription start site (TSS) (Lin et al., 2015).

Examples of using the CRISPR-Cas9 system to impart immunity have been reported (Sapranauskas et al., 2011). By integrating a section of DNA specific to antibiotic resistant bacteria to a CRISPR loci, it is possible to allow Cas9 to target these sections of invasive genomes. Not only is this a method of effectively eliminating antibiotic resistant bacteria, if the target sequence is specific enough the system will only destroy the target bacteria leaving non-resistant bacteria alone.

1.4 Using CRISPR in Humans

In a medical context, the use of CRISPR to cure human disease has been of great interest. The concept of gene therapy is not a new one. Using various viruses as a vector to transform cells that would otherwise have certain pathologies has been done before in humans. The concept of such a therapy goes back to the 1970's (Friedmann & Roblin, 1972). Although there have been numerous breakthroughs in the field, the introduction of the CRISPR-Cas9 system is of special note. The ability to rapidly target and create DSBs in a gene of interest and subsequent repair using HDR could open doors for possible therapies.

Many human pathologies could be cured with an edit to a single gene. As a proof-of-principle it has been shown in mice that injecting a zygote with Cas9 mRNA, sgRNA and an oligonucleotide with homologies to the desired sequence and correction for the gene of interest is enough to cure mice of cataracts caused by the gene crystalline gamma c (Savić & Schwank, 2016). This method proves effective in curing a disease prenatally. The use of CRISPR to cure pathologies postnatally has been researched as well. In 2016 scientists at Sichuan University's West China Hospital began human trials involving CRISPR. This trial involves taking patient cells and knocking out the PD-1 gen which is responsible for slowing the immune response. Then re introducing the transformed cells back into the patient. There are even treatments that have finished trial periods and are available to the public. A cure for spinal muscular atrophy has been engineered with the help of CRISPR/Cas9 in humans. The cure itself is called Zolgensma is a single injection designed for children that replaces the dysfunctional *SMN1* gene (Gibbons et al., 2019).

Chapter 2
Development of an F1 Visible Screen for Loss of Function Mutations Using dCas9^{VPR}

2.1 Introduction

CRISPR/Cas9 has been used for several applications in the model genetic organism *Drosophila*. First and foremost, CRISPR/Cas9 has been used to create DSBs when guided to a sequence of interest based on a designed gRNA. This can result in mutations in a sequence of interest producing small insertions or deletions due to non-homologous end joining (NHEJ) (Matsumoto et al., 2020). Moreover, when an oligonucleotide is introduced that contains sequences homologous to the sequences flanking the guide target sequence, integration of the said oligonucleotide into the genome can occur through HDR (Matsumoto et al., 2020). Other applications of the CRISPR/Cas9 system in *Drosophila* have been developed, including converting Cas9 to a transcriptional activator. This is done by mutating the catalytic sites of the Cas9 endonuclease effectively rendering it unable to make DSBs. Such a catalytically inactive Cas9 is known as dead Cas9 (dCas9) (Lin et al., 2015). This dCas9 has been further modified by fusing it to domains that can activate transcription. This is the VPR transcriptional activation domain and the resulting fusion protein is known as dCas9^{VPR} (Lin et al., 2015).

The expression of dCas9^{VPR} under the control of the Gal4 system, permits overexpression of endogenous genes through co-expression of a specific gRNA in a tissue specific manner. There are currently more than 1600 transgenic stocks that express different gRNAs under a ubiquitin promoter. A publicly available resource for such purposes is the TRiP transgenic stock collection from the Bloomington Drosophila Stock Center. While some of these stocks have been designed for targeted gene knock down (these are known as TRiP-KO lines or TKO lines), others have been designed for targeted overexpression of a gene of interest (GOI) (Zirin et al., 2020). In general, these stocks are known as TRiP-OE or simply

TOE lines and express 2 gRNAs that are targeted to regions upstream of a gene's transcriptional start site (TSS)(Zirin et al., 2020). By combining TOE lines with the UAS-dCas9^{VPR} and a Gal4 driver it is possible to over express a gene of interest in a tissue specific manner. This approach of ectopic gene expression or gene overexpression differs from the traditional Gal4/UAS method in that the GOI that is being overexpressed is an endogenous allele.

2.1.1 Hindsight

The gene *hnt* encodes the C_2H_2 zinc finger transcription factor, Hindsight, that is required at multiple times and in multiple tissues during *Drosophila* development. The Hnt protein is 1893 amino acids in length and includes 14 C_2H_2 zinc finger domains (Wilk et al., 2000). Processes known to require HNT include germ band retraction and maintenance of the extraembryonic membrane known as the amnioserosa, tracheal formation, and differentiation of the pupal retinal epithelium (Farley et al., 2018). The gene *hnt* also has a human homologue called *Ras-Responsive Binding Element 1* (*RREB-1*), which has been implicated in a number of human pathologies such as thyroid, colorectal and pancreatic cancer (Farley et al., 2018). As a proof of principle for this new screening strategy, we aimed to be able to quickly recover numerous new alleles of *hnt* with the expectation that some of these new alleles may shed some light on the function and regulation of this important gene.

Using this new system of gene overexpression we were examining the phenotypes associated with overexpression and ectopic expression of *hnt*. Similar to previous observations of *hnt* overexpression using the Gal4/UAS system, we found that dCas9^{VPR} based overexpression results in the disruption of the development of the compound eye. Unlike the Gal4/UAS system for *hnt* overexpression, however, the severity of this phenotype when using the dCas9^{VPR} proved to be sensitive to the number of functional *hnt* alleles. Based on this observation we found that dCas9^{VPR} overexpression could be used as a basis for identifying loss of function mutations of *hnt* in heterozygotes.

2.1.2 Background of the Development of a System to Overexpress *hnt* using dCas9^{VPR}

The following components are needed for targeted overexpression in the *Drosophila* eye using the VPR-based overexpression system. The first component is the *glass* mediated response element (GMR), which drives GAL4 expression in the developing eye imaginal disc (Li et al., 2012). The next component is the dCas9^{VPR} with controlled expression due to an upstream UAS. The third and final component is the TOE-GS00052 transgene which ubiquitously expresses two guide RNAs that target sequences directly upstream of *hnt* (183-203 and 110-129 bp from the transcription start site respectively). In this system, GAL4 is expressed in the developing eye imaginal disc which then activates the expression of UAS-dCas9^{VPR}. This causes overexpression of *hnt* due to the ubiquitously expressed TOE-GS00052. The overexpression of *hnt* through this method is temperature sensitive: when a homozygous wildtype *hnt* allele is subjected to this system at 25˚C, it leads to lethality due to the overexpression of *hnt*. However, the same genotype is viable at 18˚C because the GAL4/UAS system is well-known to be temperature sensitive and more efficient at 25˚C (Duffy, 2002). At 18˚C, the dCas9^{VPR} mediated overexpression of *hnt* results a visible phenotype (Figures 2-1 and 2-2), which comprises a red ring around the eye and drastic developmental defects in the pupal eye.

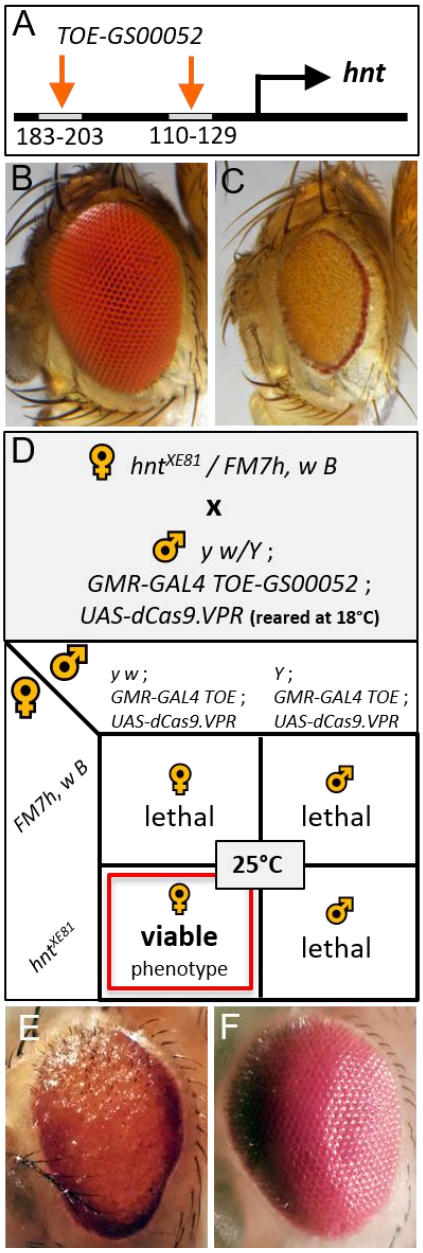

Figure 2-1 Overexpression of hnt in the developing eye disrupts eye formation in a temperature and dosage sensitive manner

(A) Shows proximity of target sequences of TOE-GS00052 to transcription start site of *hnt* (B) Wildtype eye (C) Eye under overexpression of *hnt* showing GMR-Gal4 > UAS-GFP-Hnt127. (D) When a *hnt* null mutation is crossed to the overexpression system, progeny that has the *hnt* null mutation hnt^{XE81} is viable

due to having fewer *hnt* copies to overexpress. (E) Overexpression phenotype of TOE-GS00052 with GMR-Gal4 > UAS-dCas9^{VPR}. (F) Supressed phenotype of (E) with co-expressed UAS-*hnt*-RNAi

Heterozygotes for different *hnt* alleles were crossed into this background of dCas9^{VPR} mediated overexpression. The following experiments were performed by Dr. Reed. The allele *hnt*^{XE81} is a nullomorphic allele which does not produce functional Hnt. Interestingly *hnt*^{XE81} is viable at 25˚C, but produced the same phenotype as seen at 18˚C in the overexpression background. A deletion of the *hnt* region was also viable at 25˚C resulting in the same phenotype as *hnt*^{XE81}. Another lethal *hnt* allele, *hnt*^{PL67}, resulted in lethal overexpression in this system, indicating that this allele does not disrupt the region including the gRNA target sequences or the protein coding sequence of the *hnt* gene. Overall, this ability to overexpress mutant alleles represents a novel approach for characterizing alleles which is similar to the traditional Mullerian system of allele classification (H. J. Muller, 1932).

We propose a novel screening strategy in which randomly induced mutations in *hnt* could be selected using an F1 visible screen.

2.2 Methods

2.2.1 Single Fly Genomic DNA Extraction

The extraction of whole genomic DNA from a single male fly was done through the following procedure. The fly was suspended in 150µl of squish buffer containing 1M Tris-Cl pH 8.0, 100 mM EDTA, 5M NaCl and 5M proteinase K. The fly was then subsequently squished to release its genetic material, followed by an incubation at room temperature for 20 minutes to remove some protein from the final product. The mixture was incubated at 95˚C for 5 minutes. The solution was diluted to 20ng/µl of DNA. Samples to be sequenced included 10µl of the DNA solution and 5µl of custom primers at a concentration of 2.0µM. Any unused DNA was stored at 4˚C.

2.2.2 Single Fly PCR

PCR primers were designed to amplify the region surrounding the TOE-GS00052 target sequences (Appendix B). By using a series of PCR reactions, the integrity of the guide regions were evaluated by amplifying the region between the two primers and comparing the amplified sequence to the wildtype. This approach was used to verify disruptions within or surrounding the guide target regions associated with newly recovered refractory chromosome. Direct PCR sequencing was performed to confirm that refractory chromosomes contained disruptions or deletions within the gRNA target sequences. Samples for PCR were prepared with 25µL FroggaBio 2X PCR Master mix Solution, 2.5µL of both forward and reverse primers at 100M, 2.5µL of sample DNA and 17.5µL of dH_2O. Using a thermocycler the samples were subjected to the following program:

1. 94˚C for 3 minutes
2. 94˚C for 45 seconds
3. 55˚C for 30 seconds
4. 72˚C for 1 minute and 30 seconds
5. Repeat steps 2-4 35 times
6. 72˚C for 10 minutes
7. 4˚C hold

2.2.3 EMS Mutagenesis

Adult males of various ages with the genotype *y w P[w$^+$]e02388b/Y* were starved for 8 hours. A solution containing 3mL of a 1% sucrose solution in dH_20 and 0.025M EMS from Sigma-Aldritch, was prepared. Parafilm was added over the vessel to prevent splashing of EMS. An 18-gauge needle was then inserted into the container to vigorously agitate the solution by repeatedly drawing and expulsion of the

mixture to create an emulsion. A 1L incubation vessel was then prepared with a single layer of kimwipes and the EMS solution was added until the kimwipe was thoroughly soaked through. Male flies were added to this vessel and allowed to feed on the EMS laced sugar solution for 24h. Flies were used in crosses immediately following this 24h period. Culture vessels and kimwipes were decontaminated in 10% solution of W/V sodium thiosulfate for 24h to deactivate EMS.

2.2.4 Complementation Tests

Newly derived *hnt* alleles were balanced using *FM7h* (Ashburner, 1989) and balanced females were crossed to hnt^{peb} males. hnt^{peb} is a temperature sensitive visible allele of *hnt* that is associated with a rough eye phenotype when reared at the restrictive temperature of 29˚C (Ashburner, 1989). If the putative new allele failed to complement hnt^{peb}, (i.e., meaning that at 29˚C a rough eye phenotype was observed), then the new allele was deemed to be a true new allele of *hnt*.

2.2.5 Sequencing *hnt* REF alleles

Sequencing of the PCR fragment containing the TOE-GS00052 target site was done by the London Regional Genomics Centre. Genomic DNA samples were amplified using primers that flanked the target site following the PCR methods outlined in Section 2.2.2. Samples were then purified through spin columns (Manufacturer) and diluted to 20ng/µL based on a nanodrop reading. The same primers that were used in the PCR reaction were diluted to 20µM in dH_2O and used for the sequencing. In each reaction, 10µl of the diluted PCR product and 5µl of the diluted primer were added to a microfuge tube and sent to the London Regional Genomics Center for sequencing. Two sets were sent out per sample, one with the forward primer of the PCR reaction and another of the reverse. This provided a means to verify the sequence and obtain a longer continuous read of the whole fragment.

2.2.6 Immunostaining and Imaging

PBS 10X solution was prepared with 1.3M NaCl, 70mM $Na_2HPO_4.7H_20$ and 30mM K_2HPO_4 in dH_20. Embroys that were 0-18 hours old were washed in PBS-TX (1X PBS containing 0.12% Triton X-100 v/v) and collecting them in filtration baskets made with 2000 micron nylon mesh. These baskets were made by cutting a 50mL Falcon tube in half width wise and removing the top of the cap. With this setup the mesh can be placed between the opening of the Falcon tube and the modified screw cap and thus be removed easily without disturbing embryos in the basket. These embryos were then suspended in a 50% commercial bleach and PBS-TX solution to dissolve the outer cuticle. The embryos were then fixed with formaldehyde in a 3.7% solution in 1X PBS (w/v) over which there was an equal volume of heptane for 20 minutes. Lower aqueous fixative layer was carefully removed and replaced with methanol. With vigorous shaking, devitalized embryos settled to the bottom of the methanol phase. The embryos were then carefully removed and rehydrated with a methanol series. Starting with 75% methanol to PBS by volume, each sample was suspended in solution and allowed to settle to the bottom of the tube. This process was repeated, decreasing the methanol percentage by 25% until embryos were submerged in 1X PBS.

Blocking solution was prepared using a PBS-TX solution with 2.5% bovine serum albumin (BSA) and 5% normal goat serum (Thermo Fisher Scientific 31872). The embryos were thoroughly washed and blocked for 1 hour after which the blocking agent was removed. The following antibodies were applied at the indicated dilutions: mouse monoclonal anti-Hnt 27B8 1G9 (1:10; from H. Lipshitz, University of Toronto) along with guinea pig polyclonal anti-Hindsight (1:1500; from H. Lipshitz, University of Toronto). Antibodies were removed and blocking agent was re-applied for 1 hour. Secondary antibodies used were: Alexa Four 488 goat anti-mouse (Cedarlane Labs) and TRITC goat anti-guineapig (1:500: Cedarlane Labs).

Micrographs were taken using a Nikon SMZ25 stereomicroscope equipped with Nikon Digital Sight Ri2 16.52 MP colour camera and processing was done using Nikon NIS-Elements Arv4.50 software (Figure 2-2. BC). Light microscopy images were taken using a compound Olympus microscope with a 5X objective (Figure 2-2. EF). The subject was manually focused on and images were taken to create slices which were then processed using ImageJ and an extended depth of focus plugin (Schneider et al., 2012).

2.3 Results

The following presents the proof-of-principle screen showing the recovery of loss-of-function mutations in *hnt* selected through an F1 screen. The section immediately below describes the various steps and crossing strategies involved in the development, execution and analysis of the results of this screen.

2.3.1 Creation of the Refractory Chromosome

The goal of this screening technique is to use dCas9VPR mediated overexpression for the selection of multiple *hnt* null or deficient mutants in an F1 visible screen. Due to the lethality of the overexpression at 25˚C, in order to better facilitate this screen, we selected chromosomes that were refractory to the Cas9 mediated overexpression. The first step was to render the guide target sequences unable to bind to the TOE-GS00052. To do this, we substituted the dCas9^{VPR} for active Cas9 in the overexpression background. This causes the guide target sequences to induce mutations in the target sequences so they are unresponsive (refractory) to the TOE-GS00052 and cannot be targeted by the overexpression system. We then combined all these elements into a single line that contains the refractory chromosome, gRNAs, GMR-GAL4 and UAS-dCas9^{VPR} element; this line is referred to as the RGV line. The steps of this process are shown in Figure 2-1. Refractory lines were created by Andrew Ma.

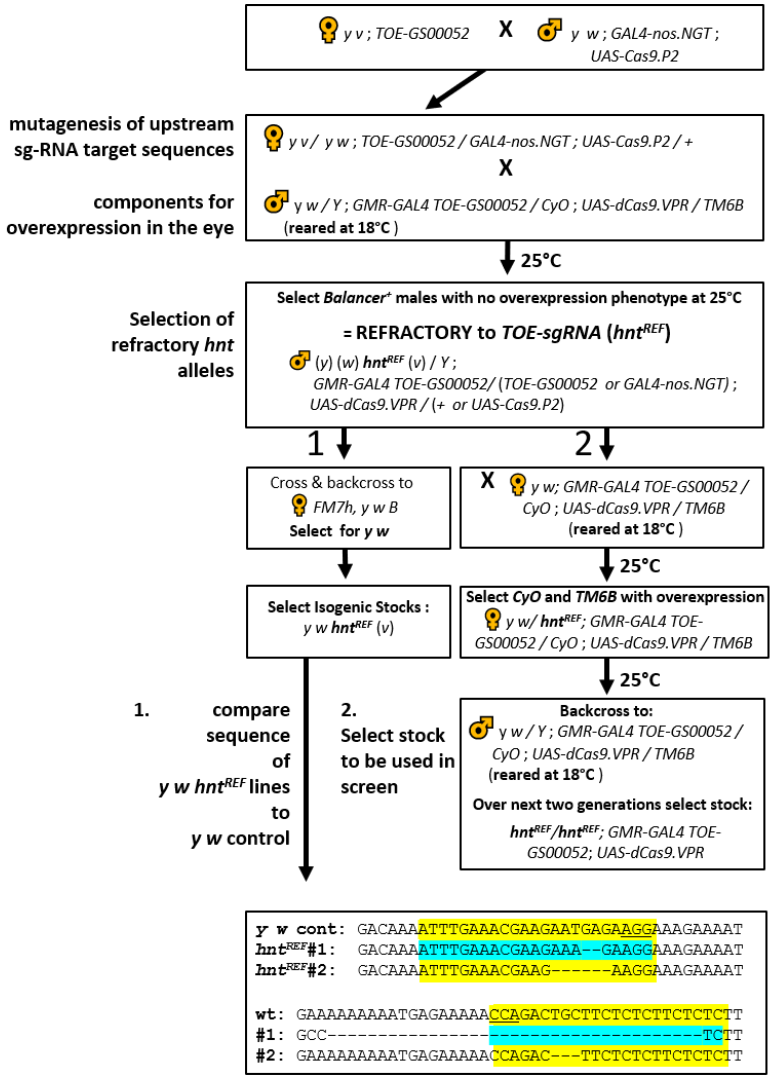

Figure 2-2 Creation of the Refractory Chromosome

Crossing scheme used in order to create refractory chromosome. Two lines were sequenced comparing them to the wildtype. Refractory stock was created and then crossed to have the UAS-dCas9^{VPR} as well as the GMR-Gal4 and the TOE-GS00052.

2.3.2 F1 Visible Screen

Adult males of various ages mutagenized with EMS were crossed to virgin RGV females. The progeny were screened for the complete absence of the overexpression phenotype. A “hit” in the screen would be any female progeny presenting a wild-type eye phenotype as seen in Figure 2-2. Individual hits

were then crossed to the RGV stock to verify that the mutation propagated through the next generation in a predictable way. All lines generated through this method also underwent complementation testing to determine which were new *hnt* alleles. The EMS-treated males carried the markers *y w P[w$^+$]e02388b*. The *P[w$^+$]e02388b* insertion is tightly linked to *hnt*, being inserted ~ 40Kb from the *hnt* coding region and is associated with the distinctive pattern of *w*$^+$ expression in the posterior eye. The *P[w$^+$]e02388b* served as a marker for isolation of the mutagenized chromosome and the establishment of isogenic stocks balanced over *FM7h* (Figure 2-3).

2.3.3 Resulting Progeny from Screen

After crossing virgin RGV females to EMS-mutagenized males, a total of ~45,000 progeny were scored. This was done over a period of 3 weeks with the help of Dr. Reed. Of these progeny 39 females were isolated that did not produce an over-expression phenotype. All of these females had a characteristic wild-type eye as illustrated in Figure 2-3, which was a striking and easily scorable phenotype in the background of the *hnt* overexpression phenotype (Figure 2-2). Of these females, 15 bred true and were established as stocks over FM7h,w. These stocks were then re-tested by crossing to RGV to make sure they continued to show a lack of overexpression. After re-testing, 8 of the previously 15 lines were found to be true *hnt* alleles based on complementation tests. The remaining 7 lines were sequenced to elucidate their refractory nature and were described as *hnt*^{REF}. Such refractory alleles are somewhat unusual in the system of *Drosophila* nomenclature as their only describable phenotype is that they are unresponsive to this particular system of dCas9^{VPR} mediated overexpression.

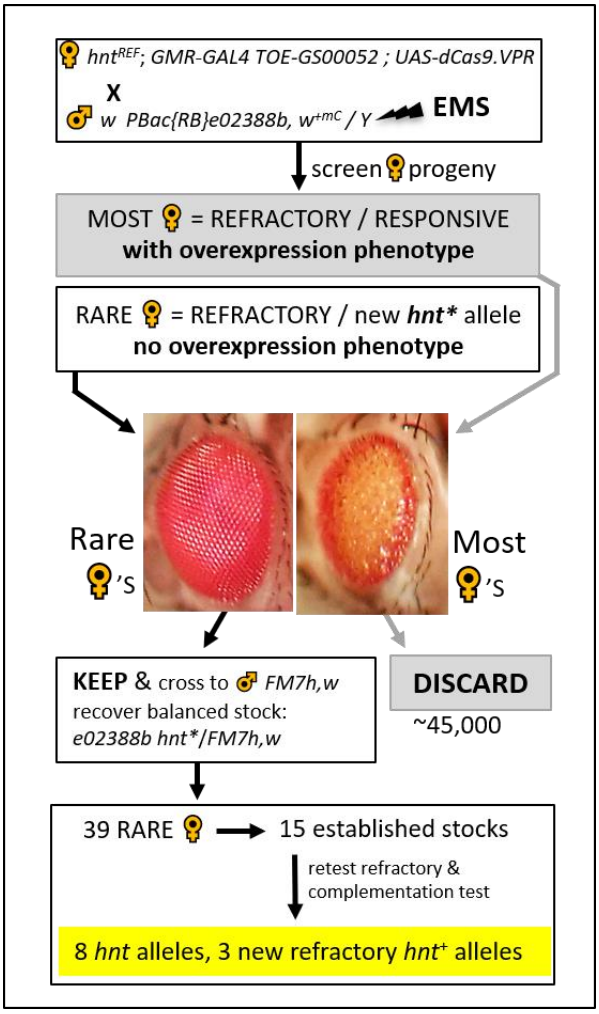

Figure 2-3 Visible Screen for loss of function Mutation Strategy

Virgin female RGV stocks were crossed to EMS mutagenized males. Resulting progeny are either responsive to overexpression or are not responsive and thus have wildtype eyes. Wildtype eyes contain the presumptive mutation that induces a loss of function allele. These flies were then bred into lines and kept balanced over FM7h,w. These were later established as 15 stocks 8 of which were true *hnt* alleles and 3 were new refractory *hnt* alleles.

2.3.4 Characterization of new *hnt* alleles

Of the 15 stocks that were recovered, 8 were *hnt* alleles and 7 were refractory chromosomes. The target regions of the new refractory lines were amplified using PCR and sequenced (Figure 2-4). Of note $hnt^{Ref\text{-}WN72}$ was found to have a single base pair change from T to C within the second sgRNA target and

hnt$^{Ref\text{-}BHR7}$ was found to contain an A to T base pair change in the first sgRNA target (Figure 2-4). The allele *hnt*$^{Ref\text{-}BHR7}$ was found to be unusual in that it can over-express *hnt* in males but not in females. All other alleles were found to contain an identical deletion which spans between and including parts of both sgRNA target sequences making them refractory.

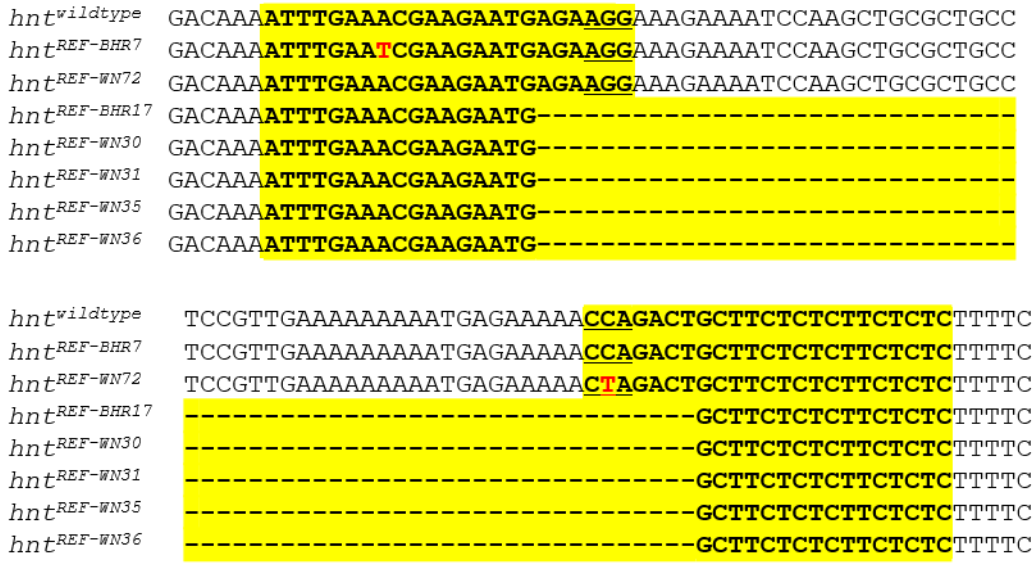

Figure 2-3 Sequencing results of new Refractory lines

Sequences of the area between TOE-GS00052 cut sites (underlined) were aligned with the wildtype reference sequence. Of note *hnt*$^{REF\text{-}BHR7}$ and *hnt*$^{REF\text{-}WN72}$ both have single base pair changes that render the sequence refractory. The latter having it in its PAM motif. The rest have an identical deletion with both PAMs missing.

The newly recovered confirmed alleles were crossed to *peb/Y; Ubi-DEcadherin-GFP* in order to visualize and quantitate the defect in eye development. By quantifying the number of cone cells within each ommatidium of the compound eye, these new alleles were ordered into an allelic series according to the severity of the defect (Figure 2-5E). The allelic series was created by Molly Chen. Some of these ommatidium contained fewer or more then the regular 4 cone cells (Figure 2-5A). Most of the mutants contained fewer cone cells but sometimes mutants managed to have 5 or 6 cone cells. These likely result

from the fusion of two neighbouring ommatidia each having less than 4 cone cells (Figure 2-5BCD). According to this analysis, all newly recovered were characterized as hypomorphic as they are not as severe as the phenotype of a complete deletion of the entire *hnt* gene within *Df(1)ED6727*. Furthermore, immunostaining was performed on these alleles. Interestingly all were positive when using a polyclonal anti-Hnt serum. When using a monoclonal ant-Hnt antibody hnt^{BHR49} and hnt^{WN52} were monoclonal negative, polyclonal positive, this indicates that the overall structure of Hnt is mostly intact but certain epitopes have been altered (Figure 2-6CC'DD'). We also discovered that hnt^{XE81}, which was previously described as a null allele stained positive for the polyclonal anti-Hnt antibody (Figure 2-6E). When imaging embryos it was noted many of the alleles were recessive lethal and failed in germ band retraction (GBR) (Table 2-1). The process of GBR is an essential step in development and embryos that do not retract do not survive (Yip, Lamka, & Lipshitz, 1997).

Allele	**α-Hnt monoclonal**	**α-Hnt polyclonal**	**GBR phenotype**
hnt^{WN29}	++	++	retracted
hnt^{BHR28}	+	+++	tail-up
hnt^{BHR49}	+	+++	unretracted
hnt^{WN52}	-	+*	unretracted
hnt^{BHR5}	+	+	unretracted
hnt^{BHR40}	+	+++	unretracted
hnt^{WN47}	++	+++	unretracted
hnt^{WN31}	-	+*	unretracted
Df(1)ED6727	_	_	unretracted

Table 2-1 Summary of Immunostaining results using anti-Hnt polyclonal and monocolonal antibodies with germ band retraction analysis

Number of (+)'s indicates a qualitative amount of signal observed during immunostaining. A (–) indicates no signal during immunostaining. A tail-up phenotype indicates a partial retraction of the germ band. (*) Indicates that anti-Hnt signal was not localized to the nuclei.

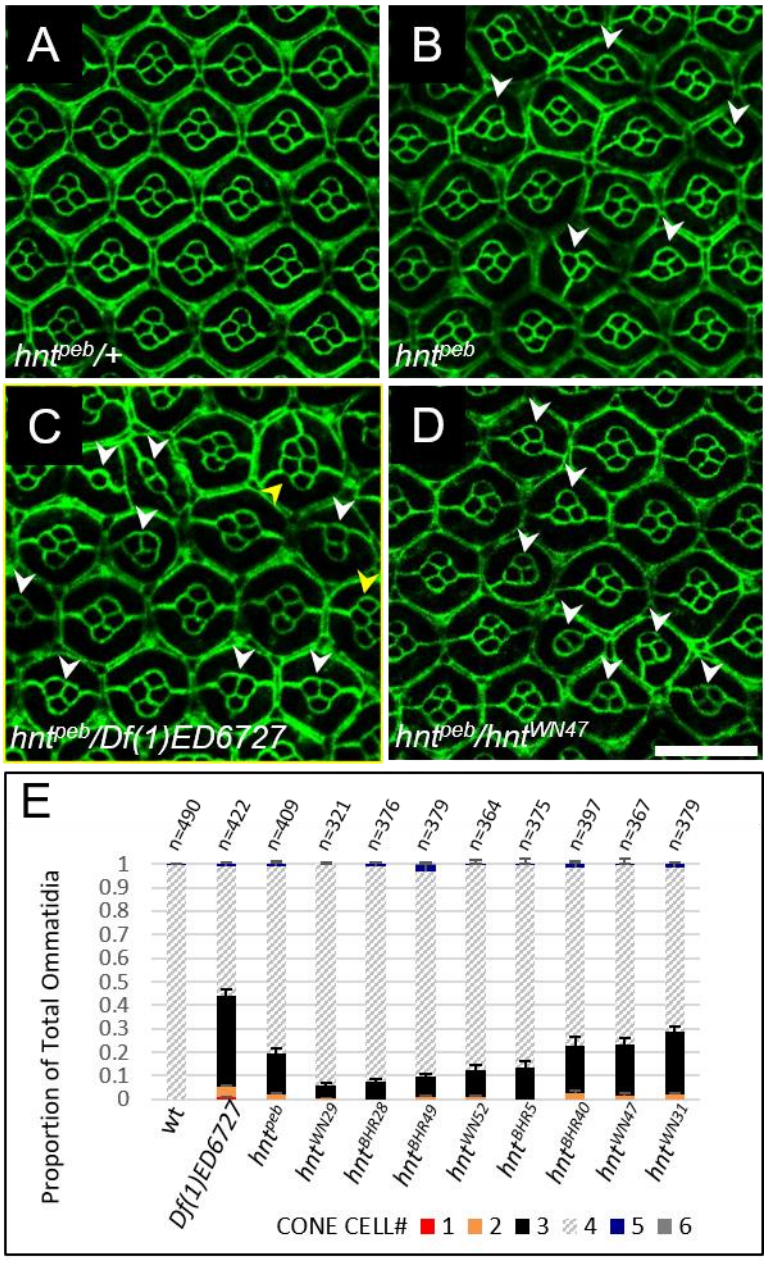

Figure 2-4 New *hnt* alleles quantified by ommatidial structure

(A) Ubi-DEcadherin-GFP expression in a control hnt^{peb}/+ shows normal ommatidial structure with 4 cone cells per ommatidium. (B) hnt^{peb}/hnt^{peb} when raised at 25 shows some disruption in the number of ommatidia as highlighted by arrows. (C) $hnt^{peb}/Df(1)ED6727$ shows both more and fewer cone cells in ommadidia. (D) $hnt^{peb}/hntWN47$ with both more and fewer cone cells in ommatidia. (E) Comparing the number of abnormal ommatidia between *wt, Df(1)ED6727* and the mutants in an allelic series. Scale bar represents 20μm.

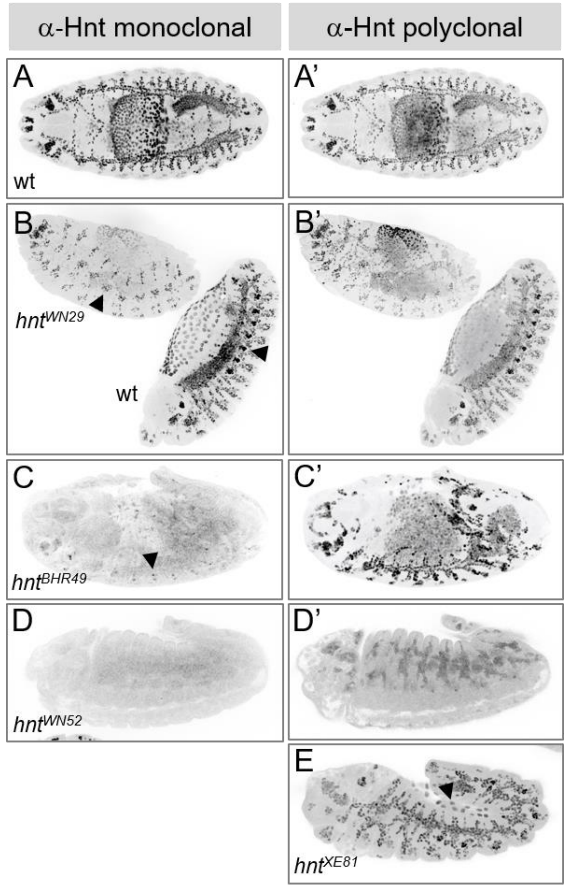

Figure 2-5 Immunostaining of Embryos using Both Hnt monoclonal and Polyclonal antibodies

(A-D) Represent anti-Hnt monoclonal staining. (A'-D',E) Represent anti-Hnt polyclonal staining. (B,B') hnt^{WN29} shown with wt sib. Shows reduced anti-Hnt signaling in both polyclonal and monoclonal antibody. Oenocyte singal was observed in wt but not in hnt^{WN29} as shown by arrows. (C, C') hnt^{BHR49} shows weak monoclonal signal but strong polyclonal. (D, D') hnt^{WN52} shows very weak signaling in monoclonal and slightly stronger non-nuclear signaling in polyclonal. (E) hnt^{XE81} with degenerated amnioserosa showing strong polyclonal signal. Experiments were performed on multiple embryos in a single experiment.

2.4 Discussion

We have shown that dCas9^{VPR} based overexpression of a gene of interest can be used as a novel method to isolate new randomly induced mutation in a gene of interest. The same approach of using dCas9^{VPR} mediated overexpression, however, is also a new tool for allele characterization. If the over-expression phenotype associated with a previously identified mutant heterozygote (m/+) is less severe when compared to the overexpression phenotype of homozygous wildtype (+), two possibilities arise. First, the new allele might not be responsive to the over-expression system, in which case it could be refractory due to a naturally occurring polymorphism disrupting the guide target sequence. Second, the mutant allele may be associated with reduced expression or function for the GOI. Such loss-of-function alleles may be complete, known as nullomorphic alleles, or partial, in which case they are known as hypomorphic alleles. Barring the first possibility of disrupted gRNA target sequences, which can be ruled out by sequencing, the dCas9^{VPR} system can be used to distinguish nullomorphic alleles from hypomorphic alleles. The expectation here is that the severity of the overexpression phenotype would relate to the total amount of gene function in the heterozygous overexpression background. That is, the overexpression phenotype of the heterozygous nullomorphic allele (m^{null}/+) would be less severe than the overexpression phenotype of the heterozygous hypomorphic allele (m^{hypo}/+). This result is supported in the allelic series (Figure 2-5) where the deletion of the *hnt* region causes the most severe phenotype when compared to *hnt^{peb}* which is considered a hypomorphic allele.

In addition, the dCas9^{VPR} mediated gene overexpression can be useful in determining if a particular mutant allele is capable of producing functional protein. If an allele for a GOI generates an overexpression phenotype that is indistinguishable from the homozygous wildtype, then it follows that this mutant allele is capable of producing functional protein. Such mutations may disrupt gene expression without disrupting

the coding sequence for the GOI. Several P element alleles of *hnt*, where the P element is inserted into a region 5' to the TSS of *hnt*, were found to behave this way in our overexpression system.

As seen in Figure 2-4, when looking at the sequence data of the refractory alleles, 5 contain a deletion between guide regions. It was later discovered that all these lines originated from a single vial. These lines are identical so we can say that we really got 8 new *hnt* alleles and 3 refractory chromosomes. EMS mutagenesis is known to create small deletions so it is possible that EMS induced the mutation in a germline stem cell in the male (Ashburner, 1989). This would result in multiple sperm carrying the mutation whereas normally during EMS mutagenesis, only mature sperm are affected (Ashburner, 1989). These sperm then fertilized multiple eggs thus showing multiple copies of the same mutation. This is not uncommon in EMS screens and is generally referred to as a 'pre-meiotic cluster' (Ashburner, 1989).

Another explanation for the recovery of the five deletion mutants from one vial would be that there had been a spontaneous reversion of dCas9 to an active form in the female germline of one fly in the RGV line. This initially seems like a reasonable explanation due to how the small deletion perfectly encapsulates both target sites. However, multiple unlikely steps would have been needed for this to happen. The reverted dCas9 would have presented a refractory somatic phenotype due to the expression of the supposedly reverted dCas9. To make this heritable we would also have to assume that there had been leaky expression of reactivated dCas9 in the female germline from which the new refractory alleles were recovered. If this reversion event was the culprit, we would assume that it would happen at a much more frequent rate, but other experiments would refute this. In addition, this explanation does not explain how these deletions were clearly on chromosomes derived from the EMS treated males, which carry the [w^+] transgene insertion e02388b which is tightly linked to the *hnt* locus. Overall, the most parsimonious explanation would be that EMS created a small deletion in the genome in the treated males. This chromosome, which was selected by its refractory nature, by chance happened to carry an EMS-induced deletion from one gRNA target sequence

to the other gRNA target sequence. Although EMS primarily produces GC to AT transitions, it is well known that EMS is quite effective at producing chromosome aberrations in Drosophila (Ashburner, 1989).

Interestingly two of the refractory chromosomes contain mutations in the guide target sequences. The alleles hnt^{WN72} and hnt^{BHR7} contain mutations in the PAM and target sequence respectively (Figure 2-4). Both alleles are refractory on their own which indicates that both guide targets are necessary for targeted overexpression.

All 8 of the new *hnt* alleles that were discovered had a very strong suppression phenotype and when re-tested, continued to show said phenotype. However, there were many supposed 'hits' that showed either intermediate phenotypes or strong phenotypes that did not show up again after re-testing. One possibility to explain these occurrences are mutations in dose-dependent second site modifiers. These modifiers could have been on any chromosome that was not the X. However, we were not able to isolate these second site modifiers due to the nature of the RGV line.

It turns out that the RGV line used was not homozygous for the same refractory chromosome. The RGV line used contained 2 different variations of the refractory chromosome with subtle differences that may effect downstream applications. When crossing this RGV line to a known *hnt* mutation, it is possible to achieve an inconsistent result. A mixture of responsive and refractory results would be observed. Although both chromosomes are refractory, they gave subtly different response. These subtle responses would normally be how we would detect certain second site modifiers thus rendering this RGV chromosome useless for this application. We have since re-created the RGV line to include a homozygous hnt^{WN35} as its refractory component. Thus, we are better able to identify second site modifiers in the future if this screen is repeated.

As mentioned, many "hits" did not breed true. The most likely explanation for this is that EMS is also well known to cause genetic mosaicism (Ashburner, 1989). This would cause mutations in the somatic

tissues of the eye but not the germline, and therefore the mutation would not be transmitted to the progeny of the selected female. When screening females we observed some inconsistency in the eye structure which would have been a result of genetic mosaicism. However, these inconsistencies were thought to be a reduced phenotype from *hnt* overexpression and were thus kept and re-tested. This later would prove false as they did not breed true.

Through creating an allelic series (Figure 2-5E), we found that none of the new alleles gained were as potent as a true *hnt* deficiency. Although we did not expect that this would happen due to the single base pair change nature of EMS mutagenesis. Creating a small point mutation has a low chance of rending a protein completely unviable. Even so, we managed to create an array of *hnt* mutations that affected the resulting ommatidium with varying degrees of severity.

It is possible that this technique of selecting new mutants by screening F1 progeny for the lack of an overexpression phenotype can be used for other GOIs. There are several requirements that must be satisfied for this technique to become viable. First, a new gRNA must be created to target just upstream of the GOI for dCas9^{VPR} to target the gene for over-expression. Second, the gene itself must have an over-expression phenotype. Fortunately, due to how the technique is established the over-expression phenotype can be lethality. Third, a viable refractory chromosome must be created. If all of these criteria are met, then this technique can work for any GOI.

There are no other documented cases where a chromosome that is refractory to dCas9^{VPR}-mediated over-expression has been created. Theoretically this should be a simple case of mutating the guide target sequences. As far as we know the ability to create this kind of chromosome may be the result of the inherent qualities of *hnt*. Until such a time when a new refractory chromosome has been created in another GOI we can only say creating a refractory chromosome is plausible.

The creation of the refractory chromosome relied heavily on the temperature sensitive dosage effect of *hnt*. However, it is possible that a single responsive copy of a GOI in the over expression background would produce lethality thus preventing the recovery of the refractory chromosome. One way to circumvent this would be to isolate a refractory chromosome over a deletion of a GOI. Another option may be to follow a similar crossing scheme as shown to achieve the refractory chromosome but reared at a lower temperature. As the Gal4 UAS system is temperature sensitive it is possible that reduced over-expression will not result in lethality.

Chapter 3

Recovering CRISPR/Cas9 Induced Translocations using the "Dubinin Effect"

3.1 Introduction

One of the major advantages of CRISPR/Cas9 genome editing technology is 'multiplexing'. Multiplexing is described as the ability to run multiple experiments concurrently in a single experiment (Lin et al., 2015). In some cases of CRISPR/Cas9 experiments this means adding multiple gRNAs to a model organism such that Cas9 will simultaneously make DSBs at multiple sites. The desired effect of this is to create mutations in more than 1 gene at a time (Chen et al., 2020). Multiplexing however, may result in undesirable chromosomal rearrangements (Chen et al., 2020). For some applications, however, the ability to recover chromosome rearrangements targeted to specific sites in the genome may be a desirable outcome. Such chromosomal rearrangements could be useful for studying aspects of gene regulation, or for creating novel rearrangements for the purpose of designing balancer chromosomes. The aim of the work presented in this chapter was to design a method of multiplexing in which the efficiency of inducing large scale chromosomal rearrangements could be determined.

Being able to create targeted chromosomal rearrangements would be a 'game changer' for genetic research. If multiplexing different gRNAs to create desired DSBs caused large scale chromosomal rearrangements it could be used to build on the established toolbox of *Drosophila* genetics and allow for different tools to be adapted into other model organisms. Creating these rearrangements can also be used in other ways in *Drosophila*. For instance, if translocations can be produced, it is possible to translocate genes from their native regulatory regions. In this process larger and larger portions of the regulatory region could be translocated away from the downstream transcribed region to find where important regulatory elements are within its promoter. One such gene that is known to have a large upstream regulatory region is *hnt*.

There are multiple conventional ways to identify if a targeted chromosomal rearrangement has occurred. These include PCR and polytene squashes (Ashburner, 1989) but these methods can be difficult to do on larger scales. In order to establish the frequency at which these rearrangements happen, a faster, more efficient way of identifying when this event occurs is needed. For this purpose we took advantage of the Dubinin effect – which is a rather obscure and unusual bit of *Drosophila* genetics (Dubinin & Sidoroff, 1934). This involves the gene *Cubitus Interruptus* (*ci*) and the mutant allele ci^1. This allele is unusual in that it is normally recessive and is associated with a visible defect in the wing vein structure, but when homologue pairing is disrupted, ci^1 becomes dominant – showing the same wing vein defect. By introducing a translocation in the region proximal to the gene *ci*, in theory it is possible to disrupt homologue paring and thus create a visible translocation marker when heterozygous to the ci^1 allele (Dubinin & Sidoroff, 1934). With this effect in mind, it is possible to quickly measure the rate at which translocations involving *ci* are recovered.

Not only do we have *Drosophila melanogaster* as a model genetic organism, but researchers frequently study aspects of evolutionary and developmental biology using other *Drosophila* species. At present the genomes of 22 sister species of *Drosophila* have been sequenced and are viewable on FlyBase (Larkin et al., 2021). Researchers can compare conserved genome regions amongst different species of *Drosophila* with the aim of identifying sites- such as enhancer elements that regulate gene expression – that are likely functional in nature. Ultimately, using CRISPR/Cas9 technology, it is now theoretically possible to create mutations in a GOI in a sibling species identical to a mutation that has been studied and characterized in *D. melanogaster*. An interesting question would be to determine if the phenotype associated with a particular genetic lesion in *D. melanogaster* would present the same phenotype in a sibling species or if it would present a different phenotype. Overall, this would permit a better understanding of the effects of genetic background on the generation of phenotypes associated with mutations. In order to

perform these types of experiments, some of the genetic tools that have been developed in *D. melanogaster* must be recreated in the sibling species. However, the genetic toolbox for these subspecies is limited. One of the largest hurdles in making a species suitable for different genetic testing is the inability to reliably keep alleles in stock. For this we would need to establish balancer chromosomes in these new species. In *D. melanogaster* balancer chromosomes have been painstakingly created over multiple decades, often through trial and error and often using random mutagenesis methods (Kaufman, 2017). However, the CRISPR/Cas9 system has been shown to be easily implemented in other model organisms. With this technology it should be much easier than before to create multiple inversions, and therefore balancer chromosomes in other species. It is therefore of general interest to determine the efficiency with which large chromosomal rearrangements can be recovered using CRISPR/Cas9 to simultaneously induce DSBs at two or more positions within the genome.

3.1.1 Cubitus Interruptus

Cubitus interruptus (ci) is a transcription factor involved in the hedgehog signaling pathway (Schwartz et al., 1995). This pathway is responsible for the proper formation of various organs. This pathway involves multiple segment polarity genes that are vital the proper formation of body segments and tissue type differentiation (Schwartz et al., 1995). Mutations in such genes can disrupt standard segment development and formation.

The wing imaginal disk of *Drosophila* can be divided into an anterior (A) and posterior (P) compartments (Ogden et al., 2004). P cells express *engrailed (en)* which signals P cells to express *hedgehog (hh)* (Ogden et al., 2004). After expressing Hh, P cells then become refractory to *hh* signaling. A cell can respond to *hh* signaling if it does not express *en* (Ogden et al., 2004). This creates a gradient in the imaginal disk allowing for the establishment of polarity. Response to Hh requires two genes *smoothened (smo)* and *cuibitus interruptus (ci)* (Ogden et al., 2004). The Ci protein itself has two different forms a repressor form

$Ci^{[rep]}$ and an activator form $Ci^{[act]}$. When Hh is present Ci establishes as $Ci^{[act]}$ and when Hh is not present then $Ci^{[rep]}$ is formed (Ogden et al., 2004). Both interact with *decapentaplegic (dpp)* which is part of the transforming growth factor β (TGFβ) family (Ogden et al., 2004). This family activates genes in a concentration depended manner (Ogden et al., 2004).

The *cubitus interruptus* (*ci*) gene is located on chromosome 4 in *D. melanogaster*. Some mutants of *ci* homozygous viable and display a phenotype of a missing or disrupted L4 and/or L5 vein in the wing (Locke & Hanna, 1996). Strong loss of function alleles of *ci* are embryonic lethal. The ci^1 allele is unusual in its behaviour. This is due to the observation that ci^1 under some circumstances behaves as a recessive but in certain backgrounds it displays dominance (Dubinin & Sidoroff, 1934). The circumstance in which ci^1 behaves as a dominant allele is due to a transvection effect, where the normally recessive *ci* phenotype is presented in a dominant fashion due to the disruption of normal homologue pairing (Locke & Hanna, 1996). In other words, when heterozygous (ie ci^1/+) and in the background of normal homologue pairing, the ci^1 allele behaves as a recessive. The same genotype in the background of a chromosomal rearrangement that disrupts pairing in the *ci* regulatory region has the effect of causing the ci^1 allele to display a dominant phenotype (Locke & Hanna, 1996). This phenotype is caused by a pairing dependent repression of *ci* expression in the posterior compartment of the developing imaginal wing disk (Locke & Hanna, 1996).

This effect has been used to recover translocations to the 4^{th} chromosome (Ashburner, 1989). If there is a translocation event involving a breakpoint in proximity to the *ci* locus, it will result in a lack of pairing between the two homologous copies of *ci* and thus produce a broken wing vein (radius incompletes) phenotype (Ashburner, 1989).

3.1.2 Translocations in other model organisms

A translocation is a type of chromosomal rearrangement were sections of non-homologous chromosomes are exchanged. When genes are translocated from their original chromosome it is possible to

lose various regulatory elements associated with the gene. This makes translocations a valuable tool to elucidate regulatory effects on various GOIs. The act of separating large sections of a chromosome physically can be used to analyze how various sections of that chromosome regulate GOIs. For instance, if a gene had a known large regulatory region, it is possible to separate that region from the GOI using a translocation.

The ability of CRISPR/Cas9 to create translocations in other model organisms is known. In 2019 researchers managed to reshuffle yeast chromosomes using CRISPR/Cas9 (Id et al, 2019). This experiment managed to randomly create fitness advantages under stressful growth conditions for the yeast (Id et al., 2019). This was all in effort to better understand the physical relationships of various sections of DNA. In 2016 new balancer chromosomes were engineered in *C. elegans* using CRISPR/Cas9. During these experiments it was discovered that translocations could also be made using CRISPR/Cas9 in *C. elegans* (Iwata et al., 2016).

In the past various forms of mutagens were used to induce translocations. These translocations would happen randomly and were very infrequent. Various other technologies to induce DSBs have been used to perform translocations in *Drosophila*. To-date CRISPR-Cas9 based translocations have not been observed.

3.2 Methods

3.2.1 Polytene Squash

Female 3rd instar larvae were harvested and washed within distilled H_2O. Each specimen was placed on a slide with a drop of 15% acetic acid solution. The salivary glands were removed and the remainder of the specimen was discarded. The acetic acid solution was blotted off and a drop of 1% orcein dye in 45% acetic acid solution (v/v) was added with a cover slip. The cover slip was gently slid back and forth to break up the salivary glands as well as break up any cellular membranes. Percussive force was then applied with

the a pen to the top of the cover slip in order to distribute polytene chromosomes. Pressure was applied to the cover slip without sliding the cover slip in any direction. The preparation was then sealed with nail polish and visualized using phase contrast microscopy.

3.2.2 Ci Screen

The main objective of this screen was to determine if it is possible to efficiently recover large scale chromosomal rearrangements using a CRISPR-based approach. This was done by attempting to translocate the distal arm of the X chromosome that includes the region distal to *hnt* to the 4th chromosome proximal to the *ci* gene. For this experiment, we obtained gRNA transgenes that target upstream of *ci* as well as downstream to act as a negative control and used our existing TOE-GS00052 to induce DNA breaks within the upstream regulatory region of *hnt*. We collaborated with the Harvard Medical School - specifically the *Drosophila* RNAi Screening Center and Transgenic RNAi Project (DRSC/TRiP) on this project who created the TOE lines that target *ci* for this project.

We used Cas9 to simultaneously induce DSBs upstream of both *hnt* and *ci*. At some frequency this should, in theory, induce a translocation. Taking advantage of the previously explained "Dubinin effect", we then crossed the progeny to ci^{1} mutants. Progeny harbouring the desired translocation should produce a visible phenotype that disrupts development of the L4 or L5 vein. Such translocations can be easily verified by polytene chromosome analysis.

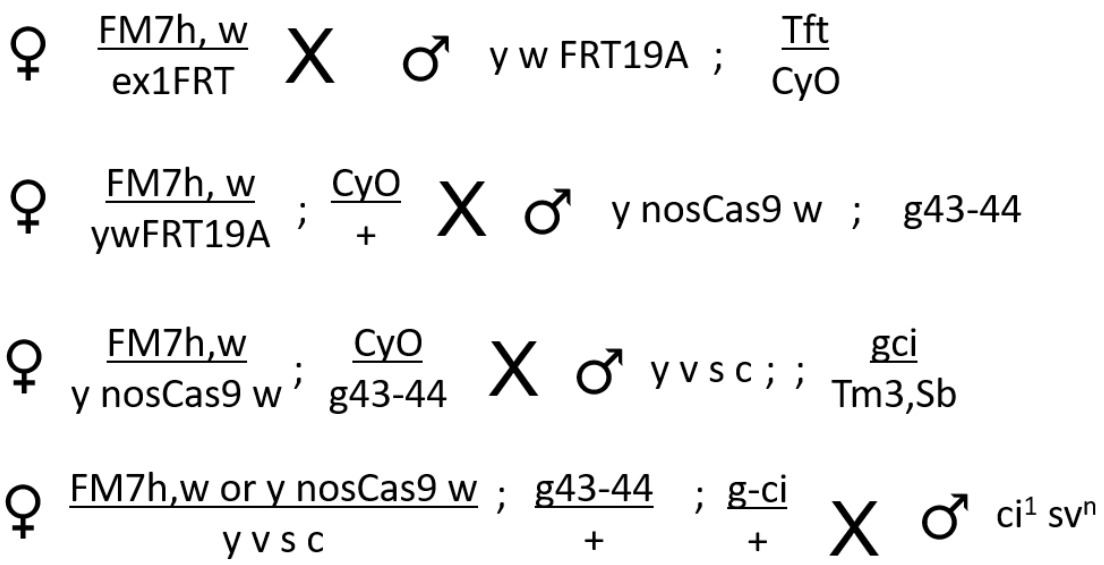

Figure 3-1 Ci Screen Crossing Scheme

Guides for both the hnt region (TOE-GS00052) and for the ci region (gci) were introduced in two separate instances using CyO and Tm3,Sb as markers respectively. The end result is that both of the guides were integrated into the same female that has the nosCas9 which allows for expression in the female germline.

3.3 Results

As shown in Table 2, after the crosses were complete, we managed to retrieve several progeny that when crossed to ci^1 males, showed a dominant ci^1 phenotype. These lines were all derived from the GP07436 guide line. These were then kept as lines and re-tested with ci^1 to see if they bred true, which they did. However, when these lines were subjected to a polythene squash test, there was no evidence of a translocation. This was true for all the lines.

ci guide used in cross	Progeny with ci^{+}	Progeny with ci^{1} phenotype
GPO7433	216	0
	67	0
	232	0
	216	0
	232	0
	317	0
GP07435	196	0
	317	0
	325	0
	67	0
	196	0
	338	0
GP07436	304	1
	116	8
	137	1
	248	12
	276	18
	304	1
	298	30
	137	1
	297	18
	298	1

Table 3-1 Summary of Translocation Screen Results

Progeny were scored based on visible ci^{1} phenotype. GP07433 is located downstream of *ci* acting as a negative control. GP07435 and GP07436 are upstream of *ci*.

3.4 Discussion

Unfortunately, no translocations were isolated using this technique. Although we recovered a ci^{11} phenotype with a loss of the L5 wing vein (Table 3-1), based on their analysis by polytene squashes none contained a translocation event. In this discussion section I explore possible reasons why this method did not work as well as highlight different adjustments to the method that may lead this approach to produce translocations in the future.

As noted previously we did see ci^{1} phenotype progeny within the screen with the GP07436 guide. This was thought to be a positive result as a ci^{1} dominant phenotype would signal a translocation due to the Dubinin effect. However, this conclusion was proven to be false based on the polytene squash data. One possible explanation for this is that the Cas9 that was guided to the *ci* region ended up creating DSBs and mutating *ci* itself due to NHEJ. This mutation could have taken the form of a small deletion that would not be detectable from cytology. A significant deletion would disrupt pairing and thus give a ci^{1} phenotype. Ultimately, we decided not to pursue the exact reasoning for having a ci^{1} phenotype as it was abundantly clear that this method was not effective in creating large scale chromosomal rearrangements which was the overall goal of the experiment.

The nosCas9 used is supposed to work by expressing Cas9 in germline cells so that the progeny are transformed instead of producing Cas9 ubiquitously. This in theory allows us to introduce the two guides separately without any Cas9 activity taking place until both guides are added. Through other side projects within the lab we have discovered that the nosCas9 is somewhat leaky. This meaning that Cas9 is not localized to the germline of the progeny and thus has the ability to create DSBs outside of its intended niche. We theorize that when Cas9 is guided to a specific site eventually enough NHEJ will be done such that the gRNA is no longer effective in directing Cas9 to the site of interest. If this was the case then by the

time the second guide was added, the other site was sufficiently mutated, and no translocations could have taken place.

Another possibility is that the translocation we were aiming to create could be associated with a dominant lethal phenotype. When a translocation takes place, multiple regulatory elements may be separated from crucial genes. The mis-regulation of such genes could lead to lethality and therefore we would not see the translocation take place.

Although many translocations have been recovered by *Drosophila* researchers over the years (Larkin et al. 2021), these were for the most part induced randomly using ionizing radiation. In such screens there would be a selection for viable translocations. It is possible that only specific translocations are allowed in the *Drosophila* genome and such attempts to isolate others would prove impossible. If we are to explore the efficacy of using Cas9 to induce a translocation, the first step may be to try to precisely re-create a known translocation. There are multiple translocations that take advantage of the Dubinin effect for screening. To create this, we would simply need replace the *hnt* guides for a guide that targets a known sight of translocation and keep the guide for ci the same.

Ultimately the reason we decided not to pursue this experiment any further in its current state was that the lack of any translocation in the number of samples taken would not be viable as a technique. The purpose of this experiment was to find a more efficient way to create translocations in a controlled manner. There are currently multiple ways to induce a translocation in the *Drosophila* system, however these techniques are extremely inefficient. The FRT based system is a way of creating translocations (Golic & Golic, 1996). FRT sites are placed randomly within the genome and using a FLPase these sites will be translocated with each other (Golic & Golic, 1996). An interesting possibility for designing large scale chromosome rearrangements would be to combine the CRISPR/Cas9 editing tool with the FRT site specific recombination. By using Cas9 to introduce cuts and providing an FRT containing template with sequences

flanking the gRNA target site, FRT sites could be delivered to any site in the genome. Controlled and selectable recombination event between any two FRT sites could then be used to generate desired chromosomal rearrangements.

Chapter 4
Future Directions

The goal of this thesis was to integrate CRISPR/Cas9 genome editing technology with traditional Drosophila genetic methods. The system of targeted gene overexpression using dCas9^{VPR} was particularly effective for driving overexpression of our GOI, *hnt*. We found that this approach could be useful for characterizing existing alleles of *hnt* as nullomorphic alleles or hypomorphic alleles. We were also able to use this system design to design a novel genetic screen for the rapid recovery of new EMS-induced alleles as an F1 visible screen. The approach was successful and resulted in the recovery of 8 new *hnt* alleles. Three new refractory lines, lines whose only apparent phenotype was not being responsive to dCas9^{VPR} mediated overexpression, were also recovered, and this was an expected side-effect of the screening strategy. A second set of experiments was aimed at the possibility of using CRISPR/Cas9 to recover large scale chromosomal rearrangements. Our strategy was to use a visible selection method to identify punitive translocations between the gene *ci* and *hnt* using what is known as the 'Dubinin effect'. This approach was not successful. During the course of these studies other Drosophila researchers have also been developing new CRISPR/Cas9 based tools, some of which are described below and could be used towards achieving one of our original aims, which was the genetic dissection of the large upstream regulatory region of the gene *hnt*.

4.1 Enhance NHEJ Pathway over HDR Pathway

In order to create a chromosomal rearrangement, the NHEJ pathway must be utilized. When a DSB takes place the two methods of repair used are NHEJ and HDR. When HDR takes place, it is more likely that the DSB is repaired normally and no chromosomal rearrangement takes place (Min et al., 2004). If NHEJ is utilized however then two blunt ends are repaired together which is how a chromosomal arrangement would take place.

Manipulation of either of these pathways could impact the effectiveness of inducing translocations. The Ku70 gene has been shown to be involved in the NHEJ pathway (Min et al., 2004). By interfering in the production of Ku subunits it is possible to reduce repair normally done through the NHEJ pathway (Min et al., 2004). When this happens, we see an increase in large and small deletions within the chromosome indicating a decrease in repair efficiency.

4.2 Large Regulatory Region

Being able to induce translocations at will within a genome could aid the analysis of cis regulatory elements. An example would be to take a GOI and translocate it away with only parts of its regulatory region. By translocating larger and larger portions of the GOI and regulatory region, it may pinpoint which regions are important to the expression of the gene. If the GOI is translocated away from an important regulatory region we should see changes in its expression.

A process has been created that serves a similar purpose called Cas9-mediated arrayed mutagenesis of individual offspring (CAMIO) (Chen et al., 2020). This is a CRISPR-based mutagenesis pipeline that creates multiple deletions in a genomic region of interest. Using an array of gRNA, it is possible to create multiple progeny with various deletions of different size (Chen et al., 2020). This is close to the effect that we would be achieving by translocating various sections of a regulatory region away. The main difference

is that by translocating elements away from a GOI we can be sure that the element has an effect on the GOI based on its position.

4.3 Balancers in Different Species

Another use of this project is to create balancer chromosomes in other *Drosophila* species. The currently available balancer chromosomes were induced through ionizing radiation and therefore the inversion locations were random. Due to their random nature, these inversions take time to isolate. By inducing inversions using Cas9 we can make more targeted inversions which opens up possibilities for creating whole new lines of balancer chromosomes in *Drosophila* melanogaster. If the Cas9 system can be adapted to other *Drosophila* species, then we can induce inversions using this system. This process can be sped up using the Cas9 system to induce specific inversions thus cutting down the development time for creating a vital genetic tool in other species.

Balancer chromosomes can be used to keep recessive lethal mutations in stock. Establishing this vital genetic tool in other species will open doors for inter-species research. One application would be to see how a mutation in a GOI can affect behaviors in various species. Being able to keep these mutations balanced in a stock in multiple species would allow other researchers to compare the effects of a gene on a genus level rather than a species level.

4.4 Insertion of FRT FLPase to Induce Inversions

One interesting idea to induce inversions is to use the CRISPR/Cas9 system to insert two FRT site into the genome and then use FLP to create a chromosomal rearrangement. Originally discovered in yeast the FLP system was adapted to *Drosophila* in 1996 (Golic & Golic, 1996). In this paper, FRT sites are randomly added into the genome and FLP is used to cause a chromosomal rearrangement between these two sites(Golic & Golic, 1996). The main limitation of this technique was the reliance on randomly inserted FRT sites. Now with the introduction of the CRISPR/Cas9 system it is possible to insert these FRT sites in

a targeted fashion by creating a DSB at a target site and allowing an oligonucleotide with the desired FRT sequence to be integrated into the genome. This system has shown that two FRT sites when exposed to FLPase and being a few kilobases away has an almost 100% success rate (Golic & Golic, 1996). Increasing this distance to several kilobases still shows a success rate of a few percent (Golic & Golic, 1996).

Another possible method for FRT insertion is the newly adapted CRISPaint [CRISPR-assisted insertion tagging] (Bosch, Colbeth, Zirin, & Perrimon, 2020). This method involves two gRNA, one targeting a plasmid and the other targeting our genomic insertion site (Bosch et al., 2020). The plasmid we would use would contain an FRT site. In this system both guides are inserted with the plasmid and Cas9. This allows both gRNA targeted sites to undergo a DSB and through NHEJ, the plasmid would integrate into the genome (Bosch et al., 2020). This would effectively create a targeted FRT site.

Appendix A
Autosynaptics

Autosynaptic chromosomes are derived when a cytologically normal chromosome goes through recombination with a chromosome that contains a pericentric inversion (Ashburner, 1989). When these chromosomes segregate from each other, it normally results in a lethal amount of aneuploidy (Ashburner, 1989). However a stock can be maintained if hereditary aneuploidy is compensated with a corresponding chromosome donated from the opposite parent. This stock is stable but can only be crossed with itself as the offspring with cytologically normal chromosomes will contain lethal amounts of aneuploidy (Ashburner, 1989).

By utilizing these stocks it is possible to create a screen where by the only possible viable progeny contain a new inversion. This new inversion would be in the autosynaptic form but it is possible through another crossover event to obtain a regular heterosynaptic chromosome. This is done by outcrossing autosynaptic females to cytologically normal males.

Detecting the Presence of an Autosynpatic Inversion

In order to detect whether an autosynaptic inversion has taken place we have decided to use a PCR method. Two sets of primers were created flanking each of the two targeted cut sites. With a cytologically normal chromosome we would expect each set of primers to make products. However, if there is an inversion we expect the forward primer of one set to create a product with the reverse primer of the other (Figure 4-1). This is a simple technique to detect the presence of an inversion without doing any sequencing. Although sequencing was done in order to verify the result (Figure 4-2).

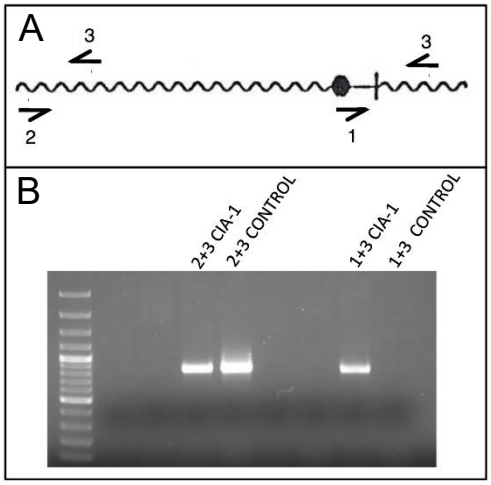

Figure 4-1 Illustrating the use of mismatched primers to interpret inversion result

Primers 2 and 3 flank a designated target site for Cas9 endonuclease activity. Primer 1 is a part of another set of primers designed for another target site on the same chromosome. (A) The result of a translocation where now 1 and 3 will produce a product where under normal circumstances they would not. (B) The result of a PCR involving a CRISPR induced allele (CIA-1) that has undergone an inversion and a wildtype control. In the wildtype control primers 1 and 3 do not create a product as they are too far away but they are brought together due to the inversion in CIA-1

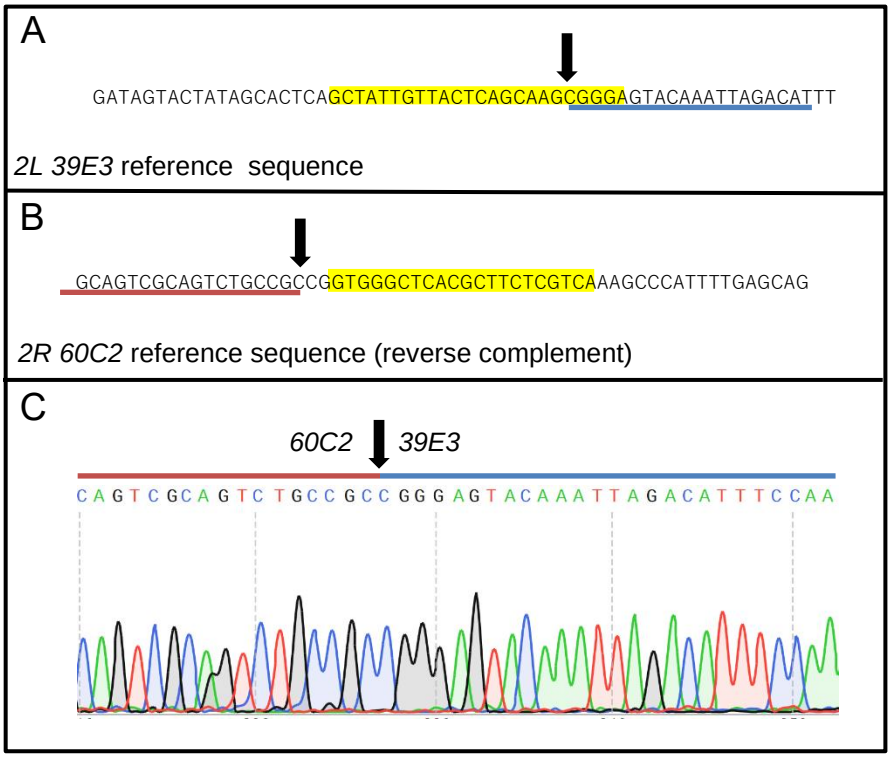

Figure 4-2 Showing inversion through sequencing data

Targeted sites for Cas9 induced inversion are highlighted in yellow. Sites that were an inversion took place is highlighted by the black arrow. (A) Shows the 2L region at 39E3 and the section that is inverted away is highlighted in blue. (B) Reverse complement of the 2R region at 60C2. (C) Sequencing results and shows the exact break point where 60C2 and 39E3 meet to form the inversion.

Appendix B

Primers used for sequencing guide region of refractory chromosome:

Hnt43-44InForw: 5’ – GGAGATTTGGTAGTCTGGCACA - 3’

Hnt43-44InRev: 5’ – AGTTGGTTTCTTGAAGCGCC – 3’

Primers used for inversion detection:

Primer 1. 5’- CAGGCACCGACAAATAAGCG – 3’

Primer 2. 5’- CGTTTCACCAACAGTACGGC – 3’

Primer 3. 5’- TCAAATTTGCCGCTCTTGCC – 3’

Bibliography

Anzalone, A. V., Koblan, L. W., & Liu, D. R. (2020). Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nature Biotechnology*, *38*(7), 824–844. Springer US.

Ashburner, M. (1989). *Drosophila A Laboratory Handbook*. Cambridge: Cold Spring Harbor Laboratory Press.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., et al. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*(5819), 1709–1712.

Bassett, A. R., & Liu, J. L. (2014). CRISPR/Cas9 and Genome Editing in Drosophila. *Journal of Genetics and Genomics*, 41(1), 7-19

Bassett, A. R., Tibbit, C., Ponting, C. P., & Liu, J. L. (2013). Highly Efficient Targeted Mutagenesis of Drosophila with the CRISPR/Cas9 System. *Cell Reports*, *4*(1), 220–228.

Bier, E., Harrison, M. M., O'connor-Giles, K. M., & Wildonger, J. (2018). Advances in engineering the fly genome with the CRISPR-Cas system. *Genetics*, *208*(1), 1–18.

Bolotin, A., Quinquis, B., Sorokin, A., & Dusko Ehrlich, S. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, *151*(8), 2551–2561.

Bosch, J. A., Colbeth, R., Zirin, J., & Perrimon, N. (2020). Gene Knock-Ins in Drosophila Using Donor Plasmids. *Genetics*, *214*(January), 75–89.

Bridges, C. B. (1919). Specific modifiers of eosin eye color in Drosophila melanogaster. *Journal of Experimental Zoology*, *28*(3), 337–384.

Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., Dickman, M. J., et al. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, *321*(5891), 960–964.

Chen, H. M., Marques, J. G., Sugino, K., Wei, D., Miyares, R. L., & Lee, T. (2020). CAMIO: a transgenic CRISPR pipeline to create diverse targeted genome deletions in Drosophila. *Nucleic acids research*, *48*(8), 4344–4356. Oxford University Press.

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., et

al. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), 602–607.

Drysdale, R. A., Crosby, M. A., Gelbart, W., Campbell, K., Crosby, M., Emmert, D., Matthews, B., et al. (2005). FlyBase: Genes and gene models. *Nucleic Acids Research*, *33*(DATABASE ISS.), 390–395.

Dubinin, N. P., & Sidoroff, B. N. (1934). Relation between the Effect of a Gene and Its Position in the System. *The American Naturalist*, *68*(717), 377–381.

Duffy, J. B. (2002). GAL4 system indrosophila: A fly geneticist's swiss army knife. *Genesis*, *34*(1–2), 1–15.

Farley, J. E., Burdett, T. C., Barria, R., Neukomm, L. J., Kenna, K. P., Landers, J. E., & Freeman, M. R. (2018). Transcription factor Pebbled/RREB1 regulates injury-induced axon degeneration. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(6), 1358–1363.

Friedmann, T., & Roblin, R. (1972). Gene Therapy for Human. *Science*, *175*(4025), 949–955.

Fujita, T., & Fujii, H. (2013). Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. *Biochemical and Biophysical Research Communications*, *439*(1), 132–136. Elsevier Inc.

Garneau, J. E., Dupuis, M. È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., et al. (2010). The CRISPR/cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, *468*(7320), 67–71.

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(39), 2579–2586.

Gibbons, M., Stratton, A., & Parsons, J. (2019). Spinal Muscular Atrophy (SMA) in the Therapeutic Era. *Current Genetic Medicine Reports*, *7*(3), 162–167. Current Genetic Medicine Reports.

Golic, K. G., & Golic, M. M. (1996). Engineering the Drosophila genome: Chromosome rearrangements by design. *Genetics*, *144*(4), 1693–1711.

Hales, K. G., Korey, C. A., Larracuente, A. M., & Roberts, D. M. (2015). Genetics on the fly: A primer on the drosophila model system. *Genetics*, *201*(3), 815–842.

Hentges, K. E., & Justice, M. J. (2004, June 1). Checks and balancers: Balancer chromosomes to facilitate genome annotation. *Trends in Genetics*. Elsevier Current Trends.

Id, F., Lu, W., & Id, N. A. (2019). Reshuffling yeast chromosomes with CRISPR /, 1–26.

Iwata, S., Yoshina, S., Suehiro, Y., Hori, S., & Mitani, S. (2016). Engineering new balancer chromosomes in C. Elegans via CRISPR/Cas9. *Scientific Reports*, *6*, 1–8. Nature Publishing Group.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA – Guided, *337*(August), 816–822.

Kaufman, T. C. (2017). A Short History and Description of Drosophila melanogaster Classical Genetics. *Genetics*, *206*(June), 665–689.

Larkin, A., Marygold, S. J., Antonazzo, G., Attrill, H., Dos Santos, G., Garapati, P. V., Goodman, J. L., et al. (2021). FlyBase: updates to the Drosophila melanogaster knowledge base. *Nucleic acids research*, *49*(D1), D899–D907. Oxford University Press.

Li, W. Z., Li, S. L., Zheng, H. Y., Zhang, S. P., & Xue, L. (2012). A broad expression profile of the GMR-GAL4 driver in Drosophila melanogaster. *Genetics and molecular research : GMR*, *11*(3), 1997–2002.

Lin, S., Ewen-Campen, B., Ni, X., Housden, B. E., & Perrimon, N. (2015). In vivo transcriptional activation using CRISPR/Cas9 in Drosophila. *Genetics*, *201*(2), 433–442.

Locke, J., & Hanna, S. (1996). engrailed gene dosage determines whether certain recessive cubitus interruptus alleles exhibit dominance of the adult wing phenotype in Drosophila. *Developmental Genetics*, *19*(4), 340–349.

Ma, D., & Liu, F. (2015). Genome Editing and Its Applications in Model Organisms. *Genomics, Proteomics and Bioinformatics*, *13*(6), 336–344. Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

Matsumoto, D., Tamamura, H., & Nomura, W. (2020). A cell cycle-dependent CRISPR-Cas9 activation system based on an anti-CRISPR protein shows improved genome editing accuracy.

Communications Biology, *3*(1), 2–3. Springer US.

McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A., & Gehring, W. J. (1984). A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature*, *308*(5958), 428–433.

Min, B., Weinert, B. T., & Rio, D. C. (2004). Interplay between Drosophila Bloom's syndrome helicase and Ku autoantigen during nonhomologous end joining repair of P element-induced DNA breaks. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(24), 8906–8911.

Muller, H. (1928). the Production of Mutations By X-rays. *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 14, p. 714).

Muller, H. J. (1932). Some Genetic Aspects of Sex. *The American Naturalist*, *66*(703), 118–138.

Nusslein-Volhard, C., & Eric, W. (1980). Mutations affecting segment number and polarity in Drosophila. *Nature*, *287*(5785), 795–801.

Ogden, S. K., Ascano, M., Stegman, M. A., & Robbins, D. J. (2004). Regulation of Hedgehog signaling: A complex story. *Biochemical Pharmacology*, *67*(5), 805–814.

Pandey, U. B., & Nichols, C. D. (2011). Human disease models in drosophila melanogaster and the role of the fly in therapeutic drug discovery. *Pharmacological Reviews*, *63*(2), 411–436.

Pourcel, C., Salvignol, G., & Vergnaud, G. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology*, *151*(3), 653–663.

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-γuided platform for sequence-specific control of gene expression. *Cell*, *152*(5), 1173–1183. Elsevier.

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., & Siksnys, V. (2011). The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic Acids Research*, *39*(21), 9275–9282.

Savić, N., & Schwank, G. (2016). Advances in therapeutic CRISPR/Cas9 genome editing. *Translational Research*, *168*, 15–21.

Schmidtmann, E., Anton, T., Rombaut, P., Herzog, F., & Leonhardt, H. (2016). Determination of local chromatin composition by CasID. *Nucleus*, *7*(5), 476–484. Taylor & Francis.

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). *NIH Image to ImageJ: 25 years of image analysis*. *Nature Methods*. 9, 671-675.

Schwartz, C., Locke, J., Nishida, C., & Kornberg, T. (1995). Analysis of cubitus interruptus regulation in Drosophila embryos and imaginal disks. *Development (Cambridge, England)*, *121*(6), 1625–1635.

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.

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.

Zhimulev, I. F., & Koryakov, D. E. (2009). Polytene Chromosomes. *Encyclopedia of Life Sciences*, (1881), 1–10.

Zirin, J., Hu, Y., Liu, L., Yang-zhou, D., Colbeth, R., & Yan, D. (2020). Large-Scale Transgenic Drosophila Resource. *Genetics*, *214*(April), 755–767.