5-2021

Molecular and Genetic Studies of robo2 Transcriptional Regulation in the Central Nervous System of Drosophila melanogaster

Muna Abdal Rahim Abdal Rhida
University of Arkansas, Fayetteville

Molecular and Genetic Studies of *robo2* Transcriptional Regulation in the Central Nervous System of *Drosophila melanogaster*

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Cell and Molecular Biology

by

Muna Abdal- Rahim Abdal- Rhida
Wasit University
Bachelor of Science in Biology, 2003
Wasit University
Master of Science in Physiology, 2010

May 2021
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Timothy A. Evans, Ph.D.
Dissertation Advisor

David S. McNabb, Ph.D.
Committee Member

Douglas D. Rhoads, Ph.D.
Committee Member

Gisela F. Erf, Ph.D.
Committee Member

Abstract

Drosophila Robo2 axon guidance receptor is a member of the evolutionarily conserved Roundabout (Robo) protein family that is involved in directing axons that cross the midline to the other side of the animal body. Robo2 roles mainly depend on two factors: The functional domains of the Robo2 protein, which is extensively studied, and the dynamic transcription of *robo2* in various subsets of cells throughout embryogenesis which is not fully understood. Thus, knowing *robo2* enhancers that transcriptionally regulate *robo2* during embryogenesis is significant. To investigate *robo2* potential enhancers, we screened 17 transgenic lines of *Drosophila* that were generated by Janelia Research Center. These lines contain 17 fragments distributed within and around the *robo2* gene. We identified six fragments that regulated *robo2* expression by the GAL4-UAS-GFP system suggesting that they were promising enhancers. Using these identified regulatory fragments in addition to three fragments generated in our lab, we built the HA-Robo2 transgenic constructs. These constructs were introduced into *Drosophila* which allowed us to test *robo2* expression and its dependent axon guidance phenotypes in the embryonic CNS. GMR28G05 and GMR28F02 fragments showed the strongest *robo2* expression in the lateral pathway. To further study these fragments, we introduced them separately or together into *robo2* null mutant background. We found that Robo2*'s* dynamic expression pattern is specified by multiple regulatory regions.

We utilized these fragments to generate and characterize an equivalent set of *robo2* transgenes expressing the axonal marker TauMyc instead of the HA-Robo2 and *hsp70* promoter instead of *robo2* promoter. The results show that GMR28F02 fragment drove strong expression of TauMyc in a subset of the lateral neurons, cell bodies, and commissural axon from which Robo2 protein is expressed.

CRISPR/Cas9 system was used to further investigate the importance of our findings. Cas9 protein and specific gRNAs were used to target and delete *robo2* potential enhancers (GMR28G05 and GMR28F02) separately or together. Applying bioinformatics tools and literature I predicted three transcription factors (Hb9, Nkx6.1, and Lhx2) that have a high probability to bind *robo2* potential enhancers.

In summary, *robo2* has potential enhancers located in the first intron and upstream of the gene, and multiple enhancers more efficiently regulated *robo2* expression in *Drosophila*.

Acknowledgments

I would like to express my deep gratitude to my academic advisor, Dr. Timothy Evans for giving me the opportunity to join his lab. His patience, motivation, enthusiasm, and immense knowledge inspired me. He provided me invaluable guidance throughout this research. It was an honor to work under his guidance and I am extremely grateful for what he has offered me. I also would like to show my sincere thanks to the rest of my dissertation committee members Dr. Douglas Duane Rhoads, Dr. David McNabb, and Dr. Erf, Gisela F. for their support and guidance.

I would like to thank CEMB program and Biological Sciences at the University of Arkansas, my sponsor MOHE Iraq, and NIH for all their supports including financial support.

I am especially thankful to my terrific husband, Ahmed for his understanding, encouragement, and support. My completion of this project could not have been accomplished without his endless support. Thanks to my wonderful children, Noor and Jaafar for bringing joy to my life. My sisters and brothers, thank you for your endorsement and support.

I am extremely grateful to my parents for their love, caring, and sacrifices to prepare me for my future.

I would like to dedicate this dissertation to my great dad who passed away a few weeks ago. May he continue to rest in heaven's peace.

Table of Contents

Abbreviations

1. Fruit fly	*Drosophila melanogaster*
2. DCC	deleted in colorectal cancer
3. VUR	vesicoureteral reflux
4. HGPPS	Horizontal gaze palsy with progressive scoliosis
5. CNS	central nervous system
6. VNC	ventral nerve cord
7. GMC	ganglion mother cell
8. MG	midline glial cells
9. ECM	extracellular membrane
10. CAMs	cell adhesion molecules
11. SDF1	stromal cell-derived factor 1
12. Ig	immunoglobulin
13. Fn	fibronectin
14. Tm	transmembrane
15. CC0-CC3	conserved cytoplasmic 0-3
16. CRISPR/Cas	clustered regularly interspaced short palindromic repeat
17. gRNA	guide RNA
18. PAM	protospacer adjacent motif
19. NHEJ	non-homologous end-joining
20. ZFNs	zinc-finger nuclease
21. TALENs	transcription activator-like effector nucleases
22. *DMD*	Duchenne muscular dystrophy gene

23. Pol I	RNA polymerase I
24. RNA PII	RNA polymerase II
25. Pol III	RNA polymerase III
26. Inr	Transcription initiator
27. Bzip	basic zipper
28. HTH	helix-turn-helix
29. bHLH	basic helix-loop-helix
30. TFBSs	transcription factor binding sites.
31. SELEX	Systematic Evolution of Ligands by Exponential Evolution
32. CHIP	Chromatin immunoprecipitation
33. PWM	position weight matrix
34. PSSM	position-specific scoring matrix
35. Ap	Apterus axon
36. rTh	rostral thalamic neurons
37. iTh	intermediate thalamic neurons
38. TCAs	Thalamocortical axons
39. Robo2	Roundabout2 protein
40. *robo2*	roundabout2 gene

Chapter one: Axon Guidance role in the central nervous system (CNS) of *Drosophila* melanogaster and Roundabout2 protein

Introduction

Significance

Animals with bilateral symmetry sense and respond to their environment through their nervous system (Comer *et al.,* 2019). During the nervous system development, the formation of a network between neuronal and non-neuronal cells is one of the most interesting processes. The key step in this process is the axons that extend precisely towards their desired synaptic targets. Researchers have extensively studied this phenomenon a few decades ago using cellular and molecular mechanisms controlling the navigation and neuronal growth (Dickson & Zou, 2010; Hidalgo & Booth, 2000; Sáchez-Soriano *et al.,* 2007; Raper & Mason, 2010; Tessier-Lavigne & Goodman, 1996). Neuronal circuit development is orchestrated by axon guidance molecules such as slit, the repellent ligand of the roundabout receptor (Robo), and netrin-1 which induces attraction by frazzled receptor (DCC in vertebrates). Individuals with CNS limitation show defects in the proteins engaged in axon guidance decisions. For example, a heterozygous mutation in DCC causes congenital mirror movements disorder. Individuals with this kind of disorder miss the ability to move both sides of the body independently, and this rare issue persists throughout the patient's life (Galléa *et al.,* 2011). Also, Robo1 disruption is implicated in communication disorder with dyslexia, a disorder characterized by trouble with reading, spelling, and pronouncing words, and writing. This disruption is caused by chromosome translocation in dyslexic individuals (Hannula-Jouppi *et al.,* 2005). Moreover, Robo2/Slit2 disorder has been involved in the vesicoureteral reflux (VUR), a complex, genetically heterogeneous developmental disorder described by the retrograding flow of urine from the bladder into the ureter (Lu et al., 2007). Furthermore, a homozygous mutation in the human *robo3* gene located on chromosome 11 causes horizontal gaze palsy with progressive scoliosis

(HGPPS) (Amoiridis et al., 2006). This mutation causes the hindbrain axons to fail to cross the midline. Individuals with Horizontal gaze palsy develop scoliosis within the first ten years of their life. (Nugent et al., 2012)

The human brain consists of around 100 billion neurons with 10,000 different cell types that have different functions such as sensory neurons, motor cells, and memory cells, making humans the most complex organism (Herculano-Houzel, 2009; Muotri & Gage, 2006). Despite being different anatomically, smaller organisms such as *Drosophila* still share several conserved cellular and genetic properties with humans (Venken et al., 2011). Being less complex than humans, *Drosophila* was utilized as one of the most powerful model organisms in biomedical science for more than a decade. The cheap, short generation time, a huge amount of progeny, and very advanced genetic tools have made *Drosophila* requisite for biological research (Tolwinski, 2017). All these qualities in *Drosophila* encouraged the researchers to use it as a model to obtain novel insight into the tools and mechanisms that can assist to suggest new treatments for various neurological diseases (Ugur *et al.,* 2016).

Development of the CNS in *Drosophila*

The early development of *Drosophila* CNS involves the differentiation of neuroblasts. Neuroblasts are produced first by splitting from a neuro-ectoderm that is housed in the ventrolateral section of the *Drosophila* embryo. They start dividing to form neurons and glia. These neuroblasts keep dividing to generate another neuroblast and ganglion mother cell (GMC). The latter (GMC) divides only one time to give rise to neurons and glial cells, the two essential cell types build up the CNS (McDonald *et al.,* 1998; Homem & Knoblich, 2012). While neurons have two processes: axons and dendrites synapsis, glial cells have only one process and no synapsis. In *Drosophila melanogaster*, the CNS contains two types of glial cells, midline glia

(MG) cells, and the lateral glia. MG cells are a subclass of neuropile glia that are expressed in the early stage of embryogenesis and reduced during the pupal stage just before turning to adult fly (Awad & Truman, 1997). They play a crucial role in embryonic signaling and wrapping the axons that cross the CNS. Once connect to the adjacent neurons, MG cells get access to perform multiple tasks, such as migration, ensheathment, subdivision of axon commissures, apoptosis, and extension of glial processes. MG cells in *Drosophila* are functionally and morphologically akin to mammalian floorplate cells (Crews, 2010) making the midline an essential point to orchestrate axon pathfinding (Crews, 2010; Kaprielian *et al.,* 2001).

Axon guidance

Information is transferred between the two halves of the nervous system through commissures, which consist of neurons that extend axons across the midline to the other side of the CNS. While some axons cross the midline once, others such as ipsilateral axons never cross the midline. Instead, they extend exclusively on their own side of the CNS (Rajagopalan, Nicolas *et al.,* 2000). At the beginning of the differentiation, neurons eject multiple projections called neurites. However, only one of these neurites develops an axon. Neurons send out their axons to reach their destination to make synaptic connections with other neuronal and non-neuronal cells. These axons can extend over very short or very long distances. So, for the axons to reach their long-distance targets, they should grow to attain this goal. The previous studies proposed two approaches for the axons to grow. The first approach suggests that axon pathways are broken into shorter intervals or segments of about a hundred micrometers long. These segments act as intermediate targets that the axons regulate the navigation by providing information that guides the axons to grow along the next segment of the trajectory (Kaprielian *et al.,* 2001). The second approach for the axon extension to a long-distance proposes that pioneer axons that project early

in the developing neurons pave the way for the later-developing axons to track the same way that the early axons established to reach their final destination (Tessier-Lavigne & Goodman, 1996). Now that these approaches simplified the task of axons navigation through a long-distance by having the long axons divided into short intervals, the question remains as to how these short intervals of axons are formed? Genetic tools, tissue culture studies, and embryological experiments show that there are four types of cues or molecules that axons respond to: short-range attractants, short-range repellents, long-range attractants (chemoattractants), and long-range repellents (chemorepellents). Roman Y. Cajal was the first one who discovered the growing tips of the axons about one hundred years ago. He anticipated that chemoattractant might mediate axon guidance decision when he proposed that target cells release chemoattractant substances to attract axons over a distance in a process similar to the chemotaxis in attracting motile cells to their targets in a response to a diffusible chemoattractant substance from the target cells (Figure 1.1).

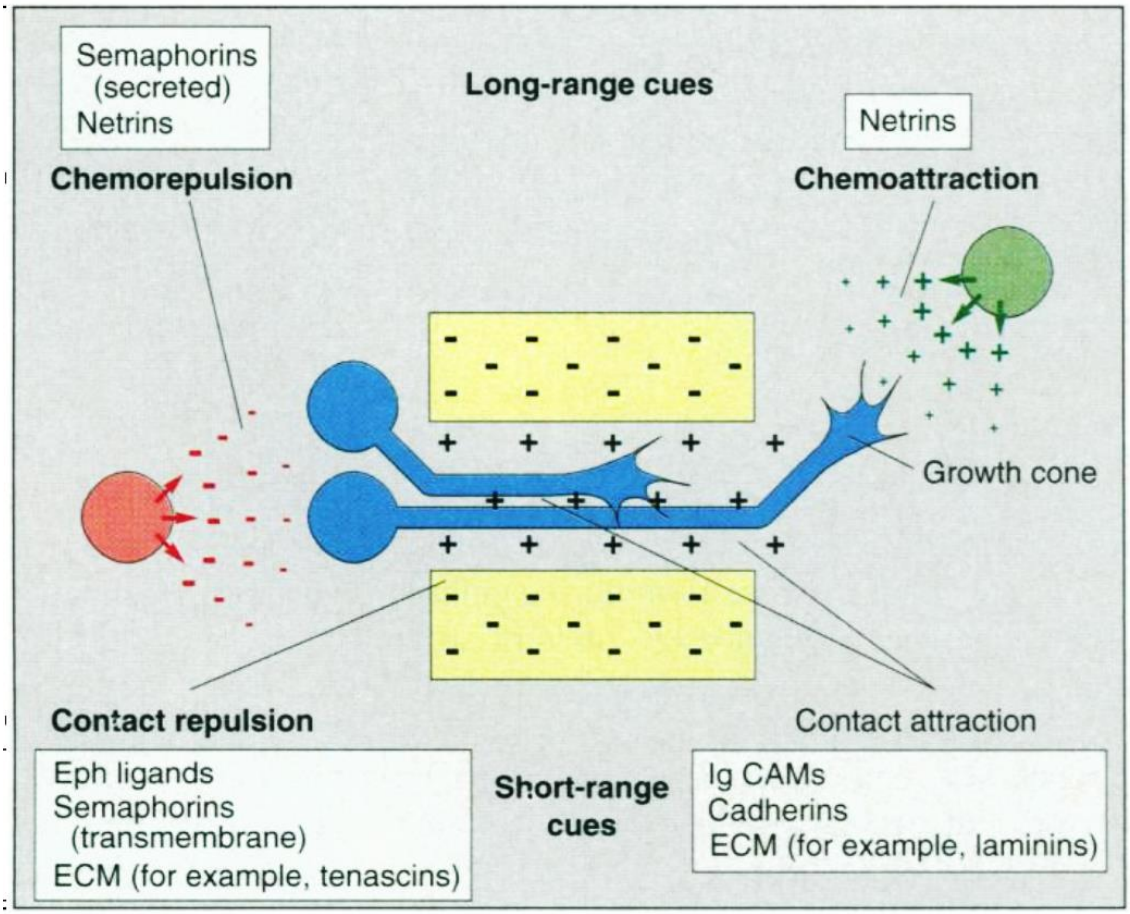


Figure 1.1. Four types of guidance forces associated with guiding growth cones: contact attraction, chemoattraction, contact repulsion, and chemorepulsion. (Top left and right), netrin ligand that is involved in both chemoattraction and chemorepulsion mechanisms. (Bottom, right), contact repulsion represented by Ephrin ligands, transmembrane semaphorins and extracellular membrane (ECM). (Bottom, left), adhesion molecules represented by immunoglobulins, cadherins, and ECMs. Adapted from (Tessier-Lavigne & Goodman, 1996).

Axon guidance molecules

Axons traverse through their environment in which different signaling molecules are displayed on or secreted into the interstitial space of extracellular matrix by neuronal and non-neuronal cells. These molecules are represented by adhesive molecules, trophic molecules, modulatory factors, tropic factors, and differentiation and morphogenic factors (Raper & Mason, 2010) (Figure 1.2).

A. Adhesive factors

For the axons to be protruded in their environment, they need suitable substrates called adhesive factors. There are two major types of adhesive cues: extracellular matrix (ECM) constituents, which are expressed on basement membrane or cellular interstices, and cell adhesion molecules (CAMs) expressed on neuronal and non-neuronal surfaces. Adhesive cells assist physical contact with the substrate required for axon extension. Laminin and fibronectin are two examples of adhesion molecules in the ECM, and cadherins and immunoglobulin family are two other examples of adhesion molecules on the cell surfaces (Raper & Mason, 2010).

B. Trophic signals

These signal molecules enhance axon outgrowth, the motility of growth cones, and neuronal survival (Reichardt, 2006); Connolly *et al.,* 1985). In vitro axonal outgrowth can be affected by the neurotrophins concentration. When they get closer to their targets, neurotrophins can behave as short-range attractants (Letourneau, 1978; Gundersen & Barrett, 1979).

C. Modulatory cues

These are signaling molecules that are not attractants or repellants on their own. However, they act by modulating the effectiveness of other guidance cues (Raper & Mason, 2010). For instance, even though these cues are not tropic, they can affect the way the axons respond to the tropic cues. Chemokines, such as stromal cell-derived factor 1(SDF1) and Laminin are examples of these types of cues. So, SDF1, or neurotrophins, for example, can reduce the axon response to the repellent Semaphorin 3A even though they are not Semaphorines on their own (Dontchev & Letourneau, 2002).

D. Tropic cues

This category of cues involves the chemotactic cues hypothesized by Roman Y. Cajal. The major roles of these cues are attraction and repulsion. These cues are responsible for creating remarkable changes in growth cone motility by influencing the cytoskeleton through intracellular signaling. Examples of these cues are netrins, semaphorins, ephrins, and slits (Luo *et al.,* 1993; Renzi *et al.,* 2000)

D.1. Netrins: They are secreted proteins that can act as both repellent and attractant in axon guidance in the midline depending on their binding to particular receptors, such as *unc-40 and unc-5* that encode the conserved transmembrane proteins (Culotti & Merz, 1998). During neurogenesis, netrins can act as either long- or short-range signals.

D.2. Semaphorins: Unlike netrins, semaphorins such as semaphorin 3A, is a secreted repellent that prevents the growing of axons from inappropriate regions. In addition to being axon guidance repellents, semaphorins act as short-range inhibitory signals. Due to their existence in the scar tissue, class 3 semaphorins play a crucial role in central nervous system injuries such as

axonal regrowth, re-vascularisation, re-myelination, and the immune response (Mecollari *et al.,* 2014).

D.3. Ephrins: They are a family of membrane-bound proteins that act as ligands of Eph receptors. Signaling of Eph/ephrin leads to the regulation of multiple biological events during embryonic development, such as cell migration, the building of tissue boundaries, segmentation, and axon guidance of growth cones. During adulthood, ephrins play a role in some processes like angiogenesis and stem cell differentiation. Based on their structure and binding to the cell membrane, ephrin ligands fall into two subclasses: ephrin-A and ephrin B. Eph receptors, in turn, are divided into two types ephAs and ephBs depending on their affinity to bind ephrin-A or ephrin B ligands (Nomenclature & Ligands, 1997).

D.4. Slit: It is a secreted extracellular matrix protein that plays an active role in neuronal development. The major function of Slit protein is acting as midline repellent by deflecting the longitudinal axons from crossing the midline of CNS. It also participates in preventing commissural axons from re-crossing the midline. The canonical receptor for Slit is Robo transmembrane proteins. In vertebrates, slit is produced by the cells in the floor plates, while in insects, including *Drosophila*, slit protein is produced by midline glia. Pioneer axon guidance depends mainly on Slit/Robo signaling (Farmer *et al.,* 2008).

E. Morphogenic and differentiation factors

These are signaling molecules that stimulate the cells to produce specific responses according to their local concentration. During the early development, morphogens are produced as gradients by specific cells and release through the tissue around these cells. These gradients are responsible for driving the differentiation of stem cells into different cell types. Eventually,

forming all the body tissues and organs. The majority of morphogens are secreted proteins that act as signaling molecules between the cells. However, some of them diffuse in the early embryonic stages such as transcription factors in *Drosophila melanogaster*.

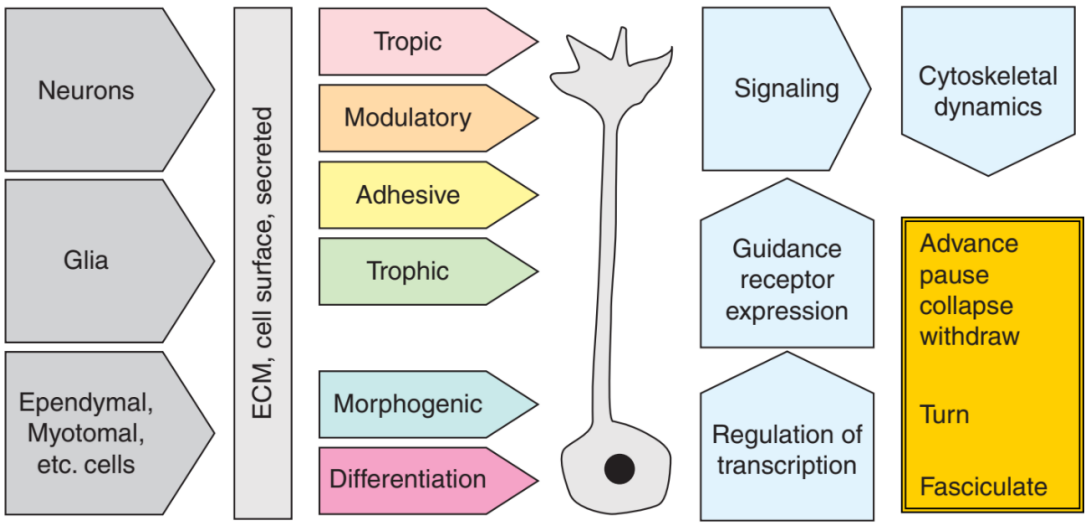


Figure 1.2. Axon guidance impacted by environmental cues: Different signaling molecules are released or displayed on the surface of the neuronal and non-neuronal cells where axons moving through. These molecules include morphogenic and differentiation factors that play important role in neurons determination, trophic, adhesive, modulatory, and tropic factors that affect the growth cones based on the guidance receptors and signaling components found in the growth cone. Growth cones motility is affected by signaling pathway resulted from specific guidance cues and according to these signaling pathways growth cones, eventually respond differently by pausing, collapsing, withdrawing, turning, or fasciculation with different axons. This image is adapted from (Raper & Mason, 2010).

Growth cones

During the nervous system development, the neuronal network establishment requires extending the axons toward their suitable postsynaptic target cells. The process of extending neurons to their axons occurs at the growth cones of the axons distal tips (Dent *et al.,* 1999; Dent & Gertler, 2003; Lowery & Vactor, 2009; Sánchez-Soriano *et al.,* 2010). Growth cones are large extensions developing or regenerating neurites looking for their appropriate synaptic targets supported by actin (Vitriol & Zheng, 2012). The morphology of growth cones looks like a human hand, it consists of two types of cytoskeletal elements which, in turn, lies in two regions, the organelle-rich central domain, which consists of microtubules, and the organelle-poor peripheral domain where actin filaments predominate and form two types of extensions: filopodia and lamellipodia. Filopodia are fine and pointed extensions, they are also called microspikes. They consist of bundles of actin filaments (F-actin) that provide the growth cones with support and shape. Filopodia are bound by a membrane rich with receptors and cell adhesion molecules which are necessary for axon growth and guidance. Unlike filopodia, lamellipodia are flat regions of actin meshwork. lamellipodia are located between two filopodia. Due to their quite dynamic merits, growth cones can respond to the surrounding environmental stimuli by branching towards the stimuli and switching their directions (Dent & Gertler, 2003)(Figure 1. 3).

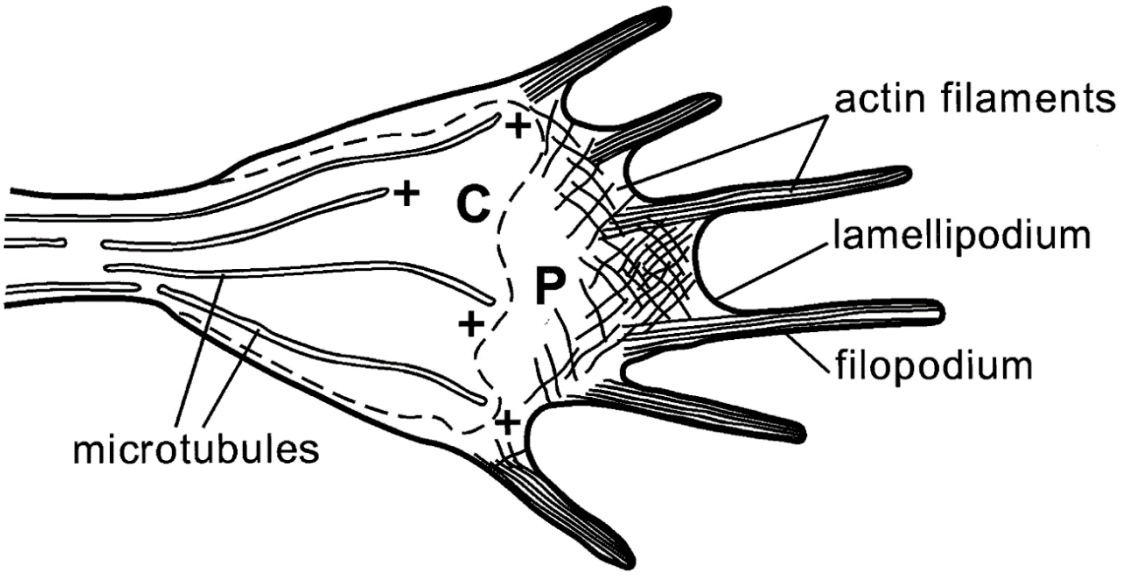


Figure 1.3. Neuronal growth cone. Growth cone is the motile human hand-like structure that navigate through the environment in response to specific cues. It consists of two cytoskeletal elements, which are distributed in two regions, the C region represents organelle-rich central domain, which involves microtubules that extend to the base of the filopodia, and P region that represents the organelle-poor peripheral domain where actin filaments predominate forming filopodia and lamellipodia. This figure is adapted from (Mueller, 1999).

Robo family

Drosophila genome screen for mutant genes that play a key role in axon guidance revealed the discovery of a special group of proteins called a roundabout. Roundabout or Robo is a family of transmembrane protein receptors that are highly conserved among many animal taxa. This family obtained its name (roundabout) from the circular-like traffic junctions, a phenotype that resulted from axon crossing and improperly re-crossing the midline in *Drosophila robo* mutant (Evans & Bashaw, 2012; Kidd *et al.,* 1998). Robo receptors play a pivotal role in nervous system development. Phylogenetic analysis shows that Robo receptors have developed from certain ancestral proteins with multiple variations occurred over time in various lineages. The *robo* gene was first discovered in *Drosophila* and has been cloned in different species such as humans and mice. In invertebrates, such as *Drosophila*, the Robo family consists of three Robo receptor members: Robo1, Robo2, and Robo3. (Rajagopalan, Vivancos *et al.,* 2000; Simpson, Kidd *et al.,* 2000). Whereas in vertebrates, such as humans, the Robo family consists of four Robo receptors: Robo1, Robo2, Robo3/Rig-1, and Robo4/magic roundabout. In *Drosophila*, *robo1* is located on chromosome 2R while *robo2* and *robo3* are located on the left arm of the second chromosome (2L). In humans, *robo1* and *robo2* are located on chromosome 3p123, while *robo3* and *robo4* are placed on chromosome 11 p24.2.

Robo protein structure

Drosophila Robo family three members (Robo1, Robo2, and Robo3) share a 5+3 conserved ectodomain protein structure of five immunoglobulin-like (Ig) domains and three fibronectin type III-like domains (FN). The cytoplasmic portion in Robo2 and Robo3 is shorter and consists of only two cytodomains (CC0, and CC1), while Robo1 consists of four cytodomains (CC0, CC1, CC2, and CC3). The cytoplasmic motifs of the Robo family are

required for interactions with downstream effectors. (Kidd *et al.,* 1998; Bashaw *et al.,* 2000) (Figure 1.4).

Robo2

Robo2 is a transmembrane protein that belongs to an evolutionarily conserved protein family called a roundabout. This family plays an active role in axon guidance. As a transmembrane protein, Robo2 structure includes mainly extracellular portion consists of five immunoglobulin-like domains and three fibronectin-like repeats, transmembrane, and a cytoplasmic domain made up of two cytoplasmic conserved motifs (CC0 and CC1). During the early stages of axon trajectory, Robo2 is expressed in the ipsilateral pioneer axon where it mediates midline repulsion. Robo2 also inhibits slit Robo1 repulsion by being expressed in the midline cells. So, Robo2 promotes midline crossing (Evans *et al.,* 2015). Furthermore, Robo2 is expressed in the lateral regions of the neuropile in the late stages of nerve cord development to enhance the formation of longitudinal axon pathways. The distinct roles of Robo2 depend in part on different functional domains within its receptor protein, and on its dynamic expression in different subsets of cells during embryogenesis in another part. Robo2 Ig1 domain is the canonical Slit ligand-binding domain, which is essential for midline repulsion. Ig1 and Ig3 domains of Robo2 affect Robo2's ability to control the lateral positions of longitudinal axons while the Ig2 domain is required for promoting midline crossing (Evans & Bashaw, 2010b). RNA seq data show that *Drosophila* Robo2 is expressed at a high-level in CNS and imaginal disc. In *Drosophila*, *robo2* expression levels are more likely organ-dependents. Studies showed that *robo2* has moderate expression in the fat body while it has a low expression in the head, salivary gland, digestive system, testis, and accessory gland. Like *robo2* expression in *Drosophila*, *robo2* expression in humans is organ dependent. It has been found that Robo2 was

expressed in a high level in the fetal brain, endocrine tissues, lung, kidney, urinary bladder, and lower in bone marrow, salivary gland, thyroid gland, skin, adipose and soft tissues, pancreas, liver, gall bladder, gastrointestinal tract, eye, and muscle tissue. (Yue *et al.,* 2006; Sasaki *et al.*, 2020)

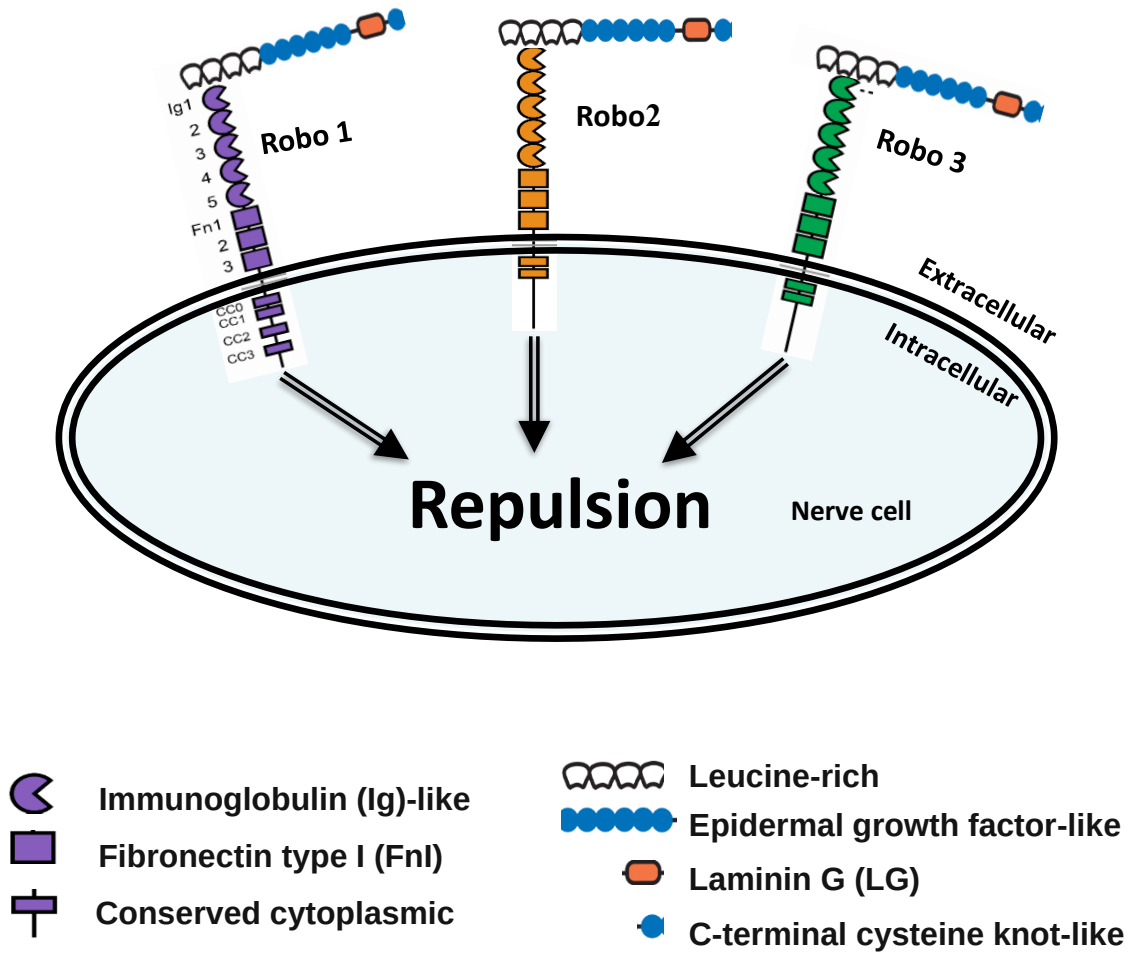


Figure 1.4.A schematic shows three robo family members (Robo1, Robo2, and Robo3) each of which consists of the 5+3 conserved ectodomain, transmembrane domain, and cytodomains, which are two in both of Robo2 and Robo3(CC0, and CC1)), and four motifs in Robo1(CC0, CC1, CC2, and CC3). Slit binds to the first immunoglobulin-like domain (Ig1) of the robo receptors to activate axon guidance repulsion.

The expression patterns and mutant phenotypes of the *Drosophila* Robo family

All Robo proteins are expressed in the embryonic CNS during the period of axon outgrowth. Kidd *et al.,* 1998a shows that Robo1 is expressed in the CNS during the formation of the pathways in a consistent pattern. While Robo2 is expressed transiently during the embryogenesis stages, Robo2 is extensively expressed during stage 12. However, its expression becomes more restricted at later stages to a specific cell subset. Robo3, in turn, is the last Robo family to be expressed in the CNS. It starts its expression at stage 13 of embryogenesis and this expression is restricted to a subset of cells since the beginning of its expression. Compared to *robo2*, *robo3* shows more expression than *robo2* in the later stages of embryogenesis. While *robo1* and *robo2* show a high degree of similarity in their expression patterns in the periphery, *robo3* shows different peripheral expressions. *robo3* is exclusively expressed in PNS neurons, while *robo2* is expressed in the epidermis as stripes in addition to the developing trachea, dorsal vessel, and muscles.

All Robos are highly expressed in the longitudinal growth cones. However, their expression is very low in the commissural growth cones. By the end of the developmental embryogenesis, each one of the three Robos emerges in a particular pattern. For example, Robo1 is expressed in the entire width of the longitudinal tracts, *robo2* shows expression on axons in the lateral two-thirds, and Robo2 is expressed on axons in the lateral one-third of the longitudinal tracts (Figure 1.5). According to their expression patterns, the three Robos divide the longitudinal tracts into three zones: a medial zone, which consists of only *robo1* expression, an intermediate zone composed of both *robo1* and *robo3*, and a lateral zone, which consists of all the three Robos (Rajagopalan, Vivancos *et al.,* 2000).

Robo1 plays a major role in regulating axon guidance in the midline. In *robo1* mutant, axons cross the midline repeatedly without staying at the midline each time they cross it (Simpson, Kidd *et al.,* 2000) (Figure 1.6, B). In *robo2* mutants, the lateral longitudinal pathway merges into the intermediate pathway in addition to the defect this mutation causes in the midline crossing represented by crossing some axons the midline at a level lower than the one that has seen in *robo1* mutant (Figure 1.6, C). In *robo3* mutants, the intermediate longitudinal pathway shows shifting medially to be fused with the medial pathway. However, the mutant phenotype does not show any midline crossing in *robo3* mutant (Figure 1.6, D) (Rajagopalan, Vivancos *et al.,* 2000).

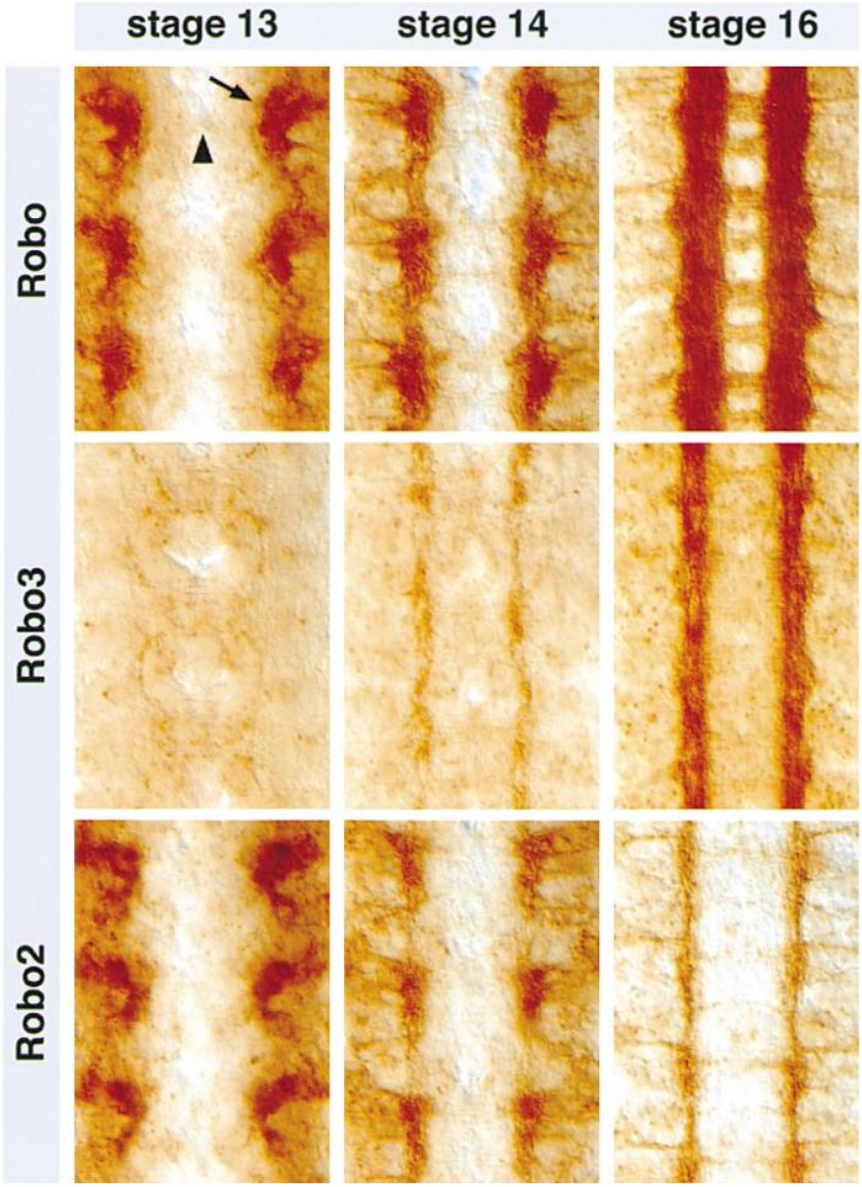


Figure 1. 5. Robo1, Robo2, and Robo3 partition the longitudinal tract into three zones. Stage 12 to early stage 13 of *Drosophila* CNS of wild type embryos. *robo1* and *robo2* share the same pattern of expression at stage early 13 with high level of protein expression in the longitudinal pathway (arrow) and no protein detectable on axons of the growth cones of commissure (arrowhead). The second panel shows Robo3 in older embryos than Robo1 and Robo2 panels. The expression of *robo3* is very low at this stage. At later stages of embryogenesis robo2 becomes more restricted to be expressed in fewer axons but its expression pattern is still close to Robo1. Robo3 is more detectable in stage 14. Stage 16 shows the three *robos* expression in the longitudinal pathways in three distinct but overlapping lateral zones. (Adapted from (Rajagopalan, Vivancos, *et al.*, 2000)).

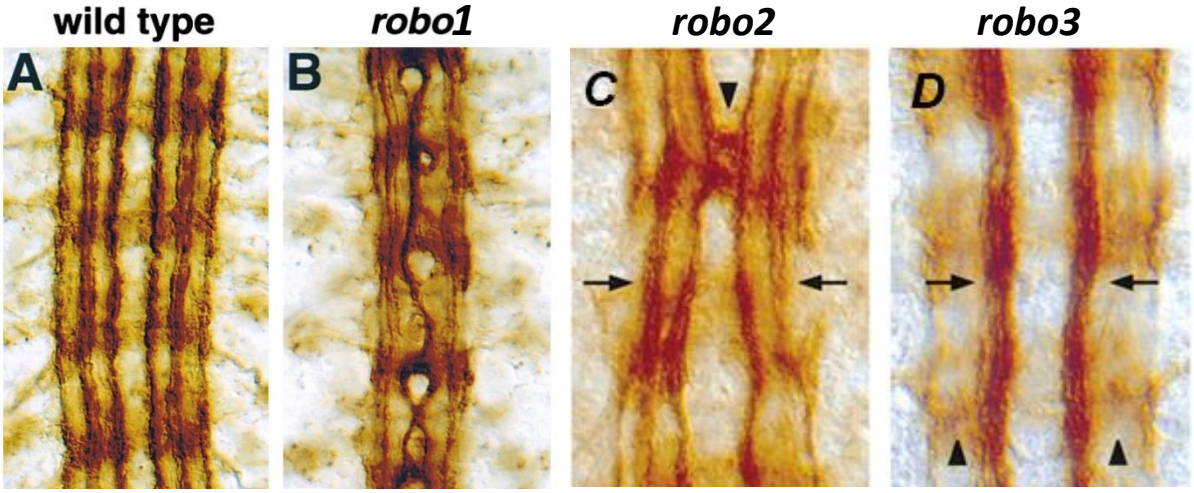


Figure 1.6. *robo* mutants. A. Wild type *Drosophila* embryo at stage 16 stained by anti-FasII antibody 1D4 shows FasII three longitudinal pathways on each side of the midline. B. *robo1* mutant shows ectopic crossing and re-crossing the midline by medial FasII pathway. However, the intermediate and lateral pathways are not affected by this mutation and they are more resemble to the wild type. C. *robo2* mutant. The lateral pathway (the third fascicle) fused with the second pathway (the intermediate) (arrows) and arrowhead shows the midline crossing. D. *robo3* mutant embryo shows shifting of the second pathway (arrowheads) to be merged with the medial pathway (arrows) and there is no midline crossing. (A and B images adapted from (Simpson, Kidd *et al.*, 2000), and C and D adapted from (Rajagopalan, Vivancos *et al.*, 2000).

Chapter Two: Transcriptional regulation of the *robo2* gene in the *Drosophila* Central Nervous System (CNS)

Abstract

Robo2 is a transmembrane protein that regulates three distinct roles in directing axon guidance in *Drosophila* central nervous system (CNS). These various roles of Robo2 depend in part on its distinct functional domains and on the dynamic transcription in different subsets of cells through embryogenesis in another part. The Robo2 structural and functional domains in *Drosophila* were studied extensively in the previous years. However, little is known about the transcriptional regulation of *robo2*. This chapter focuses on testing multiple regulatory regions around the *robo2* gene that represent potential enhancers for *robo2*. We utilized 17 transgenic lines generated by Janelia GAL4 research campus in addition to three transgenic lines generated in our lab where each transgenic line possesses one of the potential enhancers (fragments) of *robo2*. Two of these fragments (GMR28GO5 and GMR28FO2) drove strong expression of *robo2* in subsets of the lateral pathway of the *Drosophila*'s Ventral Nerve Cord (VNC), while the other two fragments (GMR28B05 and GMR28E07) enhanced *robo2* expression in the midline glia. Noticeably, individual fragments were not sufficient to fully rescue *robo2* expression in *robo2* null mutant background. Although one of these fragments (GMR28G05) rescued the midline repulsion of the *robo2* in a mutant background, its expression was not enough to rescue the lateral pathway defect. In this study, we found that a combination of two fragments is enough to rescue *robo2* expression in both midline and the lateral pathway.

Introduction

The nervous system of vertebrates and non-vertebrates is essential for processing information and sensing the environment through the basic five senses. Any developmental defects in the central nervous system (CNS), including proteins involved in axon guidance, can cause neurological dysfunction such as neurodegeneration, epilepsy, stroke, dementia, traumatic brain injury, and brain tumors. Despite anatomical differences in the CNS of vertebrates like humans and invertebrates like *Drosophila*, they still share several conserved cellular and genetic properties. For instance, there are around113 different types of neurons have been found in *Drosophila* visual system, and a very similar number of these neurons have been identified in vertebrates (Fischbach & Dittrich, 1989; Dacey & Packer, 2003; Venken *et al.,* 2011). However, there is a huge number of neurons can be found in vertebrates that are missing in flies. This reduction in complexity and the feasibility of studying the nervous system in *Drosophila* helped the researchers to use *Drosophila* as a model to obtain novel insight into the tools and mechanisms that can assist to suggest new treatments for various neurological diseases (Ugur *et al.,* 2016).

During development, the nervous system relies on axon guidance cues to decide whether axon fibers cross the midline to the other side of the nervous system or stay on the same side of the body. These cues can act as attractants or repellents, and axonal growth cones play an important role in directing these axons to their targets. In *Drosophila*, a transmembrane protein family can act as receptors to the repellants. *Drosophila* has three transmembrane proteins called roundabout (Robo) proteins (Robo1, Robo2, and Robo3), all of which are responsible for axonal midline repulsion. However, one of these proteins, Robo2, has three distinct roles. In addition to

axonal repulsion, Robo2 can mediate pro-crossing the midline and promotes the formation of the longitudinal pathway (Rajagopalan, Vivancos *et al.,* 2000; Simpson, Bland *et al.,* 2000)

Robo2 is a member of an evolutionarily conserved transmembrane protein family called roundabout that acts as axon guidance receptors. Robo2 protein consists of two main portions. The extracellular portion consists of five immunoglobulin-like (Ig) domains and three fibronectin-like (Fn) domains, while the cytoplasmic portion consists of two short, conserved motifs (CC0 and CC1) (Figure 2.1). During the early stages of axon trajectory, Robo2 is expressed in the ipsilateral pioneer axon where it mediates midline repulsion. Robo2 also inhibits slit Robo1 repulsion by being expressed in the midline cells. Furthermore, Robo2 is expressed in the lateral regions of the neuropile in the late stages of nerve cord development to enhance the formation of the longitudinal axon pathways. Robo2 has been studied in the previous years; however, it is not fully understood how Robo2 can regulate diverse axon guidance outcomes. The distinct roles of Robo2 depend in part on different functional domains within its receptor protein, and on its dynamic expression in different subsets of the cells during embryogenesis in another part.

As our understanding increases about the role of enhancer elements in gene expression, the notion of gene regulation becomes more sophisticated. Even though these regulatory elements were described previously, the mechanisms by which these regulatory elements work still controversial. Genetic variation leads to phenotypic variants and potential diseases. These variations are located outside of the genes and more likely in the cis-regulatory elements. Therefore, understanding these mechanisms is significant in this field (Pennacchio *et al.,* 2013).

Transcriptional enhancers are short cis-acting DNA sequences that control gene expression (Benabdallah & Bickmore, 2016). The human genome contains approximately hundreds of

thousands of enhancers. Enhancers can function in a position and orientation-independent manner from their target promoters. Although they regulate genes in cis, enhancers can be in different positions. They can be found upstream, downstream, or within the introns of a gene. Some enhancers were found very close to their target promotors, such as SV40, which was the first discovered enhancer. SV40 enhancer is 72 bp tandem in a sequence that is located ~100 bp upstream of the viral core promoter (Benabdallah & Bickmore, 2016). Studies showed that deletion of SV40 enhancer can eliminate gene expression and loss the viral activity (Benoist & Chambon, 1981; Gruss *et al.,* 1981). Other enhancers were found far away from their promoter, such as but not limited to β-globin (LCR), which can be found tens of kilobases from its target promoter, while Shh limb enhancer (ZRS), developmental enhancer, can be located hundreds or 1000 kb from its target gene (Figure 2.2). This chapter focuses on transcriptional regulation of *Drosophila robo2* to address the question of how individual genes can lead to different axon guidance outcomes. To examine the location of the regulatory regions of the *robo2* gene, I tested genetic fragments (potential enhancers) in the first intron and upstream of the *robo2* gene that could be potentially involved in the regulation of *robo2* expression. *Drosophila* transgenic lines called Janelia GAL4 lines were used in this experiment. These lines have been generated by Janelia Research Campus to investigate the location of enhancers in many neuronal genes including *robo2*. Seventeen of these lines have been created by cloning fragments of *robo2* and these fragments are inserted upstream of the yeast transcriptional activator GAL4 (Pfeiffer *et al.,* 2008). Each one of the seventeen lines was crossed to another transgenic line containing a GAL4 responsive GFP (UAS-TauMycGFP) to create a protein reporter system (Figure 2.3).

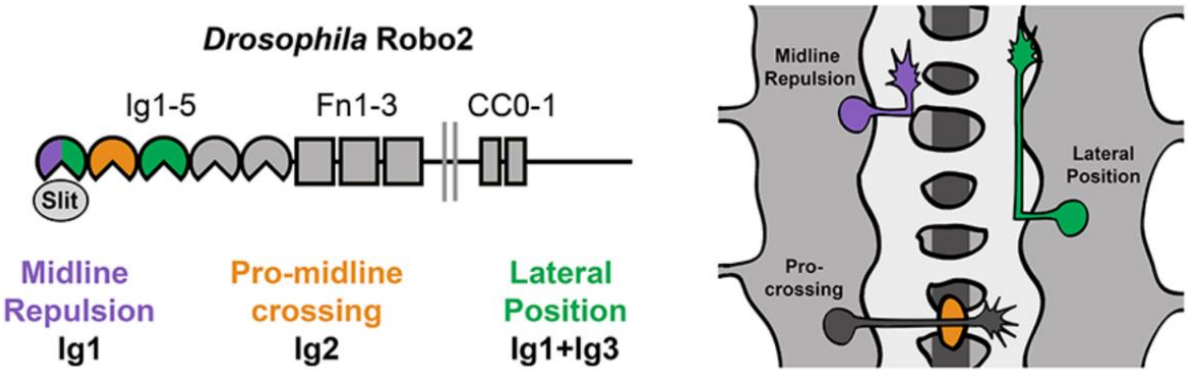


Figure 2.1: Structure and function of Robo2 receptor. Left panel, Robo2 schematic structure, Robo2 is a transmembrane protein consists of extracellular portion contains five Ig-like domains and three fibronectin-like domain, and cytoplasmic portion consists of two short, conserved motifs (CC0, and CC1). Right panel, schematic of the nerve cord of *Drosophila*. It shows the three distinct axon guidance decisions controlled by *robo2* in preventing midline crossing by binding to slit (purple), mediating longitudinal pathway formation in the lateral neuropile (green), and antagonizing slit-Robo1 repulsion and promoting midline crossing(orange).

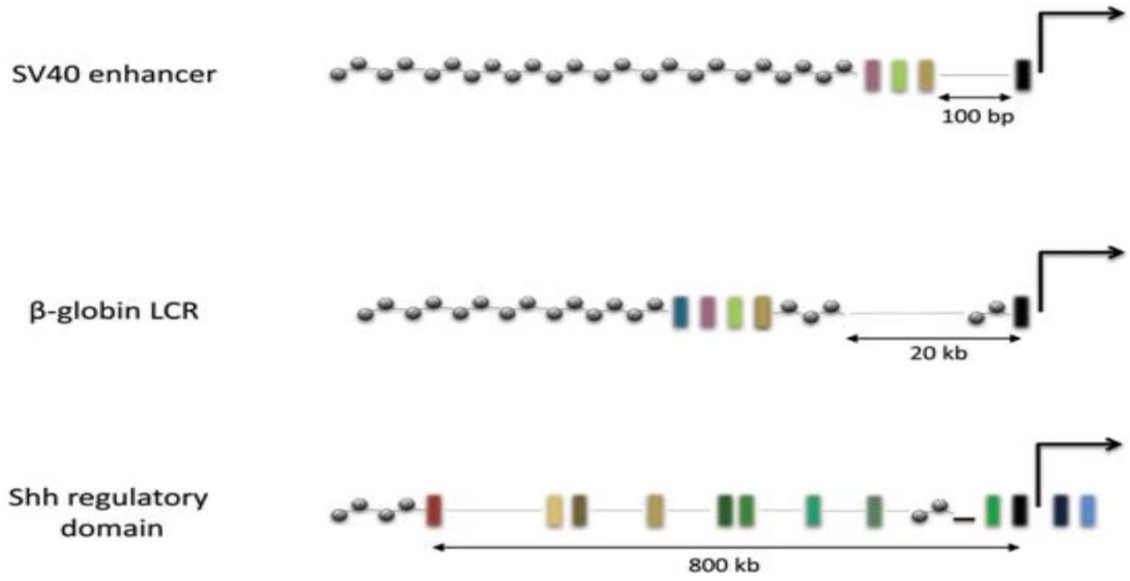


Figure 2.2. Different distances between enhancers and their promoters. (Top) SV40 early promoter as an example of an enhancer that is located a few tens of bp from their target promoter. (Middle) β-globin locus control region (LCR) as an example of enhancers that are found tens of kilobases from their target gene. (Bottom) Shh limb enhancer (ZRS) as an example of developmental enhancers that are located hundreds or 1000 kb from their target promoter. This picture is adapted from (Benabdallah & Bickmore, 2016)

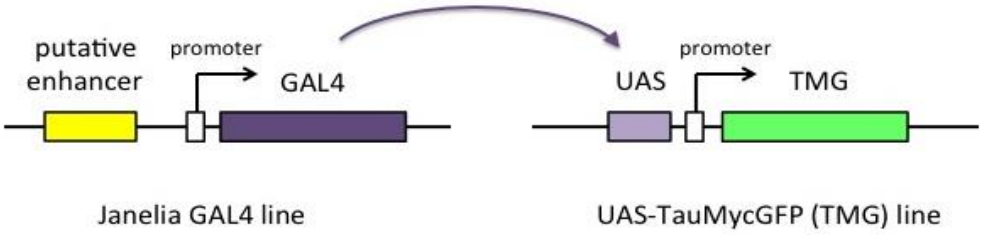


Figure 2.3: Janelia GAL4 lines. A schematic shows crosses between UAS-GFP and Janelia GAL4 lines containing *robo2* fragments. The expression patterns of the GFP is directed by the presence or absence of the enhancers in each *robo2* fragment. The *robo2* transcription factors in the CNS bind to the *robo2* putative enhancers present in each fragment upstream of the GAL4. As GAL4 protein produces, it binds to UAS enhancer and promotes GFP expression. The GFP expression reflects the expression governed by *robo2* transcriptional elements.

Material and Methods

Molecular biology

Rescue construct synthesis: All Janelia GAL4 fragments were amplified by using specific primers (**Table 4)** and fly genomic preparations using Phusion Flash PCR Master Mix (Thermo Scientific), Both PUAST vector and Janelia fragments were cut by AscI restriction enzyme. Gibson Assembly synthesis reaction or Gibson assembly Ultra Kit (SGI-DNA, Inc. La Jolla, CA 92037 USA, Cat no.:GA1200-10) were used to fusing Janelia fragments with the P10-UAST vector backbone to form a rescue construct vector, which is transformed into competent E. coli cells and recovered through mini preparation (Qiaprep spin miniprep kit (250), Cat. no. 27106). All the constructs were confirmed by PCR followed by sequencing to ensure that all the fragments were inserted correctly into P10-UAST vector. Rescue constructs including Janelia GAL4 fragments and *robo2* HA-tagged cDNA were inserted at the attB site located at 28E7 position on the left arm of chromosome 2. The transgenic lines resulted from this cloning carrying rescue constructs were crossed to Sco/CyOwg flies which are utilized as genetic balancers.

Immunohistochemistry

Drosophila embryos were collected, dechorionated, and fixed as previously described in (N. H. Patel, 1994) and stored in methanol at -20°C. Fixed embryos were rehydrated in PBT (1XPBS, and 0.1% triton) and blocked in PBT+ 5% NGS for 30 minutes at room temperature. Primary antibody was added and incubated at 4 °C overnight, then washed and incubated with PBT for 30 minutes at room temperature. Secondary antibody was added and incubated for 1 hour at room temperature and washed 3 times with PBT (1XPBS, and 0.1% triton) and then rinsed with PBS 10% and mounted in 70% glycerol/PBS. *Drosophila* embryos were genotyped

by using balancer markers or using epitope-tagged transgenes. Embryos were dissected to get the nerve cords of desired genotypes in different stages. Nerve cords were mounted in 70% glycerol/PBS. Leica SP5 confocal microscope was used for imaging and Fiji/Image J and photoshop software (Schindelin *et al.,* 2009) were used to process the fluorescent confocal stacks.

Combining GMR28F02 and GMR28G05 fragments and Generation of CRISPR-modified alleles recovery

GMR28F02 and GMR28G05 fragments were amplified from genomic DNA preparation using specific primers (Table 4) and amplified fragments were extracted from the gel and purified using (QIAquick Gel Extraction Kit (250) Cat. No. 28706). Gibson assembly master mix used to assemble these fragments in P10-UAST vector backbone that has HA-tag to form a rescue construct plasmid. The assembled vector was transformed into competent E. coli cells and recovered through minipreparation (QIAprep Spin Miniprep Kit (250), Cat. No. 27106). The construct was confirmed by PCR and sequencing. This rescue construct was inserted at attB site located at 28E07 on the left arm of chromosome 2 of wild-type fly. The G0 flies from this cloning were crossed to Sco/CyOwg to make balanced stocks. Flies from this combination were crossed to *robo2* null mutant flies. The G0 flies from this cross were individually crossed as adults to Sco/CyOwg. F1 males were then crossed individually to Sco/CyOwg virgin females. After 3 days, F1 males were removed from crosses and tested by PCR with primers to confirm any changes in the *robo2* locus. F2 progeny from positive F1 crosses were used to generate balanced stocks.

Results

robo2 fragments as putative enhancers reported by GFP expression

Embryos resulted from crossing the 17 lines of *robo2* Janelia GAL4 lines to the UAS-GFP lines were stained by anti-GFP antibody to label GFP product translated under the control of the *robo2* putative enhancers. All the 17 lines were screened by scoring the lines according to the expression pattern of the GFP in and outside of the CNS, in the Ventral Nerve Cord (VNC), and the expression in multiple cell-types of interest that normally express Robo2: commissural axons, midline cells, and longitudinal axons (Figure 2.4). Preliminary results show that six of the seventeen lines were identified to drive the GFP expression in different axons of VNC. Two of these 6 fragments drove the expression of GFP in the lateral axons of VNC (GMR28F02 and GMR28G05). Therefore, GMR28F02 and GMR28G05 were selected as putative enhancers driving Robo2's lateral pathway. Two other fragments drove the GFP expression in the midline glial cells (GMR28E07 and GMR28B05). So, they were selected as putative enhancers driving *robo2* pro-crossing role and considered fragments that have a potential role in midline repulsion (Figure 2.5). Worth mentioned that two other fragments GMR28C04, GMR28D10 were noted to drive GFP expression; However, the expression pattern was hard to define accurately (Figure 2.5).

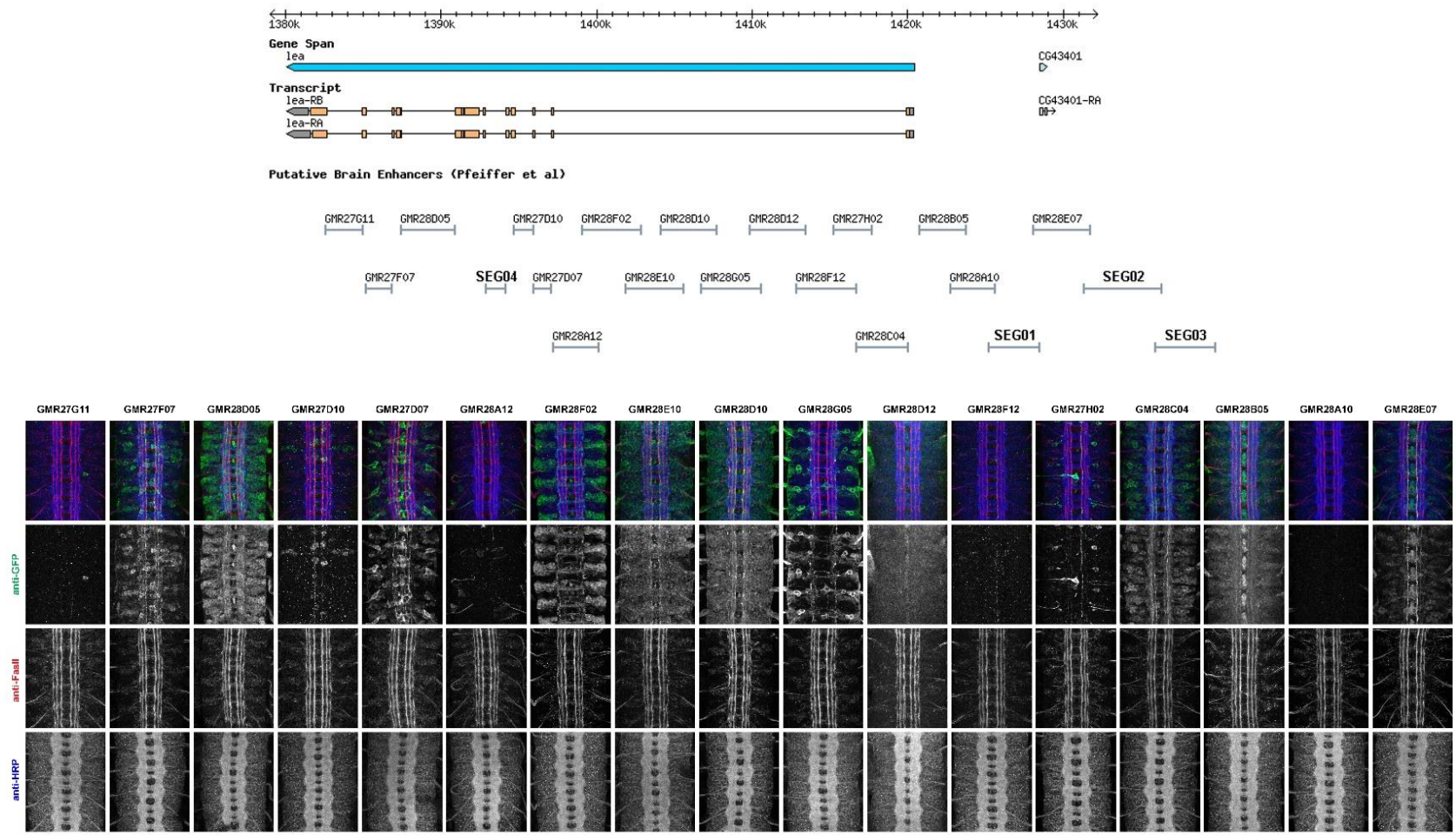


Figure 2.4. Top, *robo2* gene and 17 fragments that represent putative enhancers for the robo2 generated by Janelia Research Campus, in addition to 3 fragments generated in out lab (SEG01, SEG02, and SEG03). **Bottom**, Reflective patterns of *robo2* expression resulted by crossing Janelia GAL4 lines to UAS-GFP line. A series of transgenic fly strains were generated by Janelia Research Campus in which the DNA sequences of the *robo2* gene were cloned upstream of GAL4 transcriptional activator. Anti GFP antibody was used to detect the GFP expression to be used as a readout to determine which fragment indicate each aspect of *robo2* expression.

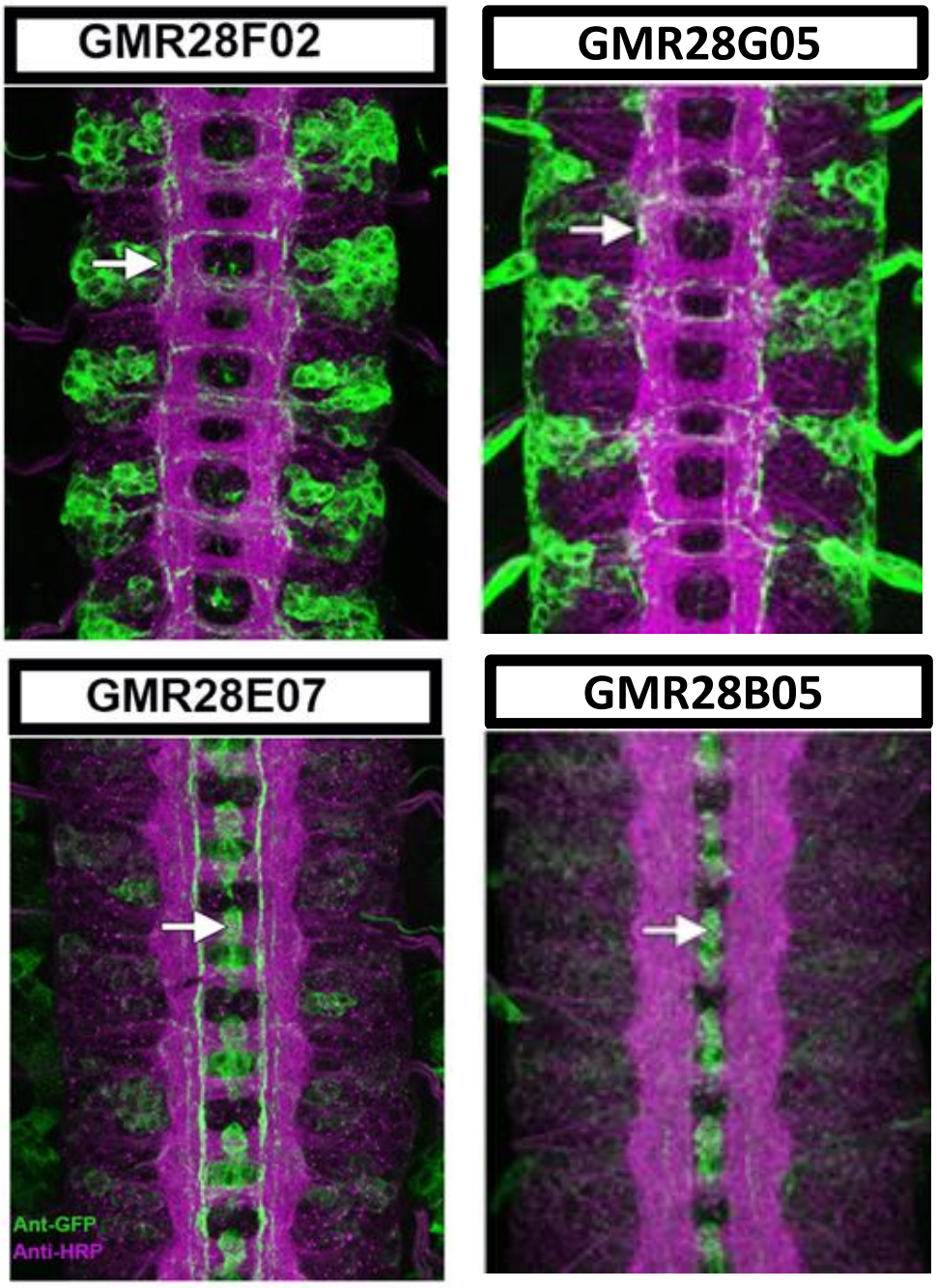


Figure 2.5.GFP expression. **(Up)** GMR28F02 and GMR28G05 lines. Embryos were stained with anti-HRP and anti-GFP, GFP is expressed in the longitudinal axons (arrows). (**Down**) GMR28E07 and GMR28B05 lines. Embryos were stained with anti-GFP and anti-HRP. GFP is expressed in the midline glia (arrows).

Generating and characterizing Robo2 rescue transgenes expressing an HA-tagged *robo2* cDNA under the control of candidate enhancer regions

From the above initial screening of a collection of non-coding DNA fragments in the first intron of *robo2* suggested that six fragments *(*GMR28F02, GMR28G05, GMR28E07, GMR28B05 GMR28C04, and GMR28D10) have the potential to serve as enhancers for *robo2* and four fragments (GMR28F02, GMR28G05, GMR28E07, and GMR28B05) among these six fragments were stronger putative enhancers. To closely characterize the *robo2* expression, all the above six fragments were introduced into the HA-Robo2 reporter transgene (Figure 2.6), (Appendix 5). The transgenic *robo2* expression was examined by using an anti-HA antibody. The results suggested that four of these fragments: GMR28F02, GMR28G05, GMR28C04, and GMR28D10 drove the *robo2* expression in different subsets of the longitudinal pathway in stage 16-17 of embryogenesis (Figure 2.7). On the other hand, two fragments (GMR28E07 and GMR28B05) did not show the *robo2* expression in the longitudinal pathway at stages 16-17 while they drove the expression of *robo2* in midline glia at stage 13 of embryogenesis (Figure 2.8).

Figure 2.6. A schematic of the *robo2* rescue construct showing the HA-Robo2 cDNA and the inserting site for Janelia GAL4 fragments (GMR28F02, GMR28G05, GMR28C04, GMR28D10, GMR28B05, and GMR28E07).

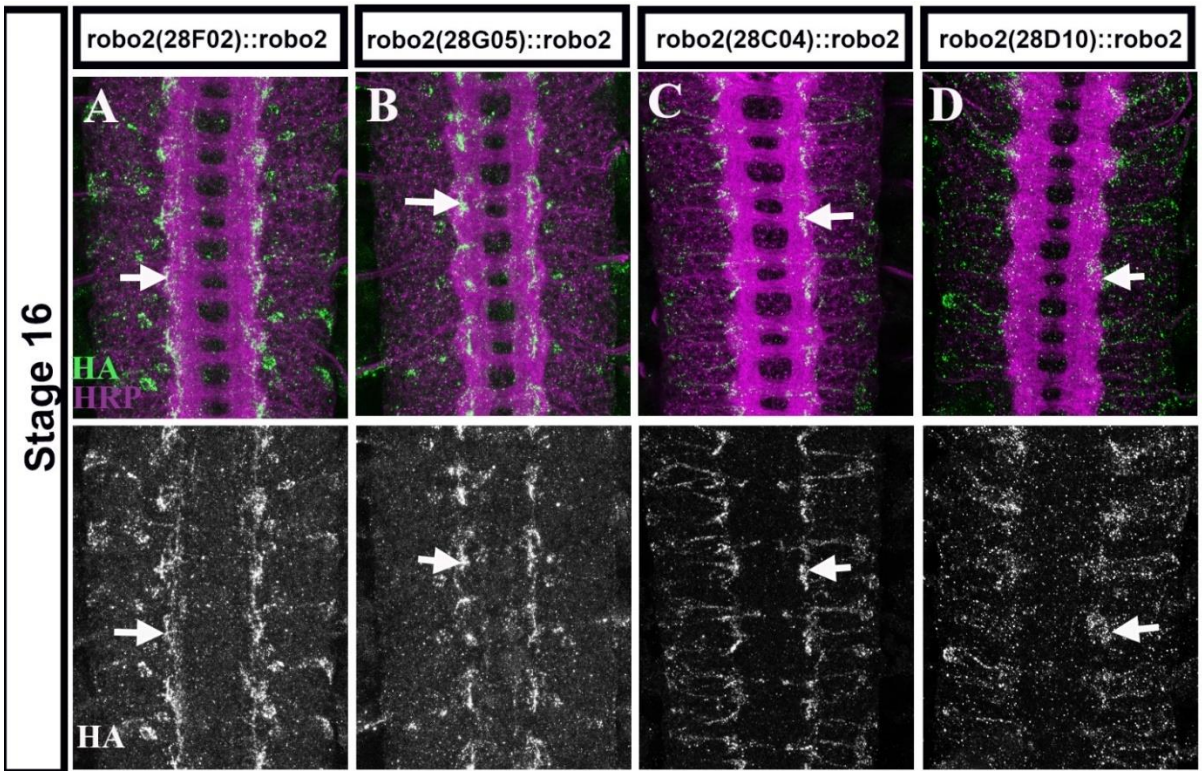


Figure 2.7: A-D. (GMR28F02)::*robo2*, (GMR28G05)::*robo2*, *(*GMR28C04)::*robo2*, and *(*GMR20D10)::*robo2* transgenic lines showing *robo2* expression in the longitudinal pathway in stage 16 embryos stained with anti-HA antibody (green), and anti-HRP antibody (magenta). Lower images show anti-HA antibody only (arrows) of the same embryos.

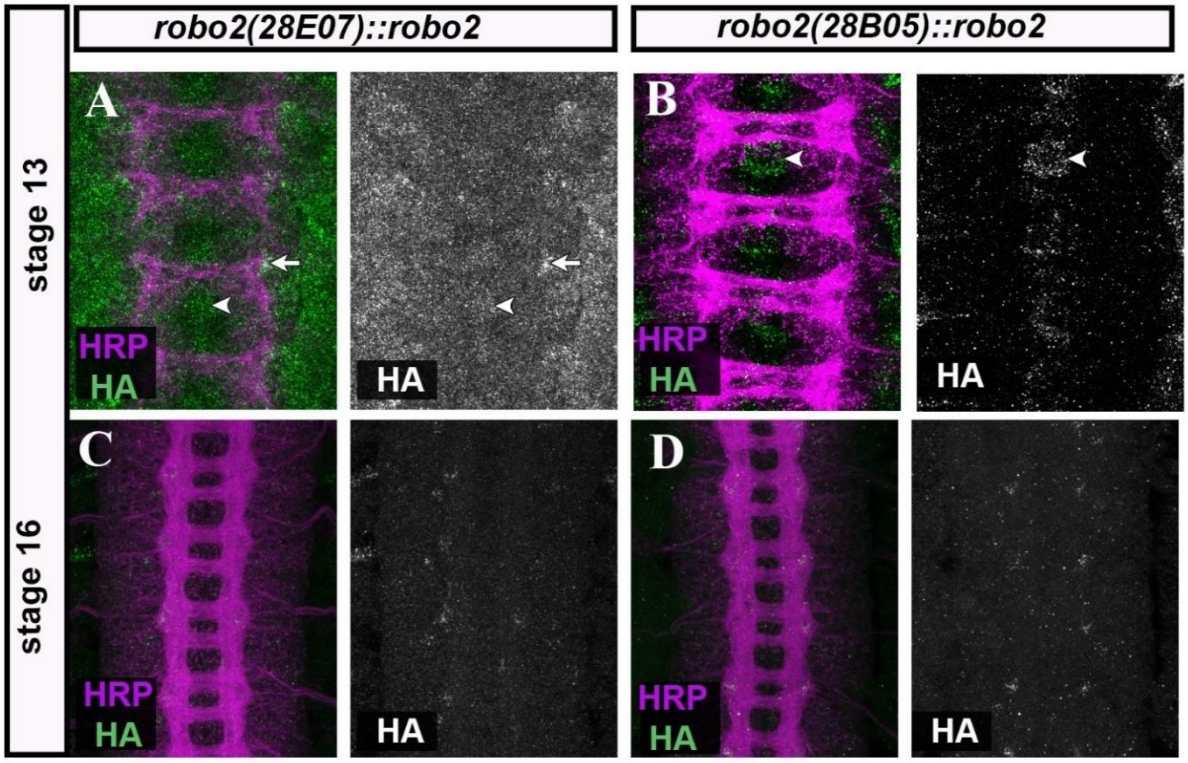


Figure 2.8.A and B. Embryos stage 13 of (GMR28E07)::*robo2* and (GMR28B05)::*robo2* expression showing midline glial cells stained by anti-HA antibody(green) (A and B, arrowheads) and some of the longitudinal axons (A, arrow). Anti-HRP antibody (magenta) stained *Drosophila* nerve cord. Lower images (C and D) show stage 16 embryos of (GMR28E07)::*robo2* and (GMR28B05)::*robo2*, no expression of *robo2* in the lateral pathway. Right images of A-D represent HA channel alone.

Characterize additional potential regulatory regions not included in the original Janelia GAL4

In addition to the DNA fragments holding suggested putative regulatory regions in the first intron of the *robo2* gene, there are potential enhancers situated upstream of the *robo2* DNA sequence, these regions are not part of the Janelia GAL4 lines. To cover these promoter-proximal regions, three additional DNA fragments (SEG01, SEG02, and SEG03) located˜20 kb upstream of the *robo2* promoter were cloned by introducing them separately into HA–Robo2 reporter transgene(Figure 2.9) (Appendix 5). The expression of these transgenes was characterized as described in figures 2.10 and 2.11. SEG01 does not show any *robo2* expression in the lateral pathway at stage16. However, it shows a slight expression at midline glia at stage 13 of embryogenesis. SEG02 and SEG03 on the other hand show *robo2* expression in different subsets of the lateral longitudinal neurons at stage 16 of the embryogenesis (Figure 2.11, A and B).

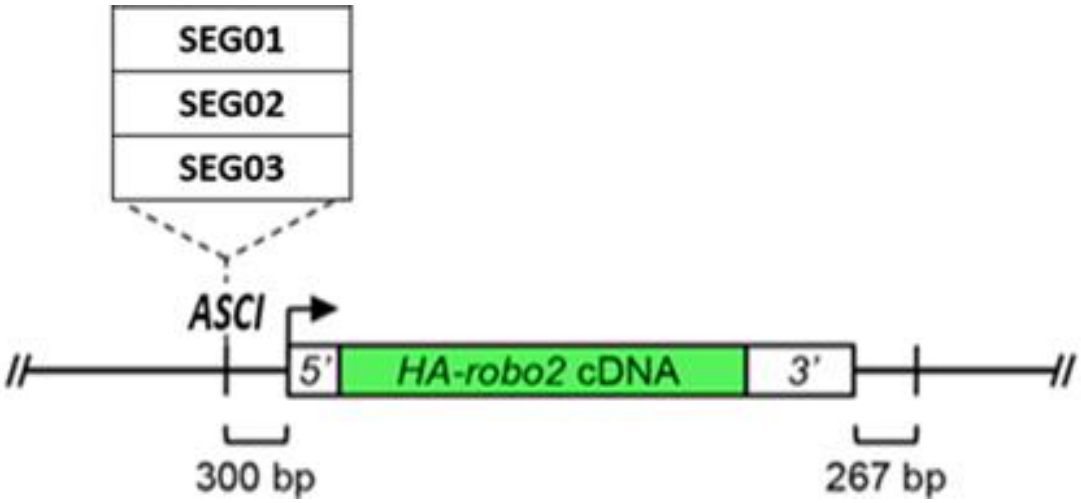


Figure 2.9. A schematic of the *robo2* rescue construct showing the HA-Robo2 cDNA and the inserting site for non-Janelia GAL4 fragments that are located upstream of the *robo2* DNA sequence (SEG01, SEG02, and SEG03)

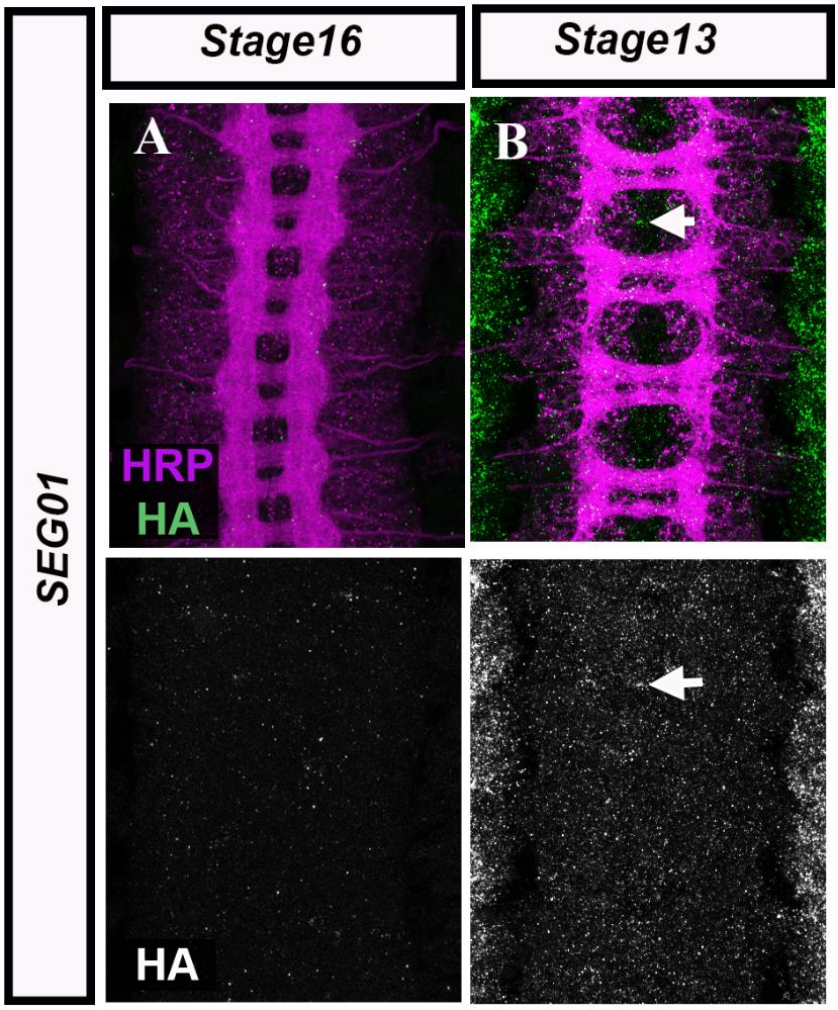


Figure 2.10. (SEG01)::*robo2*. **A.** Stage 16 embryo stained with anti-HRP antibody(magenta), and anti-HA antibody(green) showing no expression of *robo2* in the lateral pathway at this stage of embryogenesis. **B.** Stage 13 shows a little (arrow) or no expression for the *robo2* at this stage in the midline glia. Lower images **s**how HA staining channel only for the both stages.

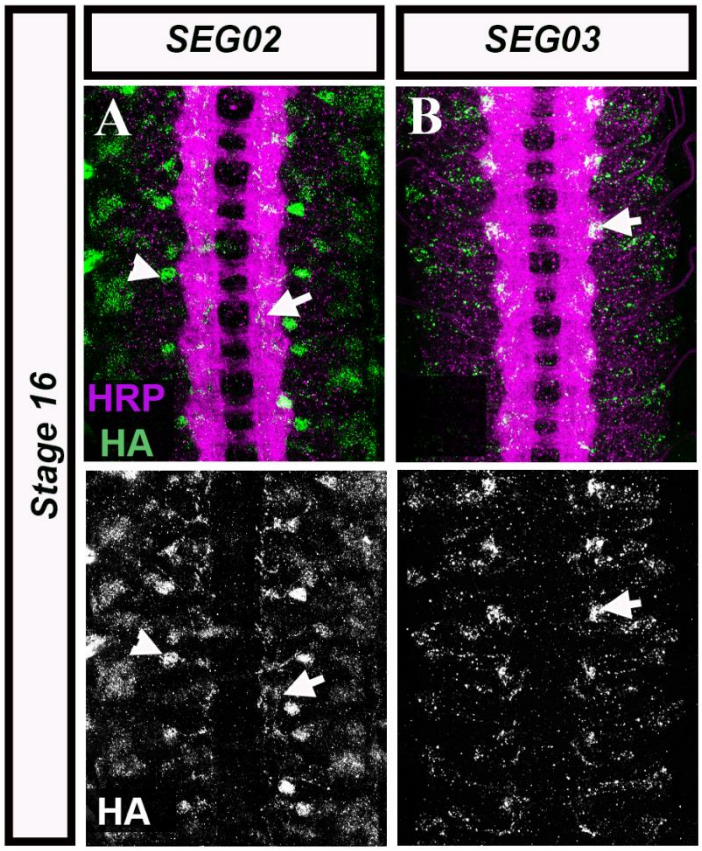


Figure.2.11. *robo2* expression patterns driven by SEG02 and SEG03. **A.** (SEG02)::*robo2* shows *robo2* expression in a subset of the longitudinal pathway (A, arrow) and some neurons outside of the nerve cord (A, arrowhead) stained with anti-HA antibody. The nerve cord is stained with anti-HRP antibody(magenta), **B.** (SEG03)::*robo2* shows *robo2* expression in a subset of the lateral pathway (B, arrow) using anti-HA antibody(green), and anti-HRP-antibody(magenta) to stain Robo2 and the nerve cord, respectively. Lower images represent HA staining channel only.

Compare exogenous HA-Robo2 expression with the full *robo2* expression (endogenous) and compare that with the Janelia GAL4 lines for each enhancer region

To investigate whether the putative enhancers would drive *robo2* expression as strong as the native enhancer (not well understood yet), GMR28F02 and GMR28G05 fragments, gave strong expression in subsets of the *robo2*'s longitudinal pathway, were chosen for this comparison. Since we do not have a good Robo2 antibody, and even if we had one, it would not work with this experiment since the rescue transgene is also expressing Robo2. To overcome such obstacles, the CRISPR-Cas9 gene-editing tool has been used to generate a line where Robo2 was labeled with Myc tag. Flies that have Myc-Robo2 were crossed to the ones that have GMR28F02::HA-Robo2 and GMR28G05::HA-Robo2 separately (**Appendix 4.4**). Embryos were collected from the progeny of this cross were stained with anti-HA and anti-Myc antibodies. The results showed that some neurons are labeled with HA and Myc antibodies which indicated that both HA-Robo2 and Myc- Robo2 were expressed. Noticeably, HA-Robo2 and Myc- *robo2* expressions were found in subsets and the entire lateral pathways, respectively (Figure 2.12). To further investigate whether these two fragments (GMR28F02 and GMR28G05) can rescue the *robo2* full expression in the lateral pathway, a recombinant line was made by crossing GMR28G05 and GMR28F02 transgenes separately to a CRISPR allele *robo2* $^{\text{myc-}\,robo2}$ line and their progeny {((GMR28G05):: *robo2*/ *robo2* $^{\text{myc-}\,robo2}$, and (GMR28F02):: *robo2*/ *robo2* $^{\text{myc-}\,robo2}$)} were crossed with each other. To make sure that both HA- and Myc- Robo2 were existed in the same fly, PCR reactions with specific primers were conducted. Positive embryos were collected and stained with anti-HA and anti-Myc. The results showed that HA still labels some of the Myc axons but not all of them. These findings suggested that the presence of one copy of each fragment on each allele is not enough to fully rescue the *robo2* expression (Figure 2.12, C).

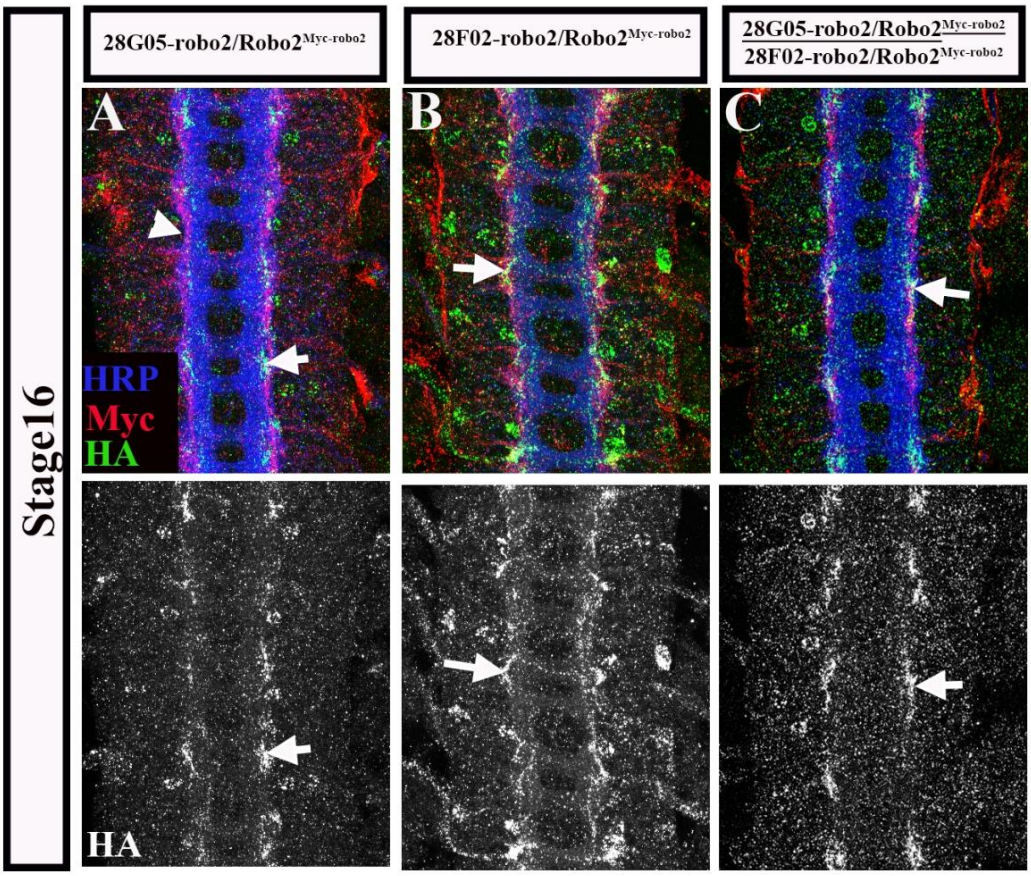


Figure 2.12. Compare the endogenous *robo2* expression with the exogenous HA-tagged *robo2*. **A.** Stage 16 embryo of (GMR28G05):: *robo2* / *robo2* $^{\text{Myc-Robo2}}$ shows anti-Myc antibody labeling all the neurons that normally express *robo2*, and anti-HA antibody represents the exogenous *robo2* driven by GMR28G05 fragment, labeling a subset of Myc (A, arrow show the region where both Myc and HA are labeled in the same embryo. (A, arrowhead shows only Myc axons) **B**. Stage 16 embryo of (GMR28F02):: *robo2* / *robo2* $^{\text{Myc-Robo2}}$ shows anti-Myc antibody labeling the neurons that normally express endogenous Robo2, and anti-HA antibody labeling the exogenous Robo2 driven by GMR28F02 fragment showing a subset of Myc labeled (B, arrows). **C.** Stage 16 embryo of genotype [[(GMR28G05):: *robo2*], [*robo2* $^{\text{Myc-Robo2}}$]/ [(GMR28F02):: *robo2*], [*robo2* $^{\text{Myc-Robo2}}$] shows the endogenous and exogenous *robo2* expression resulted by a single copy of each of GMR28G05 and GMR28F02 fragments. Lower images show HA channel only.

Then we asked whether the *robo2* expression driven by the rescue constructs have identical or different patterns than the ones which led by the original Janelia GAL4 lines. The analysis of 17 transgenic lines generated by Janelia GAL4, where GAL4 was expressed under the control of the *robo2* DNA fragments sequence revealed that some of these fragments drove the expression in the longitudinal pathway and some of them showed the expression in the midline glial cells. (GMR28F02):: *robo2*, (GMR28G05):: *robo2*, (GMR28C04):: *robo2*, and (GMR28D10):: *robo2* of the rescue construct showed distinct expression patterns in different subsets of the longitudinal pathway at stage 16 of the embryogenesis and this is consistent with what Janelia GAL4 lines showed (Figure 2.13). On the other hand, while Janelia GAL4 lines GMR28E07 and GMR28B05 showed expression in the midline glia at the stage16 of the embryogenesis, rescue constructs transgenes of the same fragments did not show the same expression at that stage. However, stage 13 embryos of these fragments showed the same expression in the midline glia in the rescue constructs transgenes.

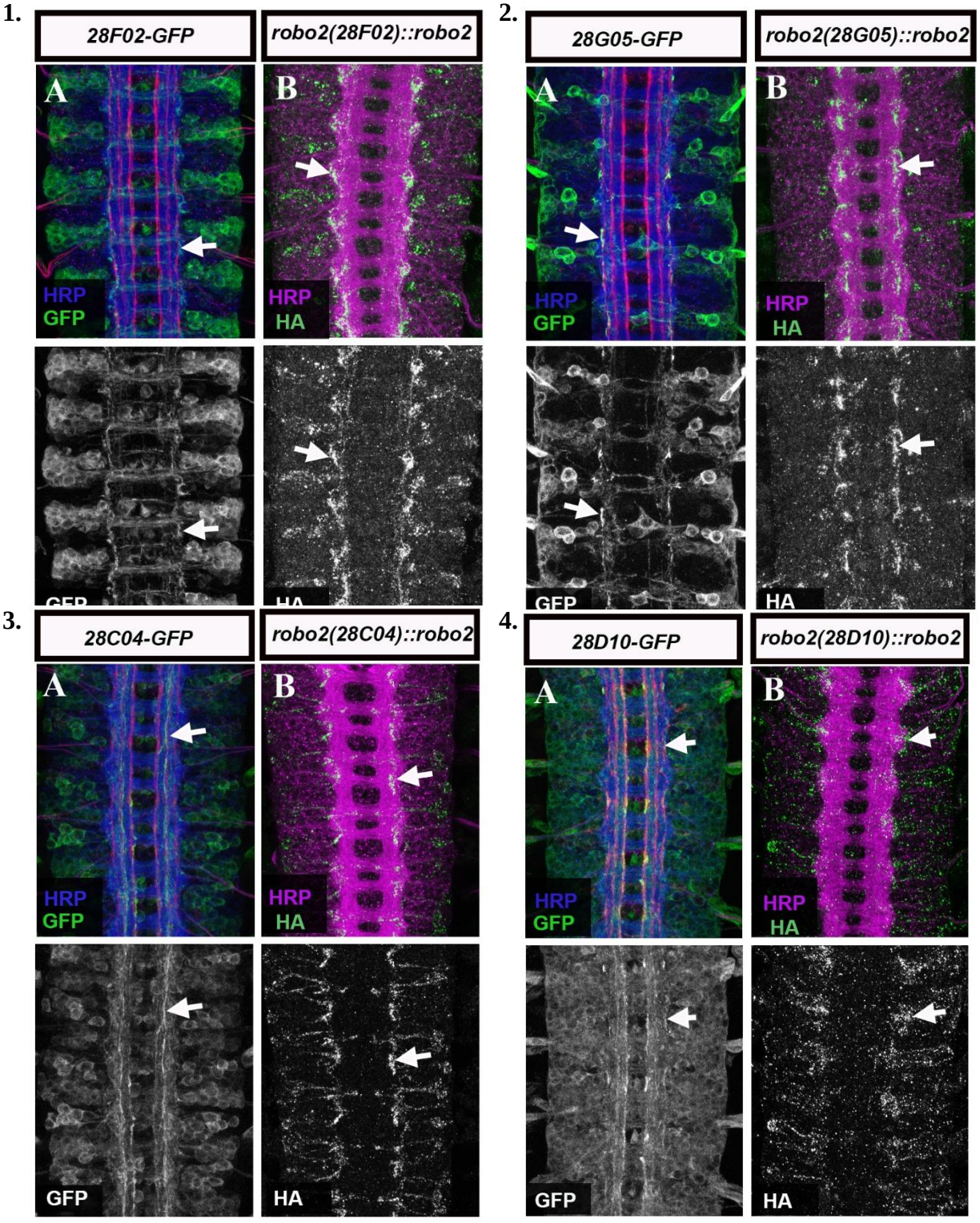

1.
28F02-GFP
robo2(28F02)::robo2
A
B
HRP
GFP
HRP
HA
GFP
HA
2.
28G05-GFP
robo2(28G05)::robo2
A
B
HRP
GFP
HRP
HA
GFP
HA
3.
28C04-GFP
robo2(28C04)::robo2
A
B
HRP
GFP
HRP
HA
GFP
HA
4.
28D10-GFP
robo2(28D10)::robo2
A
B
HRP
GFP
HRP
HA
GFP
HA

Figure 2.13.

Figure 2.13. (Cont.) Comparison of *robo2* expression driven by the DNA fragments in the original Janelia GAL4 and the rescue construct transgens.**1-4.** *robo2* expression driven by GMR28F02, GMR28G05, GMR28C04, and GMR28D10 fragments respectively in the Janelia GAL4 **(A**) and rescue construct **(B).** The lateral pathway is stained with anti-GFP antibody in A and anti- HA-antibody in B, and the nerve cord is stained with anti-HRP (blue in A) and (purple in B). Lower images represent the GFP and HA channels only.

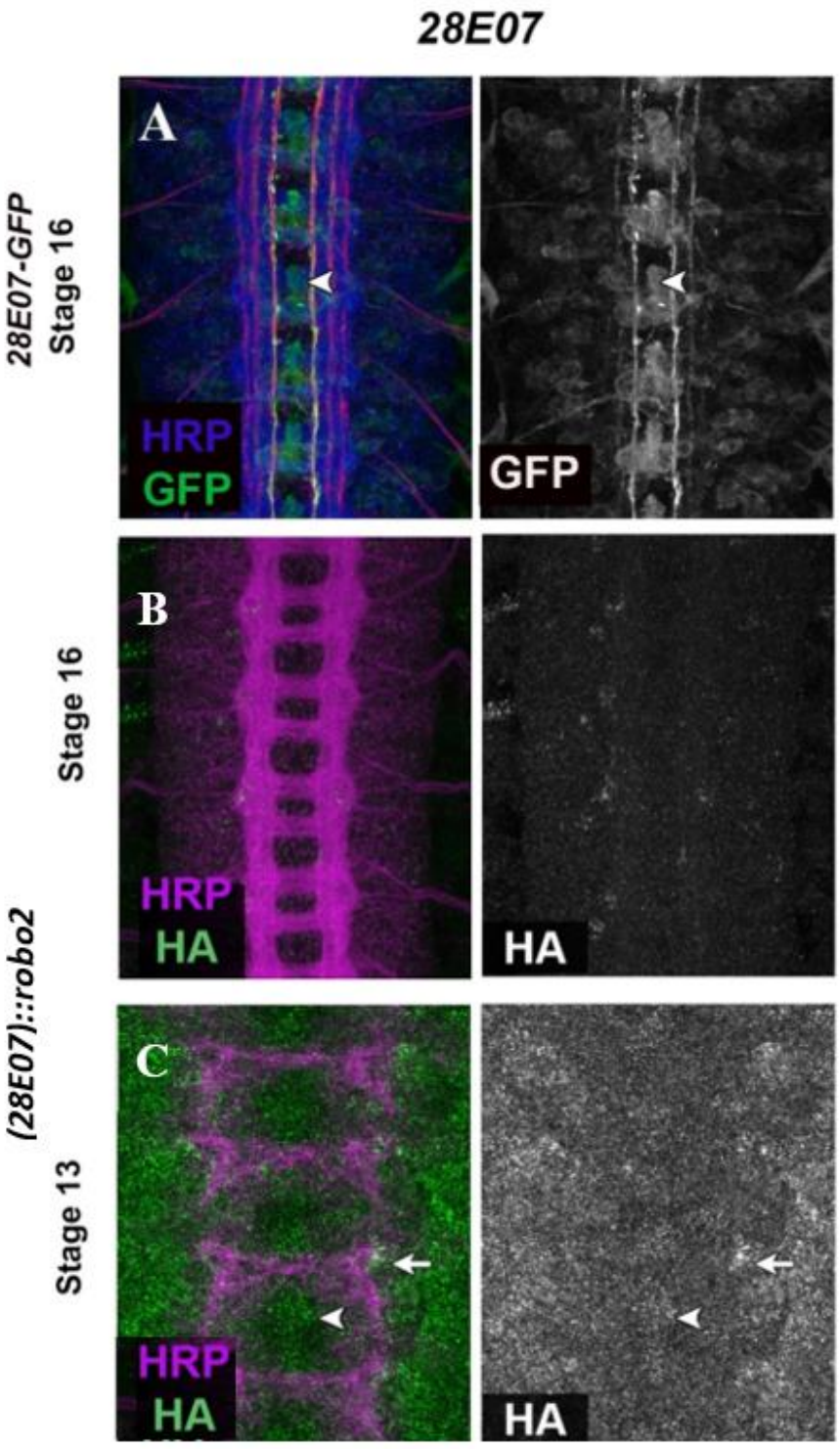

28E07
28E07-GFP
Stage 16
A
HRP
GFP
GFP
Stage 16
B
HRP
HA
HA
(28E07)::robo2
Stage 13
C
HRP
HA
HA

Figure 2.14

Figure 2.14. (Cont.) Compare *robo2* expression driven by GMR28E07 fragment in the original Janelia GAL4 and the rescue construct transgenes. **A**. Stage 16 embryo represents *robo2* expression driven by GMR28E07 fragment in the midline glia of Janelia GAL4 line stained with anti-GFP (green) and the nerve cord is stained with anti-HRP (blue). (**B**) Stage 16 embryo shows very little or no expression of *robo2* driven by GMR28E07 fragment of the rescue construct transgene in this stage of embryogenesis. **C.** Stage 13 embryo represents *robo2* expression driven by GMR28E07 fragment in the midline glia of rescue construct transgene stained with anti-HA antibody(arrowhead) and a subset of the longitudinal pathway (arrow). The nerve cord is stained with anti-HRP (purple). Right image in A represents the GFP channel only and the right images in B and C represent HA channel alone.

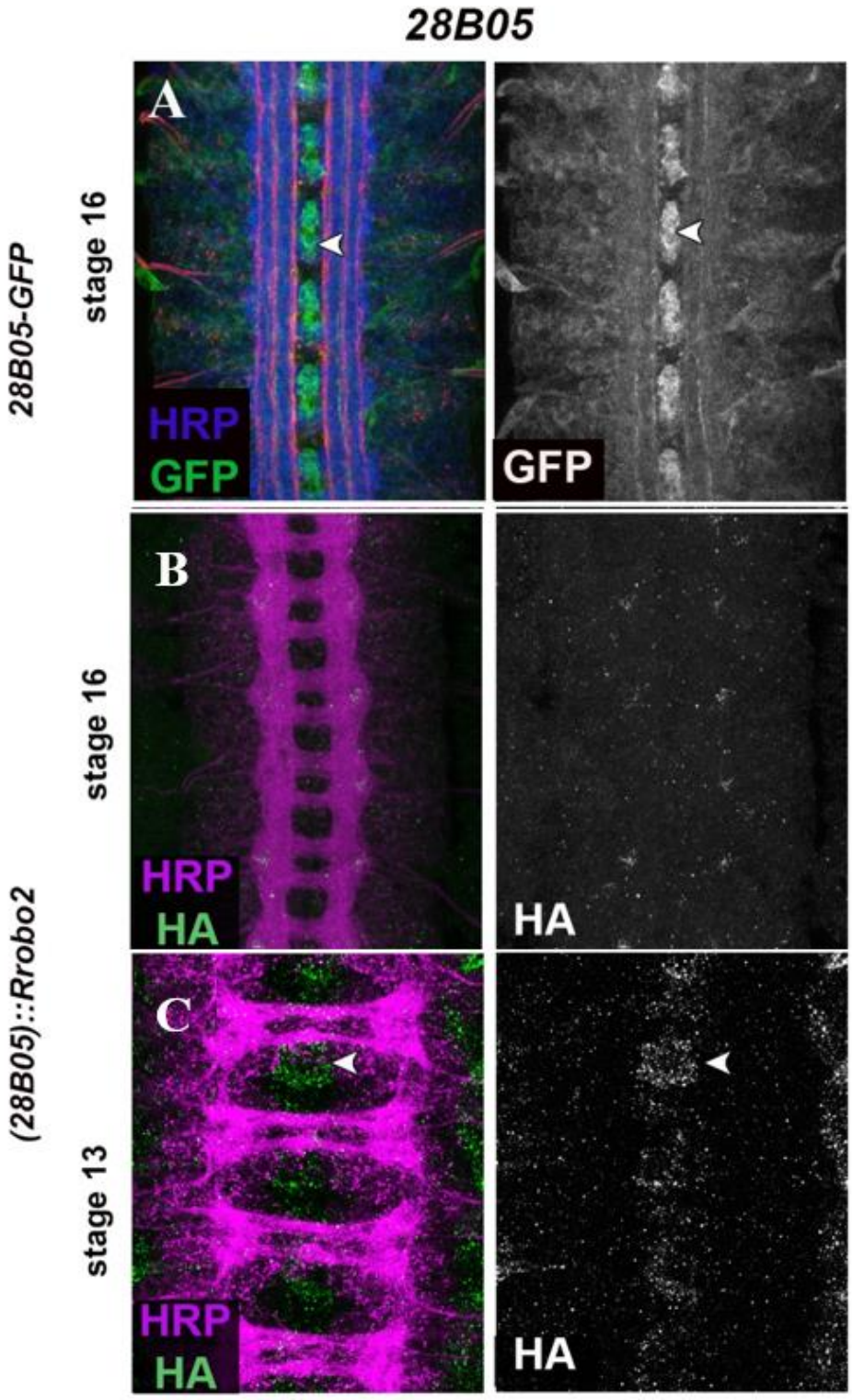

28B05
28B05-GFP
stage 16
A
HRP
GFP
GFP
B
stage 16
HRP
HA
HA
(28B05)::Rrobo2
stage 13
C
HRP
HA
HA

Figure 2.15

Figure 2.15. (**Cont.)** Comparison of *robo2* expression driven by GMR28B05 fragment in the original Janelia GAL4 and the rescue construct transgenes. **A**. Stage 16 embryo represents *robo2* expression driven by GMR28B05 fragment in the midline glia of Janelia GAL4 line stained with anti-GFP (green, arrowhead) and the nerve cord is stained with anti-HRP (blue). (**B**) Stage 16 embryo shows very little or no expression of *robo2* driven by GMR28B05 fragment in the rescue construct transgene in this stage of embryogenesis. **C**. Stage 13 embryo represents *robo2* expression driven by GMR28B05 fragment in the midline glia of rescue construct transgene stained with anti-HA antibody (green, arrowhead) and the nerve cord is stained with anti-HRP (purple). Right image in A represents the GFP channel only and the right images in B and C represent HA channel alone.

Combinatorial enhancers and *robo2* full expression

GMR28G05 and GMR28F02 transgenic lines show strong expression of the *robo2* in subsets of the longitudinal lateral axons. The previous experiment described early in this chapter shows involving one copy of each fragment on each allele is not enough to drive the full expression of the *robo2* in the longitudinal pathway. Therefore, we asked whether combining the two fragments (GMR28G05 and GMR28F02) as one big fragment would recover the full expression of the *robo2*. Gibson assembly approach was used to combine GMR28G05 and GMR28F02 fragments and introduce them to the HA-Robo2 reporter transgene. This reporter was introduced to *Drosophila* to generate a new transgenic line. Homozygous embryos of the progeny of this line were collected and stained with anti-HRP and anti HA antibodies. The results revealed that the number of cell bodies is doubled when the two fragments (GMR28G05 and GMR28F02) were combined and integrated on the same chromosome in tandem compared to the few cell bodies that were labeled in the (GMR28G05):: *robo2* and (GMR28F02):: *robo2* separately (Figure 2.16, C, arrowheads). However, the expression of the *robo2* in the longitudinal pathway was not fully recovered (Figure 2.16, C, arrows).

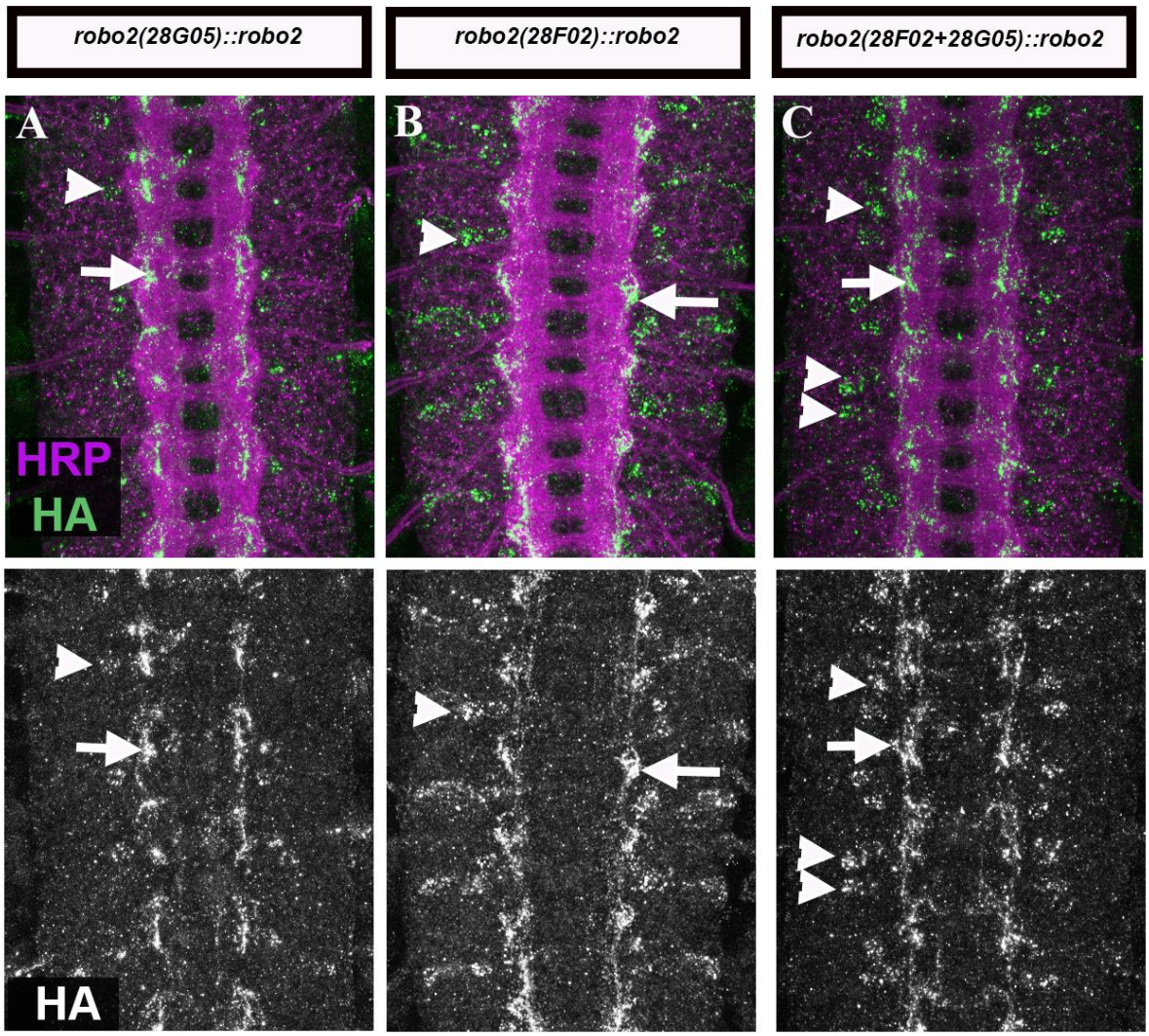


Figure 2.16. Combinatorial expression of *robo2* enhancers (GMR28G05+GMR28F02). **A** and **B** show *robo2* expression which was led by GMR28G05 and GMR28F02 separately. Anti-HRP antibody labels *Drosophila* nerve cord and anti-HA antibody labels subsets of the *robo2* in the longitudinal pathway (arrows) and some few neurons around the nerve cord (arrowheads). **C**. shows the tandem of (GMR28G05+GMR28F02) fragments on the same chromosome of the same embryo. Although the number of the labeled neurons with anti HA- antibody was increased, the longitudinal pathway still shows the *robo2* expression in subset of the lateral axons. Lower images show the HA channel alone.

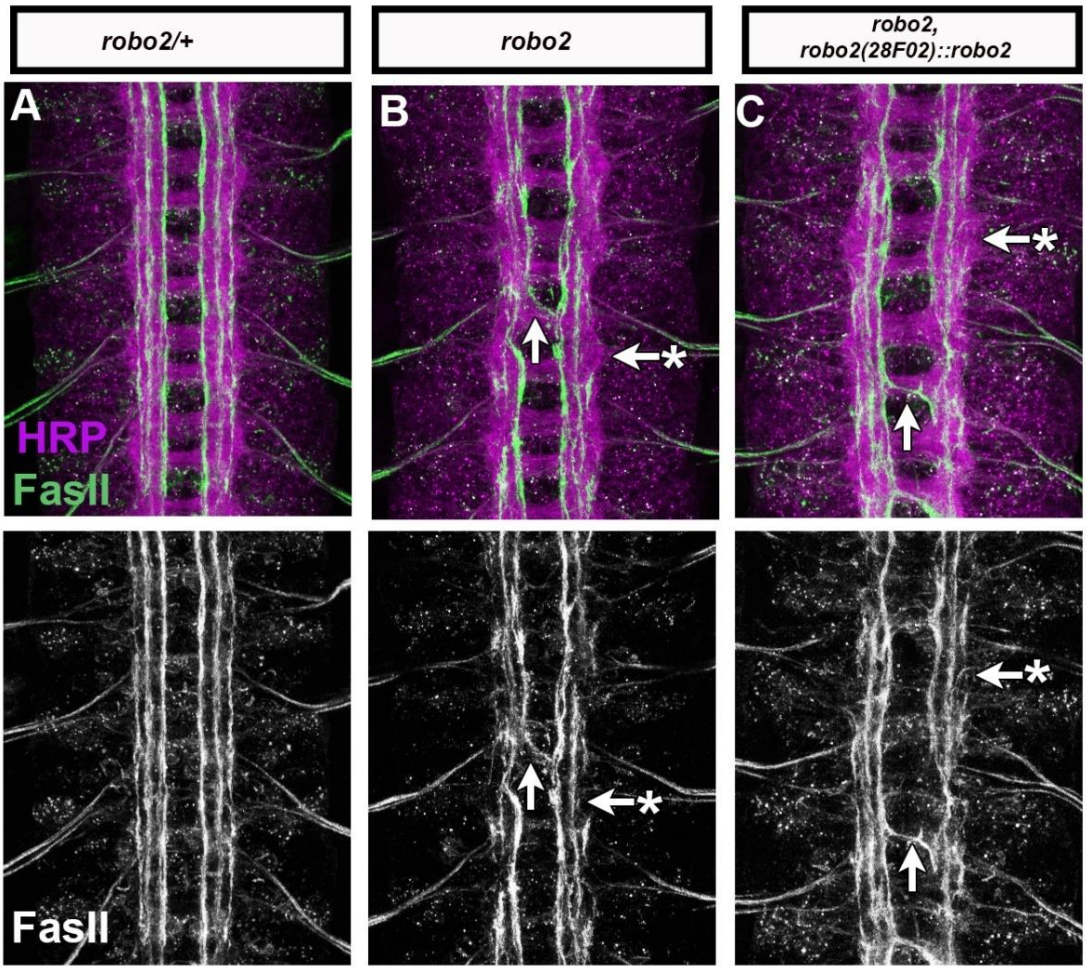
The rescue of axon guidance decision of *robo2* mutant background

This study shows that some of the DNA sequences surrounding the *robo2* promoter are potential enhancers that can regulate *robo2* expression in specific patterns. We hypothesized that these DNA regions that drove the expression of the *robo2* in the longitudinal pathway should be able to rescue the lateral pathway defects in the *robo2* mutant background. The same hypothesis can be applied to ensure that the fragments driving *robo2* expression in the ipsilateral neurons can rescue the midline ectopic crossing defects caused by the loss of the Robo2. To test this hypothesis, the HA-Robo2 transgenes (GMR28G05):: *robo2* and (GMR28F02)::rboo2 were introduced to *robo2* null mutant background (**Appendix 4)**, where small deletions occurred in the *robo2* coding region due to imprecise excision of a p-element transposon upstream of the *robo2* signal sequence. These deletions were designed to behave like genetic null alleles (Simpson, Bland *et al.,* 2000). Midline ectopic crossing and the lateral pathway defects were quantified comparing to the wild-type and *robo2* mutant embryos. Data show that GMR28F02 enhancer does not rescue the *robo2* expression neither in the midline ectopic crossing nor in the longitudinal pathway. In other words, the expression driven by GMR28F02 was not sufficient to recover the defect that occurred in the midline glia and the lateral pathway in stage 16 in *robo2* mutant embryos. Statistical analysis shows the differences between the *robo2* mutant carries the *robo2* (GMR28F02):: *robo2* transgene and *robo2* mutant in both rescuing the midline ectopic crossing and the lateral pathway were not statistically significant (p= 0.2223) and (p=0.075), respectively (Figure 2.17, C, and bottom bar graphs). *robo2*^{123}/*robo2*^{135} embryos showed 22.1% instances of the lateral pathway formation defects, while *robo2*^{123}, [GMR28F02-*robo2*]/*robo2*^{135}, [GMR28F02-*robo2*] embryos, which are expected to show rescue of Robo2's lateral positioning role, showed 18.6% instances of the lateral pathway defects indicating that the instances of the

lateral positioning defects in *robo2*^{123}, [GMR28F02-*robo2*]/*robo2*^{135}, [GMR28F02-*robo2*] showed a non-significant statistical difference from the 22.1% defects of hetero-allelic *robo2* mutants, in other words, the presence of the GMR28F02- *robo2* transgenic modification was not sufficient to rescue Robo2's lateral positioning role. (Table 2.1). Also, *robo2*^{123}/*robo2*^{135} embryos showed 17.1% instances of midline ectopic crossing, and the instances of midline crossing rescue in *robo2*^{123}, [GMR28F02-*robo2*]/*robo2*^{135}, [GMR28F02-*robo2*] showed 22.9% which was statistically not different from the 17.1% defects of hetero-allelic *robo2* mutants, indicating that the presence of the GMR28F02- *robo2* transgenic modification was not sufficient to rescue Robo2's midline ectopic crossing role occurred by hetero-allelic *robo2* mutants.

GMR28G05 on the other hand showed a rescuing for the midline ectopic crossing (P=0.0004). However, its expression was not enough to rescue the lateral pathway defects since there were no significant differences in the lateral longitudinal pathway between *robo2* transgene carrying GMR28G05 (*robo2* (GMR28G05):: *robo2*) and *robo2* mutant background (p=0.0813) (Figure 2.18, C, and bottom bar graphs). Homozygous-allelic *robo2* mutants of genotype *robo2*^{135}/*robo2*^{135} showed instances of the midline ectopic crossing of 40.06% while *robo2*^{135}, [GMR28G05-*robo2*]/*robo2*^{135}, [GMR28G05-*robo2*] showed instances of 7.14% in the midline ectopic crossing, indicating that the instances of ectopic midline crossing in *robo2*^{135}, [GMR28G05-*robo2*]/*robo2*^{135}, [GMR28G05-*robo2*] showed a significant statistical difference from the 40.06% defects of homo-allelic *robo2* mutants (Table 2.2). Instances of the lateral pathway defect in Homozygous-allelic *robo2* mutants were 20.7%while *robo2*^{135}, [GMR28G05-*robo2*]/*robo2*^{135}, [GMR28G05-*robo2*] showed 6.42% instances in the lateral pathway, which was not significantly different from the 20.7% of homo-allelic *robo2* mutants (Table 2.2).

In addition to the single fragment rescue experiment, I performed crosses between the line carrying the combination of GMR28G05 and GMR28F02 ((GMR28G05+GMR28F02):: *robo2)* and *robo2* mutant background line to obtain embryos of homozygous-allelic genotype for the *robo2* mutations *robo2*^{135}, [GMR28G05+GMR28F02-*robo2*]/ *robo2*^{135}, [GMR28G05+GMR28F02-*robo2*]. These embryos were stained, imaged, and scored for defects in the formation of the lateral axon pathways and ectopic axons crossing the midline (Figure 2.18, D). Homozygous-allelic *robo2* mutants of genotype *robo2*^{135}/*robo2*^{135} were also scored in the same way. *robo2*^{135}/*robo2*^{135} embryos showed 20.7% instances of the lateral pathway formation defects, while *robo2*^{135}, [GMR28G05+GMR28F02-*robo2*]/*robo2*^{135}, [GMR28G05+GMR28F02-*robo2*] embryos, which are expected to show rescue of Robo2's lateral positioning role, showed 4.45% instances of the lateral pathway defects indicating that the instances of the lateral positioning defects in *robo2*^{135}, [GMR28G05+GMR28F02-*robo2*]/*robo2*^{135}, [GMR28G05+GMR28F02-*robo2*] showed a significant statistical difference from the 20.7% defects of homozygous-allelic *robo2* mutants, demonstrating that the presence of the combinatorial transgenic modification was sufficient to rescue Robo2's lateral positioning role (Table 2.2). Also, *robo2*^{135}/*robo2*^{135} embryos showed 40.06% instances of midline ectopic crossing, and the instances of midline crossing rescue in *robo2*^{135}, [GMR28G05+GMR28F02-*robo2*]/*robo2*^{135}, [GMR28G05+GMR28F02- *robo2*] showed 10.72% which was statistically different from the 40.06% defects of homozygous-allelic *robo2* mutants, indicating that the presence of the combinatorial transgenic modification was sufficient to rescue Robo2's midline ectopic crossing role occurred by the homozygous-allelic *robo2* mutants.


robo2/+
robo2
robo2,
robo2(28F02)::robo2
A
B
C
HRP
FasII
FasII

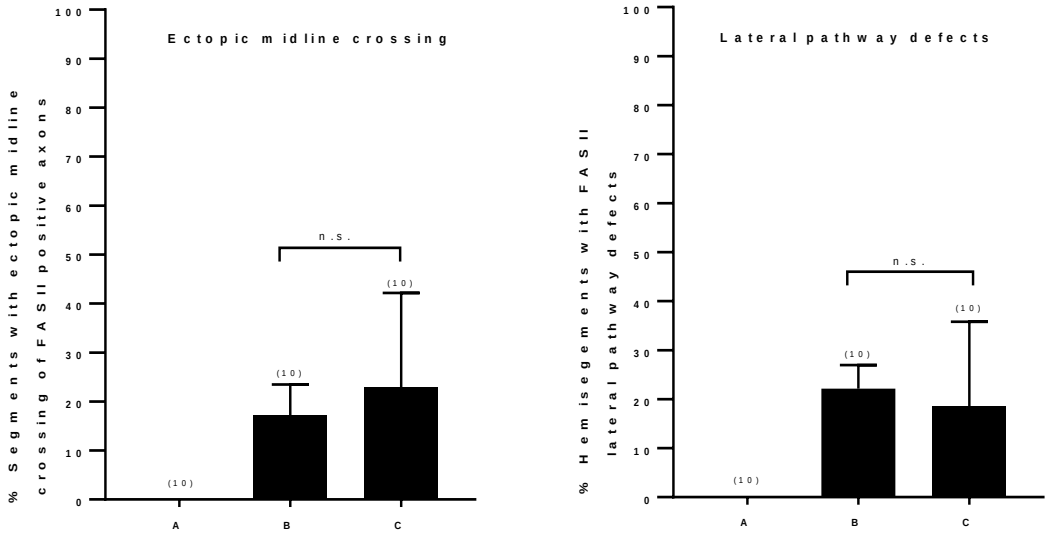
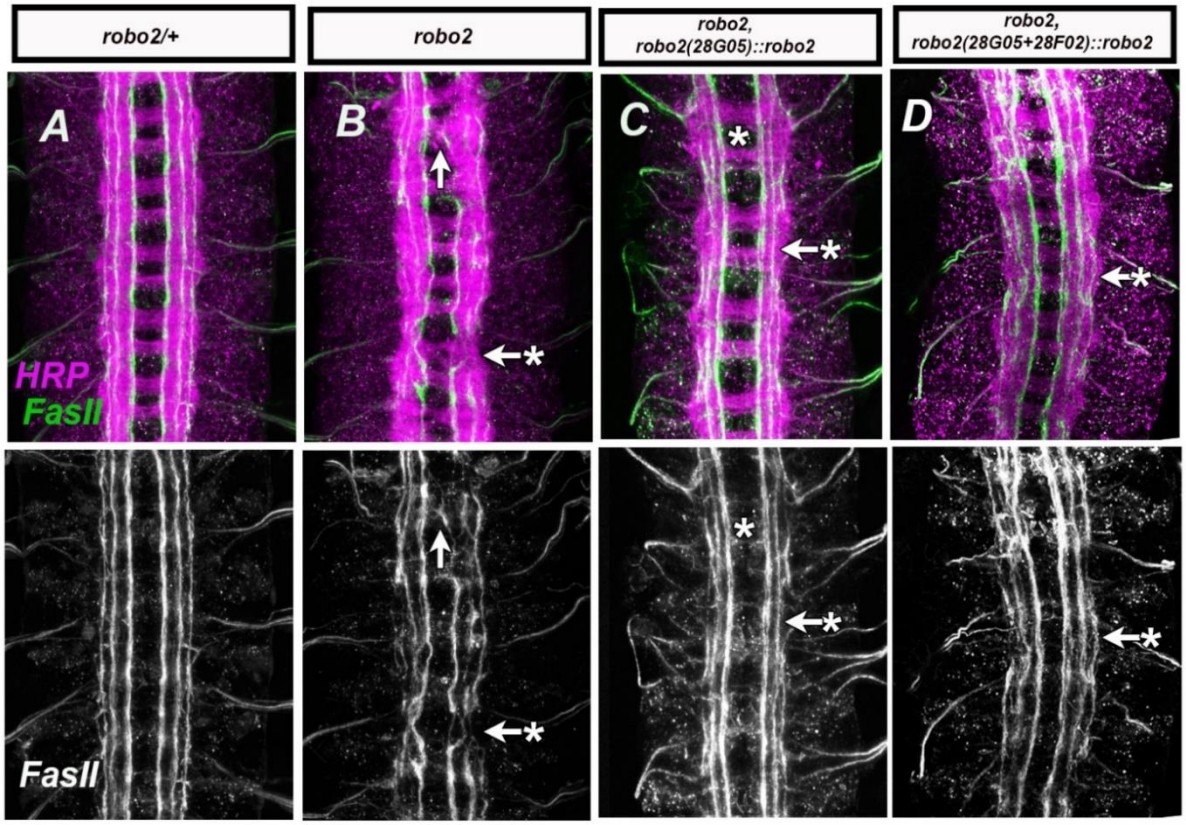

Ectopic midline crossing
% Segments with ectopic midline crossing of FASII positive axons
n.s.
(10)
(10)
(10)
A
B
C
Lateral pathway defects
% Hemisegements with FASII lateral pathway defects
n.s.
(10)
(10)
(10)
A
B
C

Figure 2.17

Figure 2.17. (Cont.) *robo2* rescue using single fragment (GMR28F02). (A-C) Stage 16 embryos stained with anti-HRP (magenta) and anti-FasII (green) antibodies. Lower images show FasII channel alone for the same embryos. FasII-positive axons in *robo2* null mutants cross the midline inappropriately in some segments (B, arrow), and show defects in the lateral pathway (B, arrow with asterisk). This phenotype is not rescued by the *robo2* (GMR28F02):: *robo2* transgene having GMR28F02 fragment that drives *robo2* expression in the lateral pathway (C, arrow with asterisk) and midline (C, arrow). Bar graphs in the bottom indicate instances of ectopic midline crossing (left) and the lateral pathway defects(right)(A-C). Error bars indicate standard error.


robo2/+
robo2
robo2,
robo2(28G05)::robo2
robo2,
robo2(28G05+28F02)::robo2
A
B
C
D
HRP
FasII
FasII

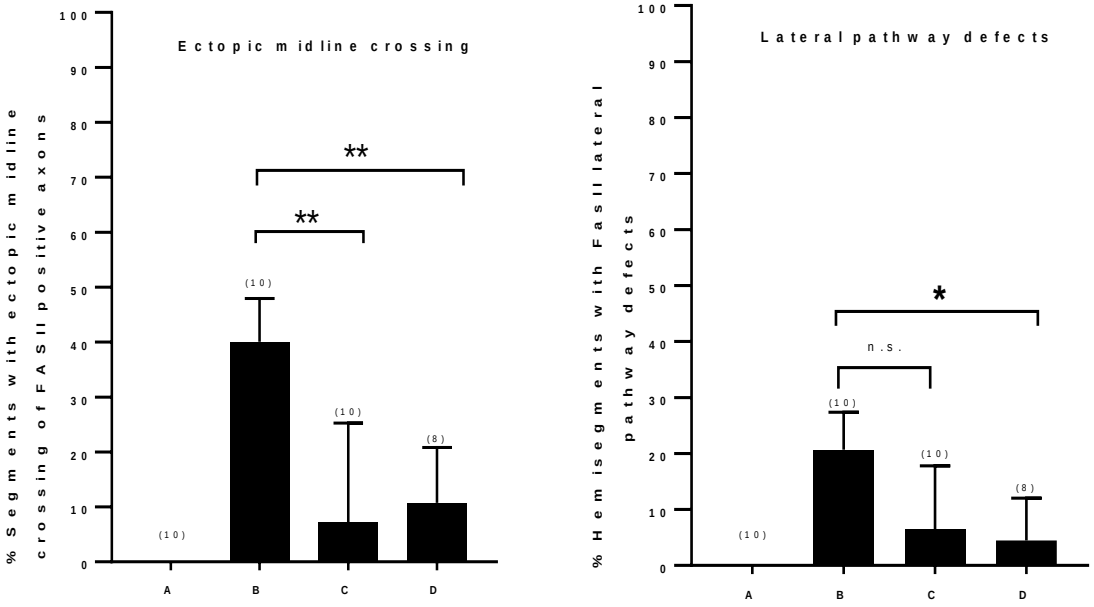

Ectopic midline crossing
% Segments with ectopic midline crossing of FASII positive axons
100
90
80
70
60
50
40
30
20
10
0
**
**
(10)
(10)
(10)
(8)
A
B
C
D
Lateral pathway defects
% Hemisegments with FasII lateral pathway defects
*
n.s.
(10)
(10)
(10)
(8)

Figure 2.18

Figure 2.18. (Cont.) *robo2* rescue using GMR28G05 and combinatorial (tandem) of GMR28G05+GMR28F02 putative enhancer regions. **(A-C**) Stage 16 embryos stained with anti-HRP (magenta) and anti-FasII (green) antibodies to label the nerve cord and axon pathways, respectively. Lower images show FasII channel alone for the same embryos. Medial FasII axons cross the midline ectopically in homozygous *robo2* null mutants (**B**, arrow), and the lateral axon pathways do not form appropriately (**B**, arrow with asterisk). This phenotype is rescued in the midline by the *robo2*(GMR28G05)::*robo2* transgene having GMR28G05 fragment (C, asterisk). However, this fragment did not rescue the lateral defect occurred by *robo2* null mutant (C, arrow, and asterisk). On the other hand, combinatorial of two fragments GMR28G05+GMR28F02 shows fully rescue of *robo2* expression in the midline and the lateral pathway(**D**). Bar graph on the left bottom indicates instances of restoring ectopic midline crossing in *robo2*(GMR28G05)::*robo2* and *robo2*(GMR28G05+GMR28F02)::*robo2* and on the right bottom indicates instances of rescuing the lateral pathway defects in the *robo2*(GMR28G05)::*robo2* and *robo2*(GMR28G05+GMR28F02)::*robo2*. Error bars indicate standard error.

Table 2.1			
Genotype	**Population size, n**	**% Ectopic Crossing**	**% Lateral Pathway Defects**
robo2	10	17.1	22.1
robo2, *robo2*(GMR28F02):: *robo2*	10	22.9	18.6

Table 2.2			
Genotype	**Population size, n**	**% Ectopic Crossing**	**% Lateral Pathway Defects**
robo2	10	40.06	20.7
robo2, *robo2*(GMR28G05):: *robo2*	10	7.14	6.42
robo2, *robo2*(GMR28G05+GMR28F02):: *robo2*	8	10.72	4.45

Generate and characterize an equivalent set of *robo2* transgenes expressing the axonal marker TauMyc instead of the HA-Robo2

Robo2 is expressed normally in a subset of longitudinal axons. To unravel the *robo2* candidate enhancers underlying the expression in these axons without relying on *robo2* expression, I generated an equivalent set of the *robo2* transgenes that express the axonal marker TauMyc. This transgene consists of *Drosophila robo2* promoter followed by TauMyc cDNA. The embryos of this transgene were collected and stained by anti-HRP and anti-Myc antibodies. However, the expression was too weak to be studied (Figure 2.20, A). *robo2* promoter was replaced by hsp70 promoter since the first one did not show a good expression. The GMR28F02 fragment was cloned into the TauMyc construct (Figure 2.19), and TauMyc expression was examined in embryos carrying this transgene. Results show that GMR28F02 fragment in *hsp70*(GMR28F02)::TauMyc drove strong TauMyc expression in a subset of the lateral neurons, it also labeled the cell bodies and axons including commissural axon segments from which Robo2 protein is expressed (Figure 2.20, B).

The previous studies show that *robo2* misexpression prevents longitudinal axons to form correctly (Simpson, Bland *et al.,* 2000). Robo2 in recent experiments has shown responsibility for regulating axon guidance in both ways, cell-autonomously and cell non-autonomously. In *robo2* mutants, FASII positive lateral axons fail to form appropriately. It is unknown whether these FASII axons express Robo2, so they show defects in their expression when *robo2* is missing. Or axons that express Robo2 normally show the same defects that FASII axons show, and whether these axons always colocalize with FASII positive axons. We utilized the generated transgenic line that expresses TauMyc in a subset of longitudinal axons that express *robo2* normally (*hsp70*(GMR28F02)::TauMyc) to test that. This transgene is introduced into *robo2* null

mutant background and TauMyc axon positions were quantified by measuring their distances from the midline in five hemi segments of each embryo in both wild-type and *robo2* null mutant. FASII axons positions were also quantified in the same way. Results show that TauMyc label axons express a significant guidance defects in *robo2* mutants (****p<0.0001). These axons show shifting to positions closer to the midline. The average distance of these axons from the midline is 9.66 µm in the *robo2* mutants while it is 13.11 µm in the wild-type. While the HRP axons are 15.23 µm in distance from the midline and TauMyc axons usually make a position of 13.11 µm in the wild-type, the mutants with the whole scaffold are down to 9.04 µm for the HRP and the TauMyc axons right now are 9.66 µm. FASII axons show multiple breaks and shifting from and between zones. They show a significant defect in *robo2* mutant (****p<0.0001). FASII average distance from the midline is 11.97 µm in the wild-type. However, their distance from the midline in *robo2* mutant is down to 7.41 µm.

I have generated another transgenic line that expresses the axonal marker TauMyc in a subset of longitudinal neurons that express Robo2 normally by using GMR28G05 (*hsp70(GMR28G05TauMyc).* Embryos from this transgenic line were collected, stained with anti-HRP, anti-FASII, and anti-Myc antibodies. Interestingly, TauMyc axons that express Robo2 normally in the longitudinal neurons were missing (Figure 2.20, D).

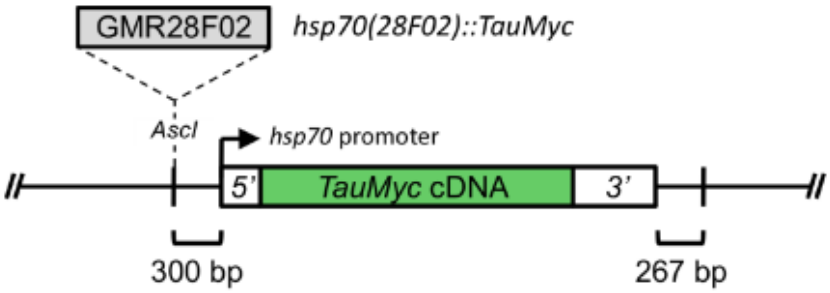


Figure 2.19. A schematic of TauMyc construct showing TauMyc cDNA with hsp70 promoter and the inserting site for GMR28F02 fragment.

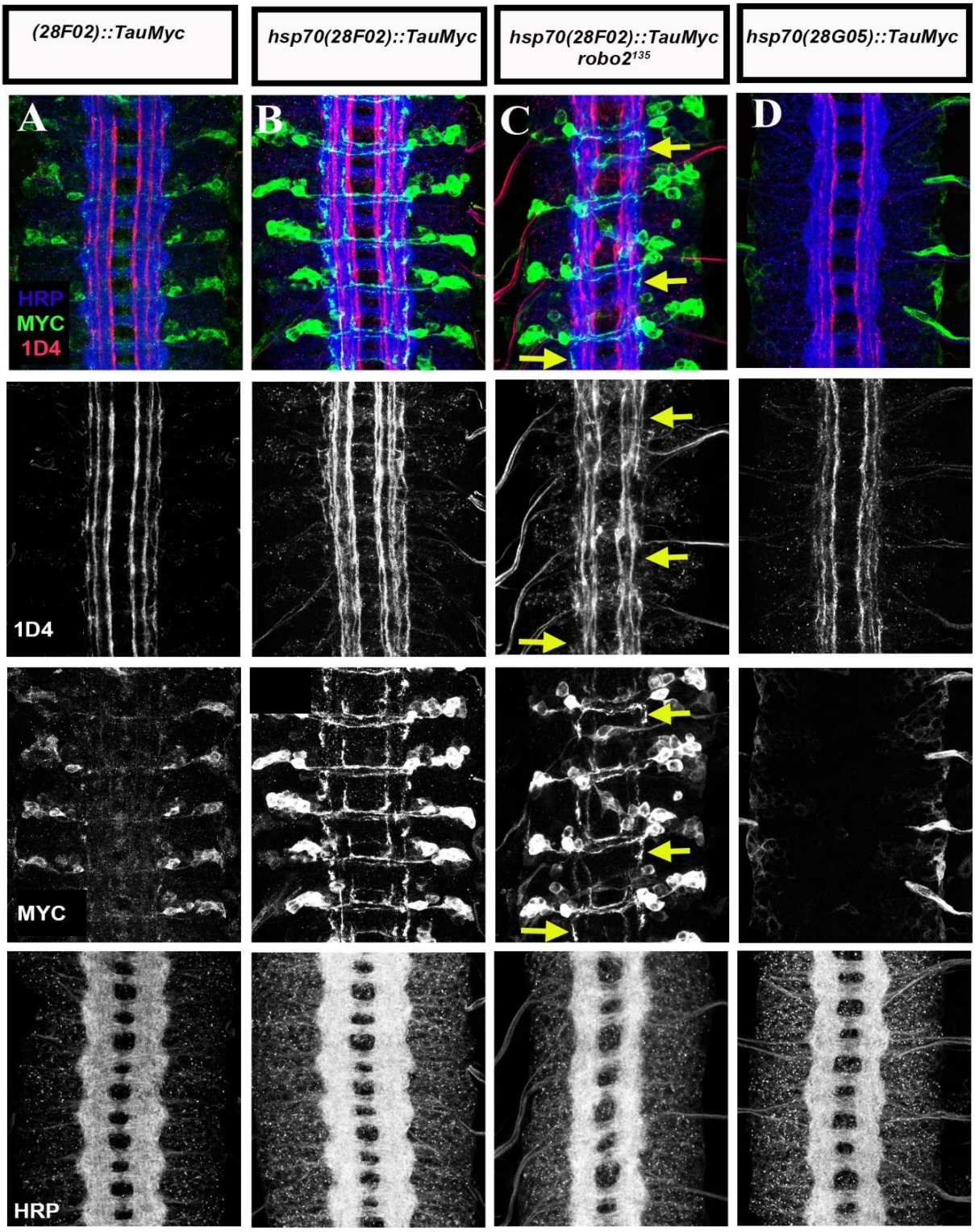

(28F02)::TauMyc
hsp70(28F02)::TauMyc
hsp70(28F02)::TauMyc robo2^135
hsp70(28G05)::TauMyc
A
B
C
D
HRP
MYC
1D4
1D4
MYC
HRP

Figure 2.20

Figure 2.20. (Cont.) Non-autonomous role of Robo2 in the formation of the lateral pathway. **A.** GMR28F02::TauMyc embryos at stage 16 shows TauMyc, an axon marker expresses Robo2 normally in the lateral pathway stained with anti-HRP (blue), anti-FASII (red), and anti-Myc (green). **B.** *hsp70(GMR28F02)::*TauMyc embryo where Robo2 promoter was replaced by *HSP70* to show TauMyc axons clearly. **C.** *hsp70(GMR28F02)::TauMyc* in *robo2* mutant background (Robo2^{135}) showing the lateral FASII pathways do not form correctly (lower channel, arrows), and they do not colocalize with TauMyc axons (arrows in C, and second lower channel of C, and third lower channel of C). TauMyc axons do not show the same defects that FASII show in *robo2* mutant background (hsp70(GMR28F02)::TauMyc /Robo2^{135}. **D.** *hsp70(GMR28G05)::TauMyc* where Robo2 promoter was replaced by HSP70 second lower channels show isolated FASII channels from the same embryos. Third layer images show isolated TauMyc channels from the same embryos. Bottom channels show separated HRP channel from the same embryos.

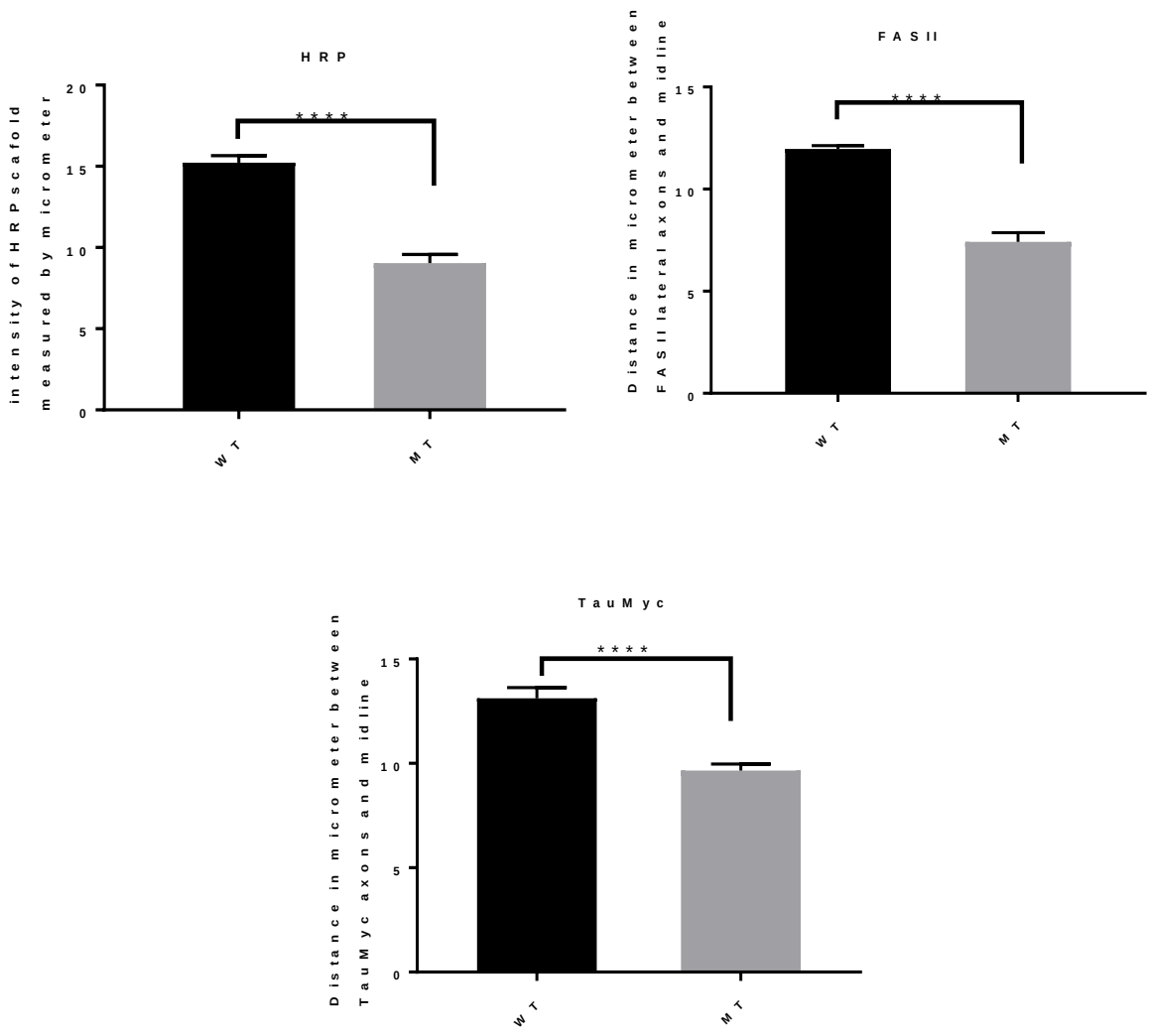


Figure 2.21. Bar Graphs showing the defect that FASII and TauMyc axons display in *hsp70(GMR28F02)::TauMyc* in *robo2* mutant background. The distance between FASII and TauMyc axons and midline was measured for five locations per embryo. Each variant was compared in +, *hsp70(GMR28F02)::TauMyc* embryos and mutants by a Student's-test. We detect a statistically significant decrease in a distance from the FASII and TauMyc axons to the midline in embryos expressing *hsp70(GMR28F02)::TauMyc* in *robo2* mutant background *(robo2*^{135}) compared to wild type embryos expressing (hsp70(GMR28F02)::TauMyc ($p<0.0001$).

Discussion

Robo2 is a member of conserved protein family that responds to Slit ligand and plays a role in several aspects of axon guidance (Evans *et al.,* 2015). The question remains as to the ways by which Robo2 acts in making such different functions. *robo2* expression has previously been shown to be transient through embryonic developmental stages (Rajagopalan, Vivancos *et al.,* 2000). While Robo2 is broadly expressed in the CNS during the early stages of development, its expression becomes more limited to axons in the lateral one-third of the longitudinal pathway in later stages. This project hypothesizes that different regulatory regions surrounding the *robo2* gene direct *robo2* expression in a specific special and temporal pattern during embryonic development, and this expression is orchestrated by gene control mechanisms through regulating regions that control *robo2* expression in specific cell types or specific developmental stages. With one or multiple enhancers controlling one aspect of *robo2*'s expression. The research in this chapter focuses on the *robo2* fragments role in driving a specific *robo2* expression in the lateral pathway and midline glial cells. The fragments tested in this chapter have been confirmed to be able to drive a specific *robo2* expression in the lateral axons and glial cells indicating that these putative enhancers exist on these fragments. For instance, GMR28F02, GMR28G05, GMR28C04, and GMR28D10 drove *robo2* expression in the lateral pathway. Two more overlapping fragments (SEG02 and SEG03) generated in our lab (not included in the original Janelia GAL4) that are located upstream of the *robo2* gene have shown expression in the lateral pathway as well. Moreover, two other fragments (GMR28E07 and GMR28B05) showed expression in the midline glia. Even though there is no overlapping between GMR28E07 and GMR28B05 fragments, the expression patterns that they show is identical in both Janelia GAL4 lines and the rescue construct transgenes at almost all stages. They both show an expression in

the midline glial cells. However, these two fragments are separated from each other by GMR28A10 fragment. Interestingly this fragment (GMR28A10) does not show any expression in Janelia GAL4 lines even if a part of GMR 28A10 sequence overlaps with GMR28B05. Figure 2.4. I suggest that using CRISPR cas-9 based gene-editing to delete the region shared between GMR28A10 and GMR28B05 from GMR28B05 fragment and then combine the remaining fragment of GMR28B05 with GMR28E07 fragment would be sufficient to drive *robo2* full expression in the midline glia.

Putting the rescue constructs in the *robo2* mutant background drove the expression of *robo2* in subsets of the lateral pathway and midline to play a role in these two positions. However, these constructs represented by (GMR28G05):: *robo2,* and (GMR28F02):: *robo2,* separately failed to rescue the whole function of Robo2 in the lateral pathway formation. Interestingly, (GMR28G05):: *robo2* construct that expected to rescue *robo2* expression in the longitudinal pathway displayed restoring of the essential Robo2 function to prevent midline ectopic crossing. To explain that, we need to look closer to the original Janelia GAL4 lines, GMR28G05-GFP line shows some of the cells are expressed in the midline, which makes it possible for (GMR28G05):: *robo2* to rescue the ectopic midline crossing. The question still arises as why transgenic constructs that led the expression in specific cell subsets were not able to completely rescue the *robo2* expression in that position? If we go back to the initial screening of Janelia GAL4 lines we can notice that multiple lines besides tested ones showed GFP expression in the lateral pathway and midline glia. In other words, not only the tested lines expressed Robo2 in different cell types and different stages, but also some other lines showed GFP expression, but that expression was hard to define accurately, which suggests that multiple fragments may be required to regulate a single aspect of *robo2* expression. Also, some of Janelia

GAL4 lines such as GMR28E07-GFP and GMR28B05-GFP that showed expression at stage 16, the same expression was missing in the rescue transgenes at the same stage. One possible explanation for that is because of the difference in the expression duration between Janelia GAL4 lines and the rescue construct transgenes. The GFP expression pattern can be seen in both early and late stages of Janelia GAL4 lines while the expression pattern was restricted to earlier stages of the embryogenesis in the rescue transgenes. Robo2 cytosolic mRNA appears for a short time before being translated to a protein. However, GFP lasts longer than the mRNA and that is why it shows the *robo2* expression at later 16 stages while the actual expression was in the stage 13 of the embryogenesis.

The rescue of the axon guidance decision of *robo2* mutant background by the combinatorial of *robo2* enhancers

Since single fragment transgenes did not show a full rescue for the longitudinal pathway of *robo2* expression, I hypothesized that combinatorial of two fragments that have expression patterns in the longitudinal pathway would rescue the *robo2* expression in the lateral pathway. GMR28G05 and GMR28F02 fragments were combined in tandem in one construct and injected into wild-type fly by Best gene, then a fly stock balancer made and crossed to the mutant background. homozygous-allelic genotype embryos with respect to the Robo2 mutations *robo2*^{135}, [GMR28G05+GMR28F02-*robo2*]/ *robo2*^{135}, [GMR28G05+GMR28F02-*robo2*] were collected, stained, imaged, and scored for defects in the formation of the lateral axon pathways and for ectopic axon crossing of the midline. Homozygous-allelic *robo2* mutants of genotype r*obo2*^{135}/*robo2*^{135} were also scored in the same way. The homozygous-allelic genotype *robo2*^{135}, [GMR28G05+GMR28F02-*robo2*]/ *robo2*^{135}, [GMR28G05+GMR28F02-*robo2*] was able to rescue the *robo2* expression not only in the lateral pathway formation but also in the midline

repulsion. The instances of the lateral positioning defects in *robo2*^{135}, [GMR28G05+GMR28F02-*robo2*]/*robo2*^{135}, [GMR28G05+GMR28F02-*robo2*] showed significant statistical difference from the defects of homozygous-allelic *robo2* mutants in both lateral and midline ectopic crossing. Together, these results confirm that multiple enhancer regions are required to restore the full *robo2* expression.

TauMyc marker transgene confirm cell non-autonomous role of Robo2

Robo2 receptor regulates the lateral positions of the axons in the VNC of *Drosophila* CNS by responding to Slit ligand. *Drosophila* wild-type embryos show three FASII positive on both sides of the midline. The previous studies show removal of *robo2* causes fusing the lateral FASII pathway with the intermediate pathway and directs the lateral axons medially (Ströhl *et al.,* 2017). Another study shows that *robo2* misexpression prevents the formation of the longitudinal pathway appropriately (Simpson, Bland *et al.,* 2000). Moreover, ectopic expression of *robo2* lead the medial axons to be extended laterally. (Simpson, Bland *et al.,* 2000; Spitzweck *et al.,* 2010) show that replacing *robo2* by *robo1* causes the lateral pathway fails to form correctly. However, HA-positive axons that express Robo1 instead of Robo2 still able to be expressed laterally. We need to know if axons expressing Robo2 normally show any defects in *robo2* mutant background and whether this defect looks like the one that FASII axons show. We also want to know the autonomy express of Robo2 by normally expressing axons. The previous studies showed that Robo2 can act cell-autonomously and non-autonomously to regulate the lateral axons. We show here that TauMyc axons that normally express Robo2 in the lateral axons act as cell non-autonomously in *robo2* mutant. However, the defects they show is different than the one that FASII show. The question remains as to whether TauMyc axons colocalize with FASII axons. The answer is no. Although they both show significant defects by being shifted

from the midline in *robo2* mutants, they still do not show the same defects and navigation in their expression. Together, these data show that FASII and TauMyc axons are not the same.

Not only GMR28F02 was cloned in the TauMyc construct, but also GMR28G05 was cloned in the same construct. Unlike *hsp70*(GMR28F02)::TauMyc, *hsp70*(GMR28G05): :TauMyc did not show any expression. We do not know if this missing of expression has been occurred due to the replacing of the endogenous promoter by hsp70 or it happened in the original transgene when we replaced the *robo2* construct by *TauMyc*. To answer this question, I suggest that looking at the original expression of the (GMR28G05::TauMyc) transgene before replacing the promoter would be more helpful, by generating a new transgene that has GMR28G05 introduced into the axonal marker TauMyc construct under the control of the *robo2* original promoter.

Conclusion

I have presented here transcriptional regulation of the *robo2* gene in *Drosophila* CNS. Screening of 17 Janelia GAL4 transgenic lines by using the GAL4-UAS system yielded six promising fragments (GMR28F02, GMR28G05, GMR28E07, GMR28B05, GMR28C04 and GMR28D10). Four of these fragments showed strong expressions, two in the longitudinal pathway (GMR28F02 and GMR28G05), and the other two showed expression in the midline glia (GMR28E07 and GMR28B05). Rescue construct for these six fragments in addition to three fragments generated in our lab (SEG01, SEG02, and SEGO3) showed *robo2* expression in subsets of the longitudinal pathway driven by 6 fragments (GMR28G05, and GMR28F02, GMR28C04, GMR28D10, SEG02, and SEG03). GMR28G05, and GMR28F02 showed the strongest expression, and two of the remainder rescue construct transgenes showed *robo2* expression in the midline glia driven by GMR28E07, and GMR28B05 with similar expression levels for both. While putting transgene of a single fragment (putative enhancer) in *robo2* mutant background did not rescue *robo2* expression, a combinatorial of two fragments showed a rescue of the *robo2* expression in both midline and longitudinal pathway.

Chapter three: Using CRISPR/Cas9 system and bioinformatics tools to characterize *robo2* enhancers.

Abstract

Regulatory regions are non-coding DNA sequences that regulate genes transcriptions through the transcription factors. One of these regulatory regions is an enhancer. Enhancers increase the transcription of genes and they can be located upstream, downstream, within the introns, or in a long distance from the gene they regulate. Multiple enhancers might be required to act in coordination to regulate a gene transcription. Results from the previous chapter suggest that Robo2 can be led by multiple enhancers. GMR28G05 and GMR28F02 fragments, which led the strongest expression of *robo2* in the longitudinal pathway, are potentially the multiple enhancers for *robo2*. To examine this finding, the CRISPR-Cas9 system was used to delete both fragments separately and together. Change in *robo2* expression in the longitudinal pathway will give insight into the role of these fragments. Also, knowing which transcription factor(s) might bind to the above fragments would help us to unveil the role(s) of these fragments regarding to *robo2* transcription. Thus, bioinformatics tools and literature were utilized in the second part of this chapter. I found three promising transcription factor candidates (Hb9, Nkx6.1 and Lhx2) that potentially bind GMR28G05 and GMR28F02 with the following DNA consensus binding motifs (TAATTA), (TTAATTG), and (TAATTA), respectively. These findings may open new avenues to understand the transcription regulation of *robo2* regarding its enhancers.

Introduction

Understanding genetic variations and their association with causing diseases is important, especially for human health. While studies show that a mutation in the protein-coding gene regions can lead to most pathogenic alterations, mutations in the non-coding sequences can affect gene regulation and pathways involved in specific diseases such as cancer (M. B. Patel & Wang, 2019). Intron sequences make up about 25% of the human genome, which is 4∼5 times the size of exons (Sakharkar *et al.,* 2004). Introns can play a protective role for the eukaryotic genome sequences by occupying 40% of the total length of the gene and eventually most of the random mutations will occur in the intron regions, keeping the protein sequences and functions safe. However, a mutation in the intronic regions can lead to tumorigenesis by affecting the splicing in direct manner and causing malignant transcript isoforms, making introns crucial for the cell health. Introns can impact protein variety by affecting alternative splicing and *Drosophila* Dscam gene represents an excellent example of increasing protein variety. It can produce more than 38,000 isoforms that can be produced from alternative splicing (Pan *et al.,* 2008). Furthermore, introns have a potential effect in enhancing gene expression and researchers prefer including introns in designing constructs in order to guarantee a higher level of expression(Clark *et al.,* 1993). Studies in mammals and yeasts confirm that genes with their introns show a higher level of expression than those without introns (Juneau *et al.,* 2006; Shabalina *et al.,* 2010). Although current genome-wide discovering revealed a wide number of functional genes in the DNA coding regions, the functional elements in the non-coding sequences of human genome are not fully uncovered. Due to the critical role of non-coding sequences of human genome in different aspects, such as being involved in the transcription factor binding, chromatin states, and modification resulted from epigenetics, more studies would be required (Li *et al.,* 2014).

This chapter focuses on testing the effect of putative enhancers located on the first intron of the *robo2* gene on its protein expression. The results from the previous chapter suggested that multiple enhancers are necessary for leading the *robo2* expression in the longitudinal pathway and two fragments (GMR28G05 and GMR28F02) located in the first intron of the *robo2* gene showed the strongest expression of *robo2* in the lateral pathway. However, it is not known whether the *robo2* expression will be affected by the absence of these two fragments. In this chapter, I used CRISPR-Cas9 to make deletions in a region of the *robo2* first intron. I deleted GMR28G05 fragment (3.8 Kb) and GMR28F02 fragment (3.7Kb). Both GMR28G05 and GMR28F02 fragments and in between fragment, the total DNA length was 11.5Kb, were deleted as well and the *robo2* gene expression will be monitored.

CRISPR-Cas9 as a precise tool in genome editing

Clustered regularly interspaced short palindromic repeats (CRISPR-Cas9) is the most recent discovery in genome editing technology that enables researchers to precisely manipulate the DNA sequence of any genome. It was first discovered as a family of the DNA sequence belongs to bacteriophages that previously infected prokaryotes. Prokaryotes used these DNA sequences to detect and destroy any DNA identical to the bacteriophage during successive infections. For this reason, these sequences are considered a defense system that prokaryotes use against viral infection. (Barrangou, 2015). Cas 9 is an enzyme guided by the guide RNA (gRNA) sequence to recognize and cleave the DNA sequence that is complementary to the gRNA sequence. Since then CRISPR-Cas9 has been the cornerstone of gene-editing in different organisms. (Zhang *et al.,* 2014). It provides a vigorous multifunctional gene-editing tool, allowing researchers to accurately manipulate genome elements, and facilitating of targeting genes. This technique has been utilized in both molecular biology and disease treatment. It was successfully used to target

important genes in many organisms and cell lines. (https://www.neb.com/tools-and-resources/feature-articles/crispr-cas9-and-targeted-genome-editing-a-new-era-in-molecular-biology).

Unlike alternative genome editing technologies such as zinc-finger nuclease (ZFNs) and transcription activator-like effector nucleases (TALENs), which depend on the use of customizable DNA-binding protein nucleases, The CRISPR-Cas9 system has a discernible benefit over these technologies. Using RNA-based CRISPR provides simplicity and flexibility for genome editing. For this reason, CRISPR has revolutionized biomedical research, and become one of the most essential approaches for genome engineering (Kim *et al.,* 1996; Smith *et al.,* 1999; Bibikova *et al.,* 2001; Boch *et al.,* 2009; Moscou & Bogdanove, 2009; Christian *et al.,* 2010; Tianfang Ge *et al.,* 2016).

The Cas9 mediated genome editing system requires only three components:1. Cas9 that can be available as a gene, mRNA, or purified protein, 2. gRNA, which can be supplied as an RNA or transcribed from the DNA template in vivo. 3. DNA donor holding the target sequence which, in turn, involves a novel sequence to be inserted or indels.

In the CRISPR-Cas9 system, cas9 makes a DNA double-strand break (DSB) and the next step involves repairing the DSBs, which occurs by one of two major repair pathways (Figure 3.1): 1. The non-homologous end-joining (NHEJ) pathway, which considered an efficient but error-prone pathway, and 2. The homology-directed repair (HDR) pathway, the less efficient but high-fidelity. Unlike HDR, NHEJ does not require a DNA template with homology to the sequences flanking the DSB. NHEJ, therefore, generates insertions or deletions during DSB repair. The efficiency of NEHJ is higher than HDR and the repairing time is shorter (tens of minutes). NHEJ is useful if there is a requirement of making a null allele (Knockout) in the gene of interest.

HDR, on the other hand, requires a presence of the DNA template homology to the DSB location flanking sequence. As a result, repairing the broken DNA strands occurs in an error-free manner. While NHEJ is active throughout the cell cycle, HDR can be used by the cell only if the homology DNA is present in the nucleus in specific phases (most likely G2 and S phase) of the cell cycle. In addition to the cell cycle, the efficiency of HDR depends on the concentration of the DNA that exists during repairing time, the homology arm length of the donor DNA, and the efficiency of the endogenous repair systems. HDR is used to make an insertion (Knockin) in the gene of interest (Figure 3.1)(Lin *et al.,* 2014; Iliakis *et al.,* 2004; Hasty *et al.,* 1991; Wolf-Dietrich Heyer, 2008; Pardo *et al.,* 2009). In this chapter, the CRISPR-Cas9 approach was used to create a deletion in *robo2* regulatory regions including GMR28G05 and GMR28F02 separately or together by non-homologous end-joining in *Drosophila*.

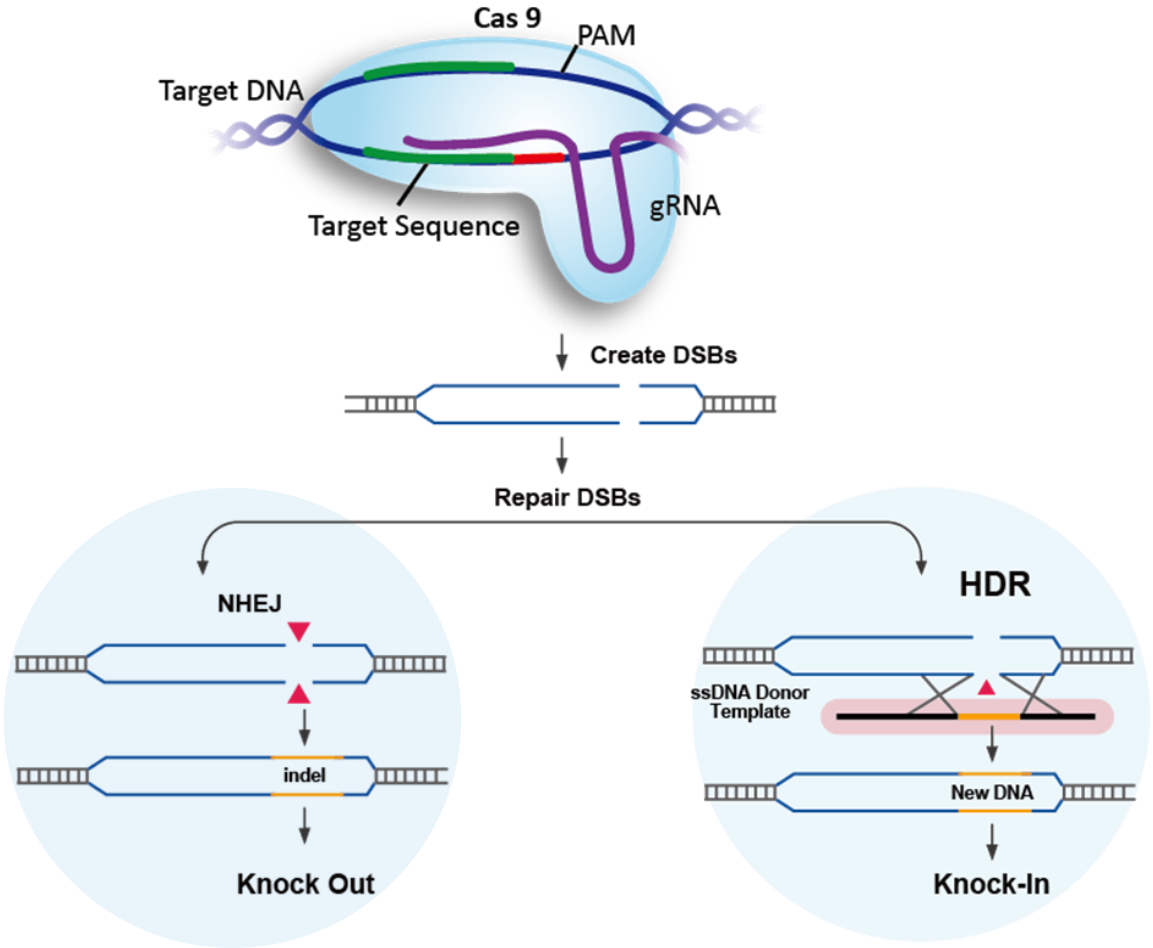


Figure 3.1: CRISPR/cas9 Genome Editing Mechanism. The complex of Cas9 and gRNA binds with the DNA close to the PAM sequence. Cas9 generates a DNA double-strand break (DSB) 3-4 bp upstream of the PAM sequence, this DSB can be repaired via NHEJ, which might result in insertion or deletion or in a frameshift that causes gene knockout, or HDR, which can cause gene knockin if the DNA donor is provided with homology in the ends. (Image adapted from https://www.genscript.com/using-crispr-for-genome-editing.html).

Transcription factors and gene expressions

Generally, the transcription of genes is controlled by different ways and one of them is the interaction of the transcription factors to specific DNA sequences. Sometimes these proteins are encoded by genes that are in a different place of the genome from the genes that they regulate; therefore, they are called trans-acting factors. When these proteins bind to their regulatory elements of the same DNA fragment that has the gene that they regulate, these regulatory elements are called cis-acting factors. The transcription of genes is mediated by the RNA polymerase enzyme that plays an important role in catalyzing the RNA synthesis and this RNA is complementary in sequence to the DNA template. There are three RNA-polymerases in eukaryotes: RNA polymerase I (Pol I),which is responsible for transcribing genes that encode the ribosomal RNA, RNA polymerase II (Pol II or RNA PII), which transcribes genes into mRNA, and RNA polymerase III(Pol III) that transcribes genes that encode the tRNA and the small nuclear RNA (Kadonaga, 2004).

Pol II is a large protein with a molecular mass of 600 KDa consists of 10-12 subunits. This protein can transcribe RNA from DNA. Despite its capability to catalyze mRNA synthesis, Pol II cannot bind to the DNA and initiate the transcription by itself without the assembly of specific proteins called general transcription factors. There are six general transcription factors that are important for class II genes transcription. These factors are TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH. These transcription factors by assembling with the Pol II make the pre-initiation complex or the basal transcriptional machinery and the latter assembles with the promoter region of the DNA (Kadonaga, 2004; Hahn, 2005).

DNA promoter consists of many regions that are incorporated in the transcription process. A TATA box region, which is essential for recognizing the RNA Pol II, an initiator (Inr) box,

which consists of a stretch of nucleotides sequence in the DNA from where the transcription lunches, and the upstream controller elements region that binds to the RNA Pol II (Schramm & Hernandez, 2002).

RNA Pol II should interact with the DNA before the transcription initiates, and for this interaction to happen, general transcription factors should bind to TATA box and Inr. The first general transcription factor that binds to the DNA template is TFIID (Kadonaga, 2004; Nikolov & Burley, 1997). TFIID consists of a TATA-binding protein (TBP), a highly conserved protein that plays a vital role in all eukaryotic transcription. TBP has an antiparallel β-sheet that can sit on the minor groove of the DNA and make a conformational changing in the DNA double-strand. TBP has also an α-helix domain, which is considered the recognition motif between the transcription factor and the DNA by binding to the DNA major groove. After binding to the DNA, TFIID starts to recruit TFIIB that can bind to TBP and GC- rich DNA sequence downstream of TATA motif at the same time. The TFIIB-TBP-DNA complex gives the order for the RNA Pol II to start the transcription and choose one of the two strands of the DNA to be the template. Binding the N- terminal domain "zinc ribbon "(a cystin-rich, zinc-binding region) of TFIIB to the RNA Pol II activates the recruitment of the RNA pol II and this, in turn, pulls the initiation complex to the RNA Pol II surface and the DNA, in this case, will easily reach to the active site. For the transcriptional initiation to start, it requires TFIIF that plays an active role in the formation of strong complex with the RNA Pol II, which recruits TFIIE, and TFIIE, in turn, helps to recruit TFIIH, the most complex structure amongst all other general transcription factors. The complex of the RNA Pol II, core promoter, and general transcription factors is called the pre-initiation complex (PIC). TFIIH has the ATPase/ helicase subunit, which hydrolyzes the ATP to melt the promoter (unwinding the DNA double-strands) by TFIIE assistance. TFIIF

captures the non-template DNA strand while the template strand goes to the RNA Pol II active site. (Hahn, 2005; (Kadonaga, 2004)

Transcription factors and enhancers

Since basal transcriptional machinery of general transcription factors and the RNA Pol II are not sufficient for gene transcription, additional proteins are needed to facilitate this process, and these proteins are called transcription factors that are able to identify and bind to the gene enhancers. The DNA looping model proposes that the basal transcriptional machinery assembles with the gene promoter and the transcription factors bind to the gene enhancer. looping out the DNA intervening allows the transcription factors to physically interact with the basal transcriptional machinery and this interaction will energize gene transcription. The interactions between the amino acid side chains of transcription factors and the DNA bases of purine and pyrimidine determine the accuracy of transcription factor binding to the DNA through the non-covalent hydrogen bonds between the amino acids and the DNA bases. The transcription factor peptide can bind to the major groove of the DNA (Rohs *et al.,* 2010). This interaction between the transcription factor and the DNA may happen electrostatically by the formation of the salt bridge with anionic phosphate groups (Luscombe, 2001).

Transcription factors are grouped in families based on sequence conservation and tertiary structure determined from X-ray crystallography and nuclear magnetic resonance spectroscopy. Members of each family have the same certain structural motifs for the DNA binding. These structures are zinc finger, basic zipper (Bzip), helix-turn-helix (HTH), basic helix-loop-helix(bHLH), and β-sheet. These motifs are composed of a specific tertiary protein structure in which its α-helix component interacts with the DNA major groove. Transcription factors are

capable of recognizing a short sequence of about 12 base pairs in most cases (Boron & Boulpaep, 2005) (Table 3-1).

Table 3.1 Transcription factors and the DNA sequences they recognize

Examples of specific transcription factors			
Factor	**Structural type**	**Recognition sequence**	**Binds as**
SP1	Zinc finger	5'-GGGCGG-3'	Monomer
AP-1	Basic zipper	5'-TGA(G/C) TCA-3'	Dimer
C/EBP	Basic zipper	5'-ATTGCGCAAT-3'	Dimer
Heat shock factor	Basic zipper	5'-XGAAX-3'	Trimer
ATF/CREB	Basic zipper	5'-TGACGTCA-3'	Dimer
c-Myc	Basic helix-loop-helix	5'-CACGTG-3'	Dimer
Oct-1	Helix-turn-helix	5'-ATGCAAAT-3'	Monomer
NF-1	Novel	5'-TTGGCXXXXXGCCAA-3'	Dimer
(G/C) = G or C **X = A, T, G or C**			

Adapted from (Boron & Boulpaep, 2005)

Predicting transcription factors that bind specific sites

Characterizing of the transcription factor binding sites (TFBSs) on the genomic DNA has become a big challenge after the publication of the complete sequence of the human genome. To identify the complete set of functional elements in these sites, many approaches have been utilized including experimental and computational comparisons of associated genomes. Transcription factors and the DNA sites that they bind are the most critical elements in any genome. The transcriptional regulatory network resulted from the interaction between the TFs and the DNA provides a big insight into the potential functions of the genes that are regulated by these DNA binding sites. Multiple methodologies have been used to detect the DNA sites that TFs bind to, such as gel shift assay, and Southern blotting of the DNA and proteins. However, these approaches are not efficient in terms of the time that they consume and being unable to cover the whole genome. Therefore, another technique was used later which is called SELEX (Systematic Evolution of Ligands by Exponential Evolution) to find the high-affinity binding sequences which is modified then to the genomic SELEX that uses the genomic library for the selections. In recent times, CHIP-chip assay, the highest throughput technique, is broadly used to identify genomic TFBSs in vivo.

In silico, many computational methods have been used to predict the proteins that can bind to the DNA and a method called a position weight matrix (PWM), or a position-specific scoring matrix (PSSM) is the most common method used to represent the degenerate sequence preferring of the DNA binding protein. In this computational method, the elements of PWM represent the scores that reflect the frequency of observing of a certain nucleotide at a certain position of specific or putative TFBSs. Despite the high accuracy in identifying of TF binding in vitro, the in vivo binding of these TFs might not occur or might not represent a direct regulatory function that

the in vitro binding shows. However, this defect does not belong to the computational techniques, but it is related to some biological facts, such as chromatin structures and competition on the binding site (Bulyk, 2004).

Material and methods

gRNA design

All the target sequences were identified using the National Center for Biotechnology Information (NCBI) website to direct cas9- mediated cleavage to target the GMR28G05 and GMR28F02 regions. CRISPR target finder website was used to design gRNAs with 100% off-target cleavage (http://tools.flycrispr.molbio.wesc.edu/targetfinder). The gRNAs were cloned into PCFD4 plasmid (port et al,2014) (Appendix 5). The following primers (Appendix3, table 2), 800 and 801 primers for GMR28G05, and 798 and 799 for GMR28F02, and 798 and 801 primers for GMR28G05+ GMR28F02, were annealed by PCR. The PCR products were inserted into BbSI- digested PCFD4 backbone using Gibson assembly Ultra Kit (Cat no.:GA1200-10). In the case of GMR28G05 gRNA an additional G nucleotide was added to the 5′ end of the GMR28G05 gRNA target sequence to facilitate the transcription from the U6-1 promoter (Addgene plasmid no. 49411). The resulting plasmid was cloned into E-coli competent cells.

Preparing the DNA for injection

A mixture of cas9 encoding the DNA plasmid (Appendix 5) and two gRNAs encoding plasmids for each deletion was used in a concentration of 500ng/µl for cas9 and 250ng/µl plasmids in a total volume of 20 µl. 2 µl of filtered food dye diluted 1:100 in distilled water was added to the 20 µl of the injection mixture. Needles used for the injection were made up of capillary tubes of 100 mm length, 1.0 mm outside diameter, 0.56 mm inside diameter borosilicate glass capillaries with filament (# 1 B100F-4 World precision Instruments Inc, Sarasota, FL). Capillaries were pulled by using P-100 Faming/ Brown Micropipette Puller (Sutter Instruments. Novato, A). A program with the following parameters was used to pull the needles: Heat 590, Pull 115, Velocity 15, Time 250, Pressure 600, and Ramp 590. Needles were

loaded with 1 µl of injection mix using an Eppendorf loader tip (Eppendorf), fitted into a needle holder attached to a syringe to facilitate injection, and controlled by a micromanipulator (Figure.3.2). The tips of the needles were opened with a micro-dissection scissor as needed.

Embryos injection

Fly stocks were amplified in advance to get enough eggs for injection. Approximately 300 mature flies of genotype $robo^{robo2}/robo2^{robo2}$ were transferred to an egg-lay cage with an apple juice plate incubated at 25˚C for at least one day prior to the injection day to get the flies accustomed. Embryos were collected on the apple juice plate each 20-30 min with the apple juice plate change each time the embryos were collected and yeast paste was spread in the center of the plate. The collected embryos were rinsed with water, and dechorionated for 60 s in 6% sodium hypochlorite, washed with tap water and lined up on a slide with a coverslip and covered with extra virgin organic olive oil. The slide with the embryos was moved to a Nikon TS 100 inverted microscope equipped with a micromanipulator. Embryos were injected by moving the stage to insert the needle in the posterior end of the embryo (Figure 3.2.A-C). The pressure was applied to the syringe until the DNA moved into the embryos which is monitored through the blue food dye. Injected embryos were thoroughly washed with ethanol 95% to remove the olive oil, then the ethanol was washed off with distilled water. The excess water on the coverslip was drained by applying the coverslip's edge onto a clean tissue. The coverslip was transferred into a food vial, pushed down into the food with the embryos anterior up until the embryos touched the food, so when the larvae hatch, they will crawl into the food. The marked food vials were placed into a humid chamber at 25 ˚C for (24-48 hr), or until the larvae hatch.

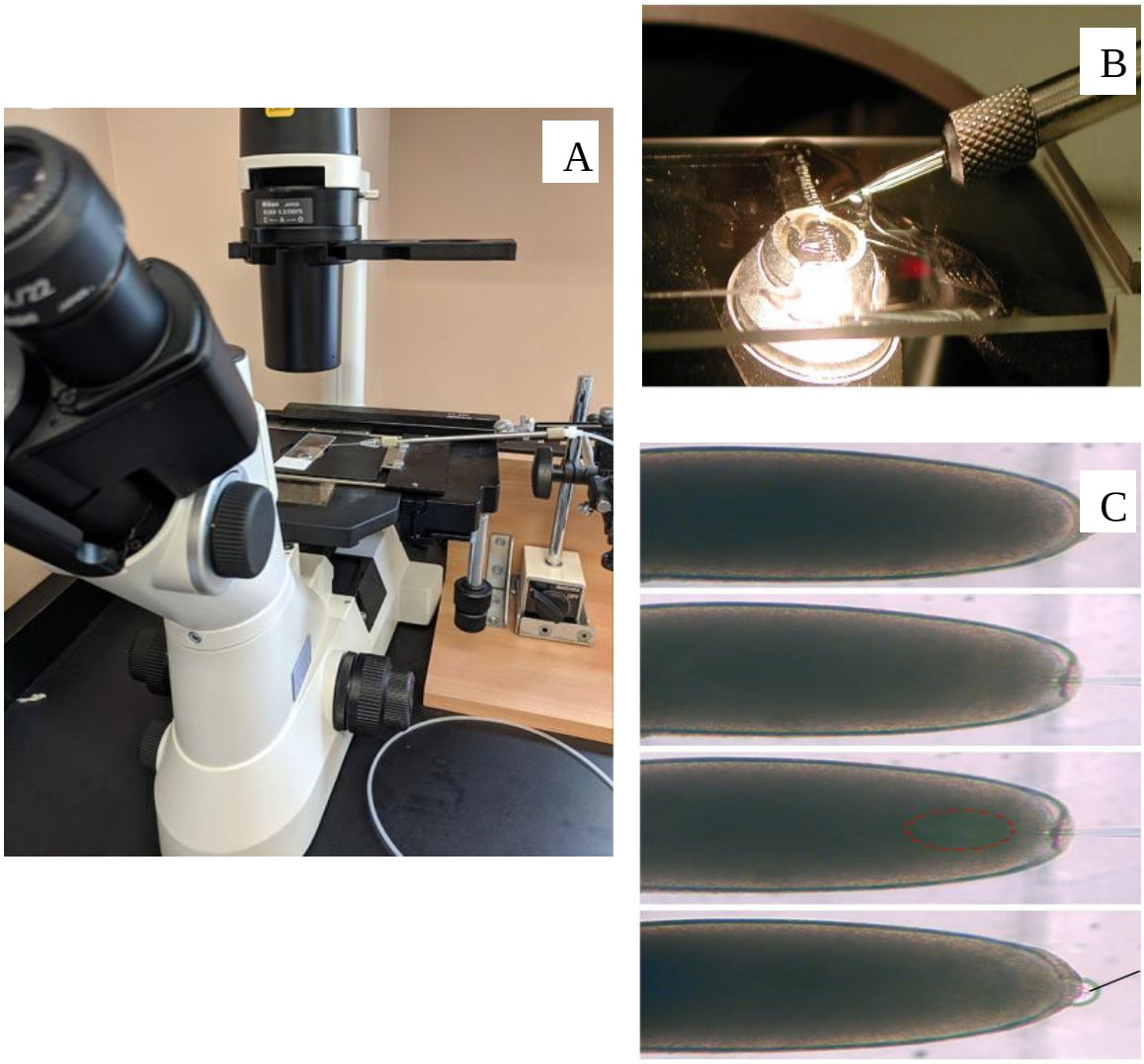


Figure 3.2. *Drosophila* embryo injection setup. **A** and **B**, (left, embryos), (right, injection needle). **C**. *Drosophila* embryos are injected by moving the stage to insert the needle in the posterior end of the embryo.

PCR analysis of the CRISPR-Cas9 deletions

After injected embryos have reached adulthood, the G0 flies were crossed 1:3 to Sco/CyOwg virgins. After 3 days of crossing, adults were removed and screened with PCR using primers 811, 812 for GMR28G04, 813, 814 for GMR28F02, and 811,814 for GMR28G04+ GMR28F02 (Appendix 3, Table 2). The progeny of positive PCR F1 will be collected to make a stock.

In silico studies for the transcription factors of *robo2* first intron putative enhancers

Depending on the predicted transcription factors taken from the literature, I used two bioinformatics websites JASPAR 2020 (http://jaspar.genereg.net/) and the footprint database (FootprintDB) website (http://floresta.eead.csic.es/footprintdb/index.php?search_entries) to investigate whether these candidate transcription factors can bind to *robo2* putative enhancers (GMR28G04 and GMR28F02). The candidate transcription factors from literature were submitted through JASPAR 2020 website to provide the ID numbers of the nominated transcription factors. These ID numbers were utilized to find the predicted consensus regions by using the FootprintDB website and the DNA logo that represents the probability of the nucleotides in each DNA transcription factor binding site. These consensus regions were aligned with *robo2* fragments sequences using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) to find the regions where these sequences are in *robo2* fragments.

Results and Discussion

In this chapter CRISPR/Cas9 was utilized to delete GMR28G05 and GMR28F02 fragments separately or together by using non-homologous end-joining (NHEJ). gRNAs were designed and inserted into PCFD4 plasmid. The constructed plasmids were confirmed by PCR (Figure 3.3) and sequencing (Figure.3.5A-C), then cloned in E. coli competent cells. These plasmids were injected alongside with Cas9 expressing plasmid into flies that have CRISPR allele where the genomic *robo2* was replaced by the HA-Robo2. Injected embryos were raised, crossed to Sco/CyOwg and offspring were collected for screening to identify the deletion in these fragments. A total of 666 embryos were injected with a mix of the GMR28G05 gRNA and the Cas9 plasmids, 216 embryos were injected with a mix of the GMR28F02 gRNA and the Cas9 plasmids, and 226 embryos were injected with a mix of the GMR28G05+GMR28F02 gRNA and the Cas9 plasmids (Table.3.2). Of the injected embryos, 21 survived of GMR28G05, 7 survived of GMR28F02, and 4 survived of GMR28G05+ GMR28F02 to adulthood. After crossing to Sco/CyOwg, potential founders of each group (GMR28G05, GMR28F02, and GMR28G05+GMR28F02) produced offspring of 16, 5, 3, respectively. The injection survival rate was 3.15%, 3.24%, and 1.76% of injected embryos to adulthood for GMR28G05, GMR28F02, and GMR28G05+GMR28F02 respectively and fertility rate found for those adults was 76.1% for GMR28G05, 71.4% for GMR28F02, 75% for GMR28G05+GMR28F02 (Table.3.2).

In general, the results of injection show that there is a low survival rate which, in turn, reduced the efficiency of the deletion. The rate of survival is comparable with another graduate student in our laboratory. Besides the low rate of survival of injected embryos, SARS-CoV-2 (COVID-19) pandemic has negatively affected my research. The university closure during the

pandemic has stopped my research and I was not able to test most of the injected flies. In future studies, I suggest that additional trials with improved techniques such as electroporation for chorionated and dechorionated *Drosophila* embryos might fix this pitfall.

Table.3.2. CRISPR Data for *Drosophila* embryo injection.

Injection mix	No. of injected embryos	%(no.) adult G0 survival	%(no.) fertile G0 survival	%(no.) G0 founders
GMR28G05	666	3.15(21)	76.1(16)	-
GMR28F02	216	3.24(7)	71.4 (5)	-
GMR28G05+ GMR28F02	226	1.76(4)	75 (3)	-

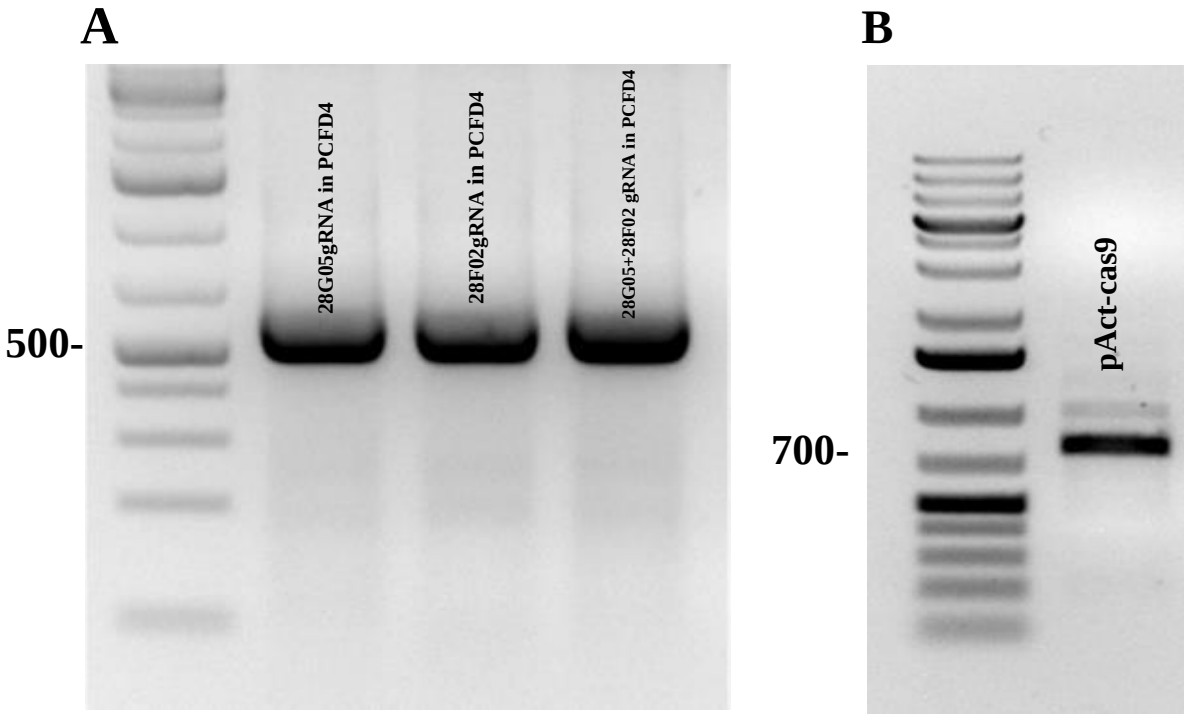


Figure 3.3 gRNA and CAS9 vectors confirmation. **A.**2µl of miniprep DNA of gRNAs in PCFD4 vector (GMR28G05 gRNA in PCFD4 595bp, GMR28F02 gRNA in PCFD4 595bp, and GMR28G05+GMR28F02 gRNA in PCFD4 600bp). **B**. 2µl of miniprep of pAct-cas9 vector 700bp.Plasmid sizes compared to 1Kb Plus DNA Ladder from Invitrogen (Carlsbad, CA).

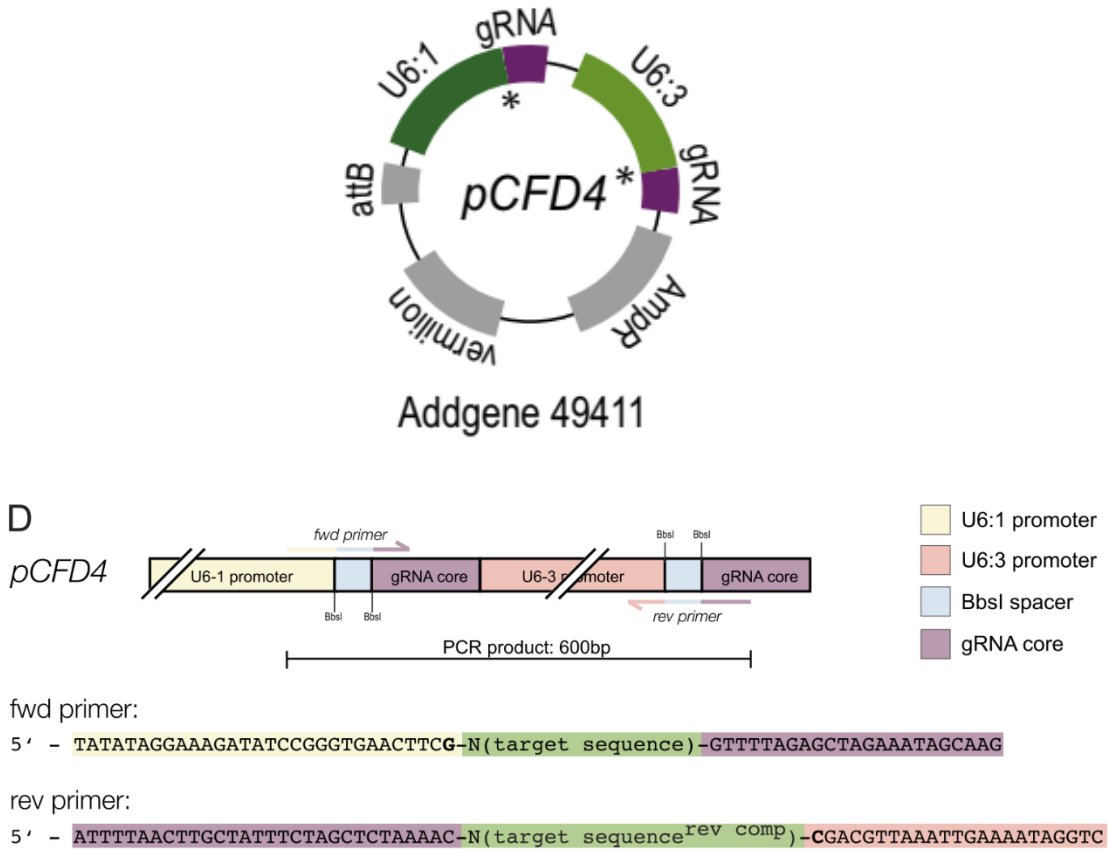


Figure3.4. Cloning tandem gRNA expression plasmid with PCFD4

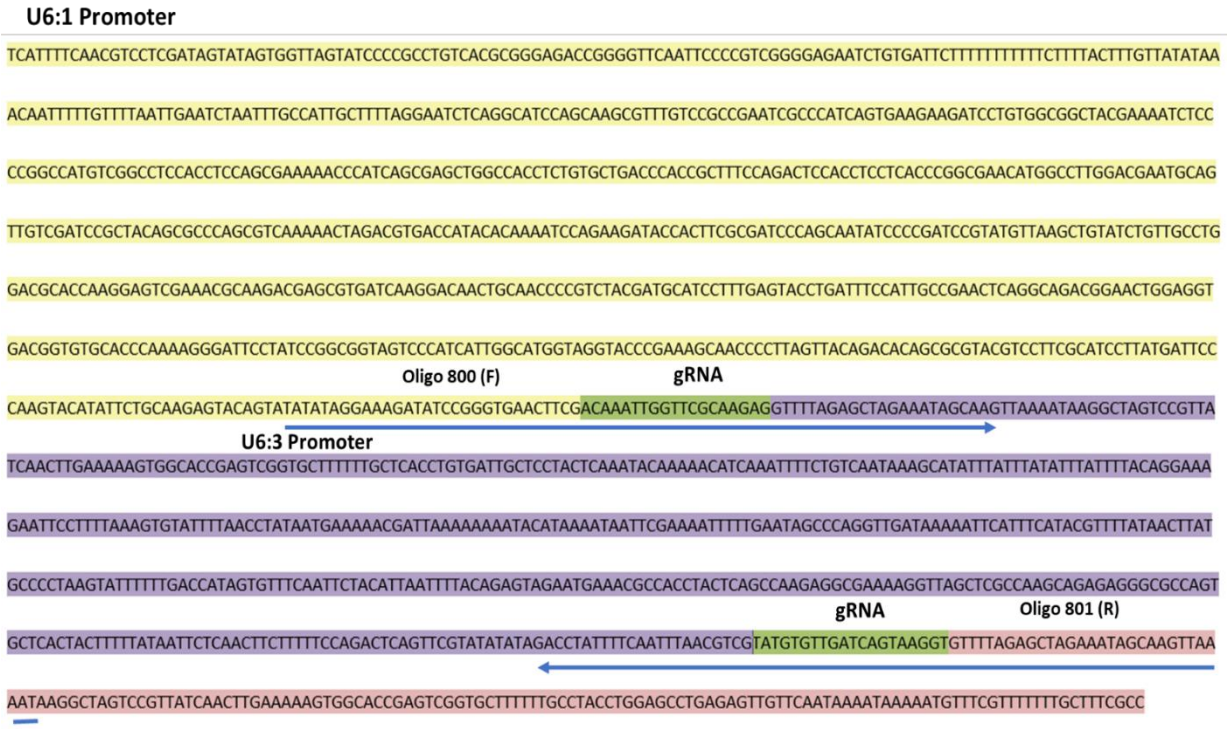


Figure 3.5A Schematic construction of GMR28G05 guide RNA in pCFD4**.**The pCFD4 plasmid was used to deliver the 5' and 3' gRNAs. Oligo 800 was used for the fwd. prime and 801 for the rev primer. gRNAs are under the control of U6:1 and U6:3 promoters (Addgene plasmid 49411 expression in PCFD4).

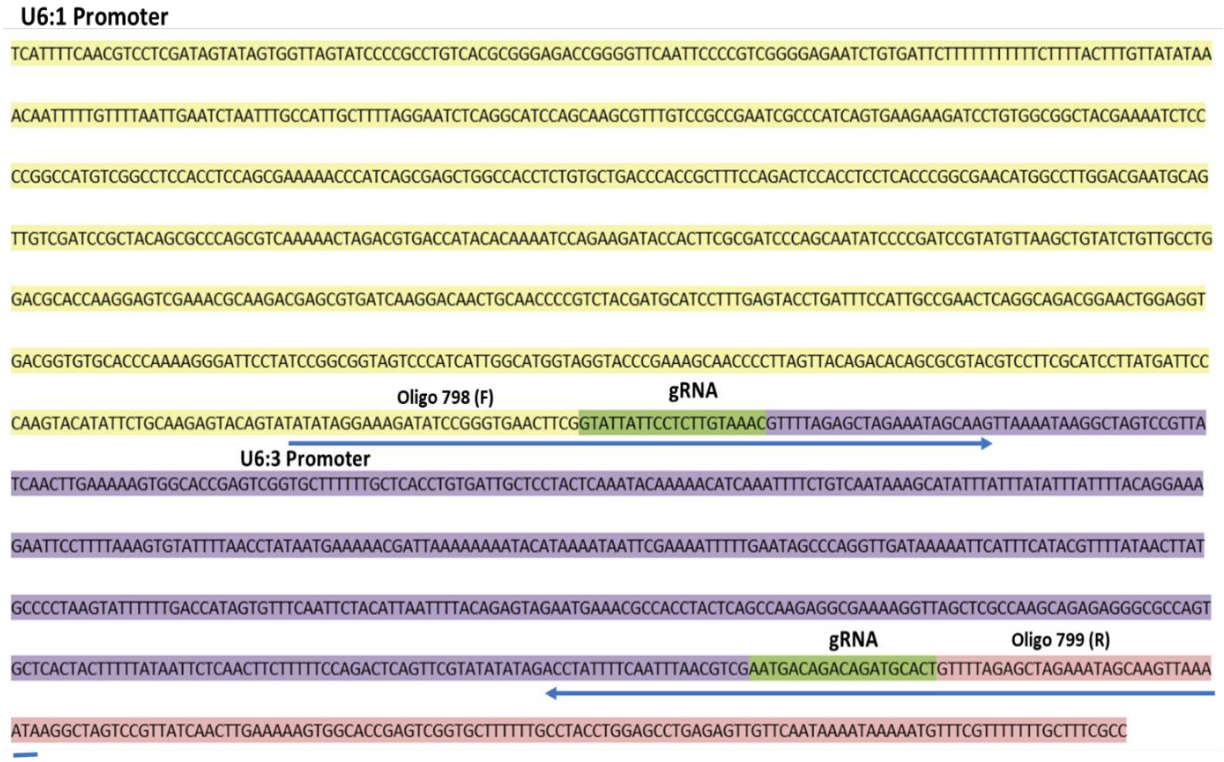


Figure 3.5B Schematic construction of GMR28F02 guide RNA in pCFD4.The pCFD4 plasmid was used to deliver the 5' and 3' gRNAs. Primer 798 was used for the fwd. primer and 799 for the rev primer. gRNAs are under the control of U6:1 and U6:3 promoters (Addgene plasmid 49411 expression in PCFD4).

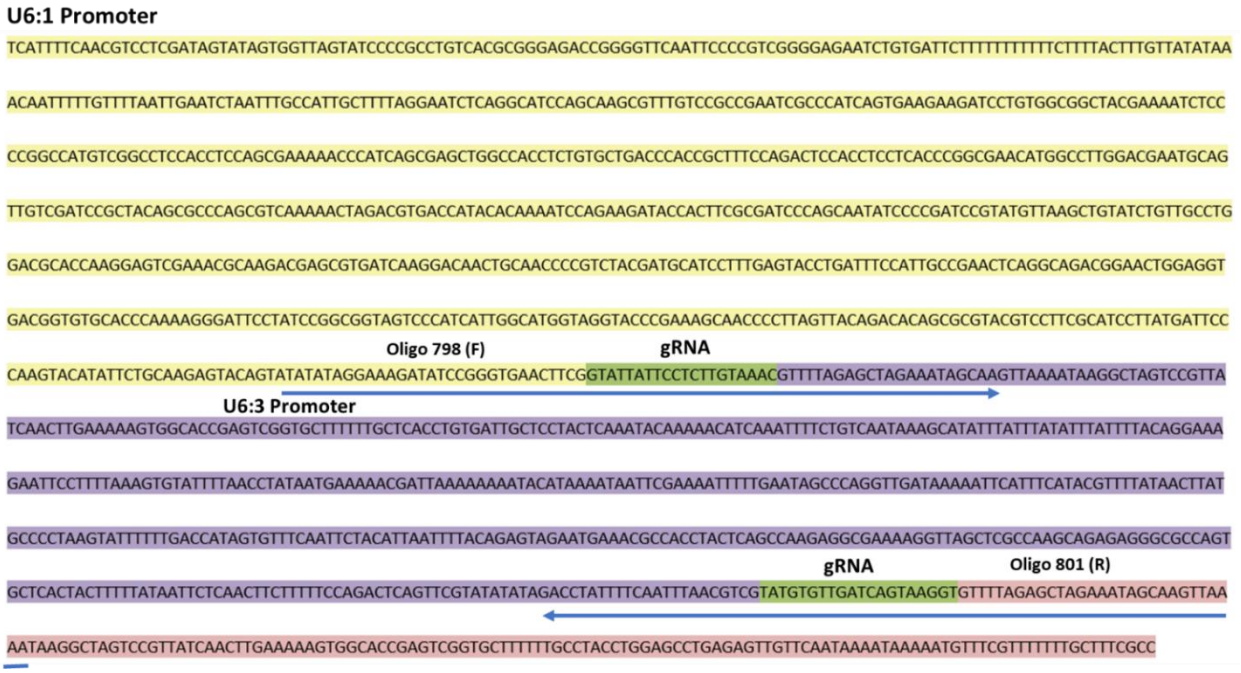


Figure 3.5C Schematic construction of GMR28G05+GMR28F02 guide RNA in pCFD4**.**The pCFD4 plasmid was used to deliver the 5’ and 3’ gRNAs. Oligo 798 was used for the fwd. primer and 801 for the rev primer. gRNAs are under the control of U6:1 and U6:3 promoters (Addgene plasmid 49411 expression in PCFD4).

Transcription factor candidates that can potentially bind to *robo2* putative enhancers (GMR28F02 and GMR28G05)

1. Hb9 transcription factor (exex in *Drosophila*)

Hb9 transcription factor (exex), also known as Extra-extra, Q9VSC2_DROME, and RE39081p, in *Drosophila* embryos, can regulate axon guidance by acting upstream of *robo2* and *robo3*. This transcription factor is necessary for *robo2* expression in ventrally extending motor neurons. In Hb9 mutants, restoring Robo2 activity plays an important role in motor axon defect rescue. Hb9 can regulate Robo2 over its conserved repressor domain and acts in parallel with NKX6.1. Robo2 and Robo3 play a key role as Hb9 effectors to regulate the nervous system development and Hb9 can regulate the mediolateral positioning of axons through Robo2 and Robo3. A previous study accomplished on Hb9+ cells that express Robo2 in which *robo2* expression was scored, *robo2* mRNA expression shows a decrease in Hb9+ mutant (Landgraf *et al.,* 1997; Santiago *et al.,* 2014).

To examine whether Hb9 could be a transcription factor that regulates *robo2* expression by binding GMR28F02 and GMR28G05, I used a combination of two bioinformatics websites (JASPAR 2020, and FootprintDB, http://floresta.eead.csic.es/footprintdb/index.php) as described in the materials and methods section. I found that the specific binding motif of Hb9 (TAATTA) is found at two locations of GMR28F02 and three locations of GMR28G05 (Figure.3.6)

These results were expected because Robo2 has been detected as an effector of Hb9 in motor neurons. Also, both Robo2 and Hb9 mutants share the same defects in the lateral positioning. Together, these data propose that Hb9 might regulate the medial and the lateral positions of a subset of interneurons by affecting Robo2. Apterus axon in the wildtype embryos projects its

fascicle in the medial FasII pathway on both sides of the midline. *robo2* overexpression in the Apterus neurons shifts their axons laterally (Evans & Bashaw, 2010; Rajagopalan *et al.,* 2000b; Simpson *et al.,* 2000a). Overexpression of Hb9 shows a similar phenotype by shifting apterus axons laterally. Importantly, overexpression of Hb9 in apterus neurons results in *robo2* upregulation significantly. In *robo2* mutants (when both copies of *robo2* were removed), overexpression of Hb9 led to full suppression in Hb9's gain of function phenotype while Apterus axons still look like a wildtype (Santiago *et al.,* 2014), which means that ectopic expression of Hb9 enhances *robo2* expression and this *robo2* expression, in turn, affects mediolateral positioning axons.

A.

B.

F02 nt 250-255:
CCCATTTTTCCTCAGCACTTAATTACGCTCATTTGGCATATATGCTCTTCATTTG

F02 nt 2770-2775:
TATGTACATTTTCTTACATTAATTATAAAGTAACTATTTTTTTTACAATACCAGT

G05 nt 1716-1721:
GACATTCCATTTTCGCTATAATTAATAAACACACGGCAAAACAATTTATCAGTT

G05 nt 2456-2461:
CTGCCAAAACGAACAACATAATTACCGTTAAGAATTGCGAGTTGCAACATCAGC

G05 nt 3527-3532:
TGCGGAGCCAAAGTTGGCTAATTAGCTATGCAAACTATGAGATACACAACACAA

Figure 3.6. Predicted DNA binding motif sequence of Hb9 in GMR28F02 and GMR28G05. **A.** Logo of the DNA binding motif sequence of Hb9 shows the probability of nucleotides in the sequence (adapted from http://floresta.eead.csic.es/footprintdb/index.php). **B.** locations and numbers of the DNA binding motif sequence of Hb9 in GMR28F02 and GMR28G05. (F02 and G05 refer to GMR28F02 and GMR28G05, respectively).

2.Nkx6.1

NKX6.1, Also known as NKX6A and HGTX, is one of the homeodomains transcription factors that belong to the NK family whose genes products are involved in specifying of the cells fate and differentiation in specific tissues. It is expressed in the ventral part of the neural tube of the mouse and chick embryos (Jørgensen *et al.,* 1999). It has an important role in the spinal cord and motor neurons specification in addition to its role in axon projecting and muscle targeting in motor neurons (Sander *et al.,* 2000; Thaler *et al.,* 1999; Garcia & Jessell, 2008; Arber *et al.,* 1999; Vallstedt *et al.,* 2001). In *Drosophila*, NKx6.1 is expressed in embryonic motor neurons that extend to the lateral or ventral body wall muscles and they are necessary for motor axons that project ventrally (Broihier & Skeath, 2002; Broihier *et al.,* 2004; Odden *et al.,* 2002). In addition to the motor neurons, NKx6.1 has shown expression in a subset of interneurons. *Drosophila* NKx6.1 can act as an activator or repressor; however, none of these activities were tested in vivo (Syu *et al.,* 2009).

To test the binding ability of NKX6.1 on GMR28F02 and GMR28G05 fragments, I followed the same procedure as in the previous transcription factor (Hb9). I found two locations of the consensus sequence (TTAATTA) of NKX6.1 in GMR28F02 while one consensus sequence (TTAATTG) of NKX6.1 was matched in GMR28G05 (Fig 3.7. B).

A.

B.

GMR28F02 nt 250-256:
CCCATTTTTCCTCAGCACTTAATTACGCTCATTTGGCATATATGCTCTTCATTTGT

GMR28F02 nt 2770-2776:
TATGTACATTTTCTTACATTAATTATAAAGTAACTATTTTTTTTACAATACCAGT

GMR28G05 nt 2562-2568:
CAGAACGTTTTGGGCTCCTTAATTGCCATTGGTAGACTCCATTTGGCATGGCCGC

Figure 3.7. Predicted DNA binding motif sequence of NKX6.1 in GMR28F02 and GMR28G05. **A**. Logo of the DNA binding motif sequence of NKX6.1 shows the probability of nucleotides in the sequence (adapted from http://floresta.eead.csic.es/footprintdb/index.php). **B.** locations and numbers of the DNA binding motif sequence of NKX6.1 in GMR28F02 and GMR28G05. (F02 and G05 refer to GMR28F02 and GMR28G05, respectively).

Homeodomain proteins compared to other transcription factors are indistinctive in their specificity to bind the DNA. These proteins recognize the consensus sequence ATTA, the DNA sequence that can be recognized by many different homeodomain proteins in vitro (Gehring *et al.,* 1994; Mannervik, 1999). Since many gene promoters contain the same core motif "TAAT" in their sequences, determining the downstream target becomes more complicated. Therefore, knowing the flanking sequence of this core motif is of utmost important goal to reduce the binding affinity. A study accomplished by (Jørgensen *et al.,* 2007) shows that using in vitro binding site selection reveals identifying the DNA sequence of NKx6.1 binding site of TTAATTG/A .

Another study conducted by (Mirmira *et al.,* 2000) showed an approach to identify the optimal DNA binding by NKX6.1 homeodomain. The coding sequence of NKX6.1 was fused to 6 tandem histidine residues (His6) at N- terminus and expressed in E-coli, then the expressed NKX6.1 protein was purified and used to select for the DNA binding from a library of 55 bp nucleotides that have a central region of 15 random nucleotides. After isolation of protein and its bound DNA, the selected DNA sequence was amplified by PCR and used for several selection rounds. Multiple PCR product sequences were aligned to give the consensus sequence of TTAATTAC. This study comes in agreement with the previous study done by Jorgensen et al with an addition of one nucleotide (C) at the 3′ of TTAATTA (TTAATTAC) sequence motif, which makes it more specific for binding specific transcription factors.

3. Lhx2 transcription factor

Lhx2is a transcription factor that regulates numerous developmental aspects in different embryonic stages including proliferation control of progenitor, specification of the cell fate of post-mitotic progeny, cell differentiation, and axon pathfinding (Chou & Tole, 2019). Lhx2 is also known as ap; GJ14977; apterous; GJ14977-PA; GJ14977-PB. According to (Subramanian *et al.,* 2011), Lhx2 plays a central role in regulating the neuron-glia cell fate decision in the hippocampus. (Marcos-Mondéjar *et al.,* 2012) showed that *robo2* mRNA expression is repressed by the LIM-HD transcription factor Lhx2 in postmitotic thalamic neurons. Lhx2 overexpression in rostral (rTh) and intermediate (iTh) thalamic neurons results in an improper invasion of Thalamocortical axons (TCAs) in improper regions of the hypothalamus which, in turn, decreasing the number of axons that reach the cortex in a process similar to that of the absence of Slit/robo signaling (Bagri *et al.,* 2002; López-Bendito *et al.,* 2007).

In silico studies that I conducted using bioinformatics websites (explained in materials and methods), I found that the DNA motif (TAATTA) of Lhx2 (Fig3.8. A) can bind two locations of GMR28F02 and three locations of GMR28G05 (Fig3.8. B).

A.

B.

GMR28F02 nt 250-257:
CCCATTTTTCCTCAGCACTTAATTACGCTCATTTGGCATATATGCTCTTCATTTGT

GMR28F02 nt 2770-2776:
TATGTACATTTTCTTACATTAATTATAAAGTAACTATTTTTTTTACAATACCAGT

G05 nt 1716-1721:
GACATTCCATTTTCGCTATAATTAATAAACACACGGCAAAACAATTTATCAGTT

G05 nt 2456-2461:
CTGCCAAAACGAACAACATAATTACCGTTAAGAATTGCGAGTTGCAACATCAGC

G05 nt 3527-3532:
TGCGGAGCCAAAGTTGGCTAATTAGCTATGCAAACTATGAGATACACAACACAA

Figure 3.8. Predicted DNA binding motif sequence of Lhx2 in GMR28F02 and GMR28G05. **A.** Logo of the DNA binding motif sequence of Lhx2 shows the probability of nucleotides in the sequence (adapted from http://floresta.eead.csic.es/footprintdb/index.php). **B.** locations and numbers of the DNA binding motif sequence of Lhx2 in GMR28F02 and GMR28G05. (F02 and G05 refer to GMR28F02 and GMR28G05, respectively).

It was expected to find several hits of the DNA binding motif of Lhx2 in GMR28F02 and GMR28G05 sequences. It has been involved in directly regulating Robo2 function in commissural neurons of the developing spinal cord of mice by binding to specific regulatory sequences of *robo2*. The phylogenetic footprinting analysis of regions 20kb upstream and 70kb downstream of *robo2* transcription initiation site shows four putative *robo2* regulatory regions that can bind Lhx2 with evolutionarly conserved regions, and CHIP assay shows that Lhx2 binds significantly to the *robo2*-region 3. These results indicate that Lhx2 acts as a transcriptional regulator for *robo2* expression in thalamic neurons of mice when it binds to the regulatory region in the *robo2* gene in vivo. Putative Lhx2 DNA binding region shows 6bp consensus sequences in *robo2*. CHIP assays experiments show a specific enhancer region (TAATTA) in *robo2* is the target sequence that Lhx2 binds in the spinal cord of E12-5 embryonic stage of thalamic tissue in vivo in genomic rat, mouse, chimpanzee, macaque, human, and dog (Marcos-Mondéjar *et al.,* 2012).

Conclusion

To sum up, this chapter presents the importance of introns in gene expression and how deleting single or multiple regulatory regions might affect *robo2* expression. The CRISPR/Cas9 system was utilized to make a deletion of GMR28G05 and GMR28F02 from *robo2* first intron separately or together. The deletion efficiency was severely reduced due to the low injection survival rate. Besides, the SARS-CoV-2 (COVID-19) pandemic has stopped my research and as a result many of the injected flies have not been tested for the deletion. With additional trials and using more effective techniques such as electroporation, it will be possible to get these fragments deleted and see the effect of this deletion on *robo2* expression. The second part of this chapter focuses on predicting the transcription factors that can potentially bind to *robo2* putative enhancers. Both bioinformatics tools and literature were utilized to find these predicted transcription factors. Hb9 (exex in *Drosophila*) with its consensus binding motif to the putative enhancers (TAATTA), Nkx6.1 with (TTAATTG) sequence, and Lhx2 with (TAATTA) binding motif were transcription factor candidates that can potentially bind to *robo2* putative enhancers GMR28F02 and GMR28G05.

Chapter Four: Conclusions and future directions

Robo2 is a transmembrane protein that regulates three distinct roles in directing axon guidance in *Drosophila* central nervous system (CNS). These different roles of Robo2 depend in part on its distinct functional domains and on the dynamic transcription in different subsets of cells through embryogenesis in another part. Robo2 structural and functional domains in *Drosophila* were studied comprehensively in the previous years. However, little is known about the transcriptional regulation of *robo2*. This study focuses on testing multiple regions in and around the *robo2* gene that could potentially be enhancers for *robo2* in *Drosophila*. To find the *robo2* putative enhancers, seventeen transgenic lines created by Janelia Research Center were utilized. Each line has one fragment of *robo2* predicted enhancer inserted upstream of the yeast transcriptional activator GAL4. The seventeen lines were crossed to another transgenic line containing a GAL4 responsive GFP (UAS-TauMycGFP) to create a protein reporter system. Six promising enhancer candidates (GMR28F02, GMR28G05, GMR28C04, GMR28D10, GMR28B05, GMR28E07) from the seventeen transgenic lines were found. Among these fragments, GMR28F02 and GMR28G05 showed strong expression in the longitudinal pathway and two other fragments (GMR28E07 and GMR28B05) showed strong expression in the midline glia. To ensure the above fragments enhance the *robo2* gene expression, I have created rescue construct transgenic lines for these six fragments in addition to three fragments created in our lab (SEG01, SEG02, and SEG03). All nine fragments were introduced into the HA-Robo2 reporter transgene. *robo2* showed an expression in all these fragments except SEG01, and GMR28F02 and GMR28G05 were the strongest expression. To further investigate whether GMR28F02, GMR28G05 fragments can rescue *robo2* expression, I introduced these fragments separately or together into *robo2* mutant background. The combination of two fragments showed better rescue of *robo2* expression in the mutant background compared

to a single fragment. These results suggest that multiple enhancer regions might be required to restore *robo2* full expression.

I also generated and characterized an equivalent set of *robo2* transgenic lines that express the axonal marker *TauMyc* instead of *robo2*. This transgene consists of GMR28F02 or GMR28G05, and hsp70 promoter followed by *TauMyc* cDNA. The results showed that unlike GMR28G05, GMR28F02 drove strong expression in subsets of the lateral neurons, cell bodies, and commissural axons. Moreover, GMR28F02::*TauMyc* was introduced into *robo2* null mutant background. The results showed that TauMyc axons that normally express Robo2 in the lateral axons act as cell non-autonomously in *robo2* mutant and they are different from FasII positive axons that are found on both sides of the midline of *Drosophila* nerve cord.

The first and the biggest intron of *robo2* was hypothesized to play an important role in *robo2* expression and our results strongly agreed with that and indicated that GMR28F02 and GMR28G05 changed *robo2* expression. The CRISPR/Cas9 system was utilized to delete these fragments separately or together. After CRISPR/Cas9 and gRNAs vectors were successfully constructed and confirmed by the PCR and the DNA sequencing, they were injected into *Drosophila* embryos. Most of these embryos were not examined for deletion because of the SARS-CoV-2 (COVID-19) pandemic, which stopped my research. Further trials and using more efficient techniques such as electroporation could facilitate deleting GMR28F02 and GMR28G05 fragments and study the effect of their deletion on *robo2* expression.

Additionally, knowing which transcription factor(s) might bind to GMR28G05 and GMR28F02 fragments would help us to reveal the role(s) of these fragments regarding to *robo2* transcription. Thus, bioinformatics tools and literature were utilized. Three promising transcription factor candidates were found (Hb9, Nkx6.1 and Lhx2) that potentially bind GMR28F02 and GMR28G05

with the following DNA consensus binding motifs (TAATTA), (TTAATTG), and (TAATTA), respectively. These findings may open new pathways to understand the transcriptional regulation of *robo2* regarding its enhancers. In the future, using Chromatin Immunoprecipitation (CHIP) assay with specific antibodies for those predicted transcription factor proteins could provide strong evidence of how these fragments act as enhancer(s) for the *robo2* gene.

References

Amoiridis, G., Tzagournissakis, M., Christodoulou, P., Karampekios, S., Latsoudis, H., Panou, T., Simos, P., & Plaitakis, A. (2006). Patients with horizontal gaze palsy and progressive scoliosis due to ROBO3 E319K mutation have both uncrossed and crossed central nervous system pathways and perform normally on neuropsychological testing. *Journal of Neurology, Neurosurgery and Psychiatry*, *77*(9), 1047–1053. https://doi.org/10.1136/jnnp.2006.088435

Arber, S., Han, B., Mendelsohn, M., Smith, M., Jessell, T. M., & Sockanathan, S. (1999). Requirement for the homeobox gene Hb9 in the consolidation of motor neuron identity. *Neuron*, *23*(4), 659–674. https://doi.org/10.1016/S0896-6273(01)80026-X

Awad, T. A., & Truman, J. W. (1997). Postembryonic development of the midline glia in the CNS of Drosophila: Proliferation, programmed cell death, and endocrine regulation. *Developmental Biology*, *187*(2), 283–297. https://doi.org/10.1006/dbio.1997.8587

Bagri, A., Marín, O., Plump, A. S., Mak, J., Pleasure, S. J., Rubenstein, J. L. R., & Tessier-Lavigne, M. (2002). Slit proteins prevent midline crossing and determine the dorsoventral position of major axonal pathways in the mammalian forebrain. *Neuron*, *33*(2), 233–248. https://doi.org/10.1016/S0896-6273(02)00561-5

Barrangou, R. (2015). The roles of CRISPR-Cas systems in adaptive immunity and beyond. *Current Opinion in Immunology*, *32*, 36–41. https://doi.org/10.1016/j.coi.2014.12.008

Bashaw, G. J., Kidd, T., Murray, D., Pawson, T., & Goodman, C. S. (2000). Repulsive Axon Guidance. *Cell*, *101*(7), 703–715. https://doi.org/10.1016/s0092-8674(00)80883-1

Benabdallah, N. S., & Bickmore, W. A. (2016). Regulatory domains and their mechanisms. *Cold Spring Harbor Symposia on Quantitative Biology*, *80*, 45–51. https://doi.org/10.1101/sqb.2015.80.027268

Benoist, C., & Chambon, P. (1981). In vivo sequence requirements of the SV40 early promoter region. *Nature*, *290*(5804), 304–310. https://doi.org/10.1038/290304a0

Bibikova, M., Carroll, D., Segal, D. J., Trautman, J. K., Smith, J., Kim, Y.-G., & Chandrasegaran, S. (2001). Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. *Molecular and Cellular Biology*, *21*(1), 289–297. https://doi.org/10.1128/mcb.21.1.289-297.2001

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., & Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, *326*(5959), 1509–1512. https://doi.org/10.1126/science.1178811

Boron, W. F., & Boulpaep, E. L. (2005). Medical Physiology: A cellular an Molecular Approach (updated edition). In *Medical Physiology: A cellular an Molecular Approach (updated edition)*.

Broihier, H. T., Kuzin, A., Zhu, Y., Odenwald, W., & Skeath, J. B. (2004). Drosophila homeodomain protein Nkx6 coordinates motoneuron subtype identity and axonogenesis. *Development*, *131*(21), 5233–5242. https://doi.org/10.1242/dev.01394

Broihier, H. T., & Skeath, J. B. (2002). Drosophila homeodomain protein dHb9 directs neuronal fate via crossrepressive and cell-nonautonomous mechanisms. *Neuron*, *35*(1), 39–50. https://doi.org/10.1016/S0896-6273(02)00743-2

Bulyk, M. L. (2004). Computational prediction of transcription-factor binding site locations. *Genome Biology*, *5*(1). https://doi.org/10.1186/gb-2003-5-1-201

Chou, S. J., & Tole, S. (2019). Lhx2, an evolutionarily conserved, multifunctional regulator of forebrain development. *Brain Research*, *1705*, 1–14. https://doi.org/10.1016/j.brainres.2018.02.046

Christian, M., Cermak, T., Doyle, E. L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A. J., & Voytas, D. F. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, *186*(2), 756–761. https://doi.org/10.1534/genetics.110.120717

Clark, A. J., Archibald, A. L., McClenaghan, M., Simons, J. P., Wallace, R., & Whitelaw, C. B. (1993). Enhancing the efficiency of transgene expression. *Philosophical Transactions of the*

Royal Society of London. Series B, Biological Sciences, *339*(1288), 225–232. https://doi.org/10.1098/rstb.1993.0020

Comer, J. D., Alvarez, S., Butler, S. J., & Kaltschmidt, J. A. (2019). Commissural axon guidance in the developing spinal cord: From Cajal to the present day. *Neural Development*, *14*(1), 1–16. https://doi.org/10.1186/s13064-019-0133-1

Connolly, J. L., Seeley, P. J., & Greene, L. A. (1985). Regulation of growth cone morphology by nerve growth factor: A comparative study by scanning electron microscopy. *Journal of Neuroscience Research*, *13*(1–2), 183–198. https://doi.org/10.1002/jnr.490130113

Crews, S. T. (2010). Axon-glial interactions at the Drosophila CNS midline. *Cell Adhesion and Migration*, *4*(1), 67–71. https://doi.org/10.4161/cam.4.1.10208

Culotti, J. G., & Merz, D. C. (1998). DCC and netrins. *Current Opinion in Cell Biology*, *10*(5), 609–613. https://doi.org/10.1016/S0955-0674(98)80036-7

Dacey, D. M., & Packer, O. S. (2003). Colour coding in the primate retina: Diverse cell types and cone-specific circuitry. *Current Opinion in Neurobiology*, *13*(4), 421–427. https://doi.org/10.1016/S0959-4388(03)00103-X

Dent, E. W., Callaway, J. L., Szebenyi, G., Baas, P. W., & Kalil, K. (1999). Reorganization and movement of microtubules in axonal growth cones and developing interstitial branches. *Journal of Neuroscience*, *19*(20), 8894–8908. https://doi.org/10.1523/jneurosci.19-20-08894.1999

Dent, E. W., & Gertler, F. B. (2003). Cytoskeletal dynamics and transport in growth cone motility and guidance. *Neuron*, *40*(2), 209–227. https://doi.org/10.1016/S0896-6273(03)00633-0

Dickson, B. J., & Zou, Y. (2010). Navigating intermediate targets: the nervous system midline. *Cold Spring Harbor Perspectives in Biology*, *2*(8), 1–16. https://doi.org/10.1101/cshperspect.a002055

Dontchev, V. D., & Letourneau, P. C. (2002). Nerve growth factor and semaphorin 3A signaling pathways interact in regulating sensory neuronal growth cone motility. *Journal of Neuroscience*, *22*(15), 6659–6669. https://doi.org/10.1523/jneurosci.22-15-06659.2002

Evans, T. A., & Bashaw, G. J. (2010a). Axon guidance at the midline: of mice and flies. *Current Opinion in Neurobiology*, *20*(1), 79–85. https://doi.org/10.1016/j.conb.2009.12.006

Evans, T. A., & Bashaw, G. J. (2010b). Functional diversification of Drosophila Robo receptors: differential immunoglobulin domain functions promote distinct axon guidance decisions. *Curr Biology*, *23*(1), 1–7. https://doi.org/10.1038/jid.2014.371

Evans, T. A., & Bashaw, G. J. (2012). Slit/Robo-mediated axon guidance in Tribolium and Drosophila: Divergent genetic programs build insect nervous systems. *Developmental Biology*, *363*(1), 266–278. https://doi.org/10.1016/j.ydbio.2011.12.046

Evans, T. A., Santiago, C., Arbeille, E., & Bashaw, G. J. (2015). Robo2 acts in trans to inhibit slit-robo1 repulsion in pre-crossing commissural axons. *ELife*, *4*(JULY 2015), 1–26. https://doi.org/10.7554/eLife.08407

Farmer, W. T., Altick, A. L., Nural, H. F., Dugan, J. P., Kidd, T., Charron, F., & Mastick, G. S. (2008). Pioneer longitudinal axons navigate using floor plate and Slit/Robo signals. *Development*, *135*(22), 3643–3653. https://doi.org/10.1242/dev.023325

Fischbach, K. F., & Dittrich, A. P. M. (1989). The optic lobe of Drosophila melanogaster. I. A Golgi analysis of wild-type structure. *Cell and Tissue Research*, *258*(3), 441–475. https://doi.org/10.1007/BF00218858

Galléa, C., Popa, T., Billot, S., Méneret, A., Depienne, C., & Roze, E. (2011). Congenital mirror movements: A clue to understanding bimanual motor control. *Journal of Neurology*, *258*(11), 1911–1919. https://doi.org/10.1007/s00415-011-6107-9

Garcia, N. V. D. M., & Jessell, T. M. (2008). Early motor neuron pool identity and muscle Nerve trajectory defined by postmitotic restrictions in Nkx6.1 activity. *Neuron*. https://doi.org/10.1016/j.neuron.2007.11.033

Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G., & Wüthrich, K. (1994). Homeodomain-DNA recognition. *Cell*, *78*(2), 211–223. https://doi.org/10.1016/0092-8674(94)90292-5

Gruss, P., Dhar, R., & Khoury, G. (1981). Simian virus 40 tandem repeated sequences as an element of the early promoter. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(2 II), 943–947. https://doi.org/10.1073/pnas.78.2.943

Gundersen, R. W., & Barrett, J. N. (1979). Neuronal chemotaxis: Chick dorsal-root axons turn toward high concentrations of nerve growth factor. *Science*, *206*(4422), 1079–1080. https://doi.org/10.1126/science.493992

Hahn, S. (2005). *Structure and mechanism of the RNA Polymerase II transcription machinery*. *11*(5), 394–403.

Hannula-Jouppi, K., Kaminen-Ahola, N., Taipale, M., Eklund, R., Nopola-Hemmi, J., Kääriäinen, H., & Kere, J. (2005). The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. *PLoS Genetics*, *1*(4), 0467–0474. https://doi.org/10.1371/journal.pgen.0010050

Hasty, P., Rivera-Pérez, J., & Bradley, A. (1991). The length of homology required for gene targeting in embryonic stem cells. *Molecular and Cellular Biology*, *11*(11), 5586–5591. https://doi.org/10.1128/mcb.11.11.5586

Herculano-Houzel, S. (2009). The human brain in numbers: A linearly scaled-up primate brain. *Frontiers in Human Neuroscience*, *3*(NOV), 1–11. https://doi.org/10.3389/neuro.09.031.2009

Hidalgo, A., & Booth, G. E. (2000). Glia dictate pioneer axon trajectories in the Drosophila embryonic CNS. *Development*, *127*(2), 393–402.

Homem, C. C. F., & Knoblich, J. A. (2012). Drosophila neuroblasts: A model for stem cell biology. *Development (Cambridge)*, *139*(23), 4297–4310. https://doi.org/10.1242/dev.080515

Iliakis, G., Wang, H., Perrault, A. R., Boecker, W., Rosidi, B., Windhofer, F., Wu, W., Guan, J., Terzoudi, G., & Panteliasc, G. (2004). Mechanisms of DNA double strand break repair and chromosome aberration formation. *Cytogenetic and Genome Research*, *104*(1–4), 14–20. https://doi.org/10.1159/000077461

Jørgensen, M. C., Ahnfelt-Rønne, J., Hald, J., Madsen, O. D., Serup, P., & Hecksher-Sørensen, J. (2007). An illustrated review of early pancreas development in the mouse. *Endocrine Reviews*, *28*(6), 685–705. https://doi.org/10.1210/er.2007-0016

Jørgensen, M. C., Vestergård Petersen, H., Ericson, J., Madsen, O. D., & Serup, P. (1999). Cloning and DNA-binding properties of the rat pancreatic β-cell-specific factor Nkx6.1. *FEBS Letters*, *461*(3), 287–294. https://doi.org/10.1016/S0014-5793(99)01436-2

Juneau, K., Miranda, M., Hillenmeyer, M. E., Nislow, C., & Davis, R. W. (2006). Introns regulate RNA and protein abundance in yeast. *Genetics*, *174*(1), 511–518. https://doi.org/10.1534/genetics.106.058560

Kadonaga, J. T. (2004). Regulation of RNA Polymerase II Transcription by Sequence-Specific DNA Binding Factors. *Cell*, *116*(2), 247–257. https://doi.org/10.1016/S0092-8674(03)01078-X

Kaprielian, Z., Runko, E., & Imondi, R. (2001). Axon guidance at the midline choice point. *Developmental Dynamics*, *221*(2), 154–181. https://doi.org/10.1002/dvdy.1143

Kidd, T., Brose, K., Mitchell, K. J., Fetter, R. D., Tessier-Lavigne, M., Goodman, C. S., & Tear, G. (1998). Roundabout controls axon crossing of the CNS midline and defines a novel subfamily of evolutionarily conserved guidance receptors. *Cell*, *92*(2), 205–215. https://doi.org/10.1016/S0092-8674(00)80915-0

Kim, Y. G., Cha, J., & Chandrasegaran, S. (1996). Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(3), 1156–1160. https://doi.org/10.1073/pnas.93.3.1156

Landgraf, M., Bossing, T., Technau, G. M., & Bate, M. (1997). The origin, location, and projections of the embryonic abdominal motorneurons of Drosophila. *Journal of Neuroscience*, *17*(24), 9642–9655. https://doi.org/10.1523/jneurosci.17-24-09642.1997

Letourneau, P. C. (1978). Chemotactic response of nerve fiber elongation to nerve growth factor. *Developmental Biology*, *66*(1), 183–196. https://doi.org/10.1016/0012-1606(78)90283-X

Li, M. J., Yan, B., Sham, P. C., & Wang, J. (2014). Exploring the function of genetic variants in the non-coding genomic regions: Approaches for identifying human regulatory variants affecting gene expression. *Briefings in Bioinformatics*, *16*(3), 393–412. https://doi.org/10.1093/bib/bbu018

Lin, S., Staahl, B. T., Alla, R. K., & Doudna, J. A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *ELife*, *3*, e04766. https://doi.org/10.7554/eLife.04766

López-Bendito, G., Flames, N., Ma, L., Fouquet, C., Di Meglio, T., Chedotal, A., Tessier-Lavigne, M., & Marín, O. (2007). Robo1 and Robo2 cooperate to control the guidance of major axonal tracts in the mammalian forebrain. *Journal of Neuroscience*, *27*(13), 3395–3407. https://doi.org/10.1523/JNEUROSCI.4605-06.2007

Lowery, L. A., & Vactor, D. Van. (2009). The trip of the tip: Understanding the growth cone machinery. In *Nature Reviews Molecular Cell Biology*. https://doi.org/10.1038/nrm2679

Lu, W., Van Eerde, A. M., Fan, X., Quintero-Rivera, F., Kulkarni, S., Ferguson, H., Kim, H. G., Fan, Y., Xi, Q., Li, Q. G., Sanlaville, D., Andrews, W., Sundaresan, V., Bi, W., Yan, J., Giltay, J. C., Wijmenga, C., De Jong, T. P. V. M., Feather, S. A., … Maas, R. L. (2007). Disruption of ROBO2 is associated with urinary tract anomalies and confers risk of vesicoureteral reflux. *American Journal of Human Genetics*, *80*(4), 616–632. https://doi.org/10.1086/512735

Luo, Y., Raible, D., & Raper, J. A. (1993). Collapsin: A protein in brain that induces the collapse and paralysis of neuronal growth cones. *Cell*, *75*(2), 217–227. https://doi.org/10.1016/0092-8674(93)80064-L

Luscombe, N. M. (2001). Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Research*, *29*(13), 2860–2874. https://doi.org/10.1093/nar/29.13.2860

Mannervik, M. (1999). Target genes of homeodomain proteins. *BioEssays*, *21*(4), 267–270. https://doi.org/10.1002/(SICI)1521-1878(199904)21:4<267::AID-BIES1>3.0.CO;2-C

Marcos-Mondéjar, P., Peregrín, S., Li, J. Y., Carlsson, L., Tole, S., & López-Bendito, G. (2012). The Lhx2 transcription factor controls thalamocortical axonal guidance by specific regulation of Robo1 and Robo2 receptors. *Journal of Neuroscience*, *32*(13), 4372–4385. https://doi.org/10.1523/JNEUROSCI.5851-11.2012

McDonald, J. A., Holbrook, S., Isshiki, T., Weiss, J., Doe, C. Q., & Mellerick, D. M. (1998). Dorsoventral patterning in the Drosophila central nervous system: The vnd homeobox gene specifies ventral column identity. *Genes and Development*, *12*(22), 3603–3612. https://doi.org/10.1101/gad.12.22.3603

Mecollari, V., Nieuwenhuis, B., & Verhaagen, J. (2014). A perspective on the role of class iii semaphorin signaling in central nervous system trauma. *Frontiers in Cellular Neuroscience*, *8*(October), 1–17. https://doi.org/10.3389/fncel.2014.00328

Mirmira, R. G., Watadaf, H., & German, M. S. (2000). β-Cell differentiation factor Nkx6.1 contains distinct DNA binding inteference and transcriptional repression domains. *Journal of Biological Chemistry*, *275*(19), 14743–14751. https://doi.org/10.1074/jbc.275.19.14743

Moscou, M. J., & Bogdanove, A. J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science*, *326*(5959), 1501. https://doi.org/10.1126/science.1178817

Mueller, B. K. (1999). Growth cone guidance: First steps towards a deeper understanding. *Annual Review of Neuroscience*, *22*(1), 351–388. https://doi.org/10.1146/annurev.neuro.22.1.351

Muotri, A. R., & Gage, F. H. (2006). Generation of neuronal variability and complexity. In *Nature*. https://doi.org/10.1038/nature04959

Nikolov, D. B., & Burley, S. K. (1997). RNA polymerase II transcription initiation: A structural view. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(1), 15–22. https://doi.org/10.1073/pnas.94.1.15

Nomenclature, U., & Ligands, T. (1997). Unified nomenclature for Eph family receptors and their ligands, the ephrins. Eph Nomenclature Committee. *Cell*, *90*(3), 403–404. https://doi.org/10.1016/S0092-8674(00)80500-0

Nugent, A. A., Kolpak, A. L., & Engle, E. C. (2012). Human disorders of axon guidance. *Current Opinion in Neurobiology*, *22*(5), 837–843. https://doi.org/10.1016/j.conb.2012.02.006

Odden, J. P., Holbrook, S., & Doe, C. Q. (2002). Drosophila HB9 is expressed in a subset of motoneurons and interneurons, where it regulates gene expression and axon pathfinding. *Journal of Neuroscience*, *22*(21), 9143–9149. https://doi.org/10.1523/JNEUROSCI.22-21-09143.2002

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., & Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics*, *40*(12), 1413–1415. https://doi.org/10.1038/ng.259

Pardo, B., Gómez-González, B., & Aguilera, A. (2009). DNA double-strand break repair: How to fix a broken relationship. *Cellular and Molecular Life Sciences*, *66*(6), 1039–1056. https://doi.org/10.1007/s00018-009-8740-3

Patel, M. B., & Wang, J. (2019). The identification and interpretation of cis-regulatory noncoding mutations in cancer. *High-Throughput*, *8*(1), 1–22. https://doi.org/10.3390/ht8010001

Patel, N. H. (1994). Imaging Drosophila Embryos and Larvae Using Antibody Probes. In *Methods in Cell Biology*.

Pennacchio, L. A., Bickmore, W., Dean, A., Nobrega, M. A., & Bejerano, G. (2013). Enhancers: five essential questions. *Nature Reviews. Genetics*, *14*(4), 288—295. https://doi.org/10.1038/nrg3458

Pfeiffer, B. D., Jenett, A., Hammonds, A. S., Ngo, T. T. B., Misra, S., Murphy, C., Scully, A., Carlson, J. W., Wan, K. H., Laverty, T. R., Mungall, C., Svirskas, R., Kadonaga, J. T., Doe, C. Q., Eisen, M. B., Celniker, S. E., & Rubin, G. M. (2008). Tools for neuroanatomy and neurogenetics in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(28), 9715–9720. https://doi.org/10.1073/pnas.0803697105

Rajagopalan, S., Nicolas, E., Vivancos, V., Berger, J., & Dickson, B. J. (2000). Crossing the midline: Roles and regulation of Robo receptors. *Neuron*, *28*(3), 767–777. https://doi.org/10.1016/S0896-6273(00)00152-5

Rajagopalan, S., Vivancos, V., Nicolas, E., & Dickson, B. J. (2000). Selecting a longitudinal pathway: Robo receptors specify the lateral position of axons in the Drosophila CNS. *Cell*, *103*(7), 1033–1045. https://doi.org/10.1016/S0092-8674(00)00207-5

Raper, J., & Mason, C. (2010a). Cellular strategies of axonal pathfinding. *Cold Spring Harbor Perspectives in Biology*, *2*(9), 1–21. https://doi.org/10.1101/cshperspect.a001933

Reichardt, L. F. (2006). Neurotrophin-regulated signalling pathways. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1473), 1545–1564. https://doi.org/10.1098/rstb.2006.1894

Renzi, M. J., Tamara Lee, W., & Raper, J. A. (2000). Olfactory sensory axons expressing a dominant-negative semaphorin receptor enter the CNS early and overshoot their target. *Neuron*, *28*(2), 437–447. https://doi.org/10.1016/S0896-6273(00)00123-9

Rohs, R., Jin, X., West, S. M., Joshi, R., Honig, B., & Mann, R. S. (2010). Origins of Specificity in Protein-DNA Recognition. *Annual Review of Biochemistry*, *79*(1), 233–269. https://doi.org/10.1146/annurev-biochem-060408-091030

Sáchez-Soriano, N., Tea, G., Whitington, P., & Prokop, A. (2007). Drosophila as a genetic and cellular model for studies on axonal growth. *Neural Development*, *2*(1), 1–28. https://doi.org/10.1186/1749-8104-2-9

Sakharkar, M. K., Chow, V. T. K., & Kangueane, P. (2004). Distributions of exons and introns in the human genome. *In Silico Biology*, *4*(4), 387–393.

Sánchez-Soriano, N., Gonçalves-Pimentel, C., Beaven, R., Haessler, U., Ofner-Ziegenfuss, L., Ballestrem, C., & Prokop, A. (2010). Drosophila growth cones: A genetically tractable platform for the analysis of axonal growth dynamics. *Developmental Neurobiology*, *70*(1), 58–71. https://doi.org/10.1002/dneu.20762

Sander, M., Paydar, S., Ericson, J., Briscoe, J., Berber, E., German, M., Jessell, T. M., & Rubenstein, J. L. R. (2000). Ventral neural patterning by. *Genes and Development*, *14*, 2134–2139. https://doi.org/10.1101/gad.820400.vitro

Santiago, C., Labrador, J. P., & Bashaw, G. J. (2014). The homeodomain transcription factor Hb9 controls axon guidance in drosophila through the regulation of robo receptors. *Cell Reports*, *7*(1), 153–165. https://doi.org/10.1016/j.celrep.2014.02.037

Schindelin, J., Arganda-Carrera, I., Frise, E., Verena, K., Mark, L., Tobias, P., Stephan, P., Curtis, R., Stephan, S., Benjamin, S., Jean-Yves, T., Daniel, J. W., Volker, H., Kevin, E., Pavel, T., & Albert, C. (2009). Fiji - an Open platform for biological image analysis. *Nature Methods*, *9*(7). https://doi.org/10.1038/nmeth.2019.Fiji

Schramm, L., & Hernandez, N. (2002). Recruitment of RNA polymerase III to its target promoters. *Genes and Development*, *16*(20), 2593–2620. https://doi.org/10.1101/gad.1018902

Shabalina, S. A., Ogurtsov, A. Y., Spiridonov, A. N., Novichkov, P. S., Spiridonov, N. A., & Koonin, E. V. (2010). Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. *Molecular Biology and Evolution*, *27*(8), 1745–1749. https://doi.org/10.1093/molbev/msq086

Simpson, J. H., Bland, K. S., Fetter, R. D., & Goodman, C. S. (2000). Short-range and long-range guidance by Slit and its Robo receptors: A combinatorial code of Robo receptors controls lateral position. *Cell*, *103*(7), 1019–1032. https://doi.org/10.1016/S0092-8674(00)00206-3

Simpson, J. H., Kidd, T., Bland, K. S., & Goodman, C. S. (2000). Short-range and long-range guidance by Slit and its robo receptors: Robo and Robo2 play distinct roles in midline guidance. *Neuron*, *28*(3), 753–766. https://doi.org/10.1016/S0896-6273(00)00151-3

Smith, J., Berg, J. M., & Chandrasegaran, S. (1999). A detailed study of the substrate specificity of a chimeric restriction enzyme. *Nucleic Acids Research*, *27*(2), 674–681. https://doi.org/10.1093/nar/27.2.674

Spitzweck, B., Brankatschk, M., & Dickson, B. J. (2010). Distinct Protein Domains and Expression Patterns Confer Divergent Axon Guidance Functions for Drosophila Robo Receptors. *Cell*, *140*(3), 409–420. https://doi.org/10.1016/j.cell.2010.01.002

Ströhl, F., Lin, J. Q., Laine, R. F., Wong, H. H. W., Urbancic, V., Cagnetta, R., Holt, C. E., & Kaminski, C. F. (2017). Single Molecule Translation Imaging Visualizes the Dynamics of Local β-Actin Synthesis in Retinal Axons. *Scientific Reports*, *7*(1), 1–6. https://doi.org/10.1038/s41598-017-00695-7

Subramanian, L., Sarkar, A., Shetty, A. S., Muralidharan, B., Padmanabhan, H., Piper, M., Monuki, E. S., Bach, I., Gronostajski, R. M., Richards, L. J., & Tole, S. (2011). Transcription factor Lhx2 is necessary and sufficient to suppress astrogliogenesis and promote neurogenesis in the developing hippocampus. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(27). https://doi.org/10.1073/pnas.1101109108

Syu, L. J., Uhler, J., Zhang, H., & Mellerick, D. M. (2009). The Drosophila Nkx6 homeodomain protein has both activation and repression domains and can activate target gene expression. *Brain Research*, *1266*, 8–17. https://doi.org/10.1016/j.brainres.2009.01.068

Tessier-Lavigne, M., & Goodman, C. S. (1996). The molecular biology of axon guidance. *Science*, *274*(5290), 1123–1133. https://doi.org/10.1126/science.274.5290.1123

Thaler, J., Harrison, K., Sharma, K., Lettieri, K., Kehrl, J., & Pfaff, S. L. (1999). Active Suppression of Interneuron Programs within Developing Motor Neurons Revealed by Analysis of Homeodomain Factor HB9 the concentration-dependent activity of Sonic hedge- hog (Shh) (Ericson *et al.,* 1997). Progenitor cells within the ventricular zone i. *Neuron*, *23*(4), 675–687. https://doi.org/10.1016/S0896-6273(01)80027-1

Tianfang Ge, D., Tipping, C., Brodsky, M. H., & Zamore, P. D. (2016). Rapid screening for CRISPR-directed editing of th drosophila genome using white coconversion. *G3: Genes, Genomes, Genetics*, *6*(10), 3197–3206. https://doi.org/10.1534/g3.116.032557

Tolwinski, N. S. (2017). Introduction: Drosophila-A model system for developmental biology. *Journal of Developmental Biology*, *5*(3), 10–11. https://doi.org/10.3390/jdb5030009

Ugur, B., Chen, K., & Bellen, H. J. (2016). Drosophila tools and assays for the study of human diseases. *DMM Disease Models and Mechanisms*, *9*(3), 235–244. https://doi.org/10.1242/dmm.023762

Vallstedt, A., Muhr, J., Pattyn, A., Pierani, A., Mendelsohn, M., Sander, M., Jessell, T. M., & Ericson, J. (2001). Different levels of repressor activity assign redundant and specific roles to Nkx6 genes in motor neuron and interneuron specification. *Neuron*, *31*(5), 743–755. https://doi.org/10.1016/S0896-6273(01)00412-3

Venken, K. J. T., Simpson, J. H., & Bellen, H. J. (2011). Genetic manipulation of genes and cells in the nervous system of the fruit fly. In *Neuron*. https://doi.org/10.1016/j.neuron.2011.09.021

Vitriol, E. A., & Zheng, J. Q. (2012). Growth Cone Travel in Space and Time: The Cellular Ensemble of Cytoskeleton, Adhesion, and Membrane. *Neuron*, *73*(6), 1068–1081. https://doi.org/10.1016/j.neuron.2012.03.005

Wolf-Dietrich Heyer. (2008). 基因的改变NIH Public Access. *Bone*, *23*(1), 1–7. https://doi.org/10.1038/jid.2014.371

Yue, Y., Grossmann, B., Galetzka, D., Zechner, U., & Haaf, T. (2006). Isolation and differential expression of two isoforms of the ROBO2/Robo2 axon guidance receptor gene in humans and mice. *Genomics*, *88*(6), 772–778. https://doi.org/10.1016/j.ygeno.2006.05.011

Zhang, F., Wen, Y., & Guo, X. (2014). CRISPR/Cas9 for genome editing: Progress, implications and challenges. *Human Molecular Genetics*, *23*(R1), 40–46. https://doi.org/10.1093/hmg/ddu125

Appendices

Appendix 1. Institutional Biosafety Committee Approval Letter

UNIVERSITY OF ARKANSAS

Office of Research Compliance

June 19, 2018

MEMORANDUM

TO:	Dr. Timothy Evans
FROM:	Ines Pinto, Biosafety Committee Chair
RE:	New Protocol
PROTOCOL #:	18042
PROTOCOL TITLE:	Genetics of neural development in insects
APPROVED PROJECT PERIOD:	**Start Date** June 14, 2018 **Expiration Date** June 13, 2021

The Institutional Biosafety Committee (IBC) has approved Protocol 18042, "Genetics of neural development in insects". You may begin your study.

If modifications are made to the protocol during the study, please submit a written request to the IBC for review and approval before initiating any changes.

The IBC appreciates your assistance and cooperation in complying with University and Federal guidelines for research involving hazardous biological materials.

1424 W. Martin Luther King, Jr. • Fayetteville, AR 72701
Voice (479) 575-4572 • Fax (479) 575-6527

Appendix 2: List of generated stocks

robo2^TauMyc /CyO,wg

(GMR28F02)::*TauMyc*/ CyO,wg

(GMR28F02)::*robo2*/ CyO,wg

(GMR28G05)::*robo2*/ CyO,wg

(GMR28D10)::*robo2*/ CyO,wg

(GMR28C04)::*robo2*/ CyO,wg

(GMR28B05)::*robo2*/ CyO,wg

(SEG01)::*robo2*/ CyO,wg

(SEG02)::*robo2*/ CyO,wg

(SEG03)::*robo2*/ CyO,wg

HSP70[GMR28F02:: *TauMyc*]/ CyO,wg

HSP70[GMR28G05:: *TauMyc*]/ CyO,wg

[GMR28F02:: *robo2*],[$robo2^{Myc\text{-}robo2}$*]/* CyO,wg

[GMR28G05:: *robo2*],[$robo2^{Myc\text{-}robo2}$*]/* CyO,wg

[GMR28F02:: *robo2*],[$robo2^{Myc\text{-}robo2}$*]/* [GMR28G05:: *robo2*],[$robo2^{Myc\text{-}robo2}$]

$robo2^{123}$/$robo2^{135}$

$robo2^{123}$,[GMR28F02::*robo2*]/ $robo2^{135}$,[GMR28F02::*robo2*]

$robo2^{135/}$ $robo2^{135}$

robo2^{135},[GMR28G05::*robo2*]/ *robo2*^{135},[GMR28G05::*robo2*]

(GMR28G05+ GMR28F02)::*robo2*

robo2^{135},[GMR28G05+ GMR28F02::*robo2*]/ *robo2*^{135},[GMR28G05+ GMR28F02::*robo2*]

Appendix 3. Primers used in this research

Table 1: Primers used to generate rescue construct transgenes.	
Primer name	**Primer sequence**
GMR28F02_F	GATTTCACTGGAACTAGGGGCACTTGGAGTTTTCTCCGTGGACATGACAC
GMR28F02_AscI_R	TGCTATTGAGTGGATGTGGGCGCGCCTGTAGCCATCTCAAGTCTAGAGTCCTCAGC
GMR28D10_F	GATTTCACTGGAACTAGGGGCACTCCCTCGAATTCCGCACTTTTAACCCTG
GMR28D10_AscI_R	TGCTATTGAGTGGATGTGGGCGCGCCATCATCCGCGCTTCTTAGGCCTCATGAGAG
GMR28G05_F	GATTTCACTGGAACTAGGGGGCATGGTTAAGGAGGAACGAAGCACTGCTTAC
GMR28G05_AscI_R	TGCTATTGAGTGGATGTGGGCGCGCCTCCTCCTCGAGATTCCGAGCATAAAACGC
GMR28C04_F	GATTTCACTGGAACTAGGGGAGGGTAGTTCTGAAGCCATTCCCCGTTAAATTTG
GMR28C04_AscI_R	TGCTATTGAGTGGATGTGGGCGCGCCAGCGGCTAAAAATAATGCCAGTCGATCGGAG
GMR28B05_F	GATTTCACTGGAACTAGGGGCTGACTTCCAGGACTAGATGGGGATTTCTTTG
GMR28B05_AscI_R	TGCTATTGAGTGGATGTGGGCGCGCCAGGGTAGTCGGGGTCGTTTATTTGGAAGTTTG
SEG01_F	GATTTCACTGGAACTAGGGGCCCATCCGGGAAAAAGGGACAGATTTTAGC
SEG01_AscI_R	TGCTATTGAGTGGATGTGGGCGCGCCGTGATACGGATTTAAGAACTATTGTGCGTTTC
GMR28E07_F	GATTTCACTGGAACTAGGGGGGGCATACCACAGTGAACTCAAAGCCTGGTC
GMR28E07_AscI_R	TGCTATTGAGTGGATGTGGGCGCGCCTCTGGTGAAATGGCCTGACACTTGACGTGGC
SEG02_F	GATTTCACTGGAACTAGGGGGCCCATTGATCACCTATCCACCCTATCCCG
SEG02_AscI_R	TGCTATTGAGTGGATGTGGGCGCGCCGATCTCGAGTGATTGGAACATAGAACATAG
SEG03_F	GATTTCACTGGAACTAGGGGGGAACCCAGCATCCCAGTTGTAATCCCAGC

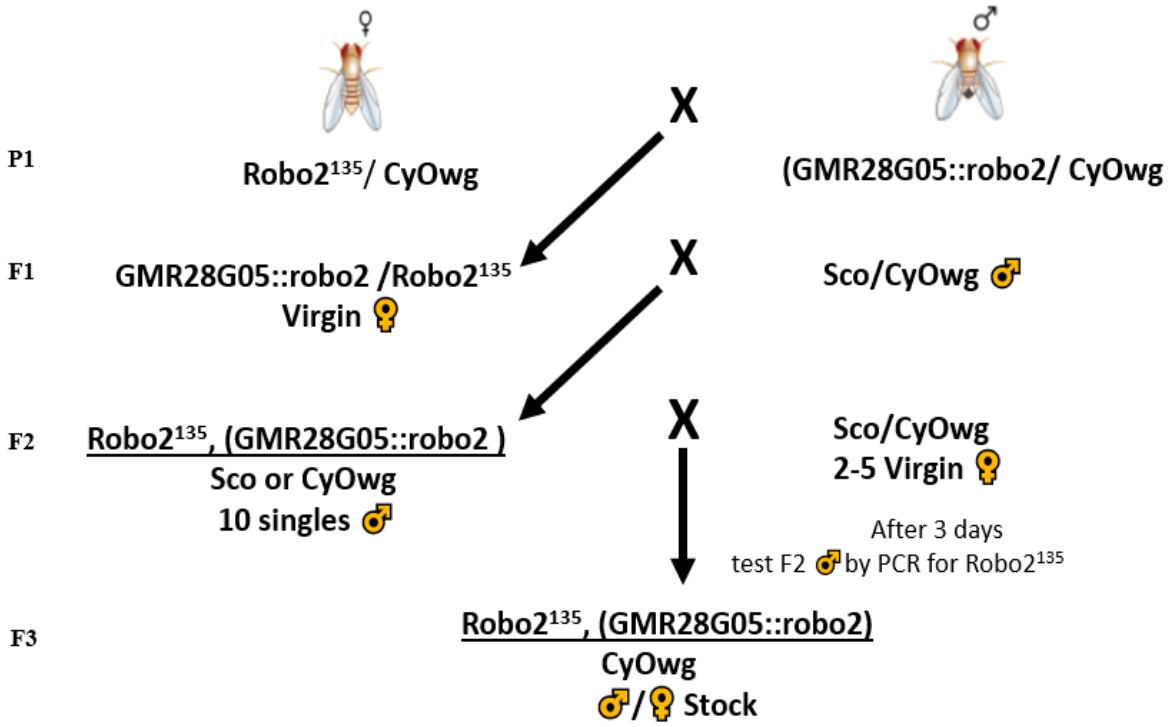
Table 1 (Cont.)

GMR28F02.R GIBSON	CTATTGAGTGGATGTGGGCGCGCCTGTAGCCATCTCAAGTCTAG
28G05.F ULTRA KIT GIBSON	GAAAATGCTTGGATTTCACTGGAACTAGGGGGGCGCGCCGCATGGTTAAG GAGGAAC
28G05.R ULTRA KIT GIBSON	CCACGGAGAAAACTCCAAGTGCCTCCTCCTCGAGATTCCGAGCATAAAAC G
28F02.F ULTRA KIT GIBSON	CTCGGAATCTCGAGGAGGAGGCACTTGGAGTTTTCTCCGTGGACATGACAC
28F02.R ULTRA KIT GIBSON	CATCATCATGTTGCTATTGAGTGGATGTGGGCGCGCCTGTAGCCATCTCAA GTCTAG

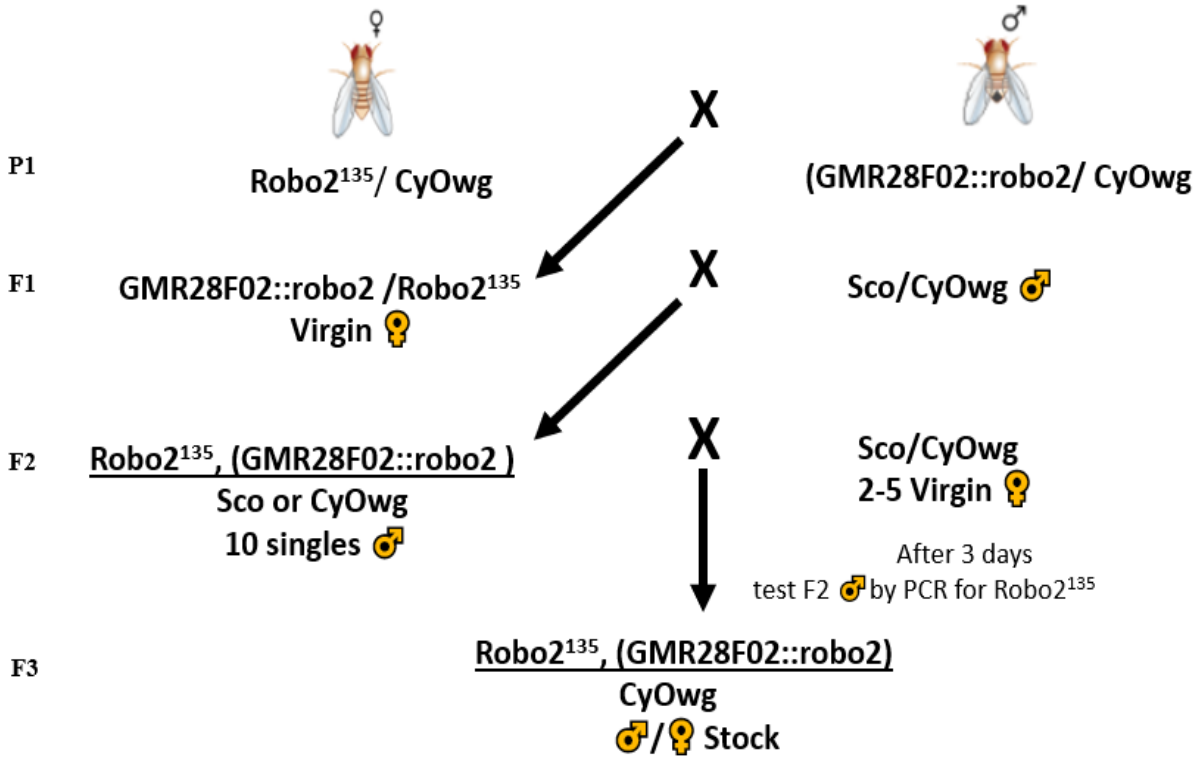
Table2: Primers used for the CRISPR CAS-9 to delete *robo2* putative enhancers

28Fo2.F gRNA	TATATAGGAAAGATATCCGGGTGAACTTCGTATTATTCCTCTTGTAAACG TTTTAGAGCTAGAAATAGCAAG
28Fo2.R gRNA	ATTTTAACTTGCTATTTCTAGCTCTAAAACAGTGCATCTGTCTGTCATTCG ACGTTAAATTGAAAATAGGTC
28Go5.F gRNA	TATATAGGAAAGATATCCGGGTGAACTTCGACAAATTGGTTCGCAAGAG GTTTTAGAGCTAGAAATAGCAAG
28Go5.R gRNA	ATTTTAACTTGCTATTTCTAGCTCTAAAACACCTTACTGATCAACACATA CGACGTTAAATTGAAAATAGGTC
CRISPR-GO5.F(811)	CTGACAGGGTCCTAAGATC
CRISPR-GO5.R(812)	GCGTTCACTGAACCTCCT
CRISPR-FO2.F(813)	GAGAGACCGACGACTGTCT
CRISPR-FO2.R(814)	CATGCCAACAACGCTCAGC
G05.N.F1	AGATTGGCTCGAAAGGCGAA
G05.N.R1	GAATCCACAATCCAGGGCCA
F02.N.F1	ATTGAATGCGGGAAGGGGAG
F02.N.R1	GTAATGAGAGCTTGCAGCGC
G05.N.F2	TCCCCTTCTTTCGACTGTGC
G05.N.R2	ACTAACAAAGCGGACTGGGG
F02.N.F2	AATGGAGCAGGCAGAGAACC

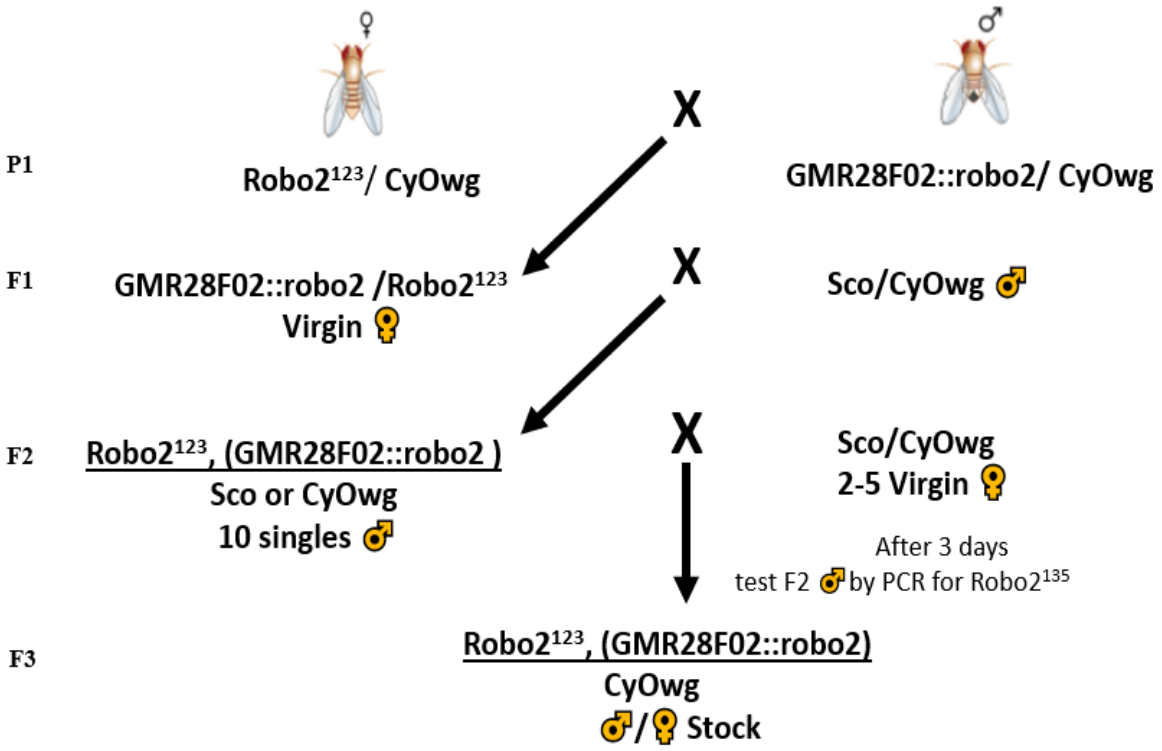
Appendix 4: Fly crosses to make stocks



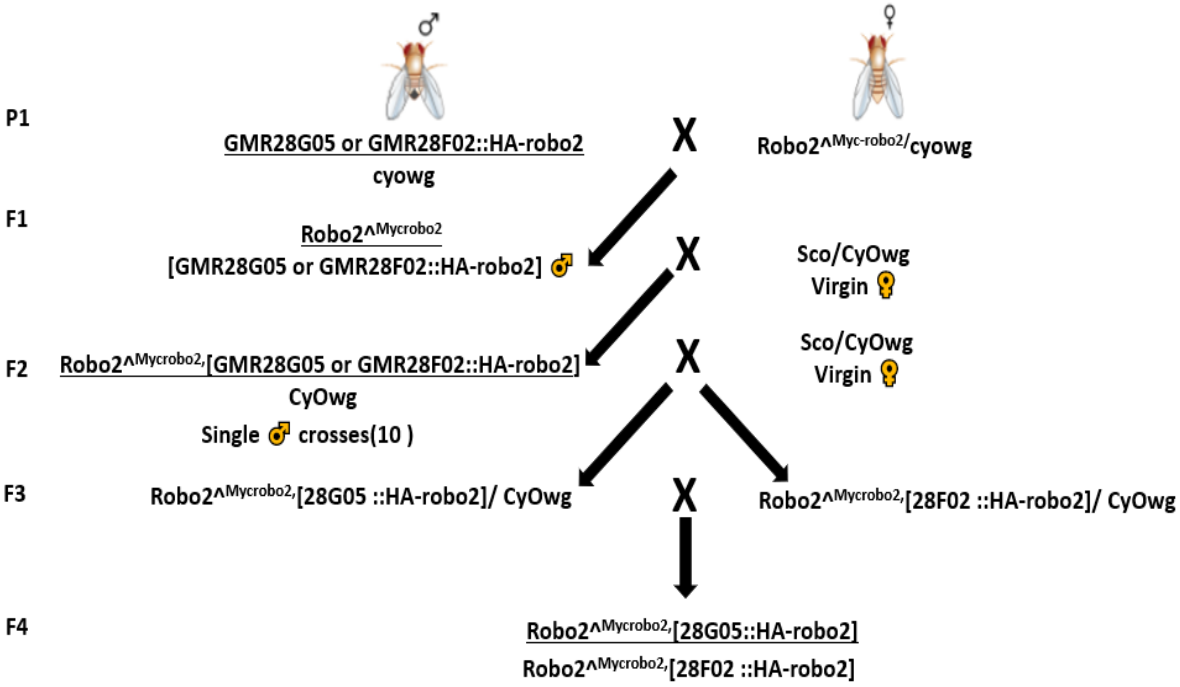
Appendix 4-1. The rescue of axon guidance decision of *robo2* mutant background: The GMR28G05::*robo2* transgenic line is generated by BestGene Inc (Chino Hills, CA). Injected (G0) individuals were crossed as adults to Sco/CyOwg to make a balanced stock. The balanced stock female flies were crossed to *robo2* mutant background males(*robo2*^{135}/CyOwg). F1 virgin females were then crossed to Sco/CyOwg males. F2 single 10 red-eyed males (positive to the transgene) were crossed individually to 3-5 virgin Sco/CyOwg females. After three days, the F2 males were removed from the crosses and tested by PCR. F3 progeny from positive F2 crosses were used to generate balanced stocks.



Appendix 4-2. The rescue of axon guidance decision of *robo2* mutant background: The GMR28F02::*robo2* is generated by BestGene Inc. (Chino Hills, CA). Injected (G0) individuals were crossed as adults to Sco/CyOwg to make a balanced stock. The balanced stock female flies were crossed to *robo2* mutant background(*robo2*^{135}/CyOwg). F1 virgin females were then crossed to Sco/CyOwg males. F2 single 10 red-eyed males (positive to the transgene) were crossed individually to 3-5 virgin Sco/CyOwg females. After three days, the F2 males were removed from the crosses and tested by PCR. F3 progeny from positive F2 crosses were used to generate balanced stocks.

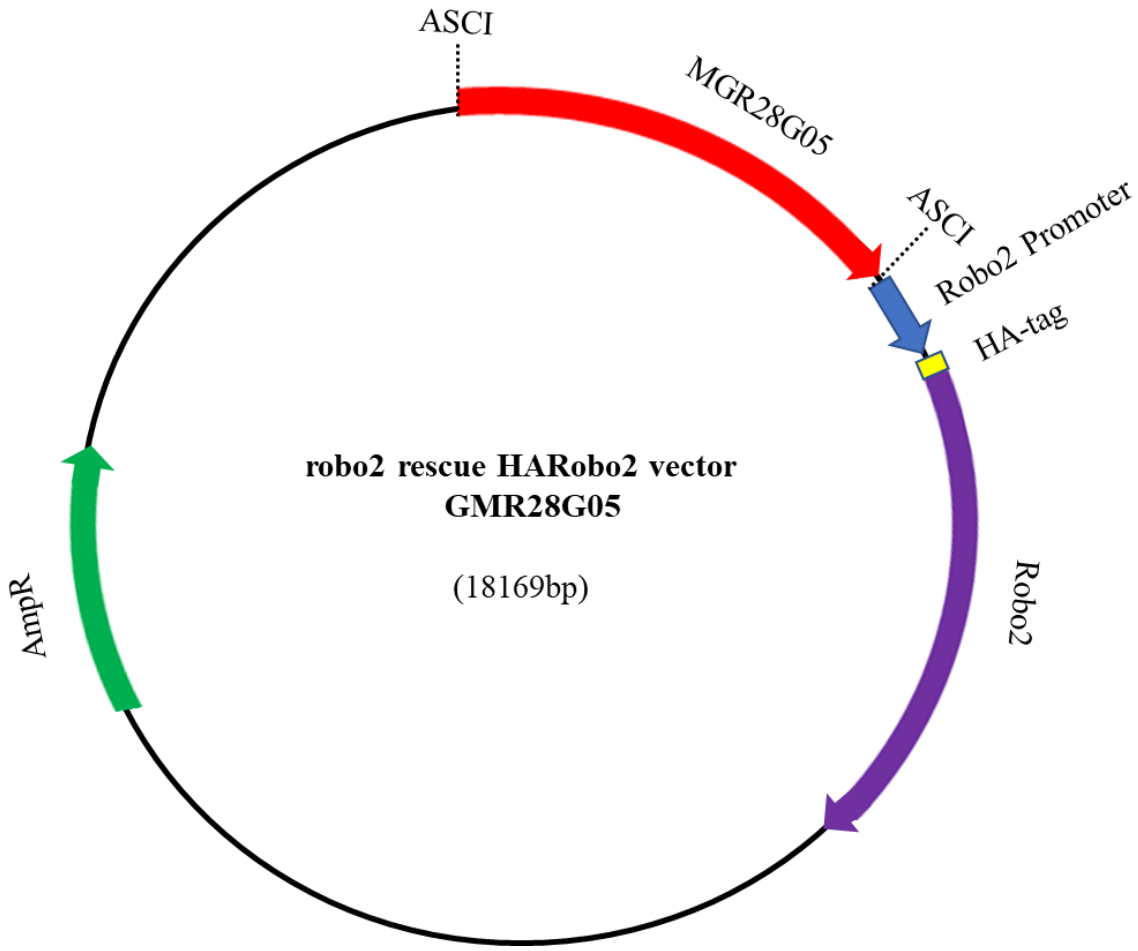


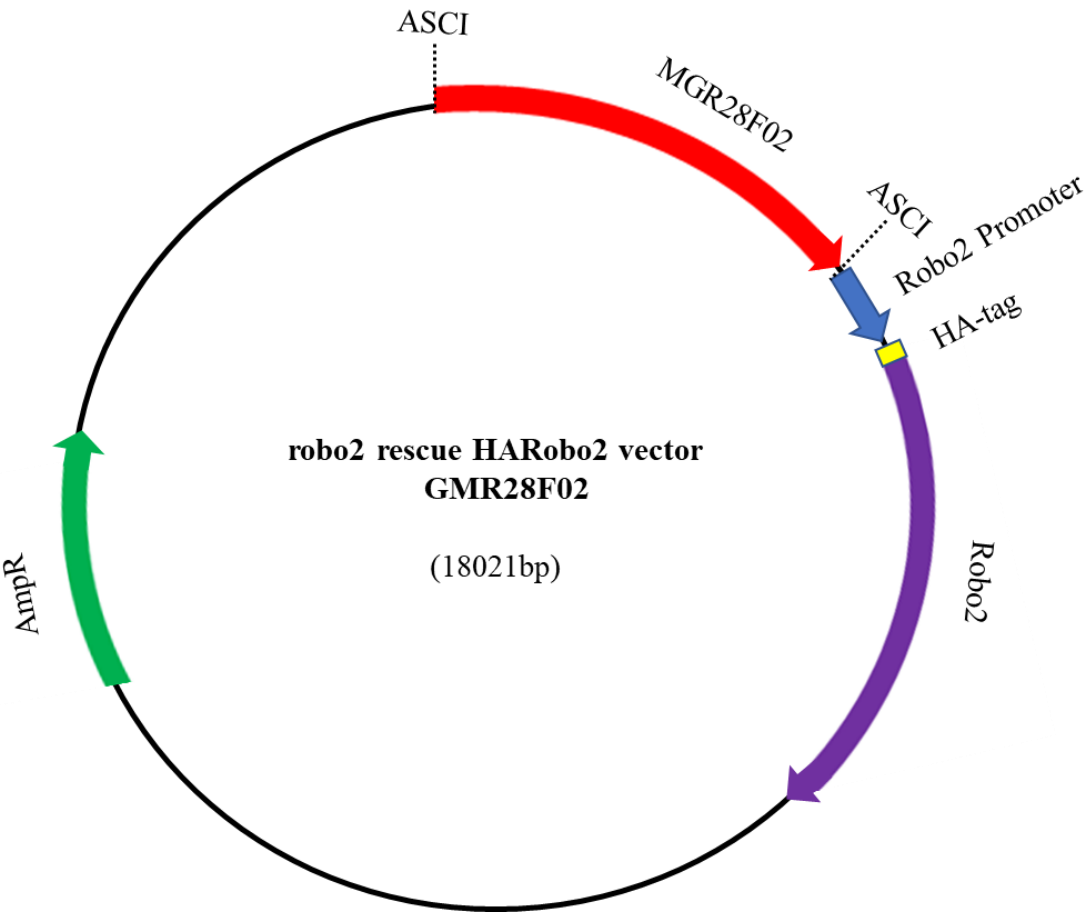
Appendix 4-3. The rescue of axon guidance decision of *robo2* mutant background: The GMR28F02::*robo2* is generated by BestGene Inc. (Chino Hills, CA). Injected (G0) individuals were crossed as adults to Sco/CyOwg to make a balanced stock. The balanced stock female flies were crossed to *robo2* mutant background(*robo2*^{123}/CyOwg). F1 virgin females were then crossed to Sco/CyOwg males. F2 single 10 red-eyed males (positive to the transgene) were crossed individually to 3-5 virgin Sco/CyOwg females. After three days, the F2 males were removed from the crosses and tested by PCR. F3 progeny from positive F2 crosses were used to generate balanced stocks.

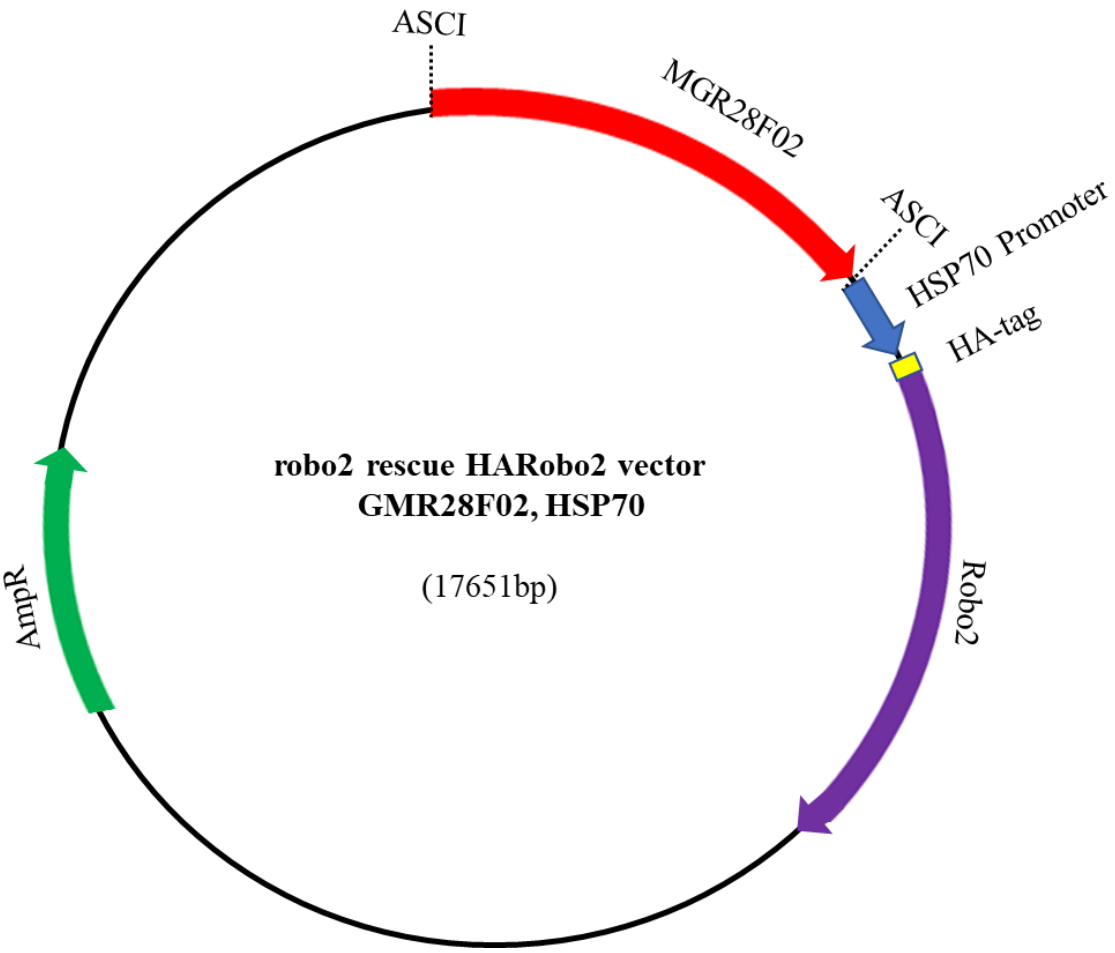


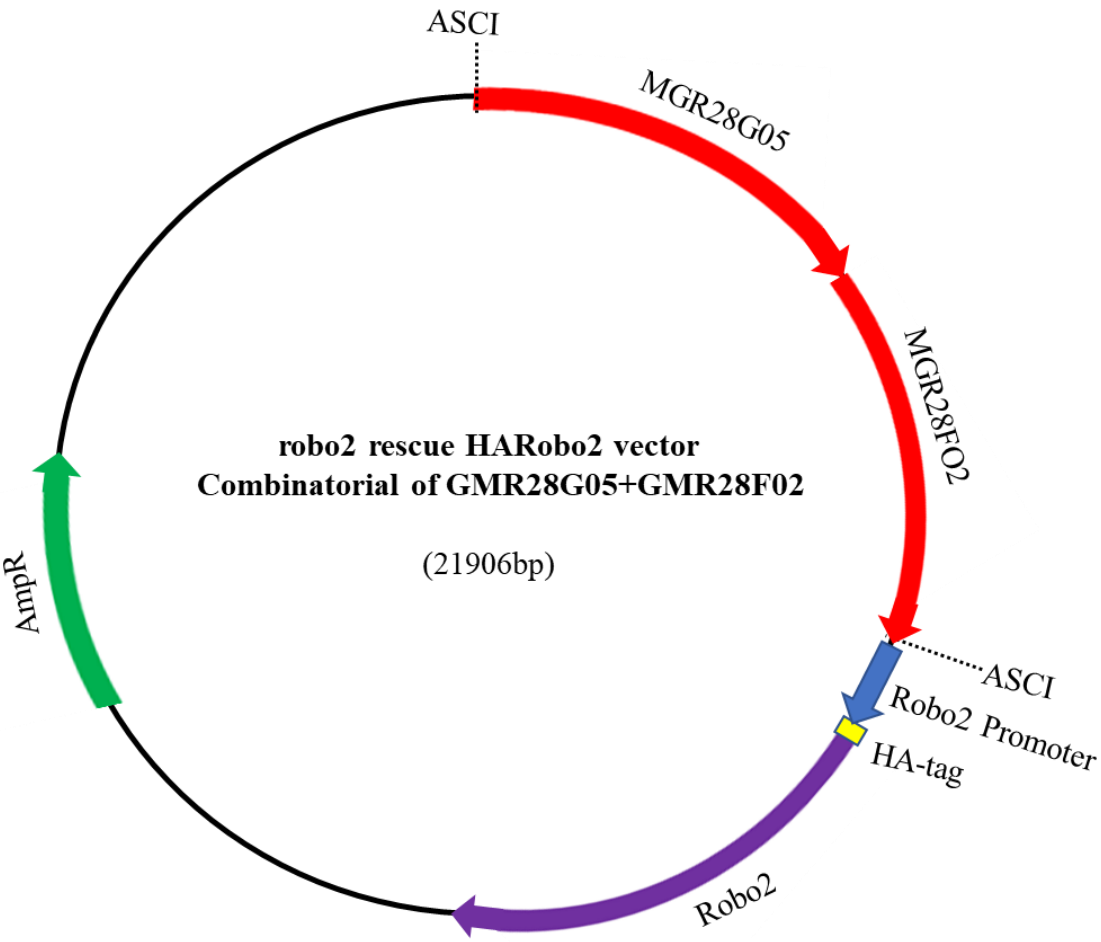
Appendix 4-4: Driving *robo2* expression by putative enhancers vs. *robo2* native enhancer. To check whether *robo2* putative enhancers could drive *robo2* expression as strong as the native *robo2* enhancer, (GMR28G05) or (GMR28F02)::HA- Robo2 transgenic lines were crossed with flies that were labeled with myc tag that represent the native *robo2* expression.

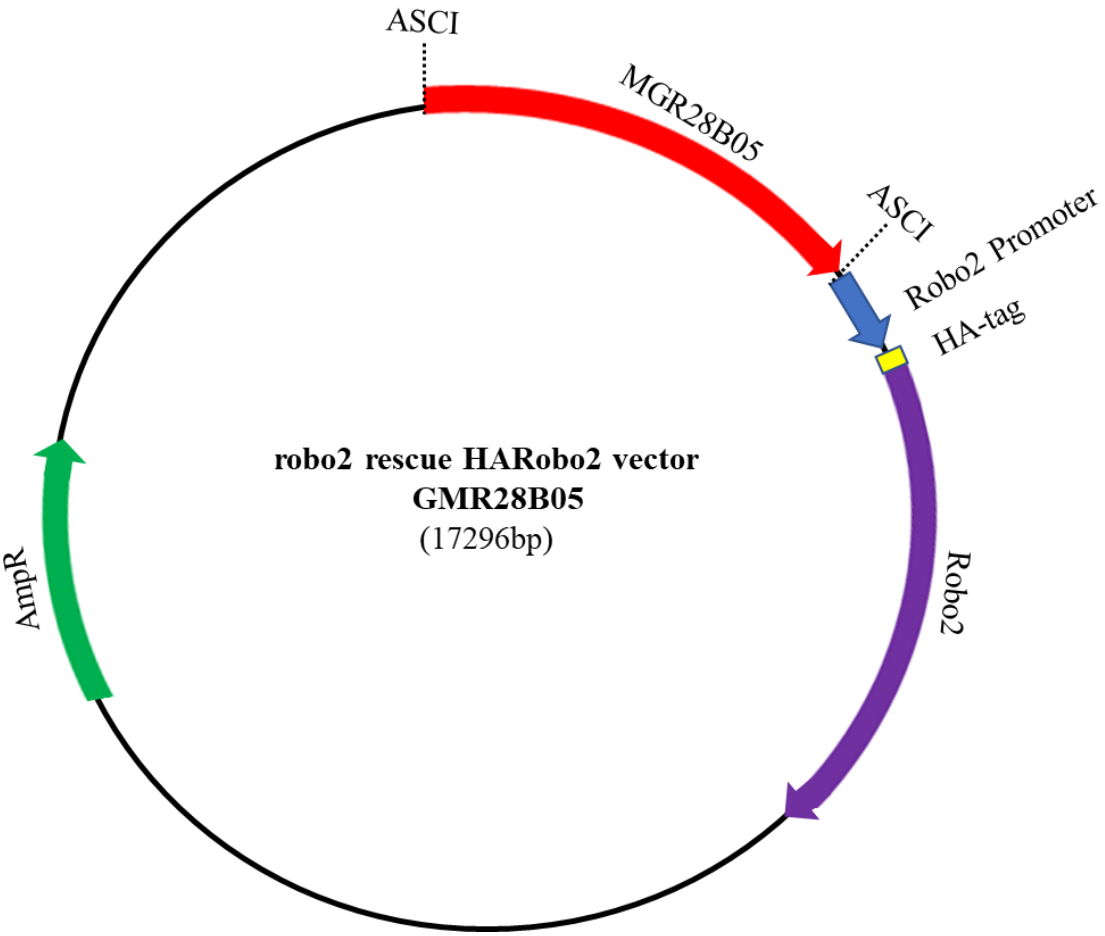
Appendix 5: Plasmid construction maps. Each map is annotated with the vector name, inserts (Janelia and non-Janelia fragments), the size of the vector, the promoter, epitope tags, the coding sequence of the gene of interest, Antibiotic resistance, and restriction enzyme locations.

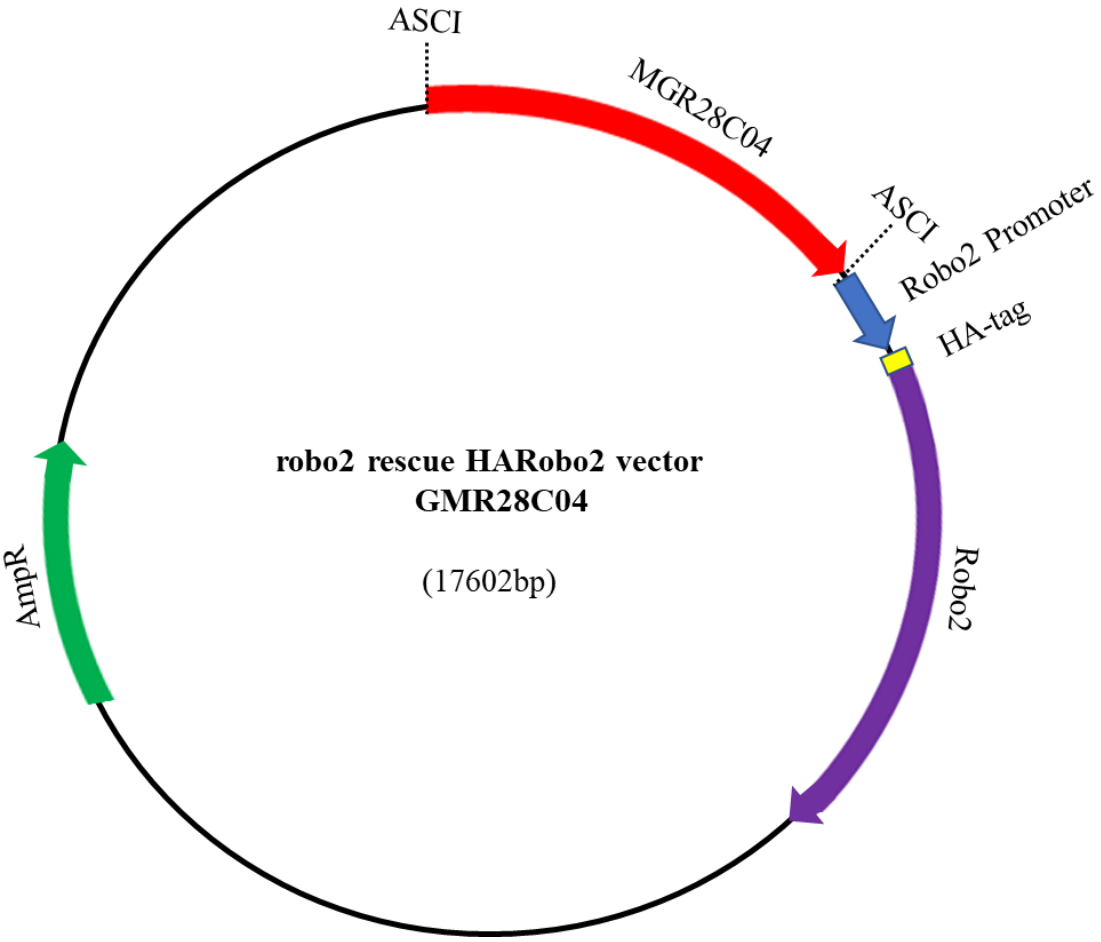


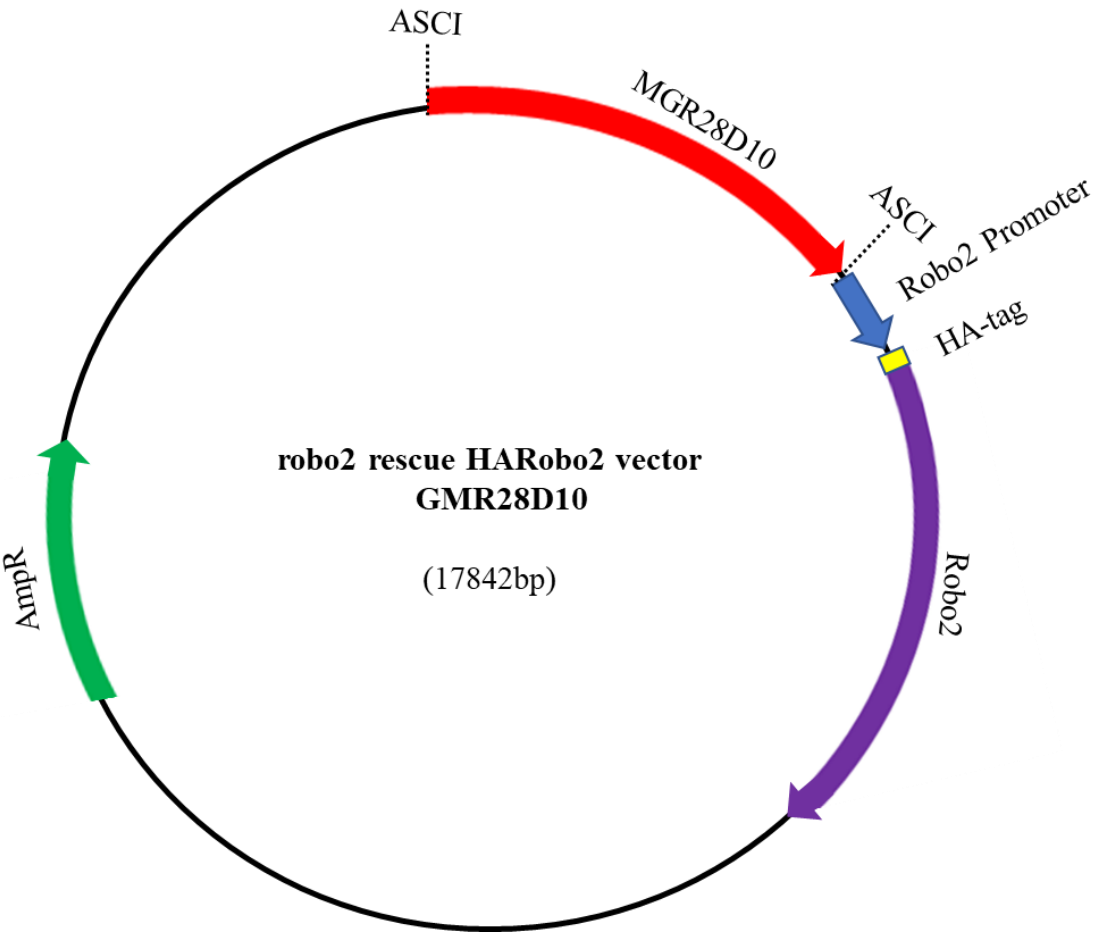

ASCI
MGR28F02
ASCI
Robo2 Promoter
HA-tag
robo2 rescue HARobo2 vector
GMR28F02
(18021bp)
Robo2
AmpR

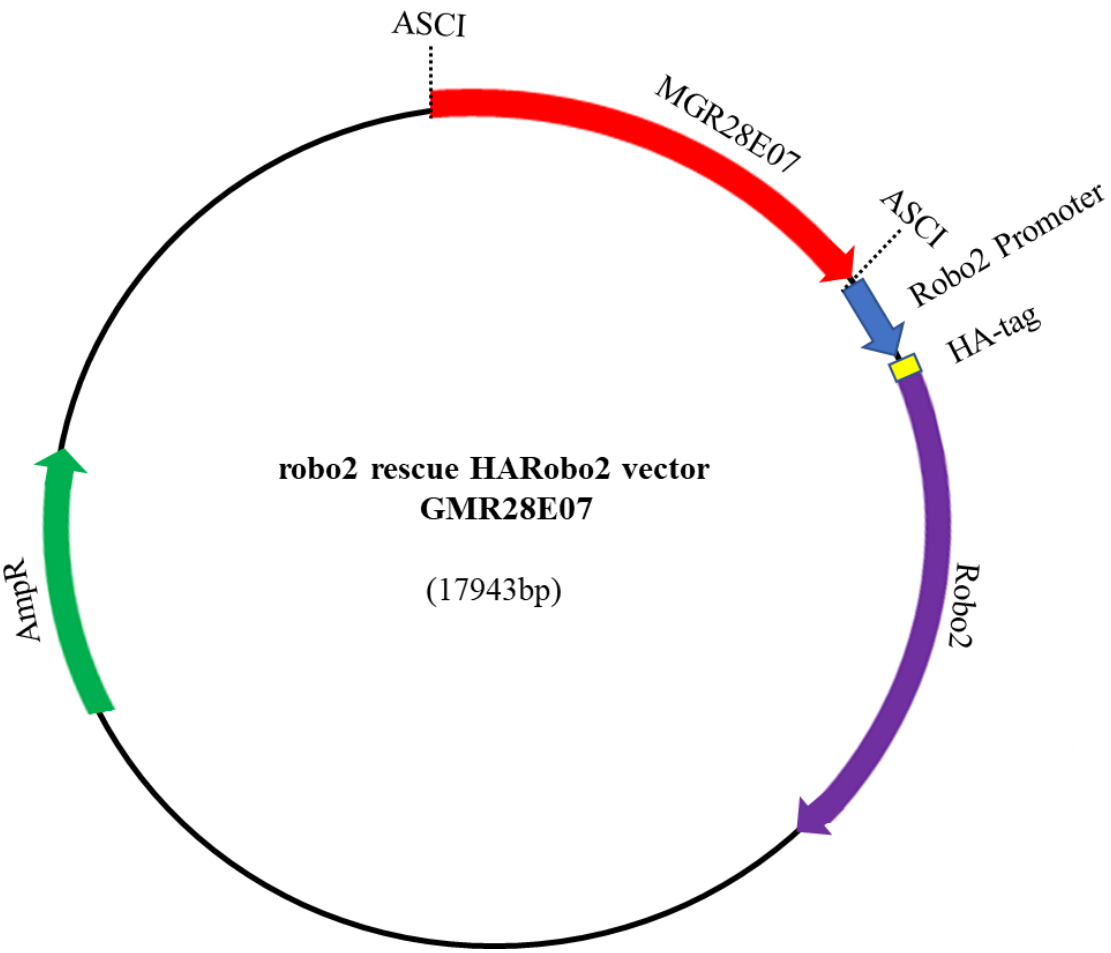

ASCI
MGR28F02
ASCI
HSP70 Promoter
HA-tag
robo2 rescue HARobo2 vector
GMR28F02, HSP70
(17651bp)
Robo2
AmpR

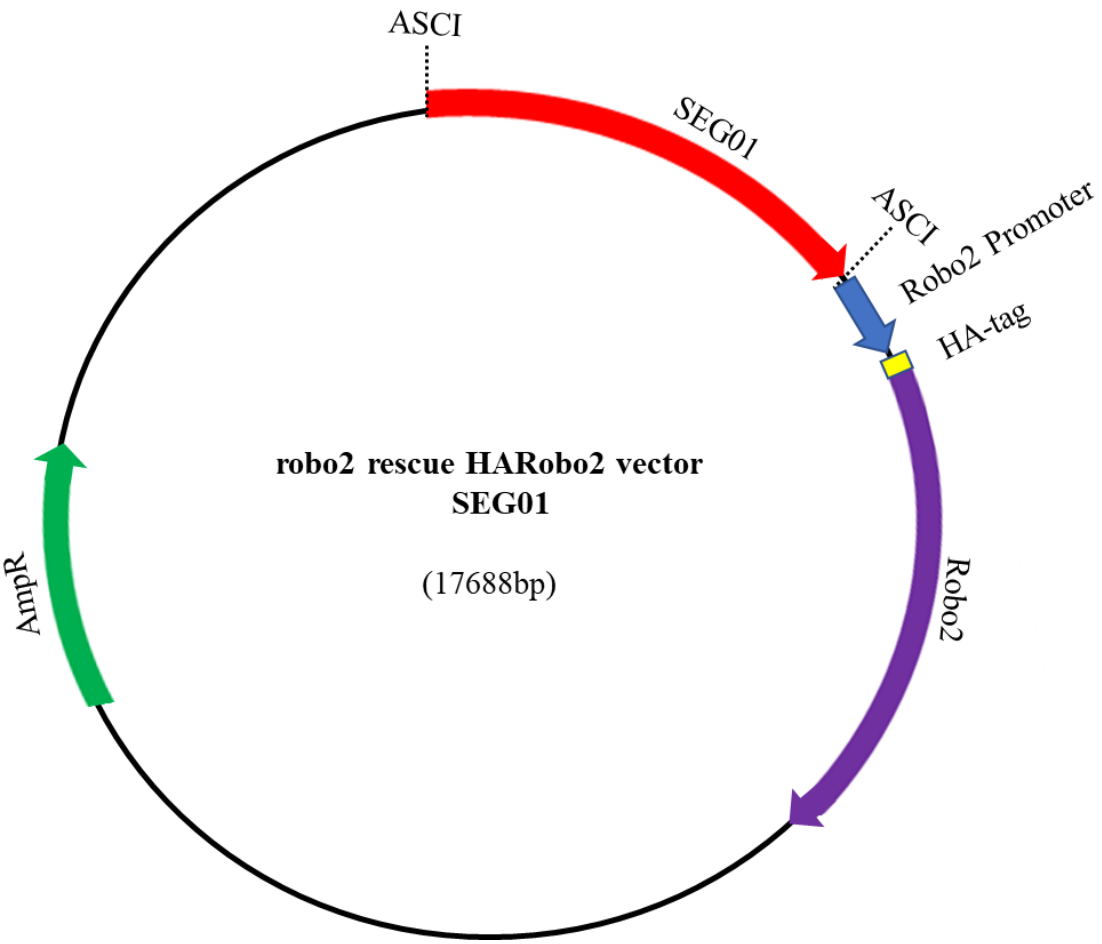

ASCI
MGR28G05
MGR28F02
ASCI
Robo2 Promoter
HA-tag
Robo2
AmpR
robo2 rescue HARobo2 vector
Combinatorial of GMR28G05+GMR28F02
(21906bp)

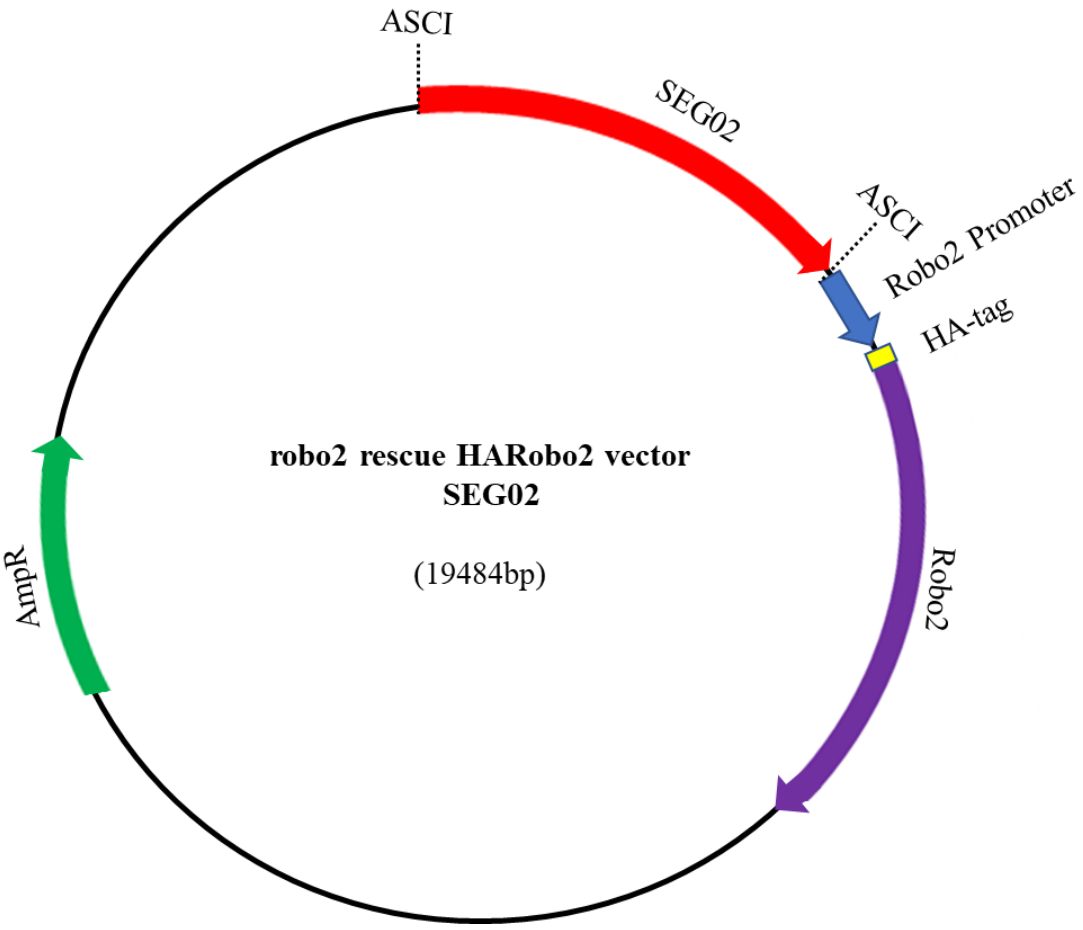

ASCI
MGR28B05
ASCI
Robo2 Promoter
HA-tag
robo2 rescue HARobo2 vector
GMR28B05
(17296bp)
Robo2
AmpR

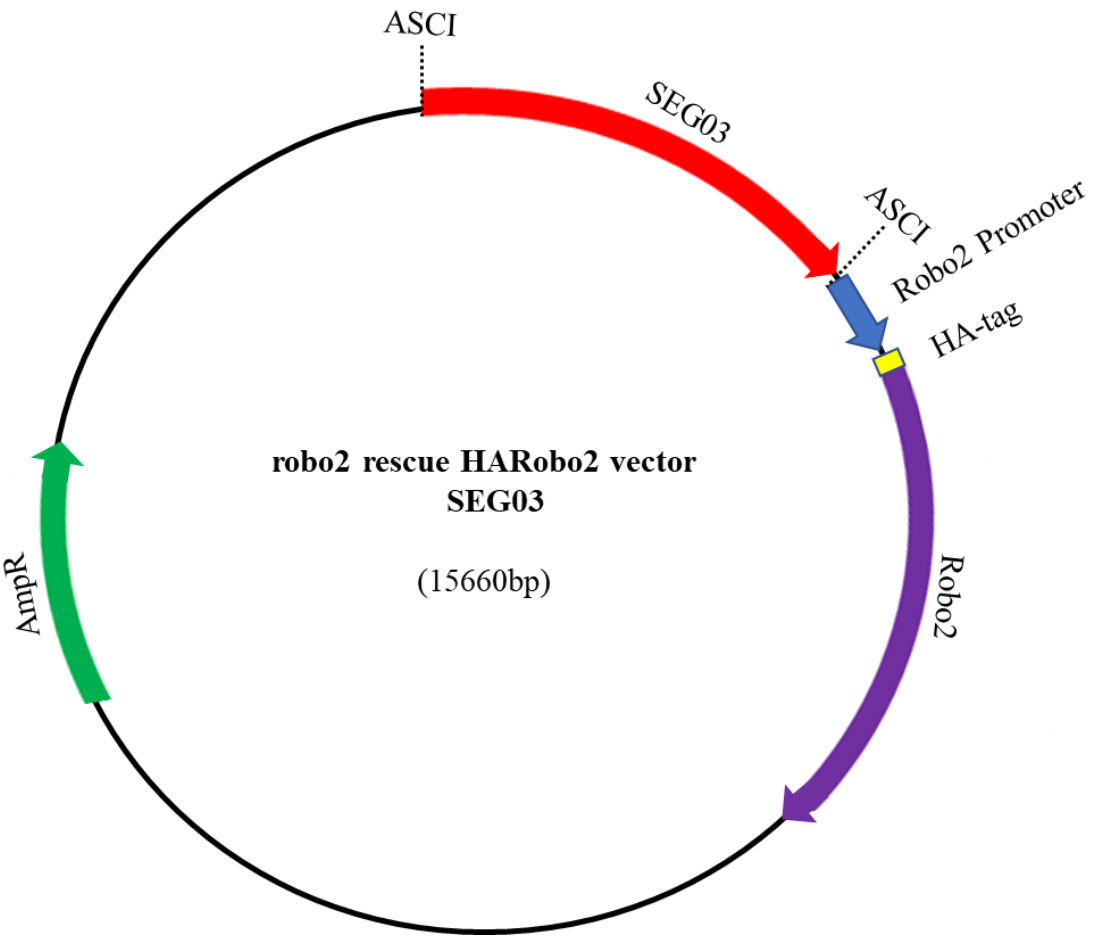

ASCI
MGR28C04
ASCI
Robo2 Promoter
HA-tag
robo2 rescue HARobo2 vector
GMR28C04
(17602bp)
Robo2
AmpR

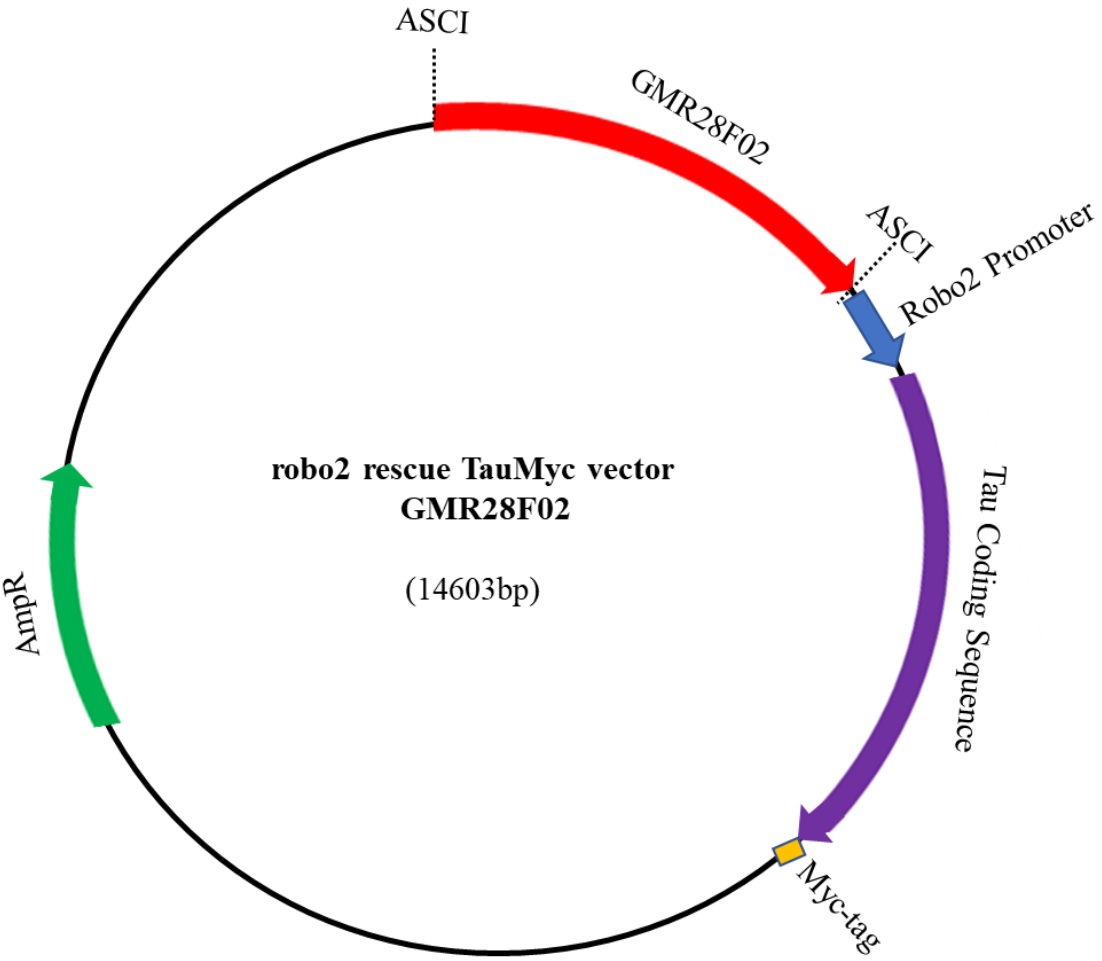

ASCI
MGR28D10
ASCI
Robo2 Promoter
HA-tag
robo2 rescue HARobo2 vector
GMR28D10
(17842bp)
Robo2
AmpR

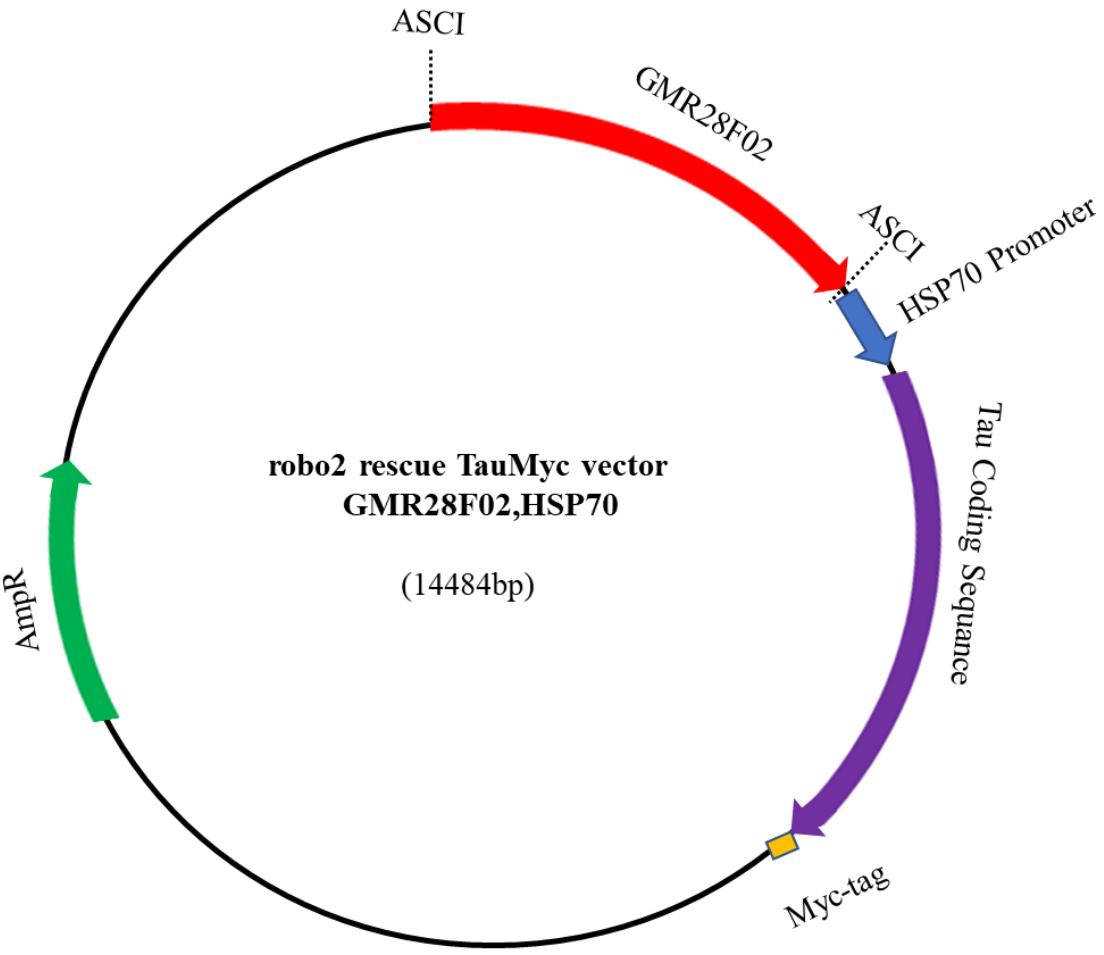

ASCI
MGR28E07
ASCI
Robo2 Promoter
HA-tag
robo2 rescue HARobo2 vector
GMR28E07
(17943bp)
Robo2
AmpR

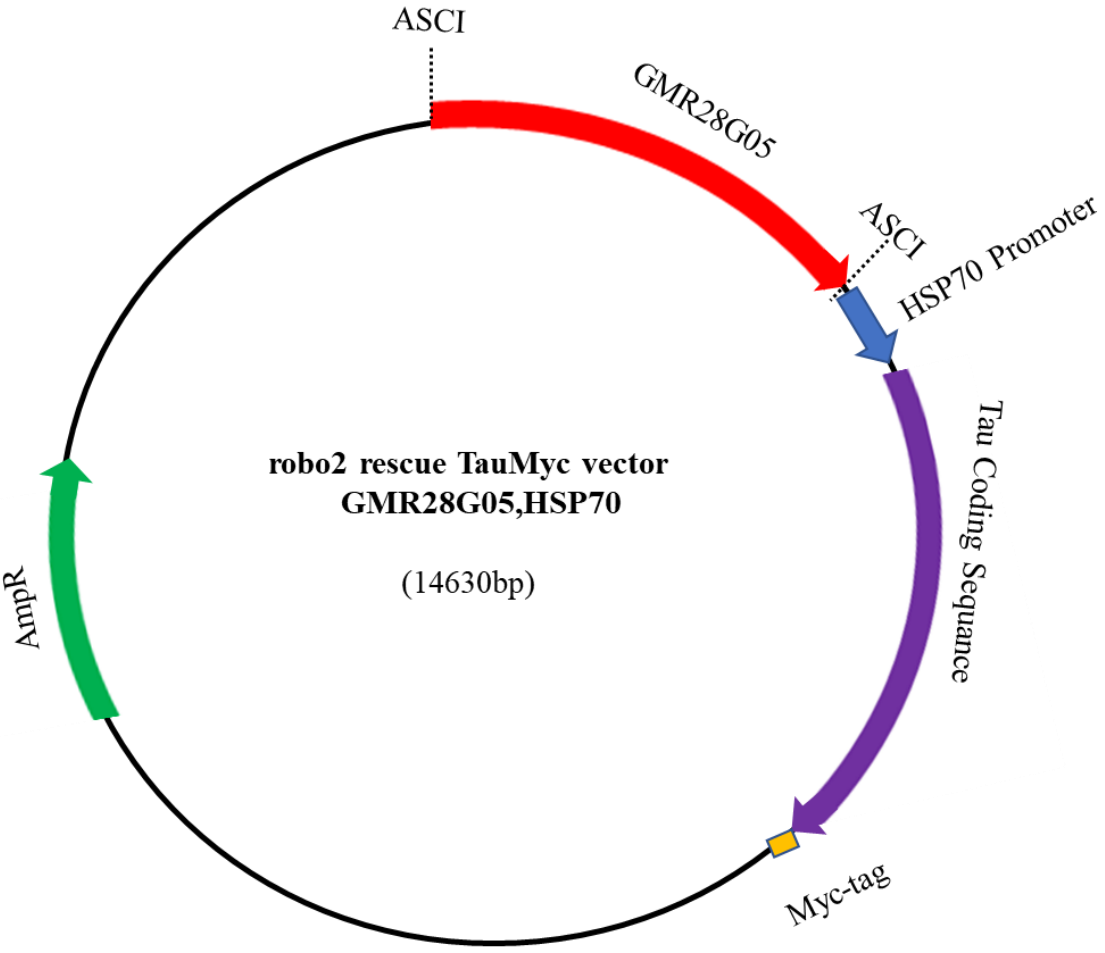

ASCI
SEG01
ASCI
Robo2 Promoter
HA-tag
robo2 rescue HARobo2 vector
SEG01
(17688bp)
Robo2
AmpR

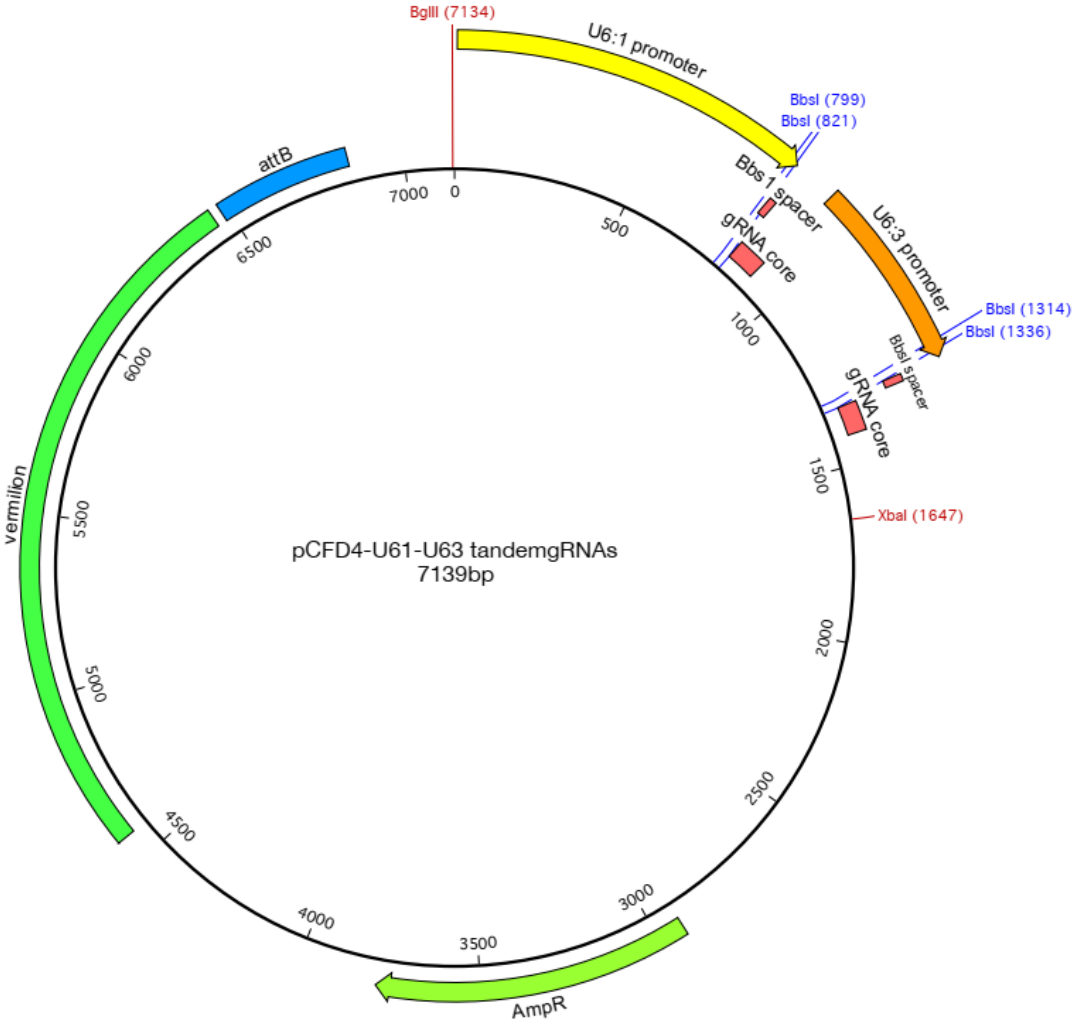

ASCI
SEG02
ASCI
Robo2 Promoter
HA-tag
robo2 rescue HARobo2 vector
SEG02
(19484bp)
Robo2
AmpR


ASCI
SEG03
ASCI
Robo2 Promoter
HA-tag
robo2 rescue HARobo2 vector
SEG03
(15660bp)
Robo2
AmpR

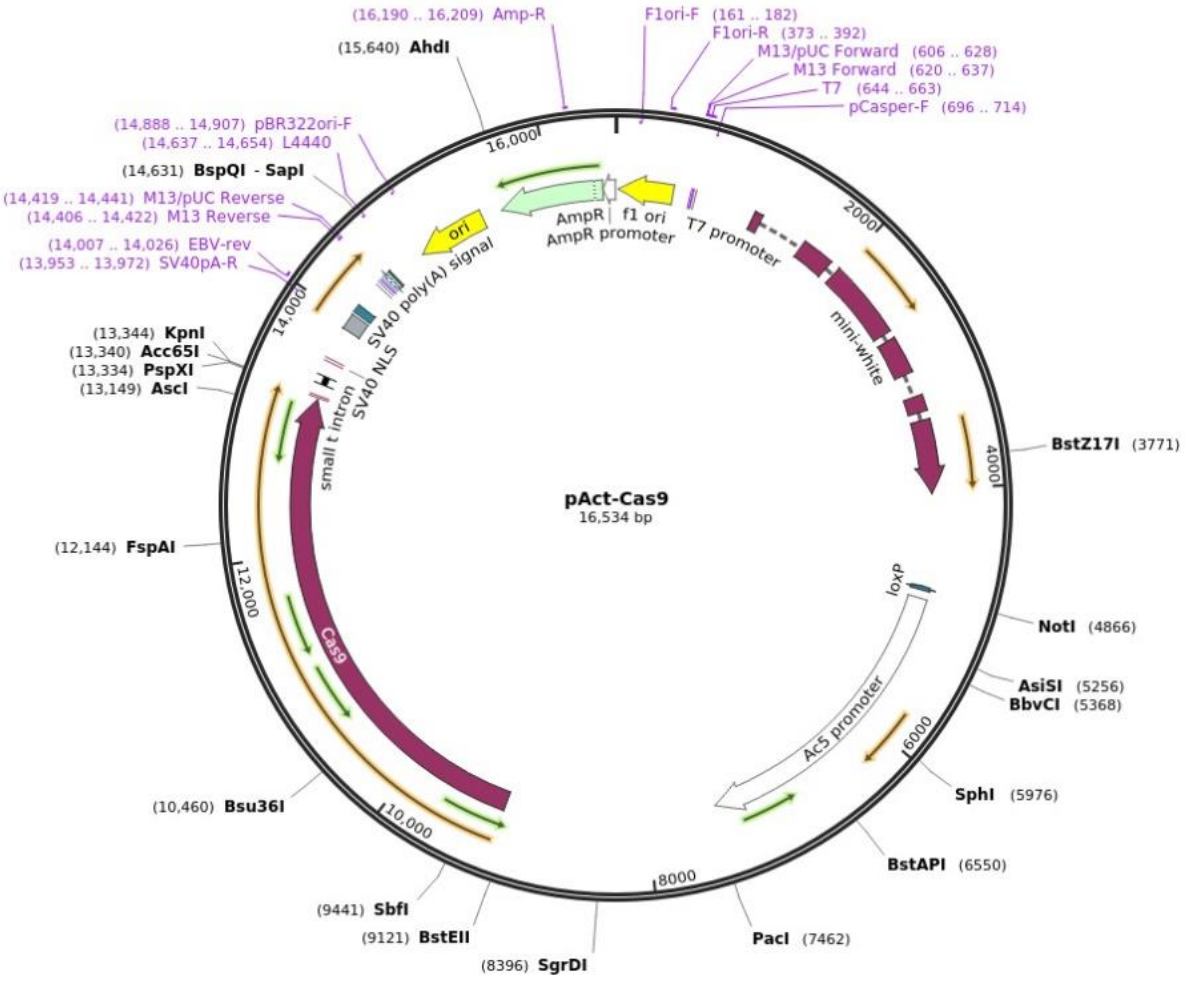

ASCI
GMR28F02
ASCI
Robo2 Promoter
Tau Coding Sequence
Myc-tag
AmpR
robo2 rescue TauMyc vector
GMR28F02
(14603bp)


ASCI
GMR28F02
ASCI
HSP70 Promoter
Tau Coding Sequance
Myc-tag
AmpR
robo2 rescue TauMyc vector
GMR28F02,HSP70
(14484bp)


ASCI
GMR28G05
ASCI
HSP70 Promoter
Tau Coding Sequance
Myc-tag
AmpR
robo2 rescue TauMyc vector
GMR28G05,HSP70
(14630bp)



Adapted from https://www.addgene.org.



Adapted from https://www.addgene.org.