Characterization of *Vps13* in *Drosophila melanogaster*

By

Kiana Kasravi, B.Sc.

A thesis submitted to the

School of Graduate Studies

in partial fulfillment of the requirement for the degree of

Master of Science

Department of Biology

Memorial University of Newfoundland

2019

St. John’s Newfoundland and Labrador

Abstract

Parkinson Disease (PD), the most common movement disorder, is a neurodegenerative disease that affects about 1% of the human population over the age of 65 years, and its prevalence increases with age. PD is clinically characterized by resting tremor, rigidity and bradykinesia as a result of the loss of dopaminergic neurons in the *substantia nigra pars compacta* in the mid-brain. Mitochondria play a significant role in PD. The PD candidate gene *VPS13C* was identified in human using whole-exome sequencing (WES). The VPS13C protein is involved in intracellular trafficking, especially lipid transportation, between the endoplasmic reticulum (ER) and other organelles, such as the late endosome/lysosome, as well as in the mitophagy pathway. I have successfully characterized the homologue of the human *VPS13C* gene in flies, *Vps13.* As a result, I have successfully modelled aspects of PD in *Drosophila melanogaster*. I found that inhibition of *Vps13* improves the longevity and locomotor ability, while overexpression of *Vps13* mostly reduced both the survival and locomotor abilities of flies. Furthermore, I have demonstrated that overexpression of *Vps13* significantly influences neurodevelopment when expressed in the developing *Drosophila melanogaster* eye. Our final set of experiments focused upon the consequences of the inhibition and overexpression of *Vps13* when expressed with the co-inhibition of *parkin,* a well-known PD gene. These co-expressions can give valuable insights into the genetic and/or functional connections between the two genes or their protein products. These findings suggest that the inhibition of *parkin* may improve the effects of *Vps13* overexpression.

Acknowledgment

I would first like to give my sincere gratitude to my supervisor, Dr. Brian E. Staveley who gave me this life-changing opportunity of being able to work for him and learn from him, for his continuous support of my M.Sc. study and research, his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me throughout my research and writing of this thesis, and especially in the times when I doubted myself. His support helped me keep going forward, and I could not have imagined doing all this work without his supervision, support and wisdom; I would as well like to thank all of Staveley's lab members, who made me enjoy every second of this journey especially Azra Hasan, who guided me with her thorough knowledge, patience, and kindness. Furthermore, I would like to thank my committee members, Dr. Jennifer Slade, and Dr. Suzanne Dufour, who provided me with constructive criticism and positive feedback throughout my project, and Dylan Goudie, who kindly assisted me in taking my SEM pictures with an attitude that made those long days fun.

I would like to thank my family, who were always willing to do anything for me. They have guided and supported me with all they had. I am everything I am because of their unconditional love, and I am forever thankful to my beloved husband, who believed in me all the way, and whose love made all the hardships easy.

Research for this thesis has been supported by graduate student fellowships from the Memorial University School of Graduate Studies. Dr. Brian Staveley's laboratory is supported by grants from the Natural Sciences and Engineering Research Council.

Table of Content

List of Figures

List of Tables

List of Abbreviations

6-OHDA- 6-hydroxydopamine

ATG-C- Autophagy-related protein C-terminal

ATP- Adenosine triphosphate

ATP13A2- ATPase Cation Transporting 13A2

Bcl2- B cell lymphoma 2

BLAST- Basic Local Alignment Search Tool

CI- Confidence interval

D. melanogaster- Drosophila melanogaster

ER- Endoplasmic reticulum

LOF- Loss of function

LRRK2- Leucine-rich repeat kinase 2

Mid rpt- Middle repeating coiled region

MPTP- 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine

mRNA- messenger RNA

MS- Multiple sclerosis

MSA- Multiple sequence alignment

N/A- Not applicable

NCBI- National Center for Biotechnology Information

NDD- Neurodegenerative disease

PARK2- Familial Parkinson disease type 2

PARK7- Parkinsonism associated deglycase

Pfam- Protein family database

pink1- PTEN-induced kinase 1

PTEN- Phosphatase tensin homologue

RNAi- RNA interference

SE- Standard error

SEM- Scanning electron microscope

SEM- Standard error for mean

siRNA- small interference RNA

SNCA - *α-synuclein*

SNPC- *Substantia nigra pars compacta*

tBLASTn- Translated Nucleotide Basic Local Alignment Search Tool

UAS- Upstream activation sequence

VPS13 mid rpt- VPS13 middle repeating coiled region

VPS13- Vacuolar protein sorting 13

VPS13C- Vacuolar protein sorting 13 homologue C

VPS35- vacuolar protein sorting 35

FBXO7- F-Box protein 7

WES- Whole-exome sequencing

Introduction

Purpose

Parkinson disease (PD) is an age-dependent, neurodegenerative disease (NDD) that has incapacitating effects on the lives of patients. Due to the increasing age of populations within the next decade, research into PD has been receiving more attention than ever before. Despite the high prevalence of Parkinson disease, our knowledge about the influential factors and details about the disease are insufficient, and there is no absolute cure for it. The goal of this study is to characterize the *Vps13* gene, a recently identified Parkinson disease candidate gene, and determine its influence on the survivorship and locomotory abilities of *Drosophila melanogaster* (*D. melanogaster*) as a model organism.

Parkinson Disease

James Parkinson was the first medical scientist to describe Parkinson disease in 1817 (Nussbaum *et al.*, 1997). In the mid-1800s, Jean-Martin Charcot separated PD from multiple sclerosis (MS) and other diseases with tremor characterization (de Lau and Breteler, 2006). However, evidence suggests that this disease existed long before (Lysia S. Frono, 1996), as it was described in Leonardo Da Vinci's transcripts sometime between 1489 and 1507, and Egyptian papyrus descriptions between 1350-1200BC.

Parkinson clinical signs can be categorized into the motor and non-motor symptoms. Motor symptoms include impaired movement (bradykinesia), rigidity and resting tremor, or any combination of these, whereas non-motor symptoms consist of mood disorders and cognitive changes. Mood disorders may involve depression, anxiety, and

irritability, and cognitive changes may include problems with focused attention and planning, slowness of thinking, language and memory difficulties, personality changes, dementia, hallucinations and delusions, which leads to substantial disabilities and early death (Lysia S. Frono, 1996). Pathological symptoms are caused by the selective degeneration of the dopaminergic neurons in the *substantia nigra pars compacta* (SNPC) in the midbrain with the appearance of Lewy Bodies (Lesage *et al.*, 2016). The term Lewy Body is associated with the abnormal aggregation of proteins (mostly α-synuclein and ubiquitin along with other proteins in PD cases) inside nerve cells that forms fibrillar cytoplasmic inclusions. Commonly, it is believed that this accumulation of protein results in cytotoxicity. However, the development and appearance of Lewy bodies are not exclusive to PD, and some PD patients do not show this type of aggregation (Verstraeten, Theuns and Van Broeckhoven, 2015a). Variable levels of α-synuclein accumulation have been observed in PD cases regarding their fundamental genetic defects.

For the most part, patients with PD show a temporarily-positive response to a Levodopa drug treatment to relieve many of the symptoms; the efficiency may decrease, along with the development of various side effects in the long-term (Marsden and Parkes, 1977). Approximately 90% of PD cases are believed to be sporadic in nature, and factors such as environment, age, and sex or any combination may contribute to the causation of PD (de Lau and Breteler, 2006; Ammal Kaidery and Thomas, 2018). Recent analysis has shown that 0.57% of the population with the age of 50 years or more are affected by PD, and a clear prevalence of the disease rises with age (Koprich, Kalia and Brotchie, 2017). As such, the prevalence of PD has been reported to be 1% among 65-year-old people (Abeliovich and Gitler, 2016) and 4 to 5% among people

aged 85 years (Farrer, 2006). Males have been reported to be 50% more susceptible than females (Swerdlow *et al.*, 2001; Farrer, 2006). However, there is no evidence supporting the sex-linked genetic variability of this observation. Rather it is suspected that environmental influences such as smoking rates, working in certain types of workplaces, including mines, and the protective effect of estrogen may play an essential role in a higher disease occurrence in males compared to females. In addition, there is possibility that geographical location may be an influential factor in the prevalence of PD (Farrer, 2006). A study showed that the prevalence of PD is lower in Asia in comparison with Europe, North America, and Australia (de Lau and Breteler, 2006). One study (Wang *et al.*, 2016), suggested that none of the four candidate genes associated with PD (*VPS13C, MIR4697, GCH1,* and *SIPA1L2*) were established to be a significant PD risk factor in the Chinese populations.

Genetic mutations are believed to be responsible for a small portion of PD cases, yet the sporadic and familial forms of PD share most of the common clinical, pathological, and biochemical features (Ammal Kaidery and Thomas, 2018). Studies of the genetics of PD guided the discovery of the importance of some organelles in the pathways related to this disease. Mitochondria are one of the most critical organelles involved in cell health and, therefore, crucial to study when investigating Parkinson disease.

Mitochondria and Parkinson disease

Mitochondria are responsible for the oxidative respiration of the cell through the synthesis of adenosine triphosphate (ATP) in aerobic cells. The structure of the mitochondrion consists of four main parts including 1) a matrix space that houses the enzymes of the citrate cycle and beta-oxidation; 2) an inner membrane responsible for

electron transport for production of ATP, the ATP synthase enzyme, and some carriers for particular metabolites; 3) a relatively permeable outer membrane; and 4) an intermembrane space. Each organelle has an independent small circular DNA which encodes a small fraction of the mitochondrial complements of biomolecules, including some of the subunits of the ATP synthase and the mitochondria's ribosomal mRNAs. Cytosolic ribosomes are responsible for the synthesis of the rest of the mitochondrial peptides that will be imported, post-translationally, to the specific intramitochondrial sites (Sherratt, 1991). In this symbiotic relationship, the cell benefits from its "powerhouse," mitochondria for oxidative respiration, while providing many of the essential proteins for the mitochondria's proper function.

Mitochondria act to establish the essential connections between organelles, both physically and functionally, which are crucial for controlling the homeostasis of cellular processes (Todkar, Chikhi and Germain, 2019). The ER-mitochondria contact sites play a role in regulating lipid transfer between the two organelles and in mitochondrial division (Daniele and Schiaffino, 2014). Morphologically, mitochondria undergo dynamic changes through the processes of continuous fusion and fission activities and maintain their morphology by creating a delicate balance between these opposing actions (Deng, Dodson, Huang, & Guo, 2008). PTEN-induced kinase 1(*Pink1*) and *parkin* are two nuclear-encoded examples of genes that control the maintenance and turnover of the mitochondria. Mutations in these genes can result in early-onset PD (B. Wang, Abraham, Gao, & Yang, 2016). Studies suggest that mutations in *Pink1* and *parkin* may cause defects in mitochondrial dynamics, as a loss of function of *Pink1* or *parkin* may be responsible for mitochondrial fission reduction and/or fusion increase in

the spermatids of *D. melanogaster* (Deng, Dodson, Huang, & Guo, 2008). Regulation of the dynamics of the mitochondria increases the longevity of cells.

Mitophagy is an extremely important process, essential for the development of the organism, cell health, and the homeostasis of mitochondria, and is of paramount importance in the control of mitochondrial quality. During this fundamentally vital process, mitochondria are selected for disintegration, which occurs as a result of engulfment of damaged and/or excess mitochondria in a double-membrane autophagosome, and subsequent fission with either a vacuole or a lysosome for the purpose of degradation (Wang and Zhang, 2018). The *Pink1/parkin* associated mitophagy pathway (Michel, Hirsch and Hunot, 2016). Mitochondrial dysfunction, when coupled with disruption in the mitophagy process in SNPC, can lead to early-onset PD (Lesage *et al.*, 2016). Recently, autophagy and mitochondrial dynamics have been extensively studied in a select number of model organism systems.

Modelling PD in Organisms

Valuable insights into the basis of PD have been obtained through the exploitation of a number of excellent animal models. Techniques that have been used in animal models include toxin-based approaches, including the treatment of animals with 6-hydroxydopamine (6-OHDA) (Breese *et al.*, 2005) and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) (T. Wichmann, 2009). Application of these techniques has been used to selectively destroy dopaminergic neurons in the brains of these animals to simulate phenotypes that attempt to mirror the symptoms of PD, to provide information regarding possible treatment approaches, and to investigate aspects of side effects related to such treatments. Although these toxin-based techniques may be fast and

informative, the utility of such approaches can be questioned as the methods that are used for killing the dopaminergic neurons may produce a dopamine-loss phenotype, but may not necessarily reflect the nature by which the neurons degenerate during the development of PD (Koprich, Kalia and Brotchie, 2017). In contrast, genetic approaches to modelling PD have resulted in *D. melanogaster* becoming an invaluable model for studying such neurodegenerative diseases.

D. melanogaster as a Model Organism

D. melanogaster has been used as a model to study human disease for over a century, beginning with Thomas Hunt Morgan, the first geneticist to utilize the study of the fruit fly to investigate the genetics and the chromosomal theory of inheritance (St Johnston, 2002; Pandey and Nichols, 2011). *D. melanogaster* has become an outstanding model organism for the genetic exploration of a wide variety of human diseases and the answering of an extensive range of biological questions. Although perhaps surprising to some, due to its very high level of functional conservation in many ways with vertebrates, the fruit fly can be used to further improve our knowledge of fundamental neurological disease mechanisms. Using *D. melanogaster* allows the determination of the possible effects of a given gene or genes through the suppression or enhancement of aspects of the disease model phenotypes since they can be examined through genetic modifier screening procedures or exploited to determine potential targets for therapeutic investigations. (Hewitt and Whitworth, 2017). Despite the significant difference between the anatomical arrangement of dopaminergic neurons in the *D. melanogaster* brain with the dopaminergic neurons in the vertebrate brain, a considerable level of

functional homology is shared (Strausfeld & Hirth, 2013) along with other advantages to be discussed.

As an extremely well-studied genetic model organism, *D. melanogaster* was the first complex organism to have its whole genome sequenced. The genome, distributed among four chromosomes, codes for more than 14,000 genes (Pandey and Nichols, 2011). Once the human whole-genome was sequenced, it became clear that 75% of the genetic-based diseases in human beings have a functional orthologue in *D. melanogaster* (Lloyd and Taylor, 2010). The level of biological conservation and physiological positions similarity between flies and humans make the organism an extremely valuable tool for the study of human disease.

Other advantages of using "the fly" as a model include: 1) they are relatively inexpensive; 2) their life cycle is fast; 3) they can produce over 100 genetically identical first-generation flies within 10 to 12 days in 25°C (Pandey and Nichols, 2011; Staveley, 2012); and 4) their life span, which is usually around 70 days, makes for easy study in laboratory settings. Importantly, the *D. melanogaster* brain structure is significantly complicated, with more than 100,000 neurons that mediate complex behaviours, such as feeding, memory, and flight navigation. Although perhaps somewhat counter-intuitive to the uninitiated, *D. melanogaster* provides a system in which genetic characterization is relatively easy to accomplish and has the biological complexity to be compared in a very favourable way to the conditions of human disease.

***UAS-Gal4* System**

Over the past few decades, many highly innovative methods have been developed to study the consequences of gene expression in cells and model organisms, especially *D. melanogaster*. The *UAS-Gal4* system is a directed method of expression that allows the control of expression of the desired gene temporally and spatially (Brand and Perrimon, 1993). The protein product of the *Gal4* gene is a gene regulator and transcription factor, which was first identified in the yeast, *Saccharomyces cerevisiae*. The protein has a DNA-binding domain and a transcription activation domain at its N- and C-terminus, As a transcription factor, this regulator directly binds to a specific set of four related, seventeen base pairs (bp) sequences, known collectively as the Upstream Activating Sequence (*UAS*), which activates transcription from a basal promoter in the downstream location (Duffy, 2002). For the most part, Gal4 is naturally inactive in *D. melanogaster*, except for being the cause of cell death under some circumstances (Kramer and Staveley, 2003). The inability to induce gene expression in the absence of ectopic *UAS* sequences is critical in the workings of the *UAS-Gal4* system as the gene located downstream of the *UAS* will not be transcribed unless the Gal4 protein is present. In order for the *UAS-GAL4* method to be employed, the *UAS* must be situated in the upstream region of the desired gene in one of the parental lines and the *Gal4* gene, directed by a tissue-specific promoter, must exist in the other parental line. With mating, these flies will then produce a critical class progeny that will have both *Gal4* and *UAS* in their genomic complement; the GAL4 protein will then be able to bind to the *UAS* sites and activate the transcription of the gene at the time and in the specific cells or tissues for which the promoter specifies. Transgenes are put under the control of the *UAS* as this has the capability of expressing transgenes at significantly higher levels

than endogenous promoters (Carter, Shieh, Carter, & Shieh, 2015) (Figure 1). Critical class progeny are then able to be assessed for many phenotypes, including gross morphologies and locomotory and survivorship abilities.

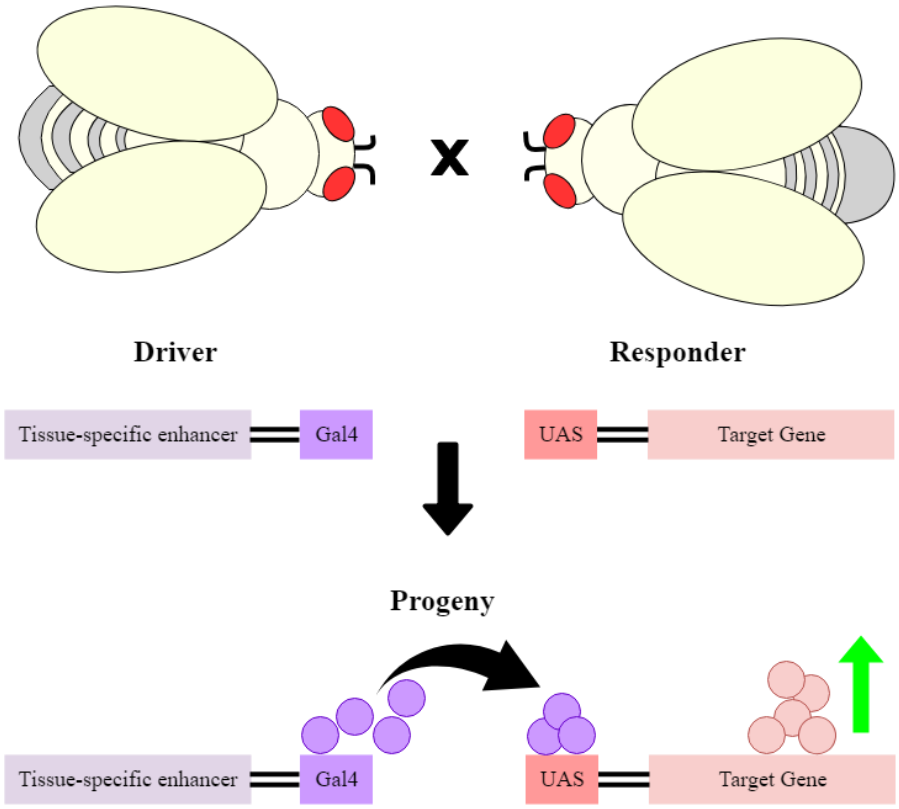


Figure 1. *UAS-Gal4* system in *Drosophila melanogaster*

After crossing transgenic group with responder group, the product of Gal4 gene (purple circles) attaches to the UAS which results in the expression of target gene (pink circles).
Generated with *draw.io*.

RNA Interference and its function

RNA interference (RNAi) is one of the cell's natural ways to silence gene expression through degrading mRNA. This can be utilized by cells in the regulation of gene expression or in protection, as it could degrade mRNA transcripts from viruses. This has become a powerful tool for scientists and enables them to selectively knockdown the expression of a gene by degrading the mRNA to a minimal level.

In this system, double-stranded RNA (dsRNA) attaches to an enzyme called Dicer, which cleaves the dsRNA to fragments with 21-23 base pair lengths that are referred to as the small interfering RNA (siRNA). The dicer-siRNA complex attracts some other cellular proteins and forms an RNA-induced complex (RISC). This complex degrades mRNAs that are complementary to the siRNA sequence and therefore knocks down the production of a specific protein (Carter & Shieh, 2010). This inhibitory process can be coupled with the UAS-Gal4 system, which facilitates the knockdown of expression of a particular gene at a specific time and tissue.

Genetics of Parkinson

It was first believed that only sporadic agents were responsible for Parkinson disease and that no genetic factor was involved (Farrer, 2006); however, Nussbaum's discovery in 1997 was a turning point in the treatments and understanding of Parkinson disease. They discovered that a missense mutation in the *SNCA* gene (encoding α-synuclein protein), located on locus *PARK1/PARK4*, could result in PD (Nussbaum *et al.*, 1997; Houlden and Singleton, 2012). Since then, the tremendous work of researchers has revealed specific genetic agents responsible for Parkinson disease. For example, a number of dominant point-mutations in *SNCA* can result in familial early-onset

Parkinson disease; furthermore, the overproduction of the SNCA protein through duplicating and triplicating the *SNCA* gene, have shown a relation to familial early-onset Parkinson disease (Khurana, Chung and Data, 2017). It has become clear that about 5-10% of the early onset PD cases are caused by genetic agents, and about 90% of the late-onset PDs are sporadic with no specific known causes (Abeliovich and Gitler, 2016). I predict that this number will increase as we learn more about the genetics of Parkinson disease.

PD is genetically very complicated, considering its numerous gene loci and the various disease risk factors (Verstraeten, Theuns and Van Broeckhoven, 2015b). Dysfunction in some cellular functions, such as vesicular trafficking, mitochondrial homeostasis, ER homeostasis, and Golgi homeostasis can lead to the sporadic and familial form of PD as well.

To date, a minimum of 25 genetic risk factors, at least 20 distinct gene loci for familial PD and 15 confirmed genes for the Mendelian forms of PD with an autosomal dominant or recessive form of inheritance have been confirmed (Verstraeten, Theuns and Van Broeckhoven, 2015b); such as *ATPase Cation Transporting 13A2* (*ATP13A2)*, *Familial Parkinson disease type 2 (PARK2)*, *Parkinsonism associated deglycase* (*PARK7)*, *Phosphatase tensin homologue [PTEN], induced kinase 1 (PINK1)*, in autosomal recessive form, *α-synuclein (SNCA)* and *Leucine-rich repeat kinase 2 (LRRK2)*,in autosomal dominant form, and *vacuolar protein sorting 35* (*VPS35) F-Box protein 7* (*FBXO7), and RAB39B.* However, the occurrence of PD, with the involvement of multiple genetic and environmental factors, is more common than mendelian forms of parkinsonism (Satake *et al.*, 2009). In 2017, using whole-exome sequencing (WES), 27 new PD candidate genes were identified, and *VPS13C*, *UHRF1BP*, *ARSB*, and

CAPS2, were some of those represented genes (Jansen *et al.*, 2017). In the following paragraph, I will focus on *Vps13C* as the gene of interest.

PARK23/VPS13C

VPS13C, vacuolar sorting-associated 13 protein (Q709C8), also known as PARK23 and KIAA1421, is localized on chromosome 15q22.2. It consists of 208kb of nucleotides, which includes 86 exons that encode 3753 amino acids (Lesage *et al.*, 2016). *VPS13C* is a member of the *VPS13* gene family, which includes *VPS13A* (involved in chorea-acanthocytosis), *VPS13B* (involved in Cohen syndrome), *VPS13C* (involved in Parkinson's Disease) and *VPS13D* (involved in altered IL-6 production) (Velayos-Baeza *et al.*, 2004). Despite having the largest number of exons among all of the *VPS13* family, *VPS13C* has the smallest genomic area and second smallest coding sequence.

Two main splice variants have been identified by Velayos and colleagues, variant 1A, and variant 2A (Velayos-Baeza *et al.*, 2004; Lesage *et al.*, 2016). Variant 1A includes exons 1 to 5 and 8 to 85 (Skipping exon numbers 6 and7). It encodes a deduced 3,710 amino acid protein, is expressed in all of its tested tissues, and shows relatively homogeneous patterns. On the other hand, variant 2A includes exons 1 to 85 (including exons number 6 and 7) which encode a deduced 3,753-amino acid protein expressed only in the brain, although at an equivalent or even higher level than variant 1A. Variant 2A contains a lysine-rich domain encoded by exon number 7.

Parkinson disease via *VPS13C* in humans can occur as a result of mutations in the *VPS13* gene, which is responsible for both neurodevelopmental and neurodegenerative disorders. In 2017 through WES and Sanger sequencing, it was predicted that a mutation

in *VPS13C* on nucleotide number 62174851, which substitutes a Cytosine with Adenine, leads to a stop codon in that RNA sequence and consequently causes loss of function in recessive forms of PD. *VPS13C* is associated with Golgi to endosome transport such as vesicular transport and protein delivery to the lysosomes (Saxena *et al.*, 2010), membrane traffic at the Golgi–endosome interface (De *et al.*, 2017), and lipid transport between mitochondria and other organelles (Dimmer and Rapaport, 2017; Kumar *et al.*, 2018), functional roles have been established in adipogenesis (Yang *et al.*, 2016) and early impairment of blood glucose homeostasis (Windholz J *et al.*, 2011). Mutation in *VPS13C* was identified in gastric and colorectal cancers with unstable microsatellites. Furthermore, *VPS13C* is involved in mitochondrion organization, through negative regulation of parkin-mediated stimulation of mitophagy in response to mitochondrial depolarization (Lesage *et al.*, 2016). The same study shows that knockdown of *VPS13C* results in the perinuclear redistribution of mitochondria and mitochondrial fragmentation.

VPS13C is also associated with a decrease in mitochondrial transmembrane potential. *VPS13C* silencing enhanced *pink1*/*parkin*-mediated mitophagy in response to mitochondrial damage. It increases respiratory cell rates as a compensatory adaptation response to the decrease in the mitochondrial membrane potential. This phenomenon could be a result of the activation of the anti-apoptotic agent, B cell lymphoma 2 (Bcl2), that is located in the inner membrane of mitochondria that increase the mitochondrial membrane potential (Petit *et al.*, 1995). ATP production in neuronal cells is mainly through mitochondrial oxidative phosphorylation. In acute mitochondrial stress, these cells cannot switch to glycolysis (Lesage *et al.*, 2016). These changes might, in the long term increase the production of reactive oxygen species and cause irreversible damage

to mitochondria with severe consequences for the cell, including cell death. Cell death in *substantia nigra pars compacta* region can cause PD.

PD-affected individuals with early-onset parkinsonism as a result of the loss of function in *VPS13C* initially show a positive response to levodopa treatment similar to PD-affected patients with mutations in their *PARK2*, *PINK1*,or *DJ-1* (Lesage *et al.*, 2016). However, in affected individuals with *VPS13C* loss of function, motor dysfunction symptoms worsen faster, and levodopa treatments lose their efficiency more quickly.

The homologue *Vps13* in *D. melanogaster*

VPS13C in humans is a well-conserved homologue of *Vps13* [FBgn0033194] in *Drosophila,* and to be more specific, *Vps13* represents an orthologous form of the entire *VPS13* family in humans (Thurmond *et al.*, 2018).

The *Vps13* gene, with the annotation symbol of CG2093, is located on chromosome 2R. It contains 13.87 kbp which starts from nucleotide number 7,566,809 to7,580,675. It has two annotated transcripts (*Vps13-RA* and *Vps13-RB*) and two polypeptide isoforms, meaning that both transcripts encode identical polypeptide sequences to each other. *Vps13-RA* has 10364 nucleotides that encode 3321 amino acids, and *Vps13-RB* owns 10480 nucleotides yet also encodes the same 3321 amino acids. Its molecular function in *D.melanogaster* is still unknown, but it is involved in vacuolar protein targeting and protein metabolic process. It is associated with lysosomal degradation pathways, as well as having an important role in protein homeostasis maintenance in the brain of an adult *Drosophila*. Vps13 protein is a peripheral membrane protein that is located at the endosomal membrane. It is most abundant in the fly's head. Mutation in

Vps13 causes several abnormalities in flies, including shortening of fly's life span and age-associated neurodegeneration. Furthermore, it creates sensitivity to proteotoxic stress and accumulated ubiquitylated proteins (Vonk *et al.*, 2017); *Vps13* mutants make good models for understanding the function of *Vps13* in the brain.

Goals and objectives

In this study, the genetics of PD is discussed with emphasis on specific gene *VPS13C* in humans and its homologue *Vps13* in *Drosophila melanogaster*. I have attempted to determine how inhibition and overexpression of *Vps13* may simulate the neurodegenerative Parkinson disease in *Drosophila melanogaster* as a model organism. I used the alteration of pattern formation in the eye as an indicator of neurodevelopment disorders. The responsible gene in that experiment could be identified as a neurodegenerative agent. I have also assessed the impact of inhibition and overexpression of *Vps13* in *D. melanogaster's* survivorship and climbing abilities and so its relationship with PD. Furthermore, assessing the effect of inhibition and overexpression of *Vps13* on inhibition of the *parkin* gene in *D. melanogaster* is of our interest

Materials and Methods

Bioinformatics Analysis

Identification of the *Drosophila* homologue of *Vps13* from the human sequence

The *D. melanogaster* homologue of *Vps13* was identified using the National Centre for Biotechnology Information's (NCBI) tBLASTn search tool (*https://www.ncbi.nlm.nih.gov/)*. The *D. melanogaster* genomic sequences were searched using the amino acid sequence of the human Vacuolar protein sorting-associated protein 13C (accession number Q709C8.1). Accession numbers were retrieved from NCBI to be used in the alignment. The *D. melanogaster* potential homologue was identified as *Vps13* (accession number NP_610299.2).

Identification of additional homologues, multiple alignments, and domain identification

Homologues of *D. melanogaster Vps13* were identified using the NCBI's Basic Local Alignment Search Tool (BLAST) with the tBLASTn function *(https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=tblastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)*. The *D. melanogaster Vps13* sequence was queried against the BLAST database. Sequences were aligned using Clustal Omega to show similarity *(https://www.ebi.ac.uk/Tools/msa/clustalo/)*. Domains were identified using Pfam (Sanger Institute) (*https://pfam.xfam.org/search/sequence)* and NCBI Conserved Domains Database (Figure:2) (*https://www.ncbi.nlm.nih.gov/cdd)*. Clustal Omega was used to

create Percent Identity Matrix of protein sequences encoded by the *Vps13* and *VPS13C* in all vertebrates and invertebrates used for the alignment and then was visualized using Python programming language version 3.7 (Figure:4). A phylogenetic tree was constructed using Clustal Omega (Figure:5)

The accession numbers for the alignment, including vertebrates and invertebrates, include *D. melanogaster Vps13 protein* (accession number NP_610299.2), *Homo sapiens* VPS13C protein (accession number Q709C8.1), *Mus musculus* VPS13C protein (accession number NP_796158.2). *Gallus gallus* VPS13C protein (accession number XP_001233000.2), *Lucilia cuprina* Vps13 (accession number XP_023292594.1), *Culex quinquefasciatus* Vps13 protein (accession number XP_001858599.1), and *Anoplophora glabripennis* Vps13 (accession number XP_018574731.1).

D. melanogaster **Culturing and Crosses**

D. melanogaster **Media**

D, melanogaster stocks and crosses were maintained on a standard media containing 5.5 g/L agar, 65 g/L cornmeal, 15 g/L yeast, and 50 ml/L fancy grade molasses diluted in water with 5 ml of 0.1g/ml methylparaben in ethanol and 2.5 ml of propionic acid. Approximately 7 ml of medium were poured per vial. The medium was prepared by Dr. Brian E. Staveley about twice a month and stored at 4 to 6°C until use.

D. melanogaster S**tocks**

All *D. melanogaster* stocks were ordered from the Bloomington *D. melanogaster* Stock Centre in Indiana University, IN, USA And Vienna *D. melanogaster* Resource Center (VDRC)

The double-balancer complex line *Ddc-GAL4-UAS; parkin RNAi* was prepared by Dr. Brian E. Staveley (Table 1).

Table 1. Genotypes of stocks used to characterize *Vps13* in this study

Genotype	Abbreviation	Expression Pattern	Balancer	Provider	Reference
Transgenic Lines					
w;GMR-GAL412	*GMR-Gal4*	Eye	…	Bloomington Drosophila Stock Center	(Freeman,1996)
w[*];P{w[+mW.hs]=GawB}D42	*D42-Gal4*	Motor neuron specific	…	Bloomington Drosophila Stock Center	Parkes et al., 1998
w1118;P{dcGAL4.L}4.3D	*Ddc-Gal4*	Neuron	…	Bloomington Drosophila Stock Center	Li et al.,2000
w;P{ple-GAL4.F}3*	*Th-Gal4*	Dopaminergic Neuron	…	Bloomington Drosophila Stock Center	Inamdar et al., 2014
UAS Lines					
w;UAS-lacZ 4-1-2	*UAS-lacZ*	…	…	Bloomington Drosophila Stock Center	Brand et al., 1994
y[1]sc[*]v[1];P{y[+t7.7]v[+t1.8]=TRiP.HMS02460}attP40	*UAS- Vps13-RNAiHMS02460*	…	CyO	Bloomington Drosophila Stock Center	Perkins et al., 2015
y[1]sc[*]v[1];P{y[+t7.7]v[+t1.8]=TRiP.HMS01715}attP40	*UAS- Vps13-RNAiHMS01715*	…	CyO	Bloomington Drosophila Stock Center	Perkins et al., 2015,
y[1]w[67c23];P{w[+mC]y[+mDint2]=EPgy2}EY09640	*UAS-Vps13 EY09640*	…	…	Bloomington Drosophila Stock Center	Bellen et al., 2004,
w[1118];P{GD14789}v29972	*UAS- Vps13-RNAiGD14789*	…	…	VDRC	Dietzl et al., 2007,
UAS Complex Lines					

w;ddcGAL4/CyO;UASparkinRNAi/TM3	*Ddc-Gal4;UAS-Parkin-RNAi*	Neuron	CyO, TM3	Staveley Lab	M'Angale and Staveley, 2016

D. melanogaster crosses

The stocks of all transgenic lines were maintained separately at room temperature (22°C±2), and the media refreshed every ten to fourteen days. Males from the desired inhibition or overexpression *UAS* lines were crossed with virgin females from lines that contained a *GAL4* transgene. To collect virgin females for crosses, newly eclosed females were isolated every 8 to 12 hours (twice a day) from stock tubes and maintained upon fresh media in groups of 3 or 4 and left for 6 to 7 days for confirmation of their virginity. They were then mated with males of desired genotypes. To do so, 3 to 5 virgin females and 2 to 3 males of desired genotypes were combined and placed in fresh media for breeding. To increase the efficiency of their breeding, the parental flies were transferred onto new media three times at 2 to 3 day intervals prior to being discarded, and the male progeny of the critical class were collected once eclosed. Male progeny of the critical class were kept at 25℃. In order to eliminate the collection of F2 male progeny with unwanted genotype, the progeny collection tubes were discarded after 18 days.

Biometric Analysis of the Compound Eye

Eye analysis of *D. melanogaster* shows the effect of specific gene manipulation upon neurodevelopment, and is accomplished by a comparison of ommatidia and bristle

numbers of each critical class to the control group. To do so, each of the desired *UAS* lines were crossed with the *GMR-GAL4* transgenic line, as described previously, and the critical class male progenies were collected and matured for 3 to 5 days in groups of no more than 20 on standard medium at 25°C. The flies were transferred into pre-labelled 1.5 ml Eppendorf tubes in groups of 10, frozen and stored at a temperature of -80°C. Flies were thawed at room temperature and mounted upon aluminum stubs. Each stub was cleaned, polished and prepared with double-sided sticky tape for immobilization of the sample flies. Using tweezers, approximately 15 to 17 flies were fixed on each stub under stereoscope so that all flies were laid onto their right side, and their left eyes were facing upwards. The stubs with fixed flies were desiccated for at least 48 hours before imaging. Scanning electron micrographs were taken of each eye using the FEI Quanta 400 Scanning Electron Microscope. The images were then analyzed using the software program ImageJ (Schneider, Rasband and Eliceiri, 2012). A total number of ommatidia and bristles and the disrupted area were determined. These data were then analyzed using GraphPad Prism 8.1.1 (GraphPad Software Inc.), where the mean ± standard error of the mean was calculated. Unpaired t-tests were used to determine significance. Results were deemed statistically significant when $p < 0.05$.

Behavioural assays

Longevity assay

The analysis of *D. melanogaster* survivorship under near-optimal conditions was conducted to determine the effect of specific gene manipulation on the lifespan compared to a control group. To accomplish this experiment, the male progeny of each critical class was collected every day and maintained on fresh medium. To reduce the

adverse effects of overcrowding, groups of no more than 20 flies per vial were kept until a sample size of 300 to 350 individuals for each cross was reached. Critical classes were aged at 25°C. Every 48 hours, the number of dead flies in each tube was recorded and surviving flies were then transfered onto the new fresh nutrient medium: if no deaths were recorded, fresh media was provided every four days. Flies were considered dead when no movement was observed. Data were analyzed, and the graph was drawn using the software GraphPad Prism 8 (Graphpad Software Inc.) using the log-rank test with significance considered at $p < 0.05$.

Climbing Assay

The climbing analysis was used to determine the effect of specific gene manipulation on the locomotory ability of flies compared to a control group. To do so, I collected seventy critical-class male progeny from desired crosses as described before within a day or two. They were then kept in separate vials as groups of 10 at 25°C. These flies were transferred onto new medium two times per week. I started scoring their climbing ability one week after their eclosion and continued scoring their locomotory ability every seven days until either the number of flies diminished to less than 10 in total or if at least 50% of the flies showed loss of climbing ability in two consecutive climbing assays. The climbing ability of fifty flies per genotype was assessed in five groups of ten. Ten trials were conducted for each cohort of ten flies per genotype, which allowed a total of 500 trials per genotype per week. The flies were scored every seven days on their ability to climb inside a 30 cm glass tube with a 1.5 cm diameter that was marked with five 2 cm sections along a buffer zone, and the levels were scored from 1, which is the closest level to the buffer zone to 5, which is the height from the last mark on the

tube above. Flies were first tapped down and then were given ten seconds to climb; after that, a score was given to each fly in the tube-based upon the sections reached. The flies' alignment was repeated ten times (trials) per climbing session. The climbing index was calculated using an equation: Climbing index = Σ nm/N. "n" indicates the number of flies at a specific level, "m" is the score of the level which is between 1 and 5 and "N" is the total number of flies climbed in that trial (Todd and Staveley, 2008). Data were then analyzed using the software GraphPad Prism 8.1.1 (GraphPad Software Inc.). A nonlinear regression curve was then used with a 95% confidence interval to interpret the graphs of the 5-climbing index as a function of time in days for each genotype. The slope for each graph represents the rate of decline in climbing ability, and the Y-intercept represents the initial climbing ability, and both parameters are calculated for each curve using GraphPad Prism 8.1.1. (GraphPad Software Inc.). Slopes of the curves were compared using a 95% confidence interval. Curves with no overlap by the 95% confidence interval were considered significantly different.

Results

Bioinformatic Analysis of *Vps13*

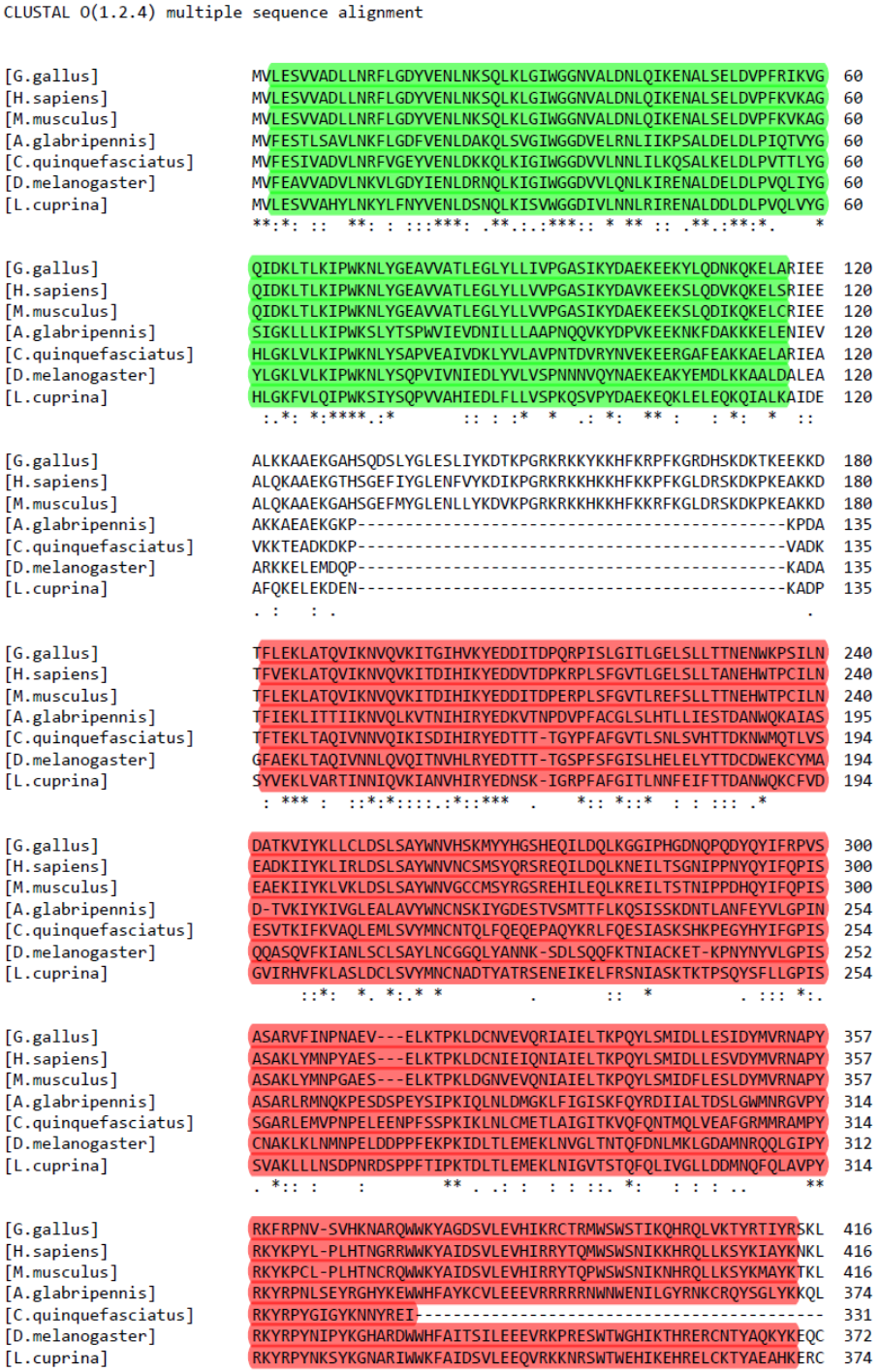
Vps13 is highly conserved among multiple species

A multiple sequence alignment (MSA) is the alignment of three or more biological sequences such as protein, DNA, or RNA. This analysis can be used to understand the molecular evolution and structural relationships between different sequences. An MSA may illustrate conserved sequences and domains as well as pointed, insertion, and deletion mutations. Differences or substitutions in a single amino acid in the multiple alignment sequences are shown as multiple characters in a single alignment column, and insertion/deletions are shown as hyphens in the sequences of the alignments (Lipman, Altschul and Kececioglut, 1989). Clustal Omega is a bioinformatic tool that is based on the (C) or (C++) programming language. The goal is to create the best pairwise alignment with the highest score. Clustal Omega uses computational algorithms, statistical models, and the factor of time, to evaluate the similarity scores of the sequences and obtain the optimal pairwise alignment. The highest score would be allocated to the highest levels of similarity, and the score decreases as the level of similarity decreases, with penalties assigned for each gap (Sievers *et al.*, 2011). Clustal Omega is a fast and promising tool for multiple sequence alignments. Multiple alignments of vertebrates *Homo sapiens* VPS13C protein (accession number Q709C8.1), *Mus musculus* VPS13C protein (accession number NP_796158.2). *Gallus gallus* VPS13C protein (accession number XP_001233000.2) and invertebrates *Lucilia cuprina* Vps13 (accession number XP_023292594.1), *Culex quinquefasciatus* Vps13 protein (accession number XP_001858599.1), and *Anoplophora glabripennis* Vps13 (accession number XP_018574731.1) was done. *D. melanogaster* and *Homo sapiens*

showed 32.31% similarity in the Vps13 protein sequence alignment. The information of these alignments was used for creating the percent identity matrix that helped visualize the percentage of similarities between all the species used in the protein alignment (Figure 2). Pfam software revealed the presence of detected domains including N-terminal region of Chorein (Chorein N) which is a leucine zipper domain (Mizuno *et al.*, 2007; Lesage *et al.*, 2016), Vacuolar-sorting associated protein 13 (VPS13), three repeating coiled region of VPS13 (VPS13 mid rpt), SHR-binding domain (SHR-BD) which is present in VPS13 proteins. Vacuolar sorting-associated 13 protein C-terminal (VPS13C) and Autophagy-related protein C-terminal domain (ATG C). Chorein N, VPS13, two VPS13 mid rpt, SHR-BD, and VPS13C domains were detected in the VPS13 protein of the invertebrates. It is essential to mention that *Flybase* identifies the VPS13 domain, starting at amino acid number 137 to amino acid number 369, which is shown in red colour in the multiple alignments, as Vacuolar protein sorting-associated protein 13, second N-terminal domain (VPS13_N2). Although *Pfam* software did not identify the ATG C domain in insects, multiple alignments showed a very high level of preserved amino acids in that area compared to vertebrates, which can lead to the conclusion that the ATG C domain has been conserved in invertebrates (Figure 2). For validating this observation, I used MSA to align Chorein-N and ATG domain of Vps13 in the fruit fly with ATG2A protein in mammalian cells and results showed high levels of conservation (Figure 3). Furthermore, Vps13 protein contains a DUF1162 domain with an unknown function (Velayos-Baeza *et al.*, 2004; Lesage *et al.*, 2016), which Pfam did not report. There are some differences between the amino acid sequences of vertebrates and invertebrates. One of the most significant differences is the absence of a second VPS13 middle repeating coiled region domain in the

invertebrates. Most of the domains are highly conserved. Since vertebrates have four different types of *VPS13* in their genome, and *D. melanogaster* has only one kind of *Vps13*, it was essential to detect the VPS13C domain in their protein sequence.

The phylogenetic tree visualizes the amount of genetic evolution and changes during time, and the numbers are representative of these changes. As shown on the tree, *Homo sapiens* (0.06607) and *Mus musculus* (0.684) had the lowest amount of genetic differences while *D. melanogaster* (0.22352), had a higher amount of genetic variation as the result of genetic evolution (Figure 5).


CLUSTAL O(1.2.4) multiple sequence alignment
[G.gallus] MVLESVVADLLNRFLGDYVENLNKSQLKLGIWGGNVALDNLQIKENALSELDVPFRIKVG 60
[H.sapiens] MVLESVVADLLNRFLGDYVENLNKSQLKLGIWGGNVALDNLQIKENALSELDVPFKVKAG 60
[M.musculus] MVLESVVADLLNRFLGDYVENLNKSQLKLGIWGGNVALDNLQIKENALSELDVPFKVKAG 60
[A.glabripennis] MVFESTLSAVLNKFLGDFVENLDAKQLSVGIWGGDVELRNLIIKPSALDELDLPIQTVYG 60
[C.quinquefasciatus] MVFESIVADVLNRFVGEYVENLDKKQLKIGIWGGDVVLNNLILKQSALKELDLPVTTLYG 60
[D.melanogaster] MVFEAVVADVLNKVLGDYIENLDRNQLKIGIWGGDVVLQNLKIRENALDELDLPVQLIYG 60
[L.cuprina] MVLESVVAHYLNKYLFNYVENLDSNQLKISVWGGDIVLNNLRIRENALDDLDLPVQLVYG 60
[G.gallus] QIDKLTLKIPWKNLYGEAVVATLEGLYLLIVPGASIKYDAEKEEKYLQDNKQKELARIEE 120
[H.sapiens] QIDKLTLKIPWKNLYGEAVVATLEGLYLLVVPGASIKYDAVKEEKSLQDVKQKELSRIEE 120
[M.musculus] QIDKLTLKIPWKNLYGEAVVATLEGLYLLVVPGASIKYDAEKEEKSLQDIKQKELCRIEE 120
[A.glabripennis] SIGKLLLKIPWKSLYTSPWVIEVDNILLLAAPNQQVKYDPVKEEKNKFDAKKKELENIEV 120
[C.quinquefasciatus] HLGKLVLKIPWKNLYSAPVEAIVDKLYVLAVPNTDVRYNVEKEERGAFEAKKAELARIEA 120
[D.melanogaster] YLGKLVLKIPWKNLYSQPVIVNIEDLYVLVSPNNNVQYNAEKEAKYEMDLKKAALDALEA 120
[L.cuprina] HLGKFVLQIPWKSIYSQPVVAHIEDLFLLVSPKQSVPYDAEKEQKLELEQKQIALKAIDE 120
[G.gallus] ALKKAAEKGAHSQDSLYGLESLIYKDTKPGRKRKKYKKHFKRPFKGRDHSKDKTKEEKKD 180
[H.sapiens] ALQKAAEKGTHSGEFIYGLENFVYKDIKPGRKRKKHKKHFKKPFKGLDRSKDKPKEAKKD 180
[M.musculus] ALQKAAEKGAHSGEFMYGLENLLYKDVKPGRKRKKHKKHFKKRFKGLDRSKDKPKEAKKD 180
[A.glabripennis] AKKAEAEKGKP---------- 135
[C.quinquefasciatus] VKKTEADKDKP---------- 135
[D.melanogaster] ARKKELEMDQP---------- 135
[L.cuprina] AFQKELEKDEN---------- 135
[G.gallus] TFLEKLATQVIKNVQVKITGIHVKYEDDITDPQRPISLGITLGELSLLTTNENWKPSILN 240
[H.sapiens] TFVEKLATQVIKNVQVKITDIHIKYEDDVTDPKRPLSFGVTLGELSLLTANEHWTPCILN 240
[M.musculus] TFLEKLATQVIKNVQVKITDIHIKYEDDITDPERPLSFGVTLREFSLLTTNEHWTPCILN 240
[A.glabripennis] TFIEKLITTIIKNVQLKVTNIHIRYEDKVTNPDVPFACGLSLHTLLIESTDANWQKAIAS 195
[C.quinquefasciatus] TFTEKLTAQIVNNVQIKISDIHIRYEDTTT-TGYPFAFGVTLSNLSVHTTDKNWMQTLVS 194
[D.melanogaster] GFAEKLTAQIVNNLQVQITNVHLRYEDTTT-TGSPFSFGISLHELELYTTDCDWEKCYMA 194
[L.cuprina] SYVEKLVARTINNIQVKIANVHIRYEDNSK-IGRPFAFGITLNNFEIFTTDANWQKCFVD 194
[G.gallus] DATKVIYKLLCLDSLSAYWNVHSKMYYHGSHEQILDQLKGGIPHGDNQPQDYQYIFRPVS 300
[H.sapiens] EADKIIYKLIRLDSLSAYWNVNCSMSYQRSREQILDQLKNEILTSGNIPPNYQYIFQPIS 300
[M.musculus] EAEKIIYKLVKLDSLSAYWNVGCCMSYRGSREHILEQLKREILTSTNIPPDHQYIFQPIS 300
[A.glabripennis] D-TVKIYKIVGLEALAVYWNCNSKIYGDESTVSMTTFLKQSISSKDNTLANFEYVLGPIN 254
[C.quinquefasciatus] ESVTKIFKVAQLEMLSVYMNCNTQLFQEQEPAQYKRLFQESIASKSHKPEGYHYIFGPIS 254
[D.melanogaster] QQASQVFKIANLSCLSAYLNCGGQLYANNK-SDLSQQFKTNIACKET-KPNYNYVLGPIS 252
[L.cuprina] GVIRHVFKLASLDCLSVYMNCNADTYATRSENEIKELFRSNIASKTKTPSQYSFLLGPIS 254
[G.gallus] ASARVFINPNAEV---ELKTPKLDCNVEVQRIAIELTKPQYLSMIDLLESIDYMVRNAPY 357
[H.sapiens] ASAKLYMNPYAES---ELKTPKLDCNIEIQNIAIELTKPQYLSMIDLLESVDYMVRNAPY 357
[M.musculus] ASAKLYMNPGAES---ELKTPKLDGNVEVQNIAIELTKPQYLSMIDFLESLDYMVRNAPY 357
[A.glabripennis] ASARLRMNQKPESDSPEYSIPKIQLNLDMGKLFIGISKFQYRDIIALTDSLGWMNRGVPY 314
[C.quinquefasciatus] SGARLEMVPNPELEENPFSSPKIKLNLCMETLAIGITKVQFQNTMQLVEAFGRMMRAMPY 314
[D.melanogaster] CNAKLKLNMNPELDDPPFEKPKIDLTLEMEKLNVGLTNTQFDNLMKLGDAMNRQQLGIPY 312
[L.cuprina] SVAKLLLNSDPNRDSPPFTIPKTDLTLEMEKLNIGVTSTQFQLIVGLLDDMNQFQLAVPY 314
[G.gallus] RKFRPNV-SVHKNARQWWKYAGDSVLEVHIKRCTRMWSWSTIKQHRQLVKTYRTIYRSKL 416
[H.sapiens] RKYKPYL-PLHTNGRRWWKYAIDSVLEVHIRRYTQMWSWSNIKKHRQLLKSYKIAYKNKL 416
[M.musculus] RKYKPCL-PLHTNCRQWWKYAIDSVLEVHIRRYTQPWSWSNIKNHRQLLKSYKMAYKTKL 416
[A.glabripennis] RKYRPNLSEYRGHYKEWWHFAYKCVLEEEVRRRRRNWNWENILGYRNKCRQYSGLYKKQL 374
[C.quinquefasciatus] RKYRPYGIGYKNNYREI---------- 331
[D.melanogaster] RKYRPYNIPYKGHARDWWHFAITSILEEEVRKPRESWTWGHIKTHRERCNTYAQKYKEQC 372
[L.cuprina] RKYRPYNKSYKGNARIWWKFAIDSVLEEQVRKKNRSWTWEHIKEHRELCKTYAEAHKERC 374

```
[G.gallus]              TLGKLSEETQRQIQDLEKKLDVFNIVLARQHAQLETIRSGQKVLKKKATEVEKKSGGWFS  476
[H.sapiens]             TQSKVSEEIQKEIQDLEKTLDVFNIILARQQAQVEVIRSGQKLRKK-SADTGEKRGGWFS  475
[M.musculus]            TQAKVSEEIQKQIQDLEKSLDVFNIILVRQQAQVEVIHSGQKLRKK-SAEAGEKR-GWFS  474
[A.glabripennis]        KKAIKGD-EKQEIENCEKILDVTSIVIIRQQIEIELQRLQQIE--------AKQDKGWFG  425
[C.quinquefasciatus]    -------------------LDLHNIVVIRQKVEFEVQKEG----KRQE---EEQKGGWFS  365
[D.melanogaster]        LSKKPSAVLTETCRLLETELDVFNLLLIRQRVNIEIAKQREAV--------PEQKSGWFS  424
[L.cuprina]             ASKKPSATVEATCTIAEQKLDLFNLILIRKRVQLEVDRLRQGEELKKE---AIAKSSWFG  431
                                          **: .::: *:: :.*  :                 .**.

[G.gallus]              GFWGRRESKKKEDE-ESFVPETINELMTPEEKAKLFTAIGYSESSYHLSLPRQYVAHIVT  535
[H.sapiens]             GLWGKKESKKK-DE-ESLIPETIDDLMTPEEKDKLFTAIGYSESTHNLTLPKQYVAHIMT  533
[M.musculus]            GFWGKKESKKR-DE-ESSVPETIDDLMTPEEKDKLFTAIGYSENAYNLALPKQYVAHILT  532
[A.glabripennis]        WIWKGSSSQQDEDSTSAVLLKQFQSEMTPAEKETMYKAIGYQENAAPTTIPKTFIAYSGA  485
[C.quinquefasciatus]    SWWGGGAKKTDDADNAADIKKQFEAAMTSEEKAKLFKAIGYQENDSPTELPEFYVAQILQ  425
[D.melanogaster]        GWGWGGGAKKDDQTTSQKLVEKFEAAMTSEEKEKMYRAIGYQENAKPTDLPESYEAIRMN  484
[L.cuprina]             GWFGRGKKNEDE---RSDLEISIQAAMTPEEKRKLHQAIGYEDNMAPLELPEHYEAIHMK  488
                                :         :   ::  **  ** .:. ****.:.     :*. : *

[G.gallus]              LKLVSTSLTIKEDKN--------------VAETLKVQIIDLSTKISQRPGAQAIKVEAKL  581
[H.sapiens]             LKLVSTSVTIRENKN--------------IPEILKIQIIGLGTQVSQRPGAQALKVEAKL  579
[M.musculus]            LKLVSTSIIIRENRN--------------VPEILRVQIIGLGTQVSQRPGAQALKIEAKL  578
[A.glabripennis]        FILRNLEIELRDDDQ-------------QIKRVLFCNLSGVGIKLEHRPAANAIKADVKI  532
[C.quinquefasciatus]    FELNQLKVSIKSEVTVGS-QERVANPNQLLERVMLLELKNVKCGVQQRPSAGAMKASLQM  484
[D.melanogaster]        FKLIALEVGLYKDERNS---SAATKDFHELPSLVLLNFSMATALITQRPAAEAISIIAGM  541
[L.cuprina]             FTLKALEIGLYDDSELCKTYGRGSVDWHSLQTLMLIKLNMTTCTVKQRPAASAINVCVGM  548
                        : *   .: : .:               :   :  ::     : :**.* *:.    :

[G.gallus]              ENWYVTGLRQENIVPSLVASIGDSRSSLLKIEFNINPEESTADQSLSIESQPVEIKYDAR  641
[H.sapiens]             EHWYITGLRQQDIVPSLVASIGDTTSSLLKIKFETNPEDSPADQTLIVQSQPVEVIYDAK  639
[M.musculus]            EHWYVTGLRQQDIVPSLVASIGDTTSSLLKIEFETNPENSPADQTLIVQSQPVEVIYDAK  638
[A.glabripennis]        DNFTIEGLQQDNFIPQLITSEIDSKGGLLEIAFETNPLDEICDQKIRLVANPVKIFYDAM  592
[C.quinquefasciatus]    QELTISGLRQGEVLPIMVRSQLEGSDTLLDVSFETSPEDKKCDQRVVVTSRPLQIIYDAE  544
[D.melanogaster]        REIKVTGLTRNDYTPLLVESKITDEFNLLEVFFETNPLDKLCDQRVKVVARPLQITYDAP  601
[L.cuprina]             KELALTGLEQEGIAPTIIKSKISDENNLLDIFFETNPLDKQCDQRIKVTARPLEIVYDAE  608
                         .  : ** :    * :: *       **.: *: .* :. .** : : :.*::: ***

[G.gallus]              TINAMVEFFQTSKGMDLERLTSATLMKLEEIKERTATGLAHIIEMRKVLDLKIHLKPSYL  701
[H.sapiens]             TVNAVVEFFQSNKGLDLEQITSATLMKLEEIKERTATGLTHIIETRKVLDLRINLKPSYL  699
[M.musculus]            TINAVVEFFQSNKGLDLEQITSATLMKLEEIKERTATGLTHIIETRKVLDLRINLKPSYL  698
[A.glabripennis]        TINKVIDIFKVPSDTVADQIQAAAGSGLSNVKQMTSTGLQFTIEKRNRLDLYVDLQAPYV  652
[C.quinquefasciatus]    TIIQLTKIFQMPRTATISQLTDAAAEKLVNIKERSATGLQYAIANHPRLELNVDISPSFI  604
[D.melanogaster]        TILALINAFQTPGDVTLSKFEDAASTKISNFKERSATGMQYMIDKKAVLDVDILLMPNIL  661
[L.cuprina]             TILKLMGVFVPQQQVNLSELEGAATLKISDFKERSATGMQYMIESHAIVEVDITFMPNIL  668
                        *:  :   *        ..:  *:   : :.*: ::**: . *  :  ::: : :    :

[G.gallus]              VVPQTGFYHENS-NLLILDFGTFQLNSINQGNSE----------ASSFSSLEEIMDKAYD  750
[H.sapiens]             VVPQTGFHHEKS-DLLILDFGTFQLNSKDQGLQK----------T-TNSSLEEIMDKAYD  747
[M.musculus]            IIPQTGFHHEKS-NLLILDFGTFQLNSKDQGAQK----------T-ANASLEEIIDKAYD  746
[A.glabripennis]        IIPHGGKYTGVE-NVLVANLGRLKMLSSGQRTNINIVKQMYQQGLGHEDIFLRLKEHCYD  711
[C.quinquefasciatus]    IVPHGGLFCSRE-AVLVVSLGKLLVQTEPRPIDQKDVHTMHEEGANQEQILQEIIRQSYD  663
[D.melanogaster]        VVPHKGVYDAGNVSLLVVSMGQVHLSSQPRRESN-KLQHLFSAGEDKDEILKTVMENAYD  720
[L.cuprina]             ILPQGGKYRPGEQSLVVIGLGEFKISSAPHRNKSKDVVTMHQGGKDQDEILQMIMEKAYD  728
                        ::*: * .   .  ::: .:* . : :  :  .                 :  :  :.**

[G.gallus]              KFDVEIKNVQLLFGRAGEDWKKAR-FQRSSTLHMLQPMDIHIQLAKSMVEKDTRMAKFKV  809
[H.sapiens]             KFDVEIKNVQLLFARAEETWKKCR-FQHPSTMHILQPMDIHVELAKAMVEKDIRMARFKV  806
[M.musculus]            KFDVEIRSVQLLFAKAEENWKKCR-FQHPSTMHILQPMDIHVELAKAMVEKDVRMAKFKV  805
[A.glabripennis]        SFVLELTDLQILIAQGDEDWRTAVKESEHSAMHLLSPLTLSVTYSKCLIMDDPRFPQNKI  771
[C.quinquefasciatus]    KFVLEIRDVQAIVATIDEDWQGTLRCNTVTEMHLLEPTSFRISAHLCVIDDDPRLAKCKI  723
[D.melanogaster]        RFTVAVDDVQMLVVRAGEPWQNALAEANSTEMHVLRPVSLKVTAALCVVDNDPRLPNIKV  780
[L.cuprina]             RYNVSIENIQIMVSKPEDDWELCMRKLIISDMHLLRPTSVLVDAELCVVDDDPRLPKTKI  788
                         : : : .:* :.    : *.       : :*:* *  . :    .:: .* *: . *:
```

```
[G.gallus]              SGGLPLVHIRVSDQKIKAIFDLIDSIPLPEMSSVSIPSTKAATIPAIPVDAKGLLTTHHL  869
[H.sapiens]             SGGLPLMHVRISDQKMKDVLYLMNSIPLPQKSSAQSPERQVSSIPIISGGTKGLLGTSLL  866
[M.musculus]            SGGLPLMHVRISDQKIKDALCLINSIPLPQKSSTPSPERQVASIPVLSGGTKALLGTSLL  865
[A.glabripennis]        TGELPSIDVRVSEARFLLLVALGTSIPLPESDVPEPQPLSKSKKS-----------SS--  818
[C.quinquefasciatus]    FGELPSVNICVTEQRVLEALSIVTSLPLPESDIEIQPAPI--AKDS------NVFSSS--  773
[D.melanogaster]        DIDLPAILVNVSEDRIFLAIKVATSIPLPEQKEPASRLTQTNS------------RSS--  826
[L.cuprina]             NILLPSINLNLTEDRVFEALRVAMSIPLPEKEKPAEQPKTNLLKSA------SQSTLM--  840
                           ** : : ::: :.   . :  *:***: .

[G.gallus]              LAEMASDSEEEYFDFEERYEPYYRALSKGEEIENTESAKEELTDLQLKFEIKEVLLELTK  929
[H.sapiens]             LDTVESESDDEYFDAEDGEPQTCKSMKGSELKKAAEVPNEELINLLLKFEIKEVILEFTK  926
[M.musculus]            LDGVESESDEEFFDAEDGDSQAARTVKASELKKAAEVPNEELVSLLLKFEIKEVVLELTK  925
[A.glabripennis]        ------M-----------------------------------ML--------------  821
[C.quinquefasciatus]    ------L-----------------------------------SL--------------  776
[D.melanogaster]        ------M-----------------------------------SI--------------  829
[L.cuprina]             ------R-----------------------------------SI--------------  843

[G.gallus]              QEKTEETVLVFDVKHLGTEATVRTFNLAAVSYLKTISLDYYEIGG-KKVPLHLISSSDKP  988
[H.sapiens]             QQKEEDTILVFNVTQLGTEATMRTFDLTVVSYLKKISLDYHEIEGSKRKPLHLISSSDKP  986
[M.musculus]            QQKEEETILVFNVTQLGTEATMRTFDLTAVSYLRKISLDYHDIKGSRKKPIHLISSSDRP  985
[A.glabripennis]        ------------------------------------LKYKELQQSAKQT----------  834
[C.quinquefasciatus]    ------------------------------------LKYLDEKQQKLTKIQR-------  792
[D.melanogaster]        ------------------------------------SNFINKEVKKIGP--S-------  843
[L.cuprina]             ------------------------------------PQFLNQDERRKTL----------  856
                                                            .: :

[G.gallus]              GLDLLKVEYIKADKNGPHFLTVFDNTEQKIQVAFSSLNILLHTEALMSAVSFLATVSPSG  1048
[H.sapiens]             GLDLLKVEYIKADKNGPSFQTAFGKTEQTVKVAFSSLNLLLQTQALVASINYLTTIIPSD  1046
[M.musculus]            GLDLLKVEYIKVDRNGPSFQTTFEKTEQTVKVAFSSLNLLLQTQALLSSLNYLTTVIPSD  1045
[A.glabripennis]        ------------------------------------------------------KDL  837
[C.quinquefasciatus]    ------------------------------------------------------PAD  795
[D.melanogaster]        ------------------------------------------------------ASG  846
[L.cuprina]             ------------------------------------------------------  856

[G.gallus]              SGS---SRETPTKEEKQQDDRTLKKVTRPFKDKDAFAFKLLARLD-AFCLNLCDEKKNIA  1104
[H.sapiens]             DQSISVAKEVQISTEKQQKNSTLPKAIVSSRDSDIIDFRLFAKLN-AFCVIVCNEKNNIA  1105
[M.musculus]            SQNTGVAKEVQAMPEK-QKNSPLQKVMVPSRDSDVIGFRLFAKLN-AFCVTVCDEKSNIA  1103
[A.glabripennis]        LPA--------------QSA---------ESDKEFVQFT---TMEAKFVMA---------  862
[C.quinquefasciatus]    TLS--------------TSD---------SVDGEVVQFI---DLEVNFVLN---------  820
[D.melanogaster]        SSA--------------SKD---------PLLDEIIQYT---SLDVNFSLG---------  871
[L.cuprina]             SAT--------------AAT---------SLTPEVVQYT---SLEVHFALK---------  881
                                            : . :     ::  * :

[G.gallus]              EIKIQGLDSSLLLQSNQTEFFARLKDVVVTDVDTRTLHKKAVSIVGDEVFSFSLVLYPYA  1164
[H.sapiens]             EIKIQGLDSSLSLQSRKQSLFARLENIIVTDVDPKTVHKKAVSIMGNEVFRFNLDLYPDA  1165
[M.musculus]            EIKIQGLDSSLSLQSKKQSLFARLENIIVTDVDPKTIHKKAVSIVGNEVFRFNLDLYPDA  1163
[A.glabripennis]        ----------------------------------------------------EMSVVINQQS  872
[C.quinquefasciatus]    ----------------------------------------------------ACSLTIYKIS  830
[D.melanogaster]        ----------------------------------------------------EINFVLFQSS  881
[L.cuprina]             ----------------------------------------------------EFSIALVRTS  891
                                                                             .. :   :

[G.gallus]              TEGEAYADMTKVDGTVSLKVGCIQVVYLHKFLVSLLTFLNNFQTAKEALSAATVQAAEKA  1224
[H.sapiens]             TEGDLYTDMSKVDGVLSLNVGCIQIVYLHKFLMSLLNFLNNFQTAKESLSAATAQAAERA  1225
[M.musculus]            TEGDSYTDMSTVDGVVALHVGCIQIVYLHKFLMSLLSFLNNFQVAKEALSAATAQAAEKA  1223
[A.glabripennis]        AIGS------------------------------------------------------  876
[C.quinquefasciatus]    GTGS------------------------------------------------------  834
[D.melanogaster]        RKCE------------------------------------------------------  885
[L.cuprina]             FDVD------------------------------------------------------  895
                           .

[G.gallus]              ATSVKDLAQRSFRLAMDIYLKAPVIVIPQSSVSFNAIVVDLGLIKVQNRFSLASPEGSLL  1284
[H.sapiens]             ATSVKDLAQRSFRVSINIDLKAPVIVIPQSSISTNAVVVDLGLIRVHNQFSLVSDEDYLN  1285
```

```
[C.quinquefasciatus]      IKIRVVAKLEDVTVELQNDKRSIAVMEVQNLTSSAVVKTSYTEVTVRLQDIILTDTNPAT  1107
[D.melanogaster]          VKVRVIANLDQVGLVLTGRKRPIAEMNVKKFVSSLIIKSSYTEVNIGLKDIQVLDLNPYT  1155
[L.cuprina]               IKMKIEANLDQVGLELTCKKRSIASLKVNHLYAGVVLKTSYTEVQIGLKDILINDMNPLN  1182
                          ...:: *:*: . : .      *:   :: : :.  :* .  :*   *::* : : :  .

[G.gallus]                IHKKAVSIMGDEVFRFQMSLYPDATAGEAYADMSKVDGKMSLKVGCIQIVYLHQFFMSLL  1712
[H.sapiens]               IHKKAVSILGDEVFRFQLTLYPDATEGEAYADMSKVDGKLSFKVGCIQIVYVHKFFMSLL  1716
[M.musculus]              IHKKAVSILGDEVFRFQMSLYPDATEGENYGDMSKVDGRLSLKVGCIQIVYVHKFFMSLL  1713
[A.glabripennis]          LHKLIISGTEGVALKAQIVINNLEG----ES--KKPDMEISAKLGGLRIVFLNCFVSNML  1154
[C.quinquefasciatus]      IHTKILSIIGDNALEVKVVLFDLEA----TSDYNSDDMRIEVTMGCARIVFLNWFVTSVL  1163
[D.melanogaster]          IHKNILSIVGKDAFNCQIVIYNKEE----TQDYNSDDMKITVDIGCMKIIFLNWFVAGVM  1211
[L.cuprina]               IHSNILSIVGKDALKCDVVLYNLDE----TRNYNSDDMKINVEIGCMKIVFVNWFVNSVL  1238
                          :*.  :*     .:. .: :             .. * .:   :*  :*:::: *. .::

[G.gallus]                NFLNNFQTAQEALNAVTVQAAEMAASSMKDFAKKSFRLLMDINLKAPVIVVPESSTSCNA  1772
[H.sapiens]               NFLNNFQTAKEALSTATVQAAERAASSMKDLAQKSFRLLMDINLKAPVIIIPQSSVSPNA  1776
[M.musculus]              SFLNNFQAAKEALSTATVQAAERAASSVKDLAQKSFRLLMDIDLKAPVITIPQSSVSPNV  1773
[A.glabripennis]          NFLNRFQEAQRQIIEASQAAAKQARENVKNAYDKATKISLDIKLKAPDIVVPVACKSFQA  1214
[C.quinquefasciatus]      AFLDNFQAAQQRIKDASAAAAEAAKNSAVEAYSKATRMKLDIKVKAPIIVIPVDSRSLKA  1223
[D.melanogaster]          NFLNNFTAAQATISQAGAAAAESARQKAMDAYETATRMKLNIRIKAPIIIVPIGSQDRNA  1271
[L.cuprina]               SFMNNFQAAQQAIANAGAMAADAAKQNAVAVYEKATRMKLNVKIKAPVIIVPIDSKSLEA  1298
                           *::.*  *:  :  .   **. * ..     ..: :: ::: :*** * :*  . . :.

[G.gallus]                LIADLGLIRVQNEFKFVSSD-ESSLPPVIDNMDVQLTHLKLSRCILSADS----------  1821
[H.sapiens]               VIADLGLIRVENKFSLVPME-HYSLPPVIDKMNIELTQLKLSRTILQASL----------  1825
[M.musculus]              VIADLGLIRVENKFSLVSVE-QLALPPVADEMSIQLTQLKLARTVLQADS----------  1822
[A.glabripennis]          LILDLGVLYLSNTFLTLETKNAEGFSAVIDELKVSLSQFQLLRVKLDQKG----------  1264
[C.quinquefasciatus]      VAMDFGHLSITNNFRDIATE-NSHGPAVIDEMKIELKDMKLSKVDVCPSEMGNLSHVS--  1280
[D.melanogaster]          LLLDLGLLELTNNTVEVAVA-EEERLAVIDEIKLQICDVKISKIVLLDGN---------E  1321
[L.cuprina]               VALDLGQLTMSNVITDISTPGSDKGPAVMDEIKLILKDMRLARVNILEESHLYQSQSLDS  1358
                          :  *:* : : *    :          * *::.: : ..:: :  :

[G.gallus]                ------------QPDSEILCPVSLTLLVQRNLAATWYHKAPTIGIKGDLKPMQIALSQD  1868
[H.sapiens]               ------------PQNDIEILKPVNMLLSIQRNLAAAWYVQIPGMEIKGKLKPMQVALSED  1873
[M.musculus]              ------------PQHDVEILKPVNMLLCIQRNLSAAWYTQIPGMEIKGELKPMQVALSQD  1870
[A.glabripennis]          -----------EKIKSEELIAPINIDIIIRRNLSASWFNAVPDIYISGKIKIIKLLLSQG  1313
[C.quinquefasciatus]      -RYGSEEIIYGIVPDQSAVLSPTSFALIMKRNLSSGWYRERPDLDISGRLKAVELNFIAT  1339
[D.melanogaster]          STVDEVDAEVGF-LSKFNMMNPMSCTLSITRNLSYTWYRDVPELNLSGRLKSIELTLFAD  1380
[L.cuprina]               DSLDVVDSNFGF-KSKVNILDPTSFTLVVKRNLSFSWYKNLPEMDISGRLRCIDLNLFMD  1417
                                      .  :: * .  : : ***:  *:   * : :.* :: :.: :

[G.gallus]                DLTVVMKILLENLGGASSQPNTMQQSIEDNQMVRKEKVPNE-------SDISEGQENIST  1921
[H.sapiens]               DLTVLMKILLENLGEASSQPSPTQSVQETV-RVRKVDVSSV-------PDHLKEQEDWTD  1925
[M.musculus]              DLTVLMKILLENLGEASSQPSPTQYAQEAA-RVKRD-TRSG-------PDYLKEQ-ELAD  1920
[A.glabripennis]          DYKMIMGILNGNLAEGQVETVEVKQSKP--AAVKGFEEHV----VHVEV----GSVCERH  1363
[C.quinquefasciatus]      DYSVIMQILSKNMTEGQDEFKKAEVKQEKSPT-SPQIHAV----SLFIDTFLCFIFCTTL  1394
[D.melanogaster]          DYALVMLVLNRNLNEGLEEFPPSEEAPQ-EAQVRPERRNS-------RAGRLSR--TVQV  1430
[L.cuprina]               DYALIMSILNRNMREGVNEFPTVVAEPPQKPRISPPKTPIDSPAKKAGDQFLEK--IRAK  1475
                          *  ::* :*  *:  .  :

[G.gallus]                SQSSAECYTVLIFDFNFESLSVTLYRSDM-NQKPLLSLHNEKLCLGELRLNLLASSGKIL  1980
[H.sapiens]               SKLSMNQIVSLQFDFHFESLSIILYNNDI-NQESGVAFHNDSFQLGELRLHLMASSGKMF  1984
[M.musculus]              PKPPGDQTVTLQFDFHFDSLSIILYNSDS-SQEPRLSFHNDSFRLGELTLHLMASAGKMF  1979
[A.glabripennis]          IHIEVKVNVFLKFTFSMDELILNLYTGSDSEVEENISSHSEDTHLAKFSIEGLSIKGRIL  1423
[C.quinquefasciatus]      EAAKQKIDTFLKFSFQIDSINIKLFTAPG-------------EGLAGFEIFCLSLQGKKL  1441
[D.melanogaster]          SPIREKIHESIKFNFQFDGVVINLMEGEG-------------AGLARFGIYFLSVKGTKL  1477
[L.cuprina]               VSHQEKISEQFKFNIQLDGVVLNLMTGVN-------------EGLARFGLYVISLKGTKL  1522
                               .    : * : :: : : *                   *. : :  ::  *  :

[G.gallus]                TDGSVDVSVKLKACTLDDLRAGIQNVTARMIDKKD-NT----------------------  2017
[H.sapiens]               KDGSMNVSVKLKTCTLDDLREGIERATSRMIDRKNDQD----------------------  2022
[M.musculus]              KDGSMNVSLKLKTCTLDDLREGIERATSRMIDKKNDQD----------------------  2017
[A.glabripennis]          SDQSIVTSVVIVNCLLDDMRKGRGDKLNRLIERSHTNQEAGD-------ST-FSSTESVA  1475
[C.quinquefasciatus]      TDGSLNTAIILCDIRLDDIRPNRENMLTRLMERRSTAEANISILSGRSDDA--ETAALTA  1499
[D.melanogaster]          DNGTLSTSVVLCNIQMDDMRSNSKSQIRQYLSRKDWVQPKLDTDEII--------DACYN  1529
```

```
[M.musculus]          ATSVKDLAQRSFRVSVDIDLKAPVIVIPQSSLSTNAVVVDLGLIRVHNRFSLVSGEDTAN  1283
[A.glabripennis]      ------------------------------------------------------------  876
[C.quinquefasciatus]  ----------------------------SSSEAYA-----------------------  841
[D.melanogaster]      -TS------------------------PDVSIEFL-----------------------  895
[L.cuprina]           -SP----------------------LSDDSADFN-----------------------  906

[G.gallus]            PPIIDKMDVQLTKLKLSRASMEEGLSHQDIQILHPINLSLSVSRNLAASWFHKLPILEIT  1344
[H.sapiens]           PPVIDRMDVQLTKLTLYRTVIQPGIYHPDIQLLHPINLEFLVNRNLAASWYHKVPVVEIK  1345
[M.musculus]          PPVIDKMEVQLTKLKLSRTAIQPGTSHPDIQLLHPINLEFFVSRNLAANWYHKVPVVEIK  1343
[A.glabripennis]      ------------------------------------------------------------  876
[C.quinquefasciatus]  TPTEEFPSSPL--------------------------------------------------  852
[D.melanogaster]      TPDGDVLPSQL--------------------------------------------------  906
[L.cuprina]           TPSQHSEESEF--------------------------------------------------  917

[G.gallus]            GYLDTMNVAVSQEDLNVFLKVLTENLGEAEQTN--AKQILQEEGRIKGMEKSTLMQSKGV  1402
[H.sapiens]           GHLDSMNVSLNQEDLNLLFRILTENLCEGTEDLDKVKPRVQETGEIKEPLEISISQDVHD  1405
[M.musculus]          GRLDSMNVSLNQEDLNLLFRILAENLGEATEDLDKGKPRIQERGETKACREVSTPQDVHT  1403
[A.glabripennis]      ------------------------------------------------------------  876
[C.quinquefasciatus]  ----------------------EQSLAQPRKS-------------------------VAF  865
[D.melanogaster]      ----------------------TENIQEPI---------------------------------  914
[L.cuprina]           ----------------------LDT--RSS---------------------------------  923

[G.gallus]            SEGILSEKRKETLNEDVINVLLNFEIKEVVITLMKQLQKERYPLHILNVLQLGTKTEIRN  1462
[H.sapiens]           SKNTLTTGVEEIRSVDIINMLLNFEIKEVVVTLMKKSEKKGRPLHELNVLQLGMEAKVKT  1465
[M.musculus]          TQGVPAARVEETRPVDIINVLLNFEIKEVVVTLMKKAERKGSPFHELKILHLGMEAKVKA  1463
[A.glabripennis]      ----------------------------------------QISELASFKVKNLECSLGQQT  897
[C.quinquefasciatus]  S---------MPFH-----------------------SSDSRKIMAFQVRQLELTMVQRT  893
[D.melanogaster]      ----------EELP-----------------------PTPPQQILSIDIRRLEAHFVSKT  941
[L.cuprina]           ----------VSLT-----------------------HQDSQKLLAFQVLQLEVYMAQRT  950
                                                          :  :.: .*      :

[G.gallus]            HELTAGAYLKKIIMRCTEITDSNGDPLCIINSSSK--TDEPLLKMEYIKADADGPDFVTT  1520
[H.sapiens]           YDMTAKAYLKKISMQCFDFTDSKGEPLHIINSSNV--TDEPLLKMLLTKADSDGPEFKTI  1523
[M.musculus]          HDMTAAAYLRNISMRCFHFPDSKGEPLRIVNTSDV--SDGILLKLLFIKADSDGPDFKTI  1521
[A.glabripennis]      FVTNVQLVLSSISLEQN----RSGEIISIISTPRTGKGSEYLFKVEFTQIDTNSPELHSK  953
[C.quinquefasciatus]  YDLKVALKLGAVTLDQFRVRNEQENILNVIQTPKYEDNDDYLFTLNYTNCKKNSPEFATK  953
[D.melanogaster]      YESVATVKLGDINLRQYDCQDSDMDVLDVIYTPKQENSSNYLFTVSCTIADKSSPEFSTK  1001
[L.cuprina]           FEMVAQAKLGAISLTSYETRDKKEEELVVIETPGFTTAGKSLLNVTFTAVDKSTPDFITK  1010
                      .   .   *  : :        . : : :: :      .  *:.:     . . *:: :

[G.gallus]            YSSTKQNIDVVFSCLDVVLHTEALISIMSFFTFSVPSGALPSTDKAP-------ENKPQT  1573
[H.sapiens]           HDSTKQRLKVSFASLDLVLHLEALLSFMDFLSSAAPFSEPSSSEKES-------ELKPLV  1576
[M.musculus]          HDNTKQKLKVSFSSLDLVLHLEALLSLMDFLSSAIPSSDSSSSEKEP-------ELKPLV  1574
[A.glabripennis]      YQSCETSLILDFGVLNITLHQEGLLSLIEFKTLLLDQISH-------LTDDQQKDRIATI  1006
[C.quinquefasciatus]  YESVEQEVGINFSTLVLLLHEEALNELIRRDRELQQF--------------PMRDRIATI  999
[D.melanogaster]      YNSTEQLVVANFEVLQIVLHQECLQRIMEVVNNFQRNLDLVLSSTRP------RDRMGSI  1055
[L.cuprina]           YKSIEQLARINFATLNVILHQENLLYIMEIANQLQRKVEKITASSKPVEIDQNKDRIASA  1070
                      :.. :      *  * : ** * *  ::                          :

[G.gallus]            EEKG---------------------S-------------------V-LRPASGSTTHDDT  1592
[H.sapiens]           GESR---------------------S-------------------IAVKAVSSNISQKDV  1596
[M.musculus]          GESR---------------------S-------------------LAIRAVPSSY-EGDA  1593
[A.glabripennis]      RSKRQQFV------------ISEGAE-T-------------VQKTVAKPKKKLQAVVVET  1040
[C.quinquefasciatus]  HEEESTATVLLHAAREKLPTILEDESAPV------------GGAAGGSKANRKRPSIVDS  1047
[D.melanogaster]      GGGDGIKR--------TLNVILEDTEEI------------MTTDQMKRRKKTRRTHVVET  1095
[L.cuprina]           GAADGFVD--------RLARIAEEAEITLERKTSISTTNASPTQTMTRRSRKTAQNVVES  1122
                                              .                               :

[G.gallus]            FELKLTANLNAFSISVCDQTCRIADIRIQGMDASVAVKTTEIEVFSRLQDIIITNVDPKT  1652
[H.sapiens]           FDLKITAELNAFNVFVCDQKCNIADIKIHGMDASISVKPKQTDVFARLKDIIVMNVDLQS  1656
[M.musculus]          FDLKITAELNAFNIFICDQKSNIAEIKIHGMDASISVKPKQTDVFARLKNIIVMNVDSLS  1653
[A.glabripennis]      IKFKLLAELQEVSVKFANDRSNISSCAIKGIVSDIIVKDTYTQVNANLQEISVIDLNPDS  1100
```

```
[L.cuprina]              MDETLTTNIVLCNMQLDDTRPSNTSEITKYLCRKDWFGSELEKQAQSSDMAYEAGEKSVK  1582
                          : :: . : :    :** * .      : : :

[G.gallus]               NDSSMIDVSYKQDKNGTEVVAILDKLYVCASMEFLLTVADFFINSMPTSSPERSTQL---  2074
[H.sapiens]              NNSSMIDISYKQDKNGSQIDAVLDKLYVCASVEFLMTVADFFIKAVPQSPENVAKET---  2079
[M.musculus]             NNSSMIDISYSQDKNGSQVDAVLDKLYVCASVEFLMTVADFFIKAMPQSPENIAKEI---  2074
[A.glabripennis]         PPKSMIDVTYQQKESDMFVDVRIFSFTIILSVDYLMKLAEFFSTAGTKKPAITTENK---  1532
[C.quinquefasciatus]     PLHSMIDVTFNMKENDMFADVKVSSFNLILSVDFLLKIQQFLQPEELVEQKALQ---AAE  1556
[D.melanogaster]         ERNFMVDVTAIIKEDDTFAEVRVRGFDLIVCIDFLLKLTTFLTLPPEENPRESVYIKPAP  1589
[L.cuprina]              DSQYMLDITAVLKQNDTVANVRISSFDLILCVDFLLKLSEFFKTPDVEEYIEIEKEKPAE  1642
                             *:*::   .:..    . :  : :  .:::*:.:  *:      .

[G.gallus]               ------------HLK----NVTSVKSKPETEAPFRPNMKVKAVIMDPEIVFVANLTSADA  2118
[H.sapiens]              ----------QILPR----QTATGKVKIEKDDSVRPNMTLKAMITDPEVVFVASLTKADA  2125
[M.musculus]             ----------QIPSR----QTAAGRVKMEKDDSVRPNMTLKAMITDPEVVFVASLTKADA  2120
[A.glabripennis]         -----QS-QTKLEE-----SKQSLPE---KQPEKESQMTINLKLEKPDIILVEHMDNIDT  1578
[C.quinquefasciatus]     E-QARRERKT---------STSNVPA----EKSEPGQMTVIIKIEQPDIILVEKMDDINC  1602
[D.melanogaster]         VSETARDTKHSIRSSAILAAQELVPVESSSHEVPNRKMNLILHIDEPDIILVENLEDLNT  1649
[L.cuprina]              VIPSTQQQSTQLKSLA---TVATVAREAKESNEAVNKMNLTLVIDEPDIILVETLKDMNT  1699
                                                      :*.:   : .*::::*  : . :

[G.gallus]               PALKVSFQCDFSLTSGKHAQRMTAQVKDFKVLACAFLREKQDRSVTKVLQPCSLVMENMM  2178
[H.sapiens]              PALTASFQCNLSLSTSKLEQMMEASVRDLKVLACPFLREKRGKNITTVLQPCSLFMEKCT  2185
[M.musculus]             PALTASFQCNLSLSTSKLEQMMEASVRDLKVLACPFLRERRGKSITTVLQPCSLFMEKCT  2180
[A.glabripennis]         NAMILNSEILVKLRFAGPHQVINGIIKDFQLYTCNYNPAHREETRGNVLYPVNISIAGST  1638
[C.quinquefasciatus]     YALILNNEIQLNVRLIGERQIIKGELKDLCLYYAEFNPERRNDTKHFVIHPCSISLNGST  1662
[D.melanogaster]         SCIIFNAQVHLNYRSINDKQIVNGQIDALKMYMCAFLPERREMTRHYILHPCVISLQGST  1709
[L.cuprina]              SSIIFNMQAKLIYRAIGEKQLINGNIDGLKMYMCSFMPERREATRHYILHPCVINLHGST  1759
                          .:  . :  .       * : . :  : :  . :   ::  .   :: *  : :

[G.gallus]               HVSGLQTVVLTVEELTIKISPIILNTVVTIMAAIKPKT-TEEDSKGAAEVPEDLWQVKPI  2237
[H.sapiens]              WASGKQNINIMVKEFIIKISPIILNTVLTIMAALSPKT-KEDGSKDTSKEMENLWGIKSI  2244
[M.musculus]             WASGKQNINIVVKEFVIKISPIILNTVMTIMAAMSPKT-KEDEWKDTPKETDNLWAVKSI  2239
[A.glabripennis]         PQDKGLHLELLITAIRLSVSPATIELLNRVMVTMTSSGTSMDENAKEQVNHPNLWEQKEF  1698
[C.quinquefasciatus]     PEGKGLHLGLNTTKIKISVSPAVIELINNALMTLTANEQVKLDETQKTVNYSDLWDIKDY  1722
[D.melanogaster]         PEEEGMHISLKLSDIIINVSPATIELLNKAMLSVSSGTMTKCAIAEESRNYSNLWHQHHF  1769
[L.cuprina]              PEEDGMHISLKTSEIIINISPATLELFNKAMQTINTAETDET-KMLEAKNYSDIWTPKKY  1818
                                : :    : :.:**  :: .   : ::.                ::*  :

[G.gallus]               DECNAWFLGVDVATEATETFK-----DHEHAVKQEKFDIVVKSVQITLECGLGHRTVPLL  2292
[H.sapiens]              NDYNTWFLGVDTATEITESFK-----GIEHSLIEENCGVVVESIQVTLECGLGHRTVPLL  2299
[M.musculus]             TDYNSWFLGVDMATEVTENFR-----DSEHPSIEENCVVAVESVQVTLECGLGHRTVPLL  2294
[A.glabripennis]         ADSDYWFLKTENALEALEYPSAEGGRIKKKNALQELCIITVPSIVFTIEAGVGNKTLPML  1758
[C.quinquefasciatus]     NPDDFWFIRPEMAEDALS------LESVCREIKEEKCMVEVPSISLIIETGLGINTIPML  1776
[D.melanogaster]         HSRTYWFTKVEQGVDALEAEQRS-VSTDNEKQKTEKCVIEIPSITLVIESGVGYYTKPLI  1828
[L.cuprina]              NEHDYWFLKAEPAEDALETLDAQQAVSTCGSVKTERCVIEIPSIMLVLEAGIGYYTTPLI  1878
                              **   : . :  .                *   : : *: . :* *:*  * *::

[G.gallus]               LAESVFSGVLKNWSSLMEVSADMSLEIHYYNETYAVWEPLIERIEGG-------KKQWSL  2345
[H.sapiens]              LAESKFSGNIKNWTSLMAAVADVTLQVHYYNEIHAVWEPLIERVEG--------KRQWNL  2351
[M.musculus]             LAESKFSGNIKNWTSLMAAAADMTLEVHYYNETHAVWEPLIERVEG--------NKPWSL  2346
[A.glabripennis]         ILETGFQGAARNWSSQLSLEASLTMQMGYYNSSLALWEPLIEPVEVE-EDGKMHFVPWEL  1817
[C.quinquefasciatus]     FIETNMQAEVSNWSSEMKINSSLRLSMSYYNQALALWEYVIEPNEVELPNGNVEHVPWEL  1836
[D.melanogaster]         SLDTRITAVFNNWSRSLTAHGSLTLNMNYYNQALAEWEPIIELNEVIGRNGVREYTPWEL  1888
[L.cuprina]              SMDTCLNAIANDWSSNLSVNGSLTLTMNYYNQSLAVWEPVIEKNEHIARDGERIFSPWEL  1938
                           :: : .   :*:  :   ..: : : ***.  * ** :**  *           *.*

[G.gallus]               KLEMKTNPVQERSL-----MPGDDFIVFPEPQTAVNISSKDTMNITISKCCLAVFSNLAK  2400
[H.sapiens]              RLDVKKNPVQDKSL-----LPGDDFI--PEPQMAIHISSGNTMNITISKSCLNVFNNLAK  2404
[M.musculus]             KLNVKKNPIQDKSL-----MPGDDFI--PEPQTAVHISSGATMNITISKSCLNVFSNLAK  2399
[A.glabripennis]         KVELSMGEQDDTLNATSPTTSESDGDSLPSPVMSIDVVSEHVLEMTVTKTCLEVLQNLSK  1877
[C.quinquefasciatus]     TLELEVDHH----E-----------DRSRDPTTRINVGSRDSLEMSVTKTCLDVFQNLGK  1881
[D.melanogaster]         KFEMGMEKVQSELE-----------DDAEQQAMHMNIHSAETLEITLSKTCLGLLSELAE  1937
[L.cuprina]              NFNLGIEKNPSEFE-----------ENKVDQTTTIKIHSEENLEMTVSKTFLDLIGTLGE  1987
                          .::                         .    :.: *   ::::::*  * ::  *.:
```

```
[G.gallus]              AFSEGTASTFDYS-FKDTAPFIVKNALGVHLQVFPSSSFRI------VNSAEKENVH---  2450
[H.sapiens]             GFSEGTASTFDYS-LKDRAPFTVKNAVGVPIKVKPNCNLRV------MGFPEKSDIF---  2454
[M.musculus]            GFSEGAASTFDYS-LKDRAPFTVKNALGVPMKVQPNRNLKV------MGSPEKSDIY---  2449
[A.glabripennis]        AFATAIGTTSVKHIRETSAPYQVLNELGEDVTILMEESSFKIAEGGSLE-----DINKSA  1932
[C.quinquefasciatus]    AFSEAIKREGIVK-TETQAPYMVQNDTGQDIKINFVGSDFVIQASHLQSGREDELI----  1936
[D.melanogaster]        AFSQAIDQNGLTK-PDIVAPYVLENDTGFDVNLNLRKGIFTLHEVHRGGTP---VGANST  1993
[L.cuprina]             AFGQAMDPNGLMK-PDVIAPYVIENDTGFDINLNFASGIFTLHECHIPNSNGTSTLNGSI  2046
                        .*. .          .  **: : *  *  : :    .

[G.gallus]              --CVESGQN---MELEYSVFEAPQRR---------RLSAL-----YRQESSIFSLSLELE  2491
[H.sapiens]             --DVDAGQN---LELEYASMVPSSQG---------NLSIL-----SRQESSFFTLTIVPH  2495
[M.musculus]            --DVGAGQH---LELDYASLEPSRQG---------KLSIL-----SRQESSLFTLTFVPY  2490
[A.glabripennis]        AVPLQLKG----D-------APANALLHL--------SKELL---VQQQEQDKFLHIKIP  1970
[C.quinquefasciatus]    --AI---EQTGSQEVTSCIVMPNGRLNLEPREKDNDLSMVTVMQDNENSATHKFLKVIVG  1991
[D.melanogaster]        LLMVAQSEEVDPSVIKTCTISTGGRAYLQTKDL----STL-----SEEDSEDYTLYVTIG  2044
[L.cuprina]             VFKNDLSCTVTADSIKCCTLSPGCKAYLQTKNL----ATI-----QNPDEEEYNIYVSVG  2097
                                                            :       . .     : .

[G.gallus]              RYEKVINVPIAKPSRRLYNIKSLSDGHS-DSIIVQIDATEGNKVITVRSPLQIKNHFSIP  2550
[H.sapiens]             GYTEVANIPVARPGRRLYNVRNPNASHS-DSVLVQIDATEGNKVITLRSPLQIKNHFSIA  2554
[M.musculus]            GYTEVASVPVARPGRRLYNVRNPSASHS-DSVLVQIDATEGNKVVTLRSPLQIKNHFSIA  2549
[A.glabripennis]        SKNCVLVLPVGRADKRFFSLNYRGDGHDNWGIISDIKVDKGSTIVTLRSVLQVYNHFCVP  2030
[C.quinquefasciatus]    DTEKELTLPVYKSDKRYFPL-YRKTDQEPWAIISEVKIEHGTTVVVLRGIVQIYNHFTAP  2050
[D.melanogaster]        DINKEIALPVSKSDTRFFNL-MRSTSHEPWGIISEVKQEYGTTKVNIHGVVSVHNHFTTG  2103
[L.cuprina]             HIQKQLVLPVSKSDKRYFSL-FRDTNQDPWGIVSEVNSEYGTTSINIHSVVNIQNHFTTP  2156
                               :*: : . * : :     .:.  .:: ::.   *.. : ::. :.: ***

[G.gallus]              FVIYKLNRD----SRLLQPIGISKPEEEFHVPLHSYRCHLYVRPTGMLEGQFRESTTNIA  2606
[H.sapiens]             FIIYKFVKN----VKLLERIGIARPEEEFHVPLDSYRCQLFIQPAGILEHQYKESTTYIS  2610
[M.musculus]            FIIYKFVKN----VKLLERIGIARPEEEFHVPLDSYRCQLYVQPAGGLEQQYTHSSTYIS  2605
[A.glabripennis]        IDVYYMTAK----GNELELIGAAQPGGYLNIPLKAVYTPTNE--LFFAISGHSITSTPYI  2084
[C.quinquefasciatus]    IYVHQFV------DHEKYLVGEVRSGGFFNVPLYFLYSDFKE--LHFSMKGYHSSAQGIS  2102
[D.melanogaster]        LNIYRRNPAPTAQCFEDIFVGRVRPGEVFHVPLHAIYAESKD--LFFSMRGYRRSVQGIS  2161
[L.cuprina]             IKVLRMNAK----TKEMILVGTVEPGKVYHVPLHAIYSEGKY--LHFAIEDYHTSVQGIK  2210
                        : :               :*  .    ::**                 .  :

[G.gallus]              WREELHRSNEVKCLLQCPATE--TNFLPLIVSTTAVPDQLNYISAHG-EEWDPAYIIHLH  2663
[H.sapiens]             WKEELHRSREVRCMLQCPSVE--VSFLPLIVNTVALPDELSYICTHG-EDWDVAYIIHLY  2667
[M.musculus]            WKEELHRSREVRCMLQCPAVE--VSFLPLIVNTVALPDELSYIGAHG-EDWDPAYVIHLY  2662
[A.glabripennis]        WKDLQTSVS-FVKLLQCSNRDFNKGKGPFIIKCIGKMEQVYFEDTVRHTMTSTCYNIHLR  2143
[C.quinquefasciatus]    WKESPNHTE-LIKSLQCDPI---KTFEPFHINAVRERHDVFHEVTSKHTMLSACYEIHLR  2158
[D.melanogaster]        WASNPSDLN-YSHQLHCDPT---NTFEPLIMNARRSKSEVYFENTNKYTLLSAFYTIHLR  2217
[L.cuprina]             WDDCPTDMN-YMRHLQCDPI---ETFEPFFINVSREKTEIYHEHSNKFKLMSAYYTLHLR  2266
                        * .          *:*          *: :.      :: .  :      .  * :**

[G.gallus]              PTLTVRNLLPYSLRYLLEGTAEARELLE--------------------------------  2691
[H.sapiens]             PSLTLRNLLPYSLRYLLEGTAETHELAE--------------------------------  2695
[M.musculus]            PPLTLRNLLPYSLRYLLEGTAETHELAE--------------------------------  2690
[A.glabripennis]        PSVIFKNFLPIDVVCCV-----DEQA--------------------------------  2164
[C.quinquefasciatus]    PPLMLRNALPIGLTISVAGCSVRRELDADLVTSTESLS---------------NASTIVG  2203
[D.melanogaster]        PPLYLRNSLPINIQVSVAGCSVRKEDGL-----------------DAQSSQRFVDRGYRK  2260
[L.cuprina]             PPVYLRNALPINITVSVAGCSVRETVSNLAVQTEESFTSTPETDTDSKKPLHQSDHSFVK  2326
                        * : .:* ** .:   :      .

[G.gallus]              ----------GSAADVFHS---------RINGEIMELVLMKYQGKNWDGHLKIHEGMPEF  2732
[H.sapiens]             ----------GSTADVLHS---------RISGEIMELVLVKYQGKNWNGHFRIRDTLPEF  2736
[M.musculus]            ----------GSSADVLHS---------RISGEIIELVLVKYLGKNWNGHFRICDTLPEF  2731
[A.glabripennis]        ------DEFEVKAGDTLQLPNIDP------GRNVLVIRLPEYLEKEWSCRYEISEEPEEF  2212
[C.quinquefasciatus]    EDYLDYGEKLLRPGELLHLPTVKTAARTATETTYIVARLVNYLEKDWSCTTEIPAQPPEF  2263
[D.melanogaster]        EDFLDYGEKPVNSGDVLHLPTVRLASKGKESKSFLVVRLVQYLEKDWSCATEIWDYTDDV  2320
[L.cuprina]             EDLLDYGEKDIEAGSVLHLPTVKLAGRGKDSKSYLVIRMIQYLERDWSCTTEISENHPDV  2386
                                   .. ::               :   : :*  ::*.   .*     :.

[G.gallus]              FSVCFTS-HSATVMTVDLYVHTRRIGSRMILSVFSPYWIINKTSRILQYRAEDTHV---K  2788
```

```
[H.sapiens]             FPVCFSS-DSTEVTTVDLSVHVRRIGSRMVLSVFSPYWLINKTTRVLQYRSEDIHV---K  2792
[M.musculus]            FLVCFSS-DTAEVMTVDLSVHVRRIGCRMELSVFSPYWLINKTSRVLQYRSEEIHV---K  2787
[A.glabripennis]        AVWTFDSYDSPTKMSLDLGMHTLNKNGSFIMSLYCPFWMLNKTGLMIGYRDSDENLNVLH  2272
[C.quinquefasciatus]    GVWTFNAYDSVEVMSLQLGVKYESRTDGLTLIVYCPFWMLNKTGLMLSYRPNDENTNILY  2323
[D.melanogaster]        ITWTFSSYDSEMKVDMDLYVKTENRHGSLMLTLFSPFWMINKTGMMLTYKSETTSVEVLY  2380
[L.cuprina]             TNWKFNSYDSDAKMSMDLCVMFEDRNGSLLITLFSPFWLINKTGEMLSYKTDSETVEVLY  2446
                            * : .:     ::* :        : : ::.*:*::***  :: *: .

[G.gallus]              HPADLRDIVLFSFKKKNIFSKNKIQLCISTSTWSSSFSLDTVGSYGCVRCSANGMEYLVG  2848
[H.sapiens]             HPADFRDIILFSFKKKNIFTKNKVQLKISTSAWSSSFSLDTVGSYGCVKCPANNMEYLVG  2852
[M.musculus]            HPADFRDIILFSFKKKNIFSKNKVQLKISTSAWSNGFSLDTVGSYGCVKCPATNMEYLVG  2847
[A.glabripennis]        HPPHFMGPILFSFNAKHFFGKKKATIRVELGDWSDKFSLDVAGSSGVISCKANDRLYQIG  2332
[C.quinquefasciatus]    HPPEYDGPILFSFREKVFFGKKKAAIRVDTGEWSEKFSLDVAGSSGVIACQANNMTYQIG  2383
[D.melanogaster]        HPPEYSGPILFTFRDKLFFDKKKASIRIDNGQWSEKIPLDVAGSVGEVICFANNQKYPVG  2440
[L.cuprina]             HPPEYNGPILFSLRDKYLFDKKSCSIRVENGEWSNKIPLDVAGSTGSVGCKANDKTYQIG  2506
                        ** .  . :**::. * :* *:.  : :. . **. : **..** * : * *..  * :*

[G.gallus]              VSIKMSSFNLTRIVTFTPFYTIANKSSLELEVGEIGPNGSFPTNKWNYISPSECLPFWPE  2908
[H.sapiens]             VSIKMSSFNLSRIVTLTPFCTIANKSSLELEVGEIASDGSMPTNKWNYIASSECLPFWPE  2912
[M.musculus]            VSIKMSSFNLSRVVTLTPFCTVANKSSLDLEVGEIASDGSIPTNKWHYVASSECIPFWPE  2907
[A.glabripennis]        VHIQLTYNNLTKQVTFTPYYVIINNAPYAIECQE----NDRPADHWLAVEPKSCSPLWPR  2388
[C.quinquefasciatus]    VHNTLTHNSLTKQIVFMPYFILINRADFDVEVQE----HLRPGDPWTRVKVEDCVPLWPK  2439
[D.melanogaster]        VHNHLTQNSLTKQITFIPFYIVCNKCHFDIELQE----QSRPADPWLHLEPNEMEPLWPR  2496
[L.cuprina]             VHNHLTQNSLTKQITFIPYYTITNKCSYVIEIQE----SSRPGDPWTKLLPNGCFPLWPK  2562
                        *   ::  .*:: :.: *:  : *..   :*  *       * : *  :  .   *:**.

[G.gallus]              NSSG--ELCVRVVGFESVSKPFLFQVPDNGTLLRLK-ELTGGLLVEVNVSQHSILISFSD  2965
[H.sapiens]             SLSG--KLCVRVVGCEGSSKPFFYNRQDNGTLLSLE-DLNGGILVDVNTAEHSTVITFSD  2969
[M.musculus]            NLSG--KLCVRVVGYEGSSKPFFYNRQDNGTLLSLE-DLNGGILVDINTAEHSTVITFSD  2964
[A.glabripennis]        SDMEDKLLRLRVEGTTENSAPFLYTESHN-TLLRLN-NKYGGVNVDIQITEGAIYINLVP  2446
[C.quinquefasciatus]    TEEN-RMLRVRAVNSTQITAPFKYTEVQC-TLLQLK-NRYGGINVDVHVTEGAIYITFTG  2496
[D.melanogaster]        NDTK-NNLVVRV--DGKITPAFDFTEVIC-TLLKLEDSKYGGINVDVQTTEGGVYITFTD  2552
[L.cuprina]             NDSD-HMMVARV--DGQRTVPFKYSEPIC-NLLQLNTNKNGGINVDVQTTEGGIYITFTE  2618
                        .      :  *.      :  * :      .** *: .  **: *::: :: .  *.:

[G.gallus]              YHEGAAPALIVNHTSWDSLRFKQSGLQEEMELKPKQVCLFAWTDPTKTRKLTWGYSQSF-  3024
[H.sapiens]             YHEGSAPALIMNHTPWDILTYKQSGSPEEMVLLPRQARLFAWADPTGTRKLTWTYAANV-  3028
[M.musculus]            YHEGSAPALIMNHTQWDVLTYKQSGSQEELVLLPGETRLFAWADPTGIRKLTWNYAANF-  3023
[A.glabripennis]        YEHGSAPALLINHTDSP-ITFWEKESVQKRILGSKHSLLYTWENPSGPRLIVWDKGNNK-  2504
[C.quinquefasciatus]    YHAGDAPALLINHTSEE-FAFKEKGDVNGKILMPSQMVLHTWINPAGDRKIVWDSGSKLV  2555
[D.melanogaster]        YKPADAPGLLINHTGKQ-IVYHEKGTKNEHILNAKSTIMYAWDDPTGPKMLVFGTN----  2607
[L.cuprina]             YNPGDAPGLLINHTRKP-IYYHEKGVDNKMVLQQRQQVLYTWSNPAGDRLLVFGEN----  2673
                        *. . **.*::***    : : :.   :   *      :.:* :*:  : :.:

[G.gallus]              -GEHDLLKDECGQFPYNANTQIHWVSFLDGRQRVLLFTDDVALVSKALQAEELEQPDQEI  3083
[H.sapiens]             -GEHDLLKDGCGQFPYDANIQIHWVSFLDGRQRVLLFTDDVALVSKALQAEEMEQADYEI  3087
[M.musculus]            -GEHDLLKDECGQFPYDANIQIHWVSFLDGRQRVLLFTDDVALVSKALQAEEMEQADHEV  3082
[A.glabripennis]        ELVEDLRKDACGEYSPTRDKKIYWASFLDGMQRVLLFTEVKMIAESAEASNLLEIIQQEI  2564
[C.quinquefasciatus]    SIENDLRRDDISEFKSPTDKDIFWVSFLNGTQRVLLVTDNANIAYEVHSSSRLDQVTQEI  2615
[D.melanogaster]        KEETDLKRDGIGEVIMQDGGKVLWVSFLDGLQRVLLFTENESIANRTESTASLQSITQSI  2667
[L.cuprina]             SLESDLRRDGIGNLTLEDNTKVYWVSFLDGLQRVILFTECEEIVTKSETSTALQSITQSI  2733
                            ** :*  .:     . .: *.***:* ***:*.*:   :.     :  ::    .:

[G.gallus]              ILSIHSLGLSLVNNENKKEISYIGITSSGVVWEEKR--KQKWRPFNQKQINLLEQAYQKY  3141
[H.sapiens]             TLSLHSLGLSLVNNESKQEVSYIGITSSGVVWEVKP--KQKWKPFSQKQIILLEQSYQKH  3145
[M.musculus]            ALSLHSLGLSLVNNENKQEVSYVGITSSGVVWEMKP--KQKWKPFSQKQIMSLEQAYSKR  3140
[A.glabripennis]        NVSIHGIGLSLINNLTRQEIMYLGIASSGIIWESCKMNGRRYKSLNQKLSMLLENAYQQY  2624
[C.quinquefasciatus]    KLEIHGVGLSLVNNSKQCDLMYIGIASSGVIWEERKR-SGRFKQMKIQETYSMESSFQQY  2674
[D.melanogaster]        DLRIHGIGLSVINNETGLDILYLGVTSSGIIWESKKVTKNRFKELTINENALLEIEYQKY  2727
[L.cuprina]             EVKIHGIGLSLINNETGVDILYLGITSSGVIWEYKKENKKRFKQLSLHDTALMEALYQEY  2793
                         : :*.:***::** .  :: *:*::***::**       ::: :. :    :*  :.:

[G.gallus]              LCKTAF---QTPGWHKLDSTTEVNFSKVPMEMRLPVRCSIRRNFLSGIQVEFKQSPHQRS  3198
[H.sapiens]             QI---S---RDHGWIKLDNNFEVNFDKDPMEMRLPIRSPIKRDFLSGIQIEFKQSSHQRS  3199
[M.musculus]            LA---S---QDRGWVKLDSNFEVNFDKVPMEMRLPIRCPIKRDFLSGIQVEFKQSPHQRS  3194
```

```
[A.glabripennis]        LLKQIAEPDEDNNIVTLDGKIVVD--FKSNLMLKPSKKKMRRTFETGLWLQMKSSPNQMH  2682
[C.quinquefasciatus]    LRDQEIGASAS-KKYFLEGEKRIEIDFTRNIVKKSTERDIRRTFYPGLWFEIKSSSHQLQ  2733
[D.melanogaster]        LVHKSVNDVQ---TYKLDNKFPIDFDLM--ILKKTVERNLRRSFYPAIWLSRKSSPFQSQ  2782
[L.cuprina]             LVDKGVHGVTD-KCYMLEGKHQINFDSL-KMVKNNVTRDIKRYFQPGVWIALNSSPFQSQ  2851
                                    *:.   ::       :       ::* *  .: .  :.*  *

[G.gallus]              LRAQLYWLQVDNQLPGSMFPVVFHPVAPPKSIALDSEPKPFIDISVITRFNEYSKVLQFK  3258
[H.sapiens]             LRARLYWLQVDNQLPGAMFPVVFHPVAPPKSIALDSEPKPFIDVSVITRFNEYSKVLQFK  3259
[M.musculus]            LRARLYWLQVDNQLPGTMFPVVFHPVAPPKSIALDSEPKPFIDVSVITRFNEYSKVLQFK  3254
[A.glabripennis]        IHAKINRLQIDNQIFDCIFPIVLAPVPPPKSVAVDSGIKPFVEVSIVQLLMKNSQIKQFK  2742
[C.quinquefasciatus]    FHAKINRIQIDNQLTDCIFPVVLAPIPPPKSVAATTEFKPFIEMSMVQRIIPHSNVKQFK  2793
[D.melanogaster]        LHVKINRIQVDNQFLDPIFPVVLAPIPPPKSVASTTSLKPFIECSMVQRIMPNSTVRQFK  2842
[L.cuprina]             IHAKINRIQLDNQLTDCIFPVVLAPIPPPKTLAKTIHFKPIIEVSIVERVVPHSNVKQYK  2911
                        ::.::  :*:***: . :**:*: *: ***::*     **::: *::  .   * : *:*

[G.gallus]              YFMVLIQEMALKIDQGFLAALSELFTPSTDPEAERQRSKLIQQDVDALNTELMET--SLT  3316
[H.sapiens]             YFMVLIQEMALKIDQGFLGAIIALFTPTTDPEAERRRTKLIQQDIDALNAELMET--SMT  3317
[M.musculus]            YFMVLIQEMALKVDQGFLGAVISLFTPTTDPEAERKRTKLIQQDIDALNTELMES--SMT  3312
[A.glabripennis]        YFKVLVQEFHIKVDLGFINAIVELLQESEQSDEE--EKKHFLTDMHLVDEGLYSHVSGES  2800
[C.quinquefasciatus]    YLRVLMQEFHVKVDLLFINEIFEMIS-SETTEAE--AKTLFAEDLKVQTQPLHAHVAIQS  2850
[D.melanogaster]        YARILIQEFLFKVDLNFLTAIAEMFA-KEVSDEA--AAKQFRQDVESIELPLSAFFEEHS  2899
[L.cuprina]             YAKMLMQEFHFKVDLVFLMAIAELFT-NTVDDER--EAKLFKDDVDTIVRPLSDLVQIQS  2968
                        *  :*:**: .*:*  *:  :  ::  .   :      . :  *:.     *       :

[G.gallus]              DVSMLSFFELFHISPIKLHLSLSLASGAEASDK-GEEMIAIHSLNLLLKSIGATLTDVDD  3375
[H.sapiens]             DMSILSFFEHFHISPVKLHLSLSLGSGGEESDKEKQEMFAVHSVNLLLKSIGATLTDVDD  3377
[M.musculus]            DMSILSFFEHFHISPVKLHLSLSLGSGGEESDKEKQEMIAIHSVNLLLKSIGATLTDVDD  3372
[A.glabripennis]        IREQKSFYDLLHFSPLKIHISFSMAAGSSPT---GQNVSTPNFLKIILQGIGVTLTDLQD  2857
[C.quinquefasciatus]    QAEVKNFYDNLHLGPLKIHVSFSMAGSES--------KALPGILSTILQGVGVTLTDIND  2902
[D.melanogaster]        LEEQKSFYDNLHLGPLKIHVSFSMAGSDT--------KALPGFLGSLVQGVGVTLTDVND  2951
[L.cuprina]             QQQQKNFYDNLHLGPMKVHVSFSMAGVDT--------SVLPGILSKLVQGVGVTLTDVND  3020
                          .  .*:: :*:.*:*:*:*:*:..                :  :::.:*.****::*

[G.gallus]              LIFKLAFFEIKYQFYKRDQLMKRAVRHYSEEFLKQMYVLVLGLDVLGNPFGLIRGLSEGV  3435
[H.sapiens]             LIFKLAYYEIRYQFYKRDQLIWSVVRHYSEQFLKQMYVLVLGLDVLGNPFGLIRGLSEGV  3437
[M.musculus]            LIFKLAYYEIRYQFYKRDQLMWSVVRHYSEQFLKQMYVLVLGLDVLGNPFGLIRGLSEGV  3432
[A.glabripennis]        VVFKLAFFERTFTFLTQKQLTSEVTSHYIGQSVKQLYVLVLGLDVIGNPYGLVLGITKGV  2917
[C.quinquefasciatus]    VVFRLAFFEREFQFLTQRQLVSECVTHYSGQAVKQLYVLVLGLDVIGNPYGLVVGFTKGV  2962
[D.melanogaster]        VVFRLAFFEREYQFFSQKQLINEITSHYTGQALKQLYVLVLGLDVLGNPYGLVVGLKKGV  3011
[L.cuprina]             VVIRLAFFEREYRFFTQEQLISEITSHYTGQALKQIYVLVLGLDVLGNPYGLVIGIKKGV  3080
                        ::::**::*  : * .: **    . **  : :**:*********:***:**: *:.:**

[G.gallus]              EAFFYEPFQGAVQGPEEFAEGIVIGVKSLLGHTVGGAAGVVSRITGSVGKGLAAITMDKE  3495
[H.sapiens]             EALFYEPFQGAVQGPEEFAEGLVIGVRSLFGHTVGGAAGVVSRITGSVGKGLAAITMDKE  3497
[M.musculus]            EALFYEPFQGAVQGPEEFAEGLVIGVRSLVGHTVGGAAGVVSRITGSVGKGLAAITMDKE  3492
[A.glabripennis]        EDLFYEPFQGAIQGPGEFAEGLALGVRSLFGHTVGGAAGAVSRITGAMGKGIAVLTFDEE  2977
[C.quinquefasciatus]    EDLFYEPFQGAIQGPGEFAEGLVLGVRSLFGHTVGGAAGAVSKITGAMGKGLAALTFDDD  3022
[D.melanogaster]        EDLFYEPFQGAIQGPGEFAEGLVLGVKSLFGHTVGGAAGAVSKITGAMGKGLAALTFDED  3071
[L.cuprina]             EDLFYEPFQGAIQGPGEFAEGLVLGVKSLVGHTVGGAAGAVSKITGAMGKGLAVLTFDEE  3140
                        * :********:*** *****:.:**:**.*********.**:***::***:*.:*:*.:

[G.gallus]              YQQKRREEMGRQPRDFSGSLAKGGKGFLRGVVGGVTGIITKPVEGAKKEGAAGFFKGIGK  3555
[H.sapiens]             YQQKRREELSRQPRDFGDSLARGGKGFLRGVVGGVTGIITKPVEGAKKEGAAGFFKGIGK  3557
[M.musculus]            YQQKRREEMGRQPKDFGDSLARGGKGFLRGVVGGVTGIITKPVEGAKKEGAAGFFKGIGK  3552
[A.glabripennis]        YQRKRRDQLNKKPATVTEGFARSGKGLVMGVVSGVTGVVTKPVSGAREQGVEGFFKGLGK  3037
[C.quinquefasciatus]    YQKKRRDALNKKPASLQEGIARSGKGLVMGVFDGVTGVFTKPISGAREDGVEGFFKGLGK  3082
[D.melanogaster]        YQKKRRQGIQNKPKNFHEGLARSSKGLVMGFVDGVTGVVTKPVTGARDNGVEGFFKGLGK  3131
[L.cuprina]             YQKKRRQNMNAKPKTFQEGMARSGKGLVMGFVDGVTGCVTKPISGAKEEGVEGFFKGLGK  3200
                        **:***: :  :*  .  .:*:..**:: *...**** .***: **:.:*. *****:**

[G.gallus]              GLVGVVARPTGGIVDMASSTFQGIQRVAESTEEVPNLRPPRFIHEDGIIRPYDRVEAEGY  3615
[H.sapiens]             GLVGAVARPTGGIVDMASSTFQGIQRAAESTEEVSSLRPPRLIHEDGIIRPYDRQESEGS  3617
[M.musculus]            GLVGAVARPTGGIIDMASSTFQGIQRVAESTEEVSSLRPPRLIHEDGIIRPYDRQESEGS  3612
[A.glabripennis]        GAVGLVARPVAGVVDFASGSLDAVKRCAESGEEICKLRPARFIPADGLVRPFNSKEAMGH  3097
[C.quinquefasciatus]    GAVGLVARPIAGVTDFASGSFDAVKRATELSDEAFRLRSPRFLHKDGIVRPYNRKEAEGN  3142
```

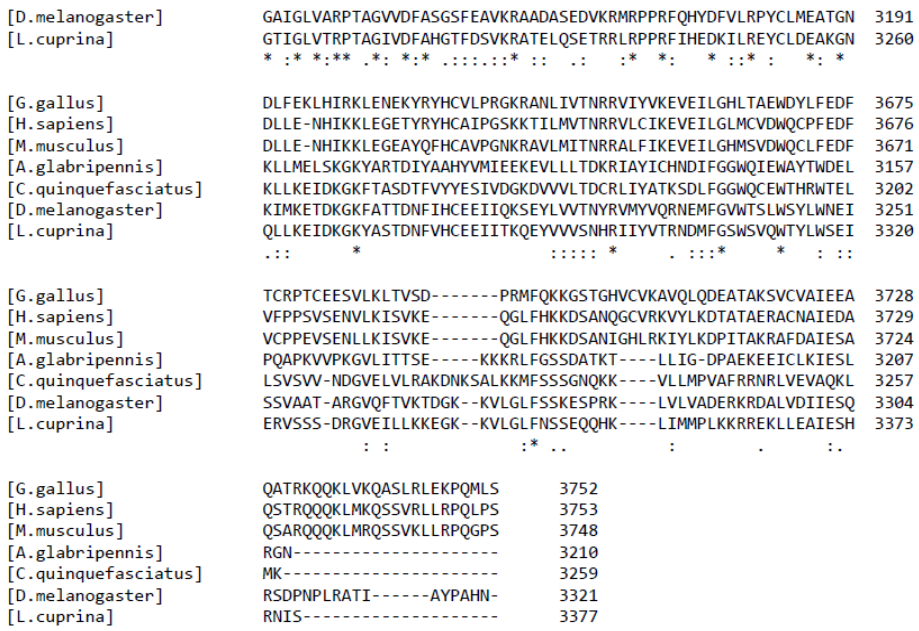


Figure 2: The Vps13 protein is highly conserved between vertebrates and invertebrates

Clustal omega multiple alignment of Vps13 protein. High lighted are N-terminal region of Chorein (Chorein N) (Green), Vacuolar-sorting associated protein 13 (VPS13) (Red), Repeating coiled region of VPS13 (VPS13 mid rpt) (Blue), SHR-binding domain (SHR-BD)(Yellow), Vacuolar sorting-associated 13 protein C-terminal (VPS13C)(Purple) and Autophagy- related protein C-terminal domain (ATG C)(Orange). "*" indicates identical amino acids in all sequences in the alignment, ":" shows conserved substitutions, and "." indicates semi-conserved substitutions. The numbers on the left side of each line represents the last amino acid number in the sequence in that row. The names of each species in the alignment are shown in brackets on the right side of the each line. Domains were obtained from Pfam.

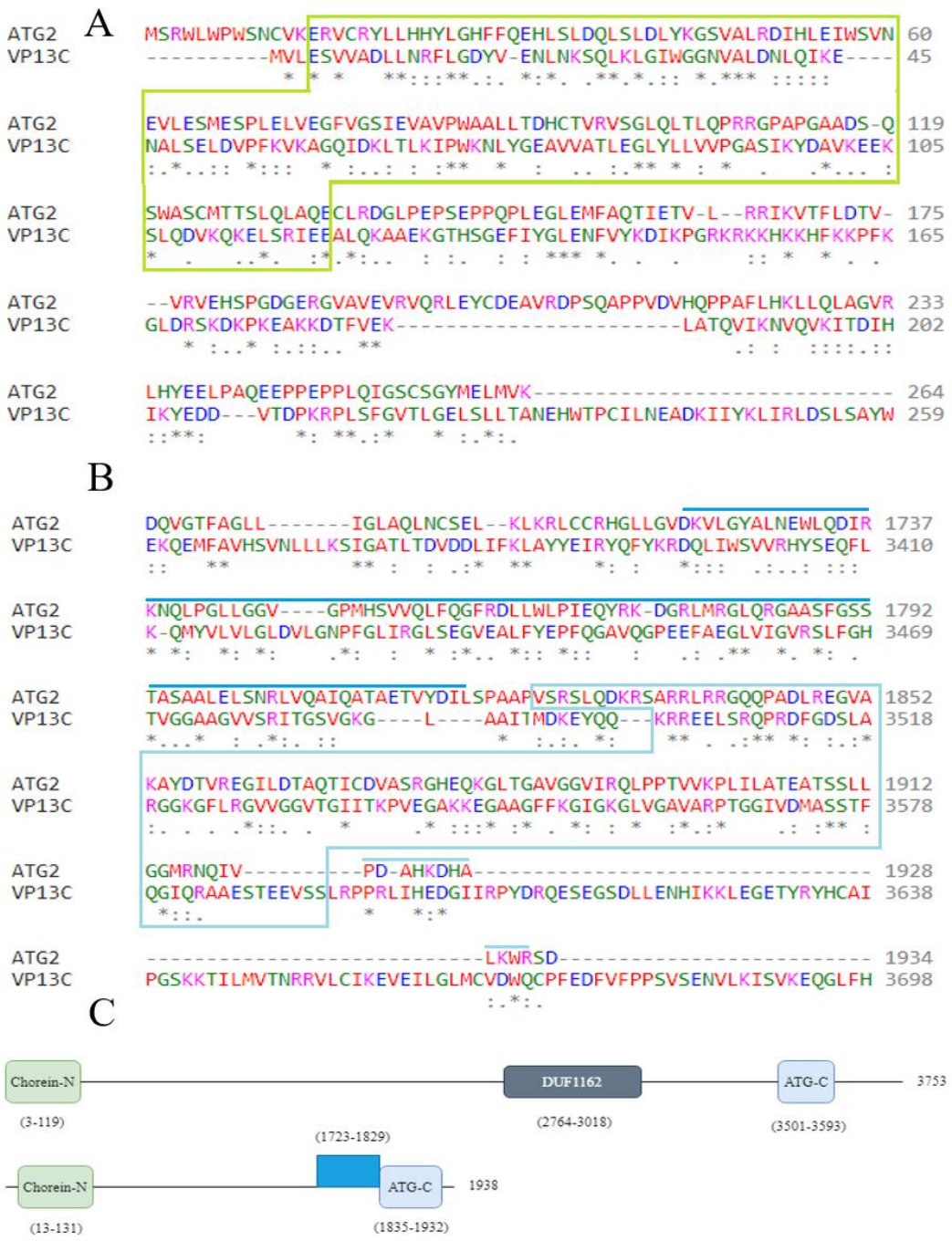


Figure 3. Chorein-N and ATG-C domain are similar in ATG2A and VPS13C proteins. Image A includes the Chorein-N domain (Green box), Image B includes the crucial sequence (Dark blue line) in the upstream of ATG-C domain (Light blue box/line).Boxes represent shared domains and lines represent domains specific to ATG2A. Image C is a schematic illustration of represented domains. "*" indicates identical amino acids in all sequences in the alignment, ":" shows conserved substitutions, and "." indicates semi-conserved substitutions. The numbers on the left side of each line represents the last amino acid number in the sequence in that row. The names of the proteins used for the alignment are shown on the right side of the each line. Multiple sequence alignment was generated with Clustal Omega. Domains were driven from Pfam. Image C was created with draw.io application.

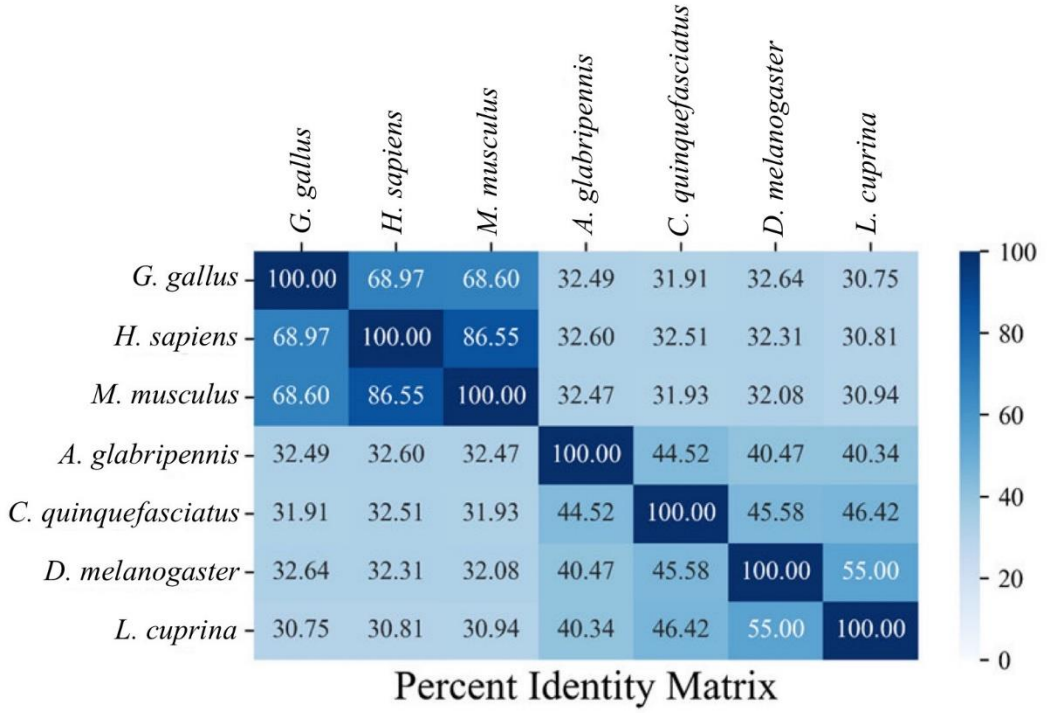


Figure 4. Percent Identity Matrix shows high level of preservation on Vps13 Protein among vertebrates and invertebrates.

The number shown in each cell represents the identity percentage of Vps13 protein comparing two species from the alignment. The intensity of the blue colour in each cell visualizes the amount of sameness in a way that the higher the identity percentage between two species the more intense the blue colour is.

Figure 5. Phylogenetic tree shows a high level of evolutionary distance between *H. sapiens* VPS13C and *D. melanogaster* Vps13.

The numbers shown on the right side of the cladogram indicate the distance values as the lengths of branches represent the genetic distance and the approximate time since the genetic divergence in the alignment. The phylogenetic tree and the numbers are produced as the output of the multiple sequence alignment which is used as an indication of the evolutionary distance between the sequences.

Eye Analysis

Effect of inhibition and overexpression of *Vps13* upon *D. melanogaster* ommatidia and bristle numbers

The *D. melanogaster* compound eye consists of about 750 to 800 spherical-shape structures that form hexagonal arrays called ommatidia and hair-like structures called bristles that together create eye units. These units are arranged in a highly consistent pattern. The viability of *D. melanogaster* is not dependent on the presence or the functionality of the eye, and genetic changes during the neurodevelopment could be easily detected and therefore analyzed. One can use the eye as an ideal model for understanding the genetics of neurodegeneration and neurodevelopment. The phenotypic alterations in the eye are mostly in shape, size, or arrangement of the units (Mishra M., 2012). In this study, I used *GMR-GAL4* to control the inhibition and overexpression of *Vps13* in the developing eye (Figure:6). Then, I counted the number of ommatidia and bristles of each compound eye using ImageJ software. Through observation, it became clear that in specific genotypes, there were significant areas that lacked bristles. To quantify this, I compared the area of the eye voided bristles to the total eye area and analyzed them through GraphPad Prism.

Analysis of Ommatidia Number

Ommatidia numbers were significantly decreased in the inhibited critical class *GMR-Gal4; UAS-Vps13-RNAi*^{HMS01715} with the average ommatidia number of 679.8 ±17.92 and the inhibited class *GMR-Gal4; UAS-Vps13-RNAi*^{HMS024460} with the average ommatidia number of 734.6 ±8.806 compared to the control group *GMR-Gal4; UAS-lacZ* with the average number of 758. While the inhibition group *GMR-Gal4; UAS-*

Vps13-RNAi^{GD14789}, with the mean ommatidia number of 746.4 ±9.667, and overexpression *GMR-Gal4; UAS-Vps13*^{EY09640}, with the average ommatidia number of 768.2 ±9.253, did not show any significant changes compared to the control. N-value for all critical classes were 10 and $p< 0.05$ considered significant (Figure 7, Table 2).

Analysis of bristle number

Bristle numbers in the inhibited group *GMR-Gal4; UAS-Vps13-RNAi*^{HMS01715} with a mean of 569.7 ±14.87 were significantly lower than the control group *GMR-Gal4; UAS-Vps13-lacZ* while bristle numbers in the inhibited group *GMR-Gal4; UAS-Vps13-RNAi*^{IHMS024460} flies with 644.6 ±20.12, and inhibited group *GMR-Gal4; UAS-Vps13-RNAi*^{GD14789} with a mean of 634.3 ±6.340, were significantly higher than the control *GMR-Gal4; UAS-lacZ* . Overexpression group *GMR-Gal4;UAS-Vps13*^{EY09640}, with an average bristle number of 510.2 ±6.740, showed a significant decrease in bristle numbers compared to *GMR-Gal4; UAS-lacZ,* N-value for all critical classes were 10 and $p<0.05$ considered significant (Figure 8, Table 2).

Analysis of Bristle Void Area

The percent of bristle void area was significantly decreased in all of the inhibited flies *GMR-Gal4;UAS-Vps13-RNAi*^{HMS01715}, *GMR-Gal4;UAS-Vps13-RNAi*^{HMS024460} and, *GMR-Gal4; UAS-Vps13-RNAi*^{GD14789} and significantly increased in the overexpression group *GMR-Gal4; UAS-Vps13*^{EY09640} compared to the control. The percent of bristle void area in all critical class groups is as follows: The control line, *GMR-Gal4; UAS-lacZ* ,14.10%, inhibition group *GMR-Gal4; UAS-Vps13-RNAi*^{HMS01715}, 8.549 ±0.5537%, *GMR-Gal4;UAS-Vps13-RNAi*^{HMS024460}, 8.607 ±0.3821% and, *GMR-Gal4;UAS-Vps13-*

RNAi^{GD14789}, 8.252 ±0.5113% and, overexpression line *GMR-Gal4;UAS-Vps13*^{EY09640}, 33.33 ±1.924%. N value for all lines was 10, and $p < 0.05$ considered significant. (Figure 9, Table 2)

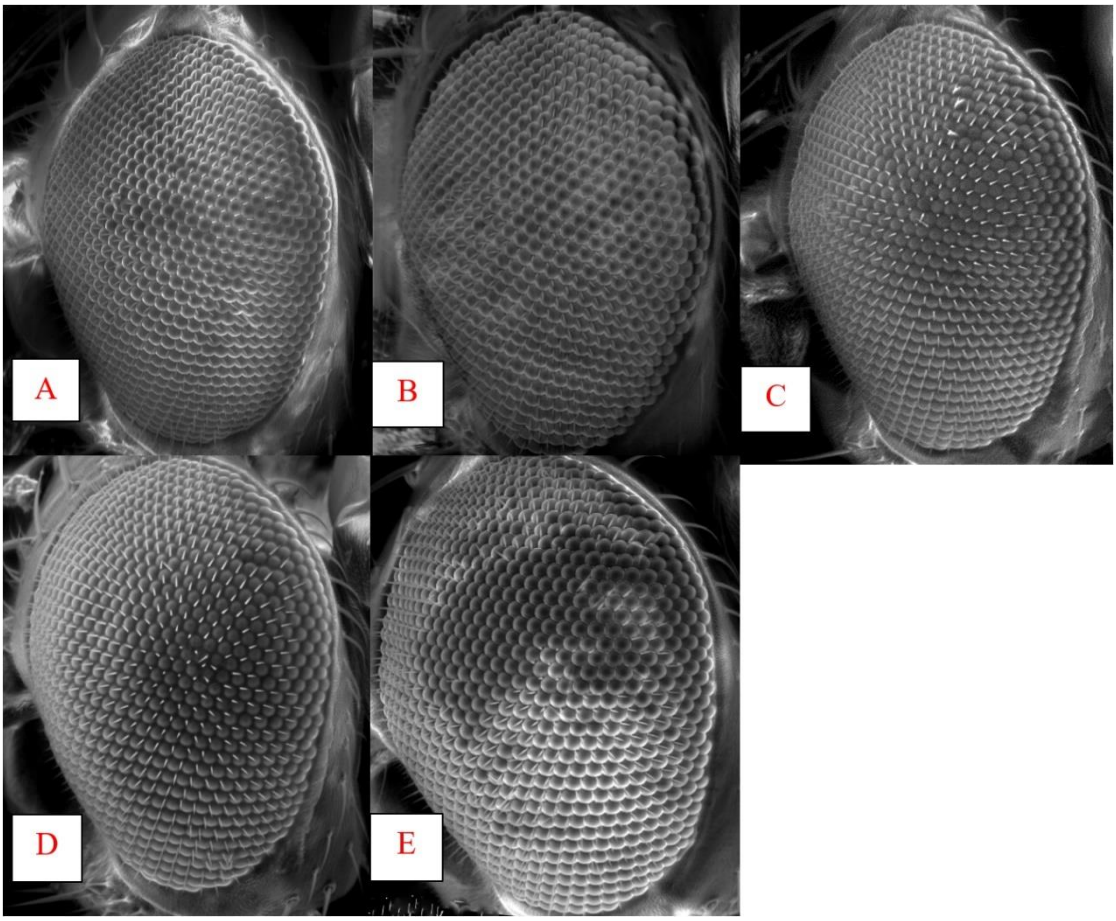


Figure 6. Bristle numbers are reduced by the overexpression of *Vps13* under control of eye-specific drivers in *D. melanogaster* compound eye.

Images were taken in 450X magnification using the FEI Quanta 400 Scanning Electron Microscope. Each image relates to one of the experimented lines as follows: A-*GMR-Gal4;UAS-lacZ*, B- *GMR-Gal4;UAS-Vps13-RNAi*^{HMS01715}, C- *GMR-Gal4;UAS-Vps13-RNAi*^{HMS024460}, D- *GMR-Gal4;UAS-Vps13-RNAi*^{GD14789}, *and* E-*GMR-Gal4;UAS-Vps13*^{EY09640}

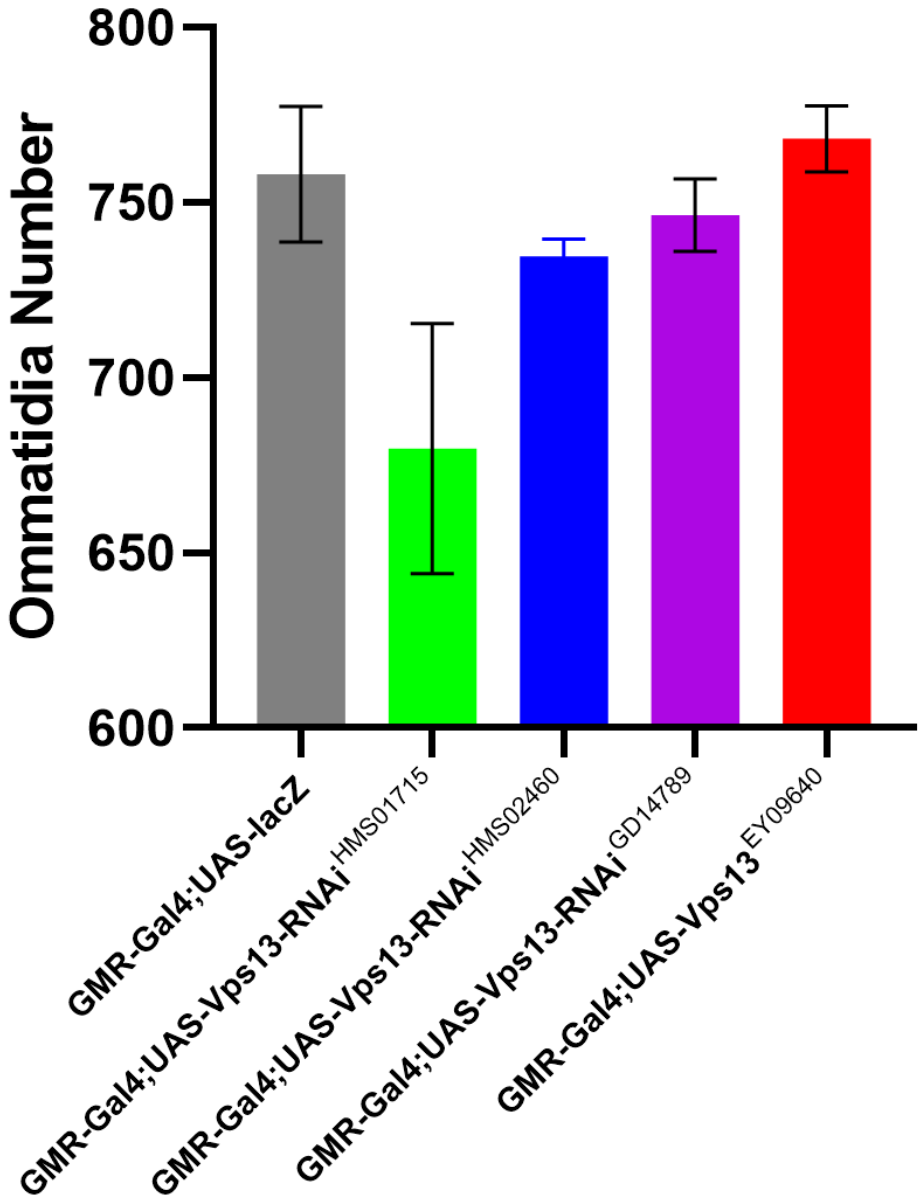


Figure 7*:* Inhibition and overexpression of *Vps13* under the influence of eye specific expression in most cases does not affect ommatidia numbers in *D. melanogaster* compound eye.

N-value is 10, $p<0.05$ considered significant. Inhibition groups that are shown with green and blue colour (*GMR-Gal4;UAS-Vps13-RNAi^{HMS01715} and GMR-Gal4;UAS-Vps13-RNAi^{HMS024460}*) showed significant decrease in their ommatidia number while inhibition group with purple colour (*GMR-Gal4;UAS-Vps13-RNAi^{GD14789}*) and the overexpression group with red colour (*GMR-Gal4;UAS-Vps13^{EY09640}*) had no significant differences with control group.

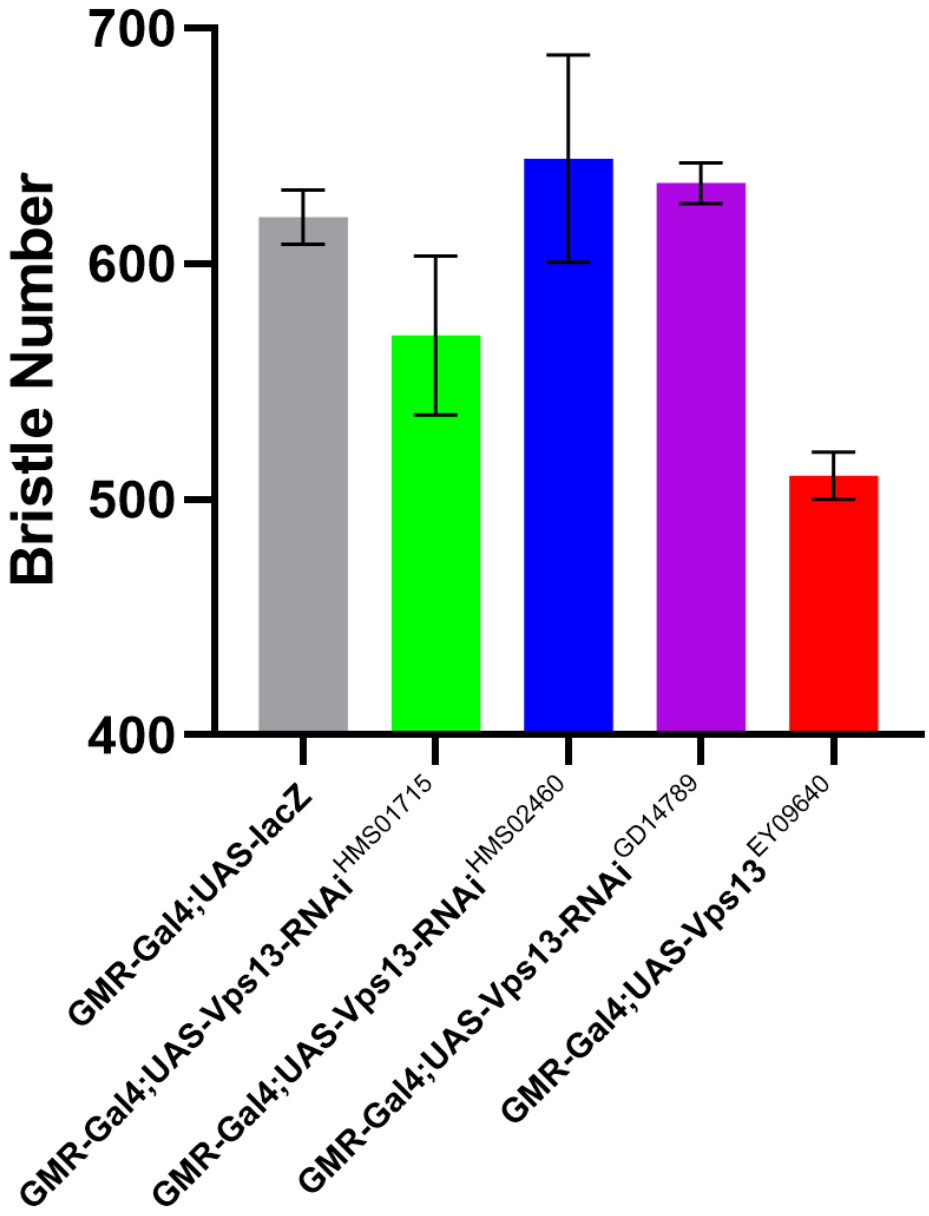


Figure 8. Overexpression of *Vps13* under the influence of eye-specific expression reduces the bristle numbers while inhibition of *Vps13* has the opposite effect on *D. melanogaster* compound eye

N-value is 10, $p<0.05$ considered significant. Overexpression group *GMR-Gal4;UAS-Vps13*^{EY09640} shows specific decrease in their bristle numbers while inhibition groups *GMR-Gal4;UAS-Vps13-RNAi*^{GD14789} *and GMR-Gal4;UAS-Vps13-RNAi*^{HMS02460} show increase in their bristle numbers compared to control.

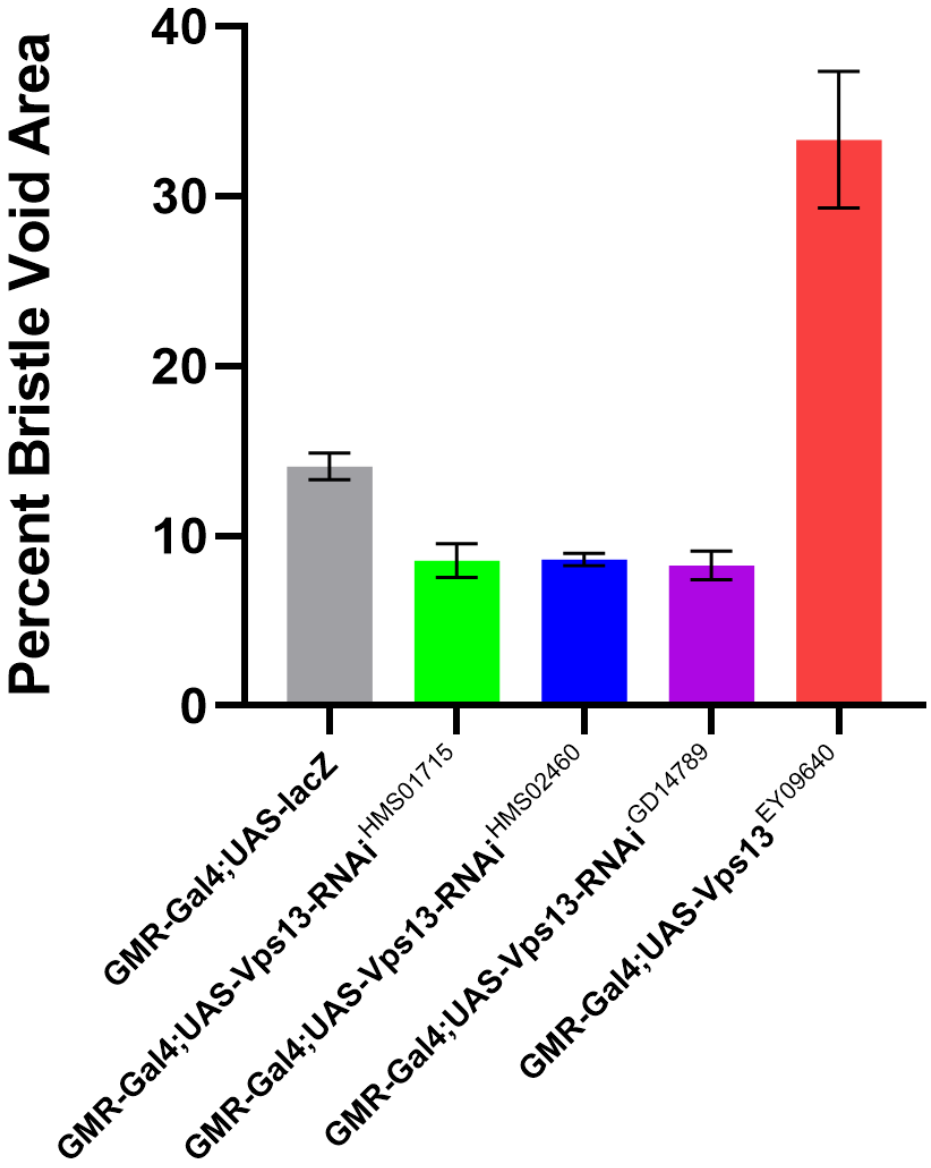


Figure 9. Overexpression of *Vps13* under the influence of eye specific expression significantly decreases the bristle-void area in *D. melanogaster* compound eye.

N-value is 10, $p<0.05$ considered significant. Overexpression of the *Vps13* in the eye (*GMR-Gal4;UAS-Vps13*^{EY09640}) significantly increases the bristle void area while inhibition of *Vps13 (GMR-Gal4;UAS-Vps13-RNAi*^{HMS01715}, *GMR-Gal4;UAS-Vps13-RNAi*^{GD14789} *and GMR-Gal4;UAS-Vps13-RNAi*^{HMS02460}) improved the bristle numbers in the eye compared to the control.

Table 2. Summary of unpaired t-test results for ommatidia number, bristle number and the percent bristle void area for inhibition and overexpression of *Vps13* in *D. melanogaster* eye.

Genotype	Sample Size	Mean ± SEM	P-value	Significant compared to control
Ommatidia				
GMR-Gal4;UAS-Vps13-lacZ	10	758.1	N/A	N/A
GMR-Gal4;UAS-Vps13-RNAi^{HMS01715}	10	679.8 ±17.92	0.0004	Yes
GMR-Gal4;UAS-Vps13-RNAi^{HMS02460}	10	734.6 ±8.806	0.0157	Yes
GMR-Gal4;UAS-Vps13-RNAi^{GD14789}	10	746.4 ±9.667	0.2418	No
GMR-Gal4;UAS-Vps13 ^{EY09640}	10	768.2 ±9.253	0.2895	No
Bristle				
GMR-Gal4;UAS-lacZ	10	619.8	N/A	N/A
GMR-Gal4;UAS-Vps13-RNAi^{HMS01715}	10	569.7 ±14.87	0.0036	Yes
GMR-Gal4;UAS-Vps13-RNAi^{HMS02460}	10	644.6 ±20.12	0.0004	Yes
GMR-Gal4;UAS-Vps13-RNAi^{GD14789}	10	634.3 ±6.340	0.0345	Yes
GMR-Gal4;UAS-Vps13^{EY09640}	10	510.2 ±6.740	<0.0001	Yes
Percent Bristle Void Area				
GMR-Gal4;UAS-lacZ	10	14.10 %	N/A	N/A
GMR-Gal4;UAS-Vps13-RNAi^{HMS01715}	10	8.549 ±0.5537 %	<0.0001	Yes
GMR-Gal4;UAS-Vps13-RNAi^{HMS02460}	10	8.607 ±0.3821%	<0.0001	Yes
GMR-Gal4;UAS-Vps13-RNAi^{GD14789}	10	8.252 ±0.5113%	<0.0001	Yes
GMR-Gal4;UAS-Vps13^{EY09640}	10	33.33 ±1.924%	<0.0001	Yes

Behavioural Analysis

Directed dopaminergic neuron-specific expression

Longevity Assay

I compared the longevity of each of the critical classes with the control group *Th-Gal4; UAS-lacZ* using Log-rank statistical analysis. P<0.05 considered significant (Figure 10, Table 3). The control group *Th-Gal4; UAS-lacZ* has the N-value of 320 and the median life span of 64 days. Inhibition group *Th-Gal4; UAS-Vps13-RNAi*^{HMS01715} with the N-value of 302 and median life span 58 days, inhibition group *Th-Gal4; UAS-Vps13-RNAi*^{HMS024460} with the N-value of 271 and median life span of 58, and overexpression line *Th-Gal4; UAS-Vps13*^{EY09640} with the N-value of 308 and median life span of 46 days showed a specific decrease in their longevity. It is noteworthy that the reduction in the survivorship ability of the overexpression line *Th-Gal4; UAS-Vps13*^{EY09640} was much more significant than the other mentioned inhibition groups. Surprisingly, inhibition group *Th-Gal4; UAS-Vps13-RNAi*^{GD14789}, with the N-value of 318 and median survival days of 72 days, showed a significant increase in its longevity.

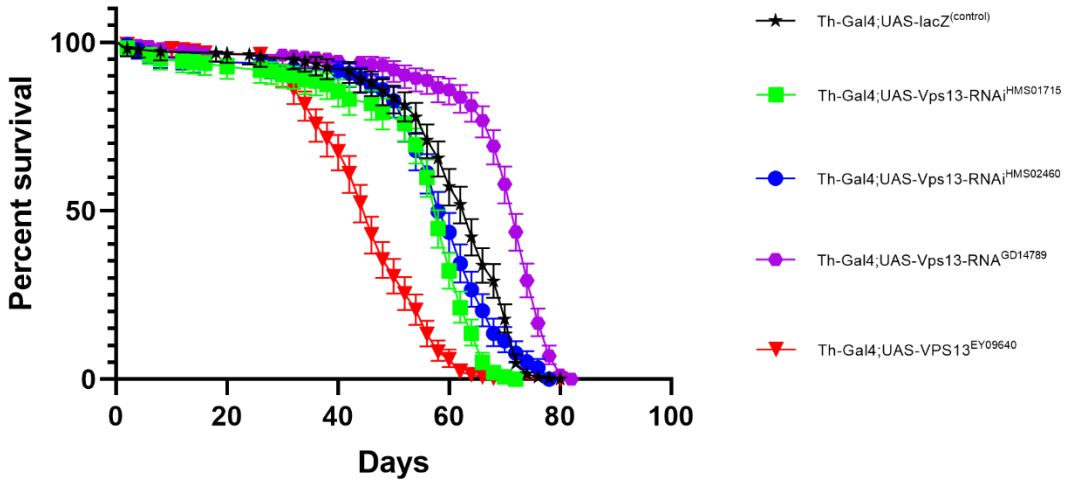


Figure 10. Overexpression of *Vps13* under the control of directed dopaminergic neuron- specific expression decreases the longevity of *D. melanogaster*.

P<0.05 considered significant. The error bars are made with a CI 95%. Inhibition group *Th-Gal4; UAS-Vps13-RNAi*^{GD14789} has the highest longevity while the overexpression group (*Th-Gal4; UAS-Vps13*^{EY09640}) shown with red colour showed the lowest longevity

Table 3. Log-rank statistical analysis of longevity of flies with directed dopaminergic neuron-specific expression with inhibition and overexpression of *Vps13*.

Survivorship Analysis					
Genotype	Number of Flies	Median Survival (Days)	Chi-square Value	P-value	Significance
Th-Gal4;UAS-lacZ	320	64	N/A	N/A	N/A
Th-Gal4;UAS-Vps13-RNAi^{HMS01715}	302	58	89.38	<0.0001	Yes
Th-Gal4;UAS-Vps13-RNAi^{HMS02460}	271	58	7.479	0.006243	Yes
Th-Gal4;UAS-Vps13-RNAi^{GD14789}	318	72	173.4	<0.000001	Yes
Th-Gal4;UAS-Vps13^{EY09640}	308	46	299.7	<0.000001	Yes

Climbing Assay

I compared the locomotor ability of each critical class with the control group using non-linear regression analysis. N-value for every critical class group is 50, $P<0.05$ considered significant (Figure 11, Table 4). All of the critical classes started with the same quality of climbing. Meanwhile, inhibition group *Th-Gal4; UAS-Vps13-RNAi*^{GD14789} had an improved climbing ability over time compared to control *Th-Gal4; UAS-lacZ* throughout the whole experiment. However, although inhibition group *Th-Gal4; UAS-Vps13-RNAi*^{HMS01715} had better climbing scores until day 35 (week five), its scores decreased significantly thereafter. Inhibition group *Th-Gal4; UAS-Vps13-RNAi*^{HMS024460} showed no statistically difference with control *Th-Gal4; UAS-lacZ*, and Overexpression group *Th-Gal4; UAS-Vps13*^{EY09640} which notably showed reduced longevity, had a significant decrease in its locomotor ability.

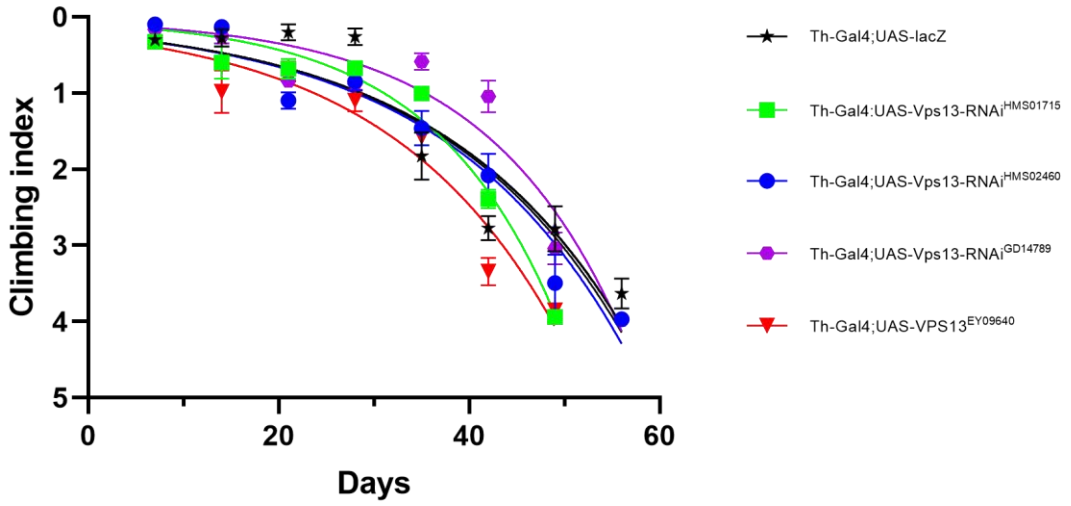


Figure 11. Overexpression of *Vps13* under the control of directed dopaminergic neuron- specific expression decreases the climbing ability of *D. melanogaster* while inhibition of *Vps13* increases it.

p<0.05 considered significant. The error bars are made a with CI 95%.

Table 4. Statistical analysis of the locomotor ability of *Drosophila melanogaster* using non-linear regression curve with directed dopaminergic neuron-specific expression with inhibition and overexpression of *Vps13.*

Climbing Analysis					
Genotype	Slope (k)	SE	95% Confidence interval (CI)	P-value	Significance
Th-Gal4;UAS-lacZ	0.05122	0.0056	0.04140 to 0.06240	N/A	N/A
Th-Gal4;UAS-Vps13-RNAi^{HMS01715}	0.07585	0.0048	0.06601 to 0.08712	0.0001	Yes
Th-Gal4;UAS-Vps13-RNAi^{HMS02460}	0.05214	0.0039	0.04486 to 0.06019	0.7957	No
Th-Gal4;UAS-Vps13-RNA^{GD14789}	0.06868	0.0058	0.05712 to 0.08214	0.0434	Yes
Th-Gal4;UAS-Vps13^{EY09640}	0.05545	0.0044	0.04701 to 0.06498	<0.0001	Yes

Directed motorneuron-specific expression

Longevity Assay

I compared the longevity of each of the critical classes with the control group using Log-rank statistical analysis. P<0.05 considered significant (Figure 12, Table 5). The control group *D42-Gal4; UAS-lacZ* has the N-value of 280 and the median life span of 57 days. Inhibition group *D42-Gal4; UAS-Vps13-RNAi*^{HMS01715} with the N-value of 328 and median life span 40 days had a significant decrease in its longevity. While inhibition group *D42-Gal4; UAS-Vps13-RNAi*^{HMS024460} with the N-value of 276 and a median life span of 58 showed no significant differences with control *D42-Gal4; UAS-lacZ*. On the other hand, inhibition group *D42-Gal4; UAS-Vps13-RNAi*^{GD14789} showed a significant increase compared to control *D42-Gal4; UAS-lacZ*. No observation of substantial differences in inhibition group *D42-Gal4; UAS-Vps13-RNAi*^{HMS024460} and the improvement of the longevity in the inhibition group *D42-Gal4; UAS-Vps13-RNAi*^{GD14789} could also be seen in the directed dopaminergic neuron-specific expression groups. However, unlike the directed dopaminergic neuron-specific overexpression group *D42-Gal4; UAS-Vps13*^{EY09640} with the N-value of 315 and a median life span of 62 days showed a significant increase in their longevity

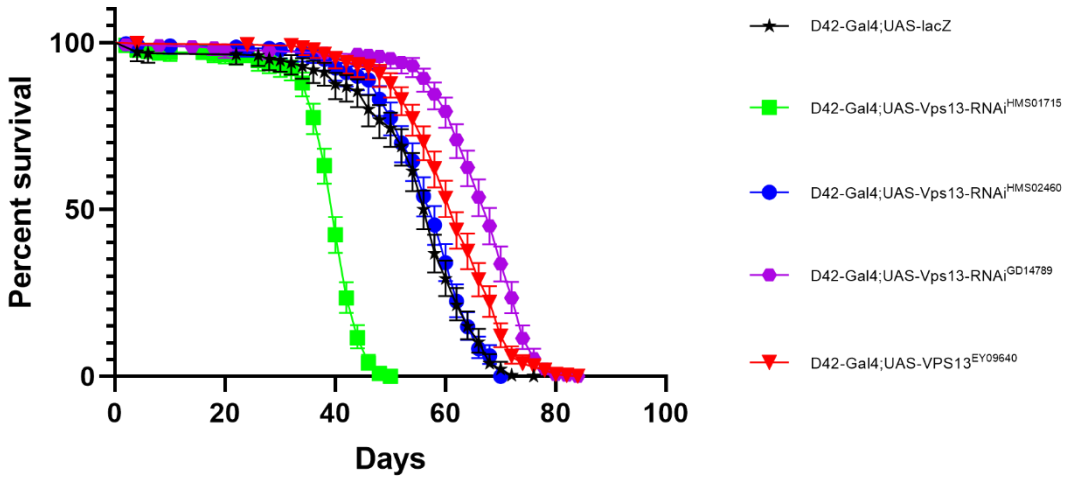


Figure 12. Inhibition and overexpression of *Vps13* under the control of directed motorneuron-specific expression increases *D. melanogaster* longevity while decreases longevity of *D42-Gal4; UAS-Vps13-RNAi^{HMS01715}*

P<0.05 considered significant. The error bars are made with a CI 95%.

Table 5. Log-rank statistical analysis of longevity of flies with directed motorneuron-specific expression with inhibition and overexpression of *Vps13.*

Survivorship Analysis					
Genotype	Number of Flies	Median Survival (Days)	Chi-square Value	P-value	Significance
D42-Gal4;UAS-lacZ	280	57	N/A	N/A	N/A
D42-Gal4;UAS-Vps13-RNAi^{HMS01715}	328	40	418.0	<0.0001	Yes
D42-Gal4;UAS-Vps13-RNAi^{HMS02460}	276	58	0.5337	0.4650	No
D42-Gal4;UAS-Vps13-RNAi^{GD14789}	315	68	227.9	<0.0001	Yes
D42-Gal4; UAS-Vps13^{EY09640}	315	62	60.21	<0.0001	Yes

Climbing Assay

I compared the locomotor ability of each critical class group with the control group *D42-Gal4; UAS-lacZ* using non-linear regression analysis. N-value for every critical class is 50, P<0.05 considered significant (Figure 13, Table 6). Overexpression line *D42-Gal4; UAS-Vps13*^{EY09640} started with lower climbing scores in the first climbing attempt compared to other critical class groups. Overexpression line *D42-Gal4; UAS-Vps13*^{EY09640} and inhibition line *D42-Gal4; UAS-Vps13-RNAi*^{GD14789} showed the same pattern as was observed in the directed dopaminergic neuron-specific experiment. A decrease is evident in the overexpression line *D42-Gal4; UAS-Vps13*^{EY09640} , and the increase in inhibition line *D42-Gal4; UAS-Vps13-RNAi*^{GD14789} climbing abilities compared to the control *D42-Gal4; UAS-lacZ*. Inhibition group *D42-Gal4;UAS-Vps13-RNAi*^{HMS01715} and *D42-Gal4;UAS-Vps13-RNAi*^{HMS024460} showed no statistically meaningful differences compared to the control.

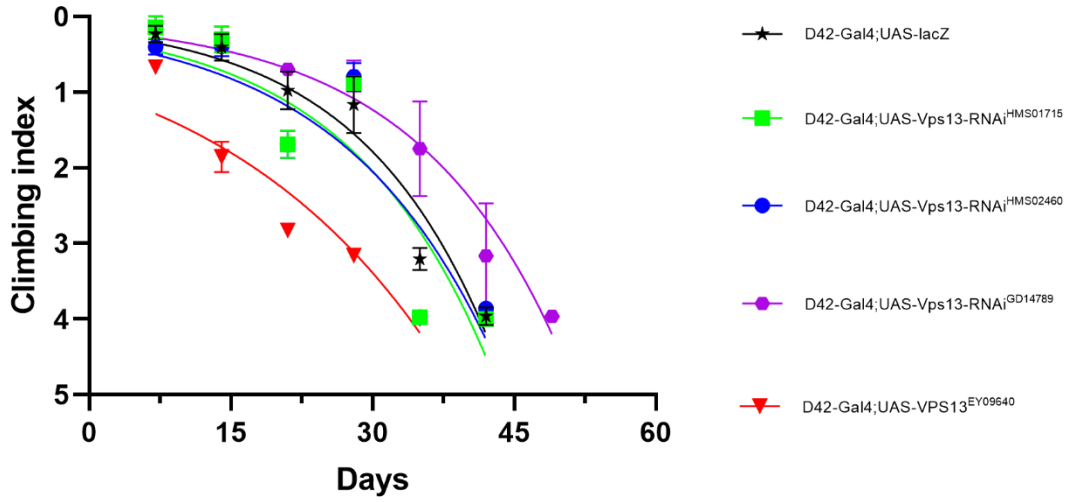


Figure 13. Overexpression of *Vps13* under the control of directed motorneuron-specific expression decreases the climbing ability of *D. melanogaster* while inhibition of *Vps13* increases it.

N-value for every critical class group is 50 and $p<0.05$ considered significant **.**The error bars are made with a CI 95%.

Table 6. Statistical analysis of the locomotor ability of *D. melanogaster* using non-linear regression curve with directed motorneuron-specific expression with inhibition and overexpression of *Vps13*

Climbing Analysis					
Genotype	Slope (k)	SE	95% Confidence interval (CI)	P-value	Significance
D42-Gal4;UAS-lacZ	0.07113	0.005058	0.06187 to 0.08149	N/A	N/A
D42-Gal4;UAS-Vps13-RNAi^{HMS01715}	0.06579	0.008606	0.05081 to 0.08362	0.3449	No
D42-Gal4;UAS-Vps13-RNAi^{HMS02460}	0.06103	0.008154	0.04657 to 0.07835	0.3818	No
D42-Gal4;UAS-Vps13-RNAi^{GD14789}	0.06450	0.004266	0.05674 to 0.07314	<0.0001	Yes
D42-Gal4; UAS-Vps13^{EY09640}	0.04212	0.003587	0.03529 to 0.04937	<0.0001	Yes

Directed neuron-specific expression

Longevity Assay

I compared the longevity of each of the critical class groups with the control group *Ddc-Gal4; UAS-lacZ* using Log-rank statistical analysis. P<0.05 considered significant (Figure 14, Table 7). The control group *Ddc-Gal4; UAS-lacZ* with the N-value of 314 and the median life span of 64 days has the highest longevity compared to experimental groups. Inhibition group *Ddc-Gal4; UAS-Vps13-RNAi*^{HMS01715} with the N-value of 311 and median life span 56 days, inhibition group *Ddc-Gal4; UAS-Vps13-RNAi*^{HMS024460} with the N-value of 331 and a median survival of 54, inhibition group *Ddc-Gal4; UAS-Vps13-RNAi*^{GD14789} with the N-value of 317 and median survival of 62 days, and overexpression group *Ddc-Gal4; UAS-Vps13*^{EY09640} with the N-value of 320 and median survival of 38 days showed a significant decrease compared to control. It is noteworthy that the overexpression group *Ddc-Gal4; UAS-Vps13*^{EY09640} had the lowest longevity, and inhibition group *Ddc-Gal4; UAS-Vps13-RNAi*^{GD14789} has the closest longevity to the control.

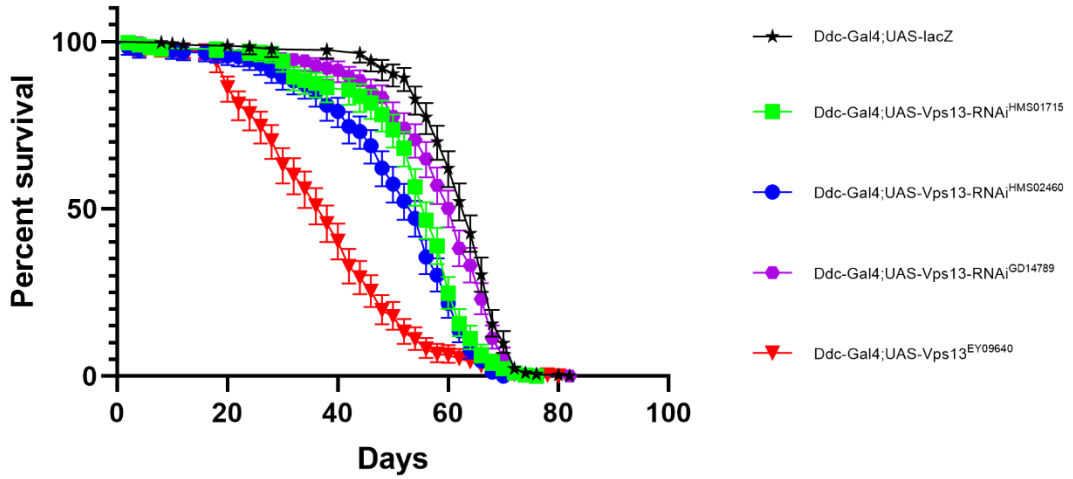


Figure 14. Inhibition and overexpression of *the Vps13* under the control of directed neuron-specific expression decreases *D. melanogaster* longevity.

P<0.05 onsidered significant. The error bars are made with a CI 95%.

Table 7. Log-rank statistical analysis of longevity of flies with directed neuron-specific expression with inhibition and overexpression of *Vps13.*

Survivorship Analysis					
Genotype	Number of Flies	Median Survival (Days)	Chi-square Value	P-value	Significance
Ddc-Gal4;UAS-lacZ	314	64	N/A	N/A	N/A
Ddc-Gal4;UAS-Vps13-RNAi^{HMS01715}	311	56	100.2	<0.0001	Yes
Ddc-Gal4;UAS-Vps13-RNAi^{HMS02460}	331	54	170.0	<0.0001	Yes
Ddc-Gal4;UAS-Vps13-RNAi^{GD14789}	317	62	11.37	0.0007	Yes
Ddc-Gal4;UAS-Vps13^{EY09640}	320	38	349.7	<0.0001	Yes

Climbing Assay

I compared the locomotor ability of each critical class group with the control group *Ddc-Gal4; UAS-lacZ* using non-linear regression analysis. N-value for every critical class group is 50, and $P<0.05$ considered significant (Figure 15, Table 8). Inhibition group *Ddc-Gal4; UAS-Vps13-RNAi*^{GD14789} started with higher climbing scores in the first climbing attempt compared to other critical class groups, and its climbing scores remained significantly higher than the control throughout the experiment. The rest of the experimental groups showed no statistically meaningful change in their locomotor abilities compared to control. Yet, overexpression group *Ddc-Gal4; UAS-Vps13*^{EY09640} climbing scores are slightly lower than the control and inhibition group *Ddc-Gal4; UAS-Vps13-RNAi*^{HMS01715} and *Ddc-Gal4; UAS-Vps13-RNAi*^{HMS024460} showed no significant difference in their locomotor ability.

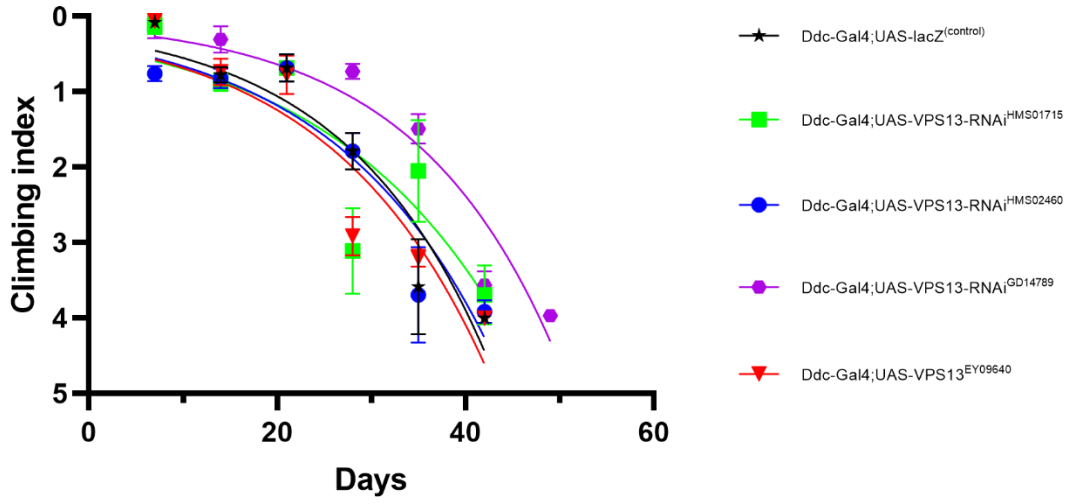


Figure 15. Inhibition of *Vps13* under the control of directed neuron specific expression increases *D. melanogaster* climbing ability.

N-value for every critical class groups in climbing assay is 50 and $p<0.05$ considered significant. The error bars as made with CI 95%.

Table 8. Statistical analysis of the locomotor ability of *D. melanogaster* using non-linear regression curve with directed neuron-specific expression with inhibition and overexpression of *Vps13*

Climbing Analysis					
Genotype	Slope (k)	SE	95% Confidence interval (CI)	P-value	Significance
Ddc-Gal4;UAS-lacZ	0.06513	0.005815	0.05479 to 0.07675	N/A	N/A
Ddc-Gal4;UAS-Vps13-RNAi^{HMS01715}	0.05240	0.008029	0.03816 to 0.06955	0.1551	No
Ddc-Gal4;UAS-Vps13-RNAi^{HMS02460}	0.05824	0.005251	0.04833 to 0.06934	0.6462	No
Ddc-Gal4;UAS-Vps13-RNAi^{GD14789}	0.06594	0.004680	0.05755 to 0.07533	<0.0001	Yes
Ddc-Gal4;UAS-Vps13^{EY09640}	0.05956	0.006719	0.04772 to 0.07271	0.4484	No

Directed neuron-specific expression with inhibition of *Parkin*

Longevity Assay

I compared the longevity of each of the lines with the control group using Log-rank statistical analysis. P<0.05 considered significant (Figure 16, Table 9). In this experiment, I determined the effect of neuron-specific inhibition and overexpression of *Vps13* along with the inhibition of *Parkin*. The control line *Ddc-Gal4;UAS-parkin-RNAi; UAS-lacZ* showed an N-value of 328 and median survival of 60 days. This number was 64 days in control group *Ddc-Gal4; UAS-lacZ*. Inhibition lines *Ddc-Gal4;UAS-parkin-RNAi;UAS-Vps13-RNAi*^{HMS01715}, with N-value 278 of and median survival of 54 days, *Ddc-Gal4; UAS-parkin-RNAi; UAS-Vps13-RNAi*^{HMS024460} with N-value of 296 and median survival of 54 days, and overexpression line *Ddc-Gal4; UAS-parkin-RNAi; UAS-Vps13*^{EY09640} with N-value of 304 and a median survival of 50 days showed a significant decrease in their longevity compared to the control *Ddc-Gal4;UAS-parkin-RNAi; UAS-lacZ*. Inhibition line *Ddc-Gal4; UAS-parkin-RNAi; UAS-Vps13-RNAi*^{GD14789} had no substantial change in its longevity. Like our results in the directed neuron-specific experiment, control *Ddc-Gal4;UAS-parkin-RNAi; UAS-lacZ* had the best median survival score, inhibition group *Ddc-Gal4; UAS-parkin-RNAi; UAS-Vps13-RNAi*^{GD14789} had the closest scores to the control *Ddc-Gal4;UAS-parkin-RNAi;UAS-lacZ,* and overexpression line *Ddc-Gal4; UAS-parkin-RNAi; UAS-Vps13*^{EY09640} had the lowest scores. The overexpression line *Ddc-Gal4; UAS-parkin-RNAi; UAS-Vps13*^{EY09640} showed an improvement in its longevity compared to the directed neuron-specific overexpression line *Ddc-Gal4;UAS-Vps13*^{EY09640}.

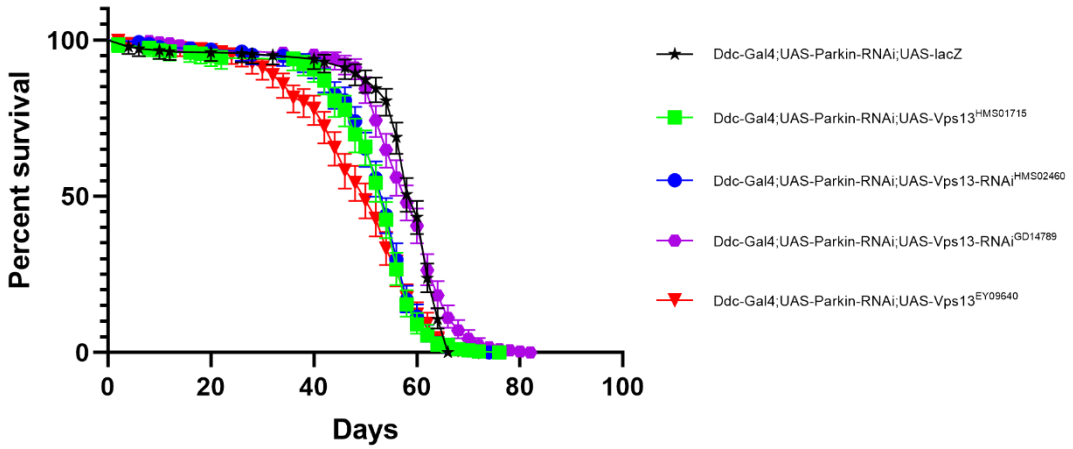


Figure 16. Directed neuron specific expression with inhibition of *Parkin* improves the effect of inhibition and overexpression of *Vps13* on *Drosophila melanogaster* longevity.

The error bars as made with CI 95%. P<0.05 considered significant.

Table 9. Log-rank statistical analysis of longevity of flies with directed neuron-specific expression with inhibition of *Parkin,* along with inhibition and overexpression of *Vps13.*

Survivorship Analysis					
Genotype	Number of Flies	Median Survival (Days)	Chi-square Value	P-value	Significance
Ddc-Gal4;UAS-Parkin-RNAi;UAS-lacZ	328	60	N/A	N/A	N/A
Ddc-Gal4;UAS-Parkin-RNAi;UAS-Vps13-RNAi^{HMS01715}	278	54	104.8	<0.0001	Yes
Ddc-Gal4;UAS-Parkin-RNAi;UAS-Vps13-RNAi^{HMS02460}	296	54	96.72	<0.0001	Yes
Ddc-Gal4;UAS-Parkin-RNAi;UAS-Vps13-RNAi^{GD14789}	296	58	2.106	0.1468	No
Ddc-Gal4;UAS-Parkin-RNAi;UAS-Vps13^{EY09640}	304	50	103.8	<0.0001	Yes

Climbing Assay

I compared the locomotor ability of each line with the control group using non-linear regression analysis. N-value for every line is 50, and $P<0.05$ considered significant (Figure 17, Table 10). The inhibition *line Ddc-Gal4; UAS-parkin-RNAi; UAS-Vps13-RNAi*^{GD14789} started with a better quality of climbing compared to all other groups and had a better locomotor ability compared to the control. Climbing ability decreased in the inhibition lines *Ddc-Gal4;UAS-parkin-RNAi;UAS-Vps13-RNAi*^{HMS01715} and *Ddc-Gal4; UAS-parkin-RNAi; UAS-Vps13-RNAi*^{HMS024460}. Surprisingly overexpression line *Ddc-Gal4; UAS-parkin-RNAi; UAS-Vps13*^{EY09640} had a significant increase in the locomotor ability compared to control while overexpression line *Ddc-Gal4;UAS-Vps13*^{EY09640} showed no substantial climbing change with its relative control group .

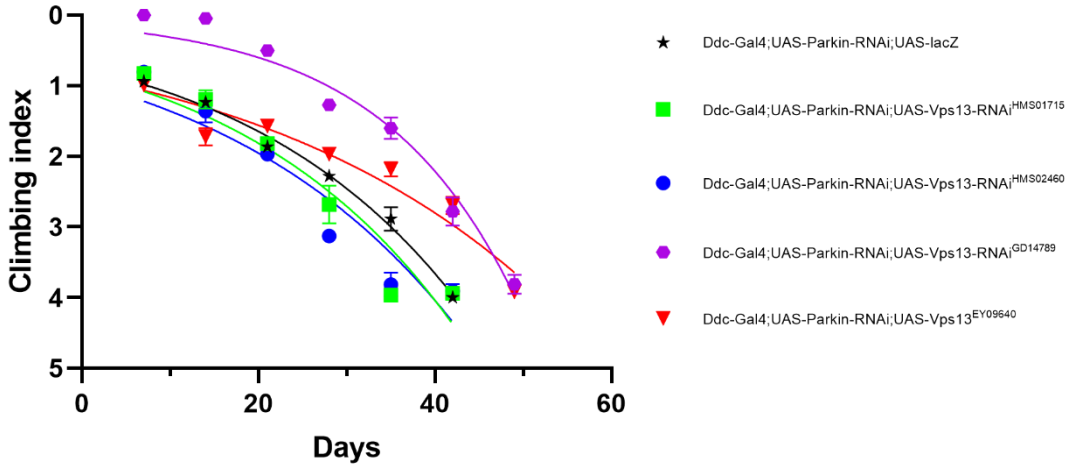


Figure 17. Directed neuron specific expression with inhibition of *Parkin* improves the effect of overexpression of *Vps13* in *D. melanogaster's* climbing ability. N-value is 50.The error bars as made with CI 95%. P<0.05 considered significant.

Table 10. Statistical analysis of the locomotor ability of *D. melanogaster* using non-linear regression curve with directed neuron-specific expression with inhibition of *Parkin*, along with inhibition and overexpression of *Vps13*

Climbing Analysis					
Genotype	Slope (k)	SE	95% Confidence interval (CI)	P-value	Significance
Ddc-Gal4;UAS-Parkin-RNAi;UAS-lacZ	0.03987	0.001577	0.03670 to 0.04316	N/A	N/A
Ddc-Gal4;UAS-Parkin-RNAi;UAS-Vps13-RNAi^{HMS01715}	0.03999	0.003317	0.03377 to 0.04664	0.0267	Yes
Ddc-Gal4;UAS-Parkin-RNAi;UAS-Vps13-RNAi^{HMS02460}	0.03630	0.003139	0.03044 to 0.04251	0.0030	Yes
Ddc-Gal4;UAS-Parkin-RNAi;UAS-Vps13-RNAi^{GD14789}	0.06512	0.004043	0.05784 to 0.07318	<0.0001	Yes
Ddc-Gal4;UAS-Parkin-RNAi;UAS-Vps13^{EY09640}	0.02924	0.001871	0.02540 to 0.03328	<0.0001	Yes

Summary of Results

Bioinformatic analysis of *Vps13* in *D. melanogaster* and *Vps13C* in *Homo sapiens* showed high levels of conservation in their domains (Figure18). The presence of the ATG-C domain, which was highly conserved among vertebrates in invertebrates, interested me in investigating the role of VPS13C in autophagy pathways. ATG2A is an autophagy-related protein containing Chroein-N and ATG-C domains that regulates autophagy and lipid droplet morphology in mammalian cells (Velikkakath *et al.*, 2012). Despite sharing the same domains with ATG2A, VPS13C does not seem to be involved in the autophagy pathways. However, VPS13A, the closest member of the VPS13 family to VPS13C, has been reported to be involved in autophagy regulation (Muñoz-Braceras, Calvo and Escalante, 2015; Yang *et al.*, 2016). Please see figure 3 for multiple alignments of ATG2A and human VPS13C protein and figure 18 for the schematic view of existing domains in these proteins.

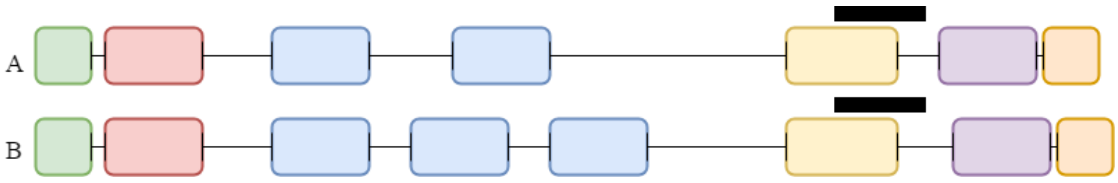


Figure 18. Comparison of *Drosophila melanogaster* Vps13 protein (A) and *Homo sapiens* VPS13C protein (B) with conserved domains. Highlighted are Chorein-N (Green), VPS13 (Red), VPS13 mid rpt (Blue), SHR-BD (Yellow), VPS13C (Purple), ATG-C (Orange), DUF1162 (Black). This image was generated with *draw.io.*

Eye analysis showed a distinct bristle void in *D. melanogaster's* eye when *Vps13* was overexpressed, while the bristle number increased when *Vps13* was inhibited (Figure 19; Table 11).

Table 11. Summary of eye analysis. Arrows pointing up, indicate an increase and arrows pointing down, indicate a decrease.

Neurodevelopment Analysis			
Genes	Ommatidia Number	Bristle Number	Bristle Void Area
Inhibition Lines			
GMR-Gal4;UAS- Vps13-RNAi^{HMS01715}	+ ↓	+ ↓	+ ↓
GMR-Gal4;UAS- Vps13-RNAi^{HMS02460}	+ ↓	+ ↑	+ ↓
GMR-Gal4;UAS- Vps13-RNAi^{GD14789}	-	+ ↑	+ ↓
Overexpression Line			
GMR-Gal4;UAS-Vps13^{EY09640}	-	+ ↓	+ ↑

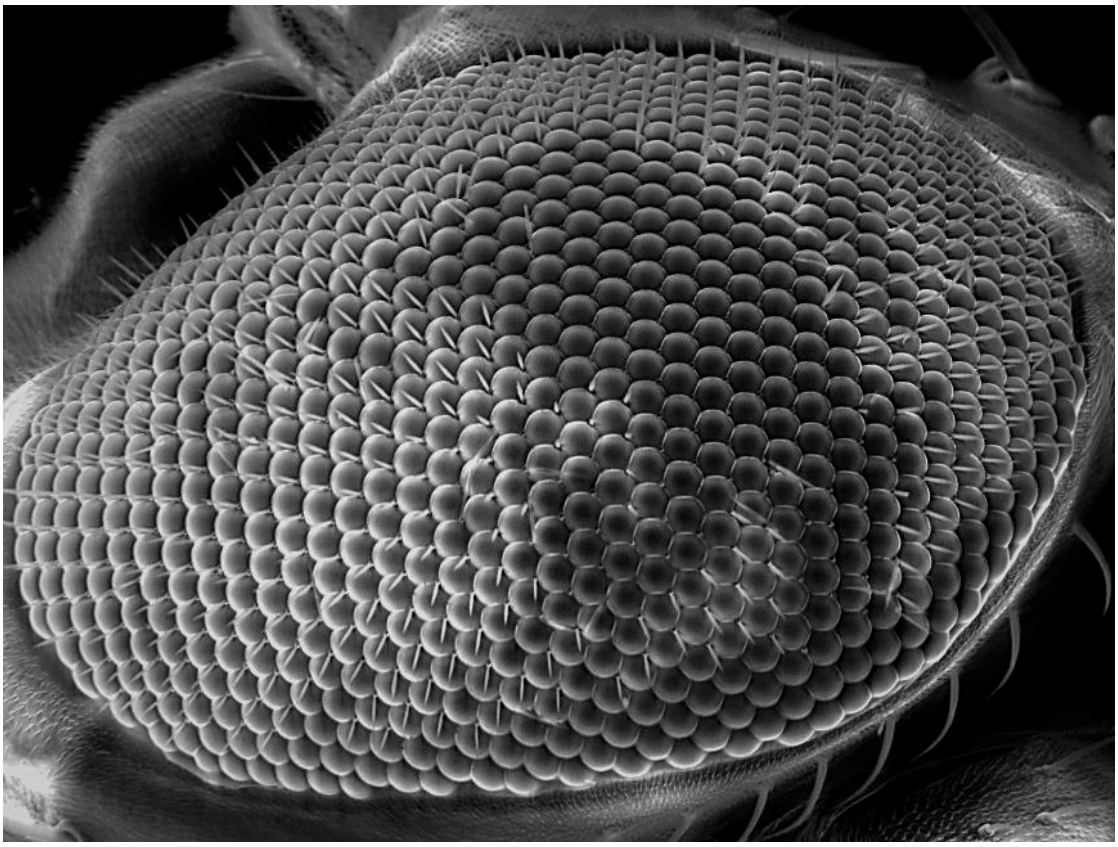


Figure 19. SEM image showing the phenotypic impact of overexpresing *Vps13* in *D. melanogaster* eye.

Ageing and Climbing Assay

I analyzed the impact of inhibition and overexpression of *Vps13* in some specific tissues, on *D. melanogaster's* ageing and climbing ability. Table 12 is a summary of my findings.

Table 12. Summary of results in survivorship and climbing abilities of *D. melanogaster*. Arrows pointing up, indicate an increase and arrows pointing down, indicate a decrease 'A' stands for Aging and 'C' stands for climbing.

Responder lines	Transgenic Lines							
Inhibition Lines	Th-Gal4		D42-Gal4		Ddc-Gal4		Ddc-Gal4;UAS-Parkin-RNAi	
	A	C	A	C	A	C	A	C
GMR-Gal4;UAS-Vps13-RNAi^{HMS01715}	+ ↓	+ ↓	+ ↓	-	+ ↓	-	+ ↓	+ ↓
GMR-Gal4;UAS-Vps13-RNAi^{HMS02460}	+ ↓	-	-	-	+ ↓	-	+ ↓	+ ↓
GMR-Gal4;UAS-Vps13-RNAi^{GD14789}	+ ↑	+ ↑	+ ↑	+ ↑	+ ↓	+ ↑	-	+ ↑
Overexpression Line								
GMR-Gal4;UAS-Vps13^{EY09640}	+ ↓	+ ↓	+ ↑	+ ↓	+ ↓	-	+ ↓	+ ↑

Discussion

Parkinson Disease (PD) is the most common movement disorder, a result of the loss of dopaminergic neurons in the *substantia nigra pars compacta* in the midbrain (Bilen and Bonini, 2005). With an incidence of 1% among 65-year-olds, the prevalence of PD increases with age (Farrer, 2006). This neurodegenerative disease is clinically characterized by the symptoms of resting tremor, rigidity, and bradykinesia. A projected increase in the average age of the population in the next decade emphasizes the importance of this study. There is no absolute cure for this debilitating disease, and dopamine replacement treatments alleviate only some of the symptoms moreover the beneficial effects are gradually lost, and intolerable side effects can be produced (Guo, 2012). Some of the risk factors, including genetic and environmental risk factors, have been identified, yet there is much to discover before a thorough understanding of PD is achieved.

Disruption in mitochondrial function, especially in the pathways related to the mitochondrial autophagy (mitophagy), and mitochondrial fission and fusion, can result in the occurrence of PD. For example, one of the genetic pathways involved in the mitochondrial integrity and maintenance through proper fission and fusion is the *Pink1-parkin* pathway, in which the Pink1 protein kinase positively regulates the parkin ubiquitin ligase to "sense" mitochondrial damage, and recruits the damaged mitochondria into a quality-control system for further actions. Mutations in these genes and those that act with them can result in PD (Guo, 2012). Only 5 to 10 percent of the PD cases are known to be genetic-based. However, scientists are discovering more about inherited forms of PD with an emphasis on the involved pathways. In 2017 whole-exome sequencing (WES) identified *VPS13C* as a PD candidate gene (Jansen *et al.*,

2017). In this set of studies, I studied the *VPS13C* orthologue, *Vps13,* in *D. melanogaster* as a model organism.

VPS13C in humans is associated with transport through the endomembrane system, from the Golgi Body to the endosomes (Saxena *et al.*, 2010), lipid transportation between mitochondria and other organelles (Kumar *et al.*, 2018), adipogenesis (Yang *et al.*, 2016), early impairment of blood glucose homeostasis (Windholz J *et al.*, 2011), mitochondrion organization, and mitophagy (Lesage *et al.*, 2016). The molecular function of *Vps13* in *D. melanogaster* is still unknown. However, studies show that *Vps13* is involved in vacuolar protein targeting, protein metabolic process, lysosomal degradation pathways, and protein homeostasis maintenance in the brain of adult *D. melanogaster* (Vonk *et al.*, 2017). In this study, I assessed the impact of inhibition and overexpression of the *Vps13* upon survivorship and locomotor ability of flies.

Bioinformatic analysis shows that the Vps13 protein is highly conserved among both vertebrates and invertebrates (Figure 2). The protein, in vertebrates and invertebrates, contain shared domains (Figure 18) including an amino-terminal region of Chorein (Chorein N), a Vacuolar-sorting associated protein 13 (VPS13) domain, a SHR-binding domain (SHR-BD), Vacuolar-sorting associated protein 13 repeating coiled regions (VPS13 mid rpt), DUF1162, a Vacuolar sorting-associated 13 protein carboxy-terminal (VPS13C) and an Autophagy- related protein C-terminal domain (ATG C). This information allowed me to pursue to the next level of my experiment, which was the analysis of the effect of *Vps13* on the development of *D. melanogaster* eye units.

Eye analysis, experimentation meant to provide an assessment of the influence of applied gene manipulation of *Vps13* upon neurodevelopment, revealed the presence of a 33.3% area void of bristles when *Vps13* was overexpressed in the developing *D.*

melanogaster eye (Figure 19). This means that overexpression of *Vps13* interfered with the process of neurodevelopment in a subtle but clear manner (Kramer and Staveley, 2003). Furthermore, inhibition of *Vps13* consistently influenced neurodevelopment of the *D. melanogaster* eye and resulted in a decrease in the area of bristle void.

These findings in the eye analysis combined with collected information in bioinformatics encouraged us to analyze the impact of inhibition and overexpression of *Vps13* in different tissues on the survival and locomotor ability of *D. melanogaster*. Please see table 12 for the summary of ageing and climbing assays. Flies expressing the *UAS-Vps13-RNAi*^{HMS01715}, construct under the control of neural-specific *Gal4* drivers showed a decrease in their longevity in all cases, and showed a decrease or no differences in their climbing ability. It is noteworthy that the only noticeable impact of *UAS-Vps13-RNAi*^{HMS01715} when it was expressed in the motorneuron tissues was on longevity. Inhibition of *Vps13* via expression of *UAS- Vps13-RNAi*^{HMS02460}, did not have a significant impact on the longevity or the climbing ability of flies. On the other hand, Inhibition of *Vps13* via *UAS- Vps13-RNAi*^{GD14789}, mostly improved survival and climbing ability of flies. Overexpression of the *Vps13* via *UAS-Vps13*^{EY09640} in *D. melanogaster* mainly reduced their survival and longevity.

In humans, VPS13C is associated with rapid progressing early-onset parkinsonism with the presence of widely distributed Lewy bodies, which affects the survival of dopaminergic neurons (Lesage *et al.*, 2016). Loss of VPS13C function increases the vulnerability of mitochondria to stress, which activates the *Pink1/parkin*-dependent mitochondrial quality control pathways and exacerbates *Pink1/parkin* response to mitochondrial depolarization. Mitochondrial depolarization partially redistributes VPS13C from mitochondrial surface to the cytosol space without significantly affecting

the transcription level of VPS13C. Pink1 and Parkin are involved in a pathway that transports damaged mitochondria directly to lysosomes in response to mitochondrial stress. Based on the VPS13C protein localization within mitochondria and its relocation to the cytosol in response to mitochondrial damage, VPS13C might be involved in this process as well. *Pink1, park2,* and *VPS13C* are engaged in a regulatory loop in which Pink1 silencing is associated with downregulation of transcription levels of VPS13C while overexpression of *Pink1* upregulates *VPS13C* transcription levels. Silencing *VPS13C* increases the *Pink1/parkin*-mediated mitophagy levels (Lesage *et al.*, 2016) and as explained earlier mitophagy regulation can affect cellular longevity.

A study in 2018 suggested that VPS13C and VPS13A transport lipids between the endoplasmic reticulum (ER) and other organelles (Kumar *et al.*, 2018). They found that mutation in VPS13C and VPS13A vitiates the lipid homeostasis of membranes in neurological disorders. Mitochondria use exogenous precursors to produce most of their lipids. These precursors are particularly exported from ER to mitochondria, and because the membranes of these two organelles are not directly connected to each other by membrane traffic, the presence of a protein-mediated transport becomes particularly necessary for these exchanges. *Kumar et al*., 2018 hypothesized that Vps13 is the protein that mediates the transfer of lipids among vacuoles and other membranes and that it possibly provides an alternative path through vacuoles for transferring lipids from ER to mitochondria. Some studies showed that spontaneous gain-of-function mutations in *Vps13* gene in yeast could rescue growth defects that were originally created by mutations in ER-mitochondria encounter structure (ERMES), and lack of both Vps13 and ERMES is lethal (Lang *et al.*, 2015); moreover, mutant yeast that lacked either ERMES or Vps13 showed defects in the integrity of their mitochondrial membrane (Park *et al.*, 2016)

which are consistent with their hypothesis. Earlier, I explained how LOF mutation in *VPS13C* resulted in mitochondrial dysfunction (Lesage et al., 2016). However, VPS13C proteins were not present at the mitochondrial contact sites, suggesting that the observed dysfunction might have been as a result of an abnormality that indirectly had influenced the intracellular lipid transport (Kumar *et al.*, 2018). They found that VPS13A proteins were anchored at the mitochondrial contact sites while VPS13C were localized on the late endosomes/lysosomes, and both VPS13A and VPS13C were attached on the lipid droplets through their N-terminals. Their observations showed that overexpression of *VPS13A* and *VPS13C* results in an expansion in the contacts between organelles. Our observation in the eye analysis followed by survival and longevity assays showed that in most cases, overexpression of *Vps13* reduces robustness in flies while inhibition of *Vps13* increased measures of climbing and longevity (Table 13). These observations could be a result of using different techniques for inhibition and overexpression of this gene. *VPS13* LOF in the Lesage study was through mutation while I used siRNA coupled with UAS-Gal4 system to inhibit the translation of the gene in a specific tissue. Furthermore, overexpression of *VPS13C* increased the contacts between organelles, however, the increase in the amount of a protein can as well be cytotoxic and lethal for cells. It should also be noted that most of the studies on VPS13C were done on *Sacaromyses cervisiae* which has less similarity to human VPS13C in terms of its protein sequence compared to *D. melanogaster.*

With the present study, I successfully characterized and identified the *D. melanogaster* homologue of *VPS13C.* Although there remains insufficient information about the exact function of this gene, its high level of conservation among vertebrates and invertebrates suggests the importance of this gene. Furthermore, consistent

evidence of the presence of Vps13 in mitochondrial regulatory pathways of mitophagy underscores its link to Parkinson disease. Further analyses should be carried out at the cellular and molecular levels, such as microarray and PCR analyses, primarily due to the involvement of *Vps13* in protein-mediated lipid transportation and mitophagy. As well, further research into the interaction of *Vps13* with other genes such as *Pink1, parkin, Rab7,* and *FBXO7* may provide valuable insight into the role of *Vps13* in the etiology of PD .

References

Abeliovich, A. and Gitler, A. D. (2016) 'Defects in trafficking bridge Parkinson's disease pathology and genetics', *Nature*, 539(7628), pp. 207–216. doi: 10.1038/nature20414.

Ammal Kaidery, N. and Thomas, B. (2018) 'Current perspective of mitochondrial biology in Parkinson's disease', *Neurochemistry International*. Pergamon, 117, pp. 91–113. doi: 10.1016/J.NEUINT.2018.03.001.

Bilen, J. and Bonini, N. M. (2005) '*Drosophila* as a Model for Human Neurodegenerative Disease', *Annual Review of Genetics*, 39(1), pp. 153–171. doi: 10.1146/annurev.genet.39.110304.095804.

Brand, A. H. and Perrimon, N. (1993) 'Targeted gene expression as a means of altering cell fates and generating dominant phenotypes', *Development*, 118(2), pp. 401–415.

Breese, G. R. *et al.* (2005) 'The neonate-6-hydroxydopamine-lesioned rat: a model for clinical neuroscience and neurobiological principles', *Brain Research Reviews*. Elsevier, 48(1), pp. 57–73. doi: 10.1016/J.BRAINRESREV.2004.08.004.

Daniele, T. and Schiaffino, M. V. (2014) 'Organelle biogenesis and interorganellar connections: Better in contact than in isolation', *Communicative and Integrative Biology*, 7(4), pp. e29587-1-e29587-6. doi: 10.4161/cib.29587.

De, M. *et al.* (2017) 'The Vps13p-Cdc31p complex is directly required for TGN late endosome transport and TGN homotypic fusion', *Journal of Cell Biology*, 216(2), pp. 425–439. doi: 10.1083/jcb.201606078.

Dimmer, K. S. and Rapaport, D. (2017) 'Mitochondrial contact sites as platforms for phospholipid exchange', *Biochimica et Biophysica Acta (BBA) - Molecular and Cell*

Biology of Lipids. Elsevier, 1862(1), pp. 69–80. doi: 10.1016/J.BBALIP.2016.07.010.

Farrer, M. J. (2006) 'Genetics of Parkinson disease: Paradigm shifts and future prospects', *Nature Reviews Genetics*, 7(4), pp. 306–318. doi: 10.1038/nrg1831.

Guo, M. (2012) 'Drosophila as a model to study mitochondrial dysfunction in Parkinson's disease', *Cold Spring Harbor perspectives in medicine*, 2(11), pp. 1–18. doi: 10.1101/cshperspect.a009944.

Hewitt, V. L. and Whitworth, A. J. (2017) *Mechanisms of Parkinson's Disease: Lessons from Drosophila*. 1st edn, *Current Topics in Developmental Biology*. 1st edn. Elsevier Inc. doi: 10.1016/bs.ctdb.2016.07.005.

Houlden, H. and Singleton, A. B. (2012) 'The genetics and neuropathology of Parkinson's disease', *Acta Neuropathologica*, 124(3), pp. 325–338. doi: 10.1007/s00401-012-1013-5.

Jansen, I. E. *et al.* (2017) 'Discovery and functional prioritization of Parkinson ' s disease candidate genes from large-scale whole exome sequencing'. Genome Biology, pp. 1–26. doi: 10.1186/s13059-017-1147-9.

Khurana, V., Chung, C. and Data, R. U. S. A. (2017) 'Cellular Discovery Platform for Neurodegenerative Diseases', *Patient Application*.

Koprich, J. B., Kalia, L. V. and Brotchie, J. M. (2017) 'Animal models of α-synucleinopathy for Parkinson disease drug development', *Nature Reviews Neuroscience*. Nature Publishing Group, 18(9), pp. 515–529. doi: 10.1038/nrn.2017.75.

Kramer, J. M. and Staveley, B. E. (2003) 'GAL4 causes developmental defects and apoptosis when expressed in the developing eye of Drosophila melanogaster', *Genetics and Molecular Research*, 2(1), pp. 43–47.

Kumar, N. *et al.* (2018) ‘VPS13A and VPS13C are lipid transport proteins differentially localized at ER contact sites’, *J Cell Biol*. Rockefeller University Press, 217(10), pp. 3625–3639. doi: 10.1083/JCB.201807019.

Lang, A. B. *et al.* (2015) ‘ER-mitochondrial junctions can be bypassed by dominant mutations in the endosomal protein Vps13’, *Journal of Cell Biology*, 210(6), pp. 883–890. doi: 10.1083/jcb.201502105.

de Lau, L. M. and Breteler, M. M. (2006) ‘Epidemiology of Parkinson’s disease’, *The Lancet Neurology*. Elsevier, 5(6), pp. 525–535. doi: 10.1016/S1474-4422(06)70471-9.

Lesage, S. *et al.* (2016) ‘Loss of VPS13C Function in Autosomal-Recessive Parkinsonism Causes Mitochondrial Dysfunction and Increases PINK1/Parkin-Dependent Mitophagy’, *American Journal of Human Genetics*, 98(3), pp. 500–513. doi: 10.1016/j.ajhg.2016.01.014.

Lipman, D. J., Altschul, S. F. and Kececioglut, J. D. (1989) ‘A tool for multiple sequence alignment (proteins/structure/evolution/dynamic programming)’, *Proc. Nail. Acad. Sci. USA*, 86(June), pp. 4412–4415. Available at: https://www.pnas.org/content/pnas/86/12/4412.full.pdf.

Lloyd, T. E. and Taylor, J. P. (2010) ‘Flightless flies: Drosophila models of neuromuscular disease.’, *Annals of the New York Academy of Sciences*, 1184, pp. e1-20. doi: 10.1111/j.1749-6632.2010.05432.x.

Lysia S. Frono, M. (1996) ‘Neuropathology of Parkinson’s Disease’, *Neuropathology and Experimental Neurology*, 55(March), p. 14.

M’Angale, P. G. and Staveley, B. E. (2016) ‘Knockdown of the putative Lifeguard homologue CG3814 in neurons of Drosophila melanogaster’, *Genetics and Molecular Research*, 15(4). doi: 10.4238/gmr15049290.

Marsden, C. D. and Parkes, J. D. (1977) ‘SUCCESS AND PROBLEMS OF LONG-TERM LEVODOPA THERAPY IN PARKINSON’S DISEASE’, *The Lancet*. Elsevier, 309(8007), pp. 345–349. doi: 10.1016/S0140-6736(77)91146-1.

Michel, P. P., Hirsch, E. C. and Hunot, S. (2016) ‘Understanding Dopaminergic Cell Death Pathways in Parkinson Disease’, *Neuron*, 90(4), pp. 675–691. doi: 10.1016/j.neuron.2016.03.038.

Mishra M., K. E. (2012) *Analysis of the Drosophila Compound Eye with Light and Electron Microscopy. In: Weber B., LANGMANN T. (eds) Retinal Degeneration. Methods in Molecular Biology, vol 935. (Methods and Protocols)*. Humana Press, Totowa, NJ. doi: https://doi.org/10.1007/978-1-62703-080-9_11.

Mizuno, E. *et al.* (2007) ‘Brain-specific transcript variants of 5′ and 3′ ends of mouse VPS13A and VPS13C’, *Biochemical and Biophysical Research Communications*. Academic Press, 353(4), pp. 902–907. doi: 10.1016/J.BBRC.2006.12.122.

Muñoz-Braceras, S., Calvo, R. and Escalante, R. (2015) ‘TipC and the chorea-acanthocytosis protein VPS13A regulate autophagy in Dictyostelium and human HeLa cells’, *Autophagy*, 11(6), pp. 918–927. doi: 10.1080/15548627.2015.1034413.

Nussbaum, R. L. *et al.* (1997) ‘Mutation in the α-Synuclein Gene Indentified in Families with Parkinson ’ s Disease Published by : American Association for the Advancement of Science Stable URL : http://www.jstor.org/stable/2892982 JSTOR is a not-for-profit service that helps scholars ,’ *Science*, pp. 25–28.

Pandey, Udai Bhan and Nichols, C. D. (2011) ‘Human Disease Models in Drosophila melanogaster and the Role of the Fly in Therapeutic Drug Discovery’, *Drug Delivery*, 63(2), pp. 411–436. doi: 10.1124/pr.110.003293.411.

Pandey, U. B. and Nichols, C. D. (2011) ‘Human Disease Models in Drosophila

melanogaster and the Role of the Fly in Therapeutic Drug Discovery', *Pharmacological Reviews*. doi: 10.1124/pr.110.003293.

Park, J.-S. *et al.* (2016) 'Yeast Vps13 promotes mitochondrial function and is localized at membrane contact sites', *Molecular Biology of the Cell*, 27(15), pp. 2435–2449. doi: 10.1091/mbc.e16-02-0112.

Petit, P. X. *et al.* (1995) 'Alterations in mitochondrial structure and function are early events of dexamethasone-induced thymocyte apoptosis.', *The Journal of Cell Biology*, 130(1), pp. 157 LP – 167. doi: 10.1083/jcb.130.1.157.

Satake, W. *et al.* (2009) 'Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease', *Nature Genetics*. Nature Publishing Group, 41, p. 1303. Available at: https://doi.org/10.1038/ng.485.

Saxena, R. *et al.* (2010) 'Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge', *Nature Genetics*. Nature Publishing Group, 42, p. 142. Available at: https://doi.org/10.1038/ng.521.

Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012) 'NIH Image to ImageJ: 25 years of image analysis', *Nature methods*. Nature Publishing Group, 9(7), p. 671.

Sherratt, H. S. (1991) 'Mitochondria: structure and function', *Revue neurologique*, 147(6–7), p. 417—430. Available at: http://europepmc.org/abstract/MED/1962047.

Sievers, F. *et al.* (2011) 'Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega', *Molecular Systems Biology*, 7(1), p. 539. doi: doi:10.1038/msb.2011.75.

St Johnston, D. (2002) 'The art and design of genetic screens: Drosophila melanogaster', *Nature Reviews Genetics*, 3(3), pp. 176–188. doi: 10.1038/nrg751.

Staveley, B. E. (2012) 'Successes of modelling parkinson disease in drosophila',

Mechanism in Parkinson's disease—models and treatments ISBN, pp. 953–978.

Swerdlow, R. H. *et al.* (2001) 'Gender ratio differences between Parkinson's disease patients and their affected relatives', *Parkinsonism & Related Disorders*. Elsevier, 7(2), pp. 129–133. doi: 10.1016/S1353-8020(00)00029-8.

T. Wichmann, M. R. D. (2009) 'MPTP Parkinsonism Model', *Encyclopedia of Neuroscience,*.

Thurmond, J. *et al.* (2018) 'FlyBase 2.0: the next generation', *Nucleic Acids Research*, 47(D1), pp. D759–D765. doi: 10.1093/nar/gky1003.

Todd, A. M. and Staveley, B. E. (2008) 'Pink1 suppresses α-synuclein-induced phenotypes in a Drosophila model of Parkinson's disease', *Genome*, 51(12), pp. 1040–1046. doi: 10.1139/G08-085.

Todkar, K., Chikhi, L. and Germain, M. (2019) 'Mitochondrial interaction with the endosomal compartment in endocytosis and mitochondrial transfer', *Mitochondrion*. Elsevier. doi: 10.1016/J.MITO.2019.05.003.

Velayos-Baeza, A. *et al.* (2004) 'Analysis of the human VPS13 gene family', *Genomics*. doi: 10.1016/j.ygeno.2004.04.012.

Velikkakath, A. K. G. *et al.* (2012) 'Mammalian Atg2 proteins are essential for autophagosome formation and important for regulation of size and distribution of lipid droplets', *Molecular Biology of the Cell*, 23(5), pp. 896–909. doi: 10.1091/mbc.e11-09-0785.

Verstraeten, A., Theuns, J. and Van Broeckhoven, C. (2015a) 'Progress in unraveling the genetic etiology of Parkinson disease in a genomic era', *Trends in Genetics*. Elsevier Current Trends, 31(3), pp. 140–149. doi: 10.1016/J.TIG.2015.01.004.

Verstraeten, A., Theuns, J. and Van Broeckhoven, C. (2015b) 'Progress in unraveling

the genetic etiology of Parkinson disease in a genomic era', *Trends in Genetics*. Elsevier Current Trends, 31(3), pp. 140–149. doi: 10.1016/J.TIG.2015.01.004.

Vonk, J. J. *et al.* (2017) 'Drosophila Vps13 Is Required for Protein Homeostasis in the Brain', pp. 1–21. doi: 10.1371/journal.pone.0170106.

Wang, L. *et al.* (2016) 'Association of four new candidate genetic variants with Parkinson's disease in a Han Chinese population', *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 171(3), pp. 342–347. doi: doi:10.1002/ajmg.b.32410.

Wang, Z. and Zhang, H. (2018) 'Mitophagy: Vps13D Couples Mitochondrial Fission and Autophagic Clearance', *Current Biology*. Cell Press, 28(2), pp. R66–R68. doi: 10.1016/J.CUB.2017.12.017.

Windholz J, Kovacs P, Tönjes A, Dittrich K, Blüher S, Kiess W, et al. (2011) 'Effects of Genetic Variants in ADCY5, GIPR, GCKR and VPS13C on Early Impairment of Glucose and Insulin Metabolism in Children.', *PLoS ONE 6(7): e2210*. doi: https://doi.org/10.1371/journal.pone.0022101.

Yang, R. Y. *et al.* (2016) 'Identification of VPS13C as a galectin-12- binding protein that regulates galectin-12 protein stability and adipogenesis', *PLoS ONE*. Public Library of Science, 11(4). doi: 10.1371/journal.pone.0153534.