2021

Identifying Genes Downstream of Mef2 that Influence Ethanol Sedation in Drosophila melanogaster

Ananya Talikoti
Virginia Commonwealth University

IDENTIFYING GENES DOWNSTREAM OF MEF2 THAT INFLUENCE ETHANOL SEDATION IN *DROSOPHILA MELANOGASTER*

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science at Virginia Commonwealth University

By

Ananya Talikoti

Bachelor of Science, North Carolina State University, 2019

Advisor: Mike Grotewiel, PhD

Associate Dean for Graduate Education

Professor, Department of Human & Molecular Genetics

Table of Contents

Acknowledgments

There are several people who have shaped my short yet rewarding time at VCU. Firstly, I'd like to thank my mentor, Dr. Grotewiel for his support, patience and guidance throughout my graduate training. I really appreciate his constant willingness to go above and beyond to help me be successful. I would not have been able to complete this thesis without his guidance.

I would also like to thank my committee members, Dr. Andrew Davies, Dr. Jim Lister and Dr. Laura Mathies for their support. A special thanks to Dr. Mathies for her guidance and willingness to help with our sequencing experiments.

Additionally, I'd like to thank all the students and faculty in the Department of Human and Molecular Genetics (especially Dr. Rita Shiang) for their encouragement. I would also like to thank Brandon Shell for his company and help in the lab.

Finally, I'd like to thank my family, friends and loved ones without whose support completing this stage of my graduate training would not have been possible.

Statement of Contributions

The primary purpose of my thesis project was to identify genes downstream of *Mef2* that may influence ethanol sedation behaviors in *Drosophila melanogaster*. Dr. Danielle Dick's lab at VCU provided us with a list of genes involved in human externalizing behaviors and a list of 38 genes that overlapped between the externalizing behavior genes and known Mef2-bound genes. These lists were used in several stages in this thesis. Dr. Silviu Bacanu at VCU previously performed a GSCAN analysis of human orthologs of known Mef2-bound genes, which was also used in Chapter 2. Dr. Michael Miles also provided us with several data sets that were referenced in Chapter 2.

GeneWiz performed the RNA-sequencing (Chapter 3) and several downstream analyses. Figures provided by them are denoted in the figure legends. Dr. Laura Mathies helped with several of the bioinformatic analyses described in Chapter 3. All experiments were designed by myself and Dr. Grotewiel. Other than figures specifically denoted as being provided by GeneWiz, I performed all of the experiments and generated the data shown.

List of Figures

CHAPTER 2

CHAPTER 3

List of Tables

CHAPTER 2

CHAPTER 3

List of Abbreviations

19-mer…….....19 nucleotide sequence

ADH…….....alcohol dehydrogenase

ALDH……..aldehyde dehydrogenase

ANOVA……analysis of variance

AUD……alcohol use disorder

AUDIT…….....alcohol use disorder identification test

AUDIT-C…….alcohol use disorder identification test for consumption

AUDIT-P…….alcohol use disorder identification test for preference

AUTS2…….....*Activator of transcription and developmental regulation*

BDSC…….....Bloomington *Drosophila* Stock Center

BMC……..Bonferroni Multiple Corrections

Bx…….....*Beadex*

CAFÉ……...……capillary feeder

cGMP……...cyclic guanosine monophosphate

CNS…….....central nervous system

CTBP……*C-terminal binding protein*

ChIP-seq……chromatin immunoprecipitation sequencing

DALY's…….....disability adjusted life years

DEG…….....differentially expressed genes

DIOPT…….....DRSC Integrative and Predictive Tool

Drat……..……*Death resistor Adh domain*

DSM-5……..Diagnostic and Statistical Manual

EGF……epidermal growth factor receptor

egl-4……..egg laying defective-4

elav…….....embryonic lethal abnormal vision

FDR…….false discovery rate

Fas2……..*Fasciclin 2*

For…….....*Foraging*

GO……..gene ontology

GOTERM_BP_DIRECT……..GO term Biological Process

GOTERM_CC_DIRECT…….....GO term Cellular Component

GOTERM_MF_DIRECT…….....GO Term Molecular Function

GS……..GeneSwitch

GSCAN…….GWAS and Sequencing Consortium of Alcohol and Nicotine

GWAS……..genome wide association study

HAP1…….....high alcohol preferring 1

HIV/AIDs……..human immunodeficiency virus/acquired immunodeficiency disease

Hr38…….hormone receptor like 38

ics…….*icarus*

iDEP…….integrated Differential Expression and Pathway Analysis

Jhe…….....*juvenile hormone esterase*

LAP1……..low alcohol preferring 1

LMO1…….....*LIN-11, Isl-1, Mec-3 Only Domain*

LOF…….....loss of function

MADS……MCM1, agamous, deficient, serum response factor

MEF2…….....*myocyte enhancer factor 2*

NA……..narrow abdomen

NCAM…….*neural cell adhesion molecule*

PCA…….....principal components analysis

PER……proboscis extension response

PI……preference index

PLR-1 phosphatase…….....*phosphatase of regenerating liver 1*

PRKG1……...*protein kinase cGMP-dependent 1*

RIN……RNA integrity number

RISC…….....RNA-induced silencing complex

RNAi……RNA interference

s-LVv…….....small lateral ventral

SNP…….single nucleotide polymorphism

sl…….....*slowpoke*

SPNS1…….....*Sphingolipid transporter 1*

SRE…….selective response to ethanol

ST50……..Sedation Time 50

S_{19}……...specificity score

Spin……...*spinster*

rRNA……..…….....ribosomal RNA

TGF-β……..transforming growth factor beta

UAS…….....upstream activating sequence

unc79……..……...*uncoordinated 79*

VDRC……...Vienna *Drosophila* Resource Center

w[A]……....white ash burner

w[VDRC]…….....white Vienna Drosophila Resource Center

Abstract

Alcohol use disorder is a global public health issue that affects millions across the world. Alcohol use disorder can result in negative physical and mental health outcomes, and currently treatment options are limited and rates of relapse are high. Identifying genes that affect aspects of ethanol behaviors in model organisms, such as *Drosophila melanogaster*, can serve to eventually develop more robust therapeutic interventions for those experiencing alcohol use disorder or other forms of alcohol dependence. Previous studies have identified a relationship between a person's initial sensitivity to alcohol and their abuse potential for the drug in later life. Therefore, we can study sedation behaviors in *Drosophila melanogaster* to better understand genes that affect alcohol sensitivity. Work in the Grotewiel laboratory has identified the gene *Mef2* as a key regulator of ethanol sedation. The major goals of these studies were to identify genes downstream of *Mef2* that produce a consistent behavioral impact on sedation when knocked down (Chapter 2), and to identify global gene expression changes when *Mef2* is knocked down (Chapter 3). We found RNAi transgenes against two genes, *spin* and *unc79* consistently and significantly increase the amount of time it takes for flies to become sedated when exposed to ethanol. Additionally, through RNA-seq studies, we identified several *Mef2* dependent differentially expressed candidates for future study in ethanol sedation. We compared whether these differentially expressed genes were shared between other gene sets of interest, finding that one set of differentially expressed genes had a significant overlap with genes previously known to bind Mef2. Overall, the studies in this thesis support a number of novel hypotheses regarding the role of *Mef2* and its downstream genes in ethanol sedation that can be explored in the future.

CHAPTER 1: INTRODUCTION

1. Overview of alcohol use in humans

1a. Alcohol use disorder and alcohol consumption patterns

Alcohol use disorder (AUD) is a pattern of problematic alcohol consumption that results in significant impairments or distress, and problems completing daily tasks. Features of AUD include an inability to limit alcohol consumption, development of cravings or tolerance to alcohol and the exhibition of withdrawal symptoms when alcohol is unavailable (NIAAA, SAMHSA 2019). Conditions such as alcohol abuse, dependence and addiction are components of AUD (WHO 2018). In 2016, 15.1 million American adults (aged 18 years or more) were estimated to suffer from AUD, while globally, an estimated 283 million adults aged 15 years or more met some criteria for AUD (WHO 2018). AUD risk varies considerably by gender: in the United States approximately 9.2 million males and 5.3 million females were affected in 2016 (WHO 2018). This is consistent with broader patterns – 237 million men and 46 million women are affected globally. AUD is also prevalent in adolescent communities. In 2018, 400,000 adolescents were predicted to suffer with AUD in some capacity (WHO 2018). Alcohol consumption patterns also vary by geographic location. Although AUD impacts some groups more frequently than others, it is clear that AUD impacts large numbers of individuals across a wide variety of populations.

Alcohol use can have significant health consequences. In 2016, alcohol use contributed to approximately 3 million deaths and 132.6 million disability-adjusted life years (DALY's) (WHO 2018). These metrics equate to 5.3% of all deaths and 5.3% of all DALY's worldwide; alcohol related mortality thus surpasses deaths from diseases such as tuberculosis, HIV/AIDS and diabetes (WHO 2018). Alcohol related deaths fall into several categories, including injuries (28.7%), digestive disease (21.3%), cardiovascular disease (19%), infectious disease (12.9%) and cancer (12.6%) (WHO 2018).

Alcohol consumption has also been linked to a plethora of diseases. For example, increased or irregular heavy alcohol consumption has been strongly associated with hypertensive heart disease and cardiomyopathy, among other types of ischemic heart diseases (Briasoulis et. al 2012, WHO 2018). Mechanisms related to ethanol metabolism, or the oxidation of alcohol (Cederbaum 2012) are causally linked to several types of liver damage, including alcoholic hepatitis and liver cirrhosis (WHO 2018; Gao & Betaller 2011). Rising rates of alcoholic liver disease could cause significant burden on health systems (WHO 2018). There is also a causal link between alcohol use and various types of cancer including esophageal, liver, colon and female breast cancer (Bagnardi et. al 2015, WHO 2018). Alcohol is thought to operate via numerous biological pathways to advance cancer growth (Bagnardi et. al 2015). Specifically, alcohol is able to cause lasting DNA damage and impede normal DNA repair processes (Bagnardi et. al, Cao et. al 2015). Alcohol related cancer risk is generally higher in females than males. This could, in part, be due to alcohol's ability to alter estrogen signaling pathways (Cao et. al 2015, Bagnardi et. al 2015, WHO 2018).

Furthermore, alcohol abuse has been linked to weakening of the immune system and therefore, increased susceptibility to infectious disease (Sarkar et. al 2015). In the upper and lower airways alcohol metabolism, or the oxidation of alcohol, facilitated by alcohol and aldehyde dehydrogenase enzymes respectively (detailed in later sections) (Cederbaum 2012) plays a role in disrupting ciliary function and weakening epithelial cells (Sarkar et. al 2015). Additionally, alcohol impairs the function of immune cells like neutrophils and alveolar macrophages. This can lead to more serious disease in drinkers vs. non-drinkers (Sarkar et. al 2015). This is relevant when considering how essential the immune system is for treatment of various diseases - for example, in cancer patients, chemotherapy is most effective when the immune system is fully functional (Sarkar et. al 2015). Additionally, immune signaling in the brain may be a contributing factor to development of AUD. Alcohol leads to neuroimmune signaling, which can further increase alcohol consumption (Sarkar et. al 2015). Furthermore,

alcohol use is associated with many mental, behavioral and neurodevelopmental disorders including anxiety disorders, depression, dementia and Alzheimer's disease (WHO 2018).

The mechanisms driving negative health consequences of alcohol consumption can be grouped into three broad categories: 1) alcohol toxicity on various tissues and organs leading to conditions such as liver disease, heart disease, cancer and immune dysfunction; 2) development of alcohol dependency altering the patient's self-control and potentially contributing to the development of mental disorders such as depression and psychoses; and 3) the psychoactive effects of intoxication (Babor et. al 2010, WHO 2018). It is clear that alcohol use is widespread and alcohol abuse can have dire effects on individual health outcomes as well as broader population health patterns. Therefore, it is essential to better understand AUD with the aim of eventually developing better diagnostic or therapeutic tools for affected individuals.

1b. AUD diagnostics and DSM definitions

Diagnostic criteria for AUD are described in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5). The DSM-5 diagnosis of AUD is based on the answers to 11 questions; to be diagnosed with AUD, a person must experience a minimum of 2 symptoms listed. The sub-classification of an individual's AUD is based on the number of listed symptoms they experienced (mild: 2-3 symptoms, moderate: 4-5 symptoms and severe: 6+ symptoms). (NIAAA 2016). Examples of questions in the DSM-5 include:

- In the past year, have you:
    - Had times when you ended up drinking more, or longer, than you intended?
    - Spent a lot of time drinking? Or being sick or getting over other aftereffects?
    - Found that drinking - or being sick from drinking - often interfered with taking care of your home or family? Or caused job troubles? Or school problems?

- Given up or cut back on activities that were important or interesting to you, or gave you pleasure, in order to drink?

The DSM 5 is the first version of the DSM that has a specific diagnosis for AUD.NIAAA 2016). Previous iterations of the DSM separated diagnoses of alcohol abuse and alcohol dependency. Under DSM-IV guidelines, individuals experiencing more than one criterion would be diagnosed with “alcohol abuse,” and those experiencing 3+ symptoms would be diagnosed with “alcohol dependency” (NIAAA 2016). While general themes are consistent between the two manuals, there are also key differences between the actual criteria: the DSM-5 removes a criterion referencing legal troubles stemming from alcohol use and adds one regarding alcohol craving as indicators of inappropriate alcohol use (NIAAA 2016).

1c. Externalizing behaviors

Defining features of externalizing behaviors include impulsivity and behavioral disinhibition (Dick et. al 2003, Barr et. al 2020). They include psychiatric, non-clinical, substance abuse (including alcohol abuse) and antisocial conditions as well as risky behaviors (Dick et. al 2003). Externalizing behaviors can generally be thought of as those that have direct consequences on the external world relative to the individual exhibiting the behaviors. This separates them from internalizing behaviors, such as anxiety or depression, that have direct effects on the patient (Barr et. al 2020). Externalizing behaviors are quite common; substance and impulse control disorders have lifetime prevalence rates of 29% and 24.8% respectively, and exhibit strong co-morbidity (Barr et. al 2020). These behaviors are responsible for causing significant social burden (Dick 2003, Barr 2020). Twin studies, and more recently, GWAS (genome wide association studies) studies in large populations have identified that all externalizing behaviors have genetic components, but the work of causally linking genetic variants to phenotypes is still incomplete (Dick 2003).

It is thought that the same sets of genes underlie several externalizing behaviors, including ethanol behaviors, as they are observed together more often than would be expected by chance (Aliev et. al 2015). Aliev et. al evaluated associations between a panel of genes and single nucleotide polymorphisms (SNPs) previously associated with alcohol related traits and various externalizing phenotypes. Examples of phenotypes analyzed include: conduct disorders, adult antisocial disorders, illicit drug use and sensation seeking behaviors (Aliev 2015). It was found that the alcohol-related behavior SNPs were significantly enriched for each externalizing phenotype of interest. Thus, there may be a shared genetic liability that predisposes an individual to a host of externalizing behaviors including alcohol-related behaviors (Aliev 2015).

1d. AUD, alcohol abuse and alcohol dependence as complex disorders

Understanding the cause of AUD and behaviors such as alcohol dependence is a crucial first step of lessening the negative consequences of alcohol abuse. Historically, twin studies have been a key way to determine the relative contribution of genes and environmental factors in the development of alcohol misuse. Twin studies are a common method to identify the source of variance in a specific trait (Kendler 1992). By comparing phenotypic correlations between monozygotic and dizygotic twins, it is possible to quantify the amount of phenotypic variance due to genetic factors, the shared environment and the unique environment (Kendler 1992).

The ACE model considers the following categories: additive genetic factors (A), the common environment (C) and the unique environment (E). If a specific category is taken out but the model still fits the data well, that factor was non-significant in its contribution to the phenotype. In early twin studies, it was found that the AE model fit twin data best, indicating that the common environment was not significant (Kendler 1992). However, more recent meta-analysis of twin and adoption studies found that the estimate of heritability of AUDs across 13 twin studies was 0.51, meaning that 51% of the risk of developing AUD is genetic (Verhulst et. al 2015, Ducci & Goldman 2008). The same genes are responsible for development of AUD in

males and females because correlations between same sex and opposite sex twin pairs were not significantly different. Contribution of the common environment was lower (0.083), but was still significant in males and females. The contribution of the unique environment was 0.39. It is well accepted that ~50% of the risk for developing AUD is genetic (Verhulst et. al 2015, Ducci & Goldman 2008).

Previous findings also suggest there is a relationship between alcohol sensitivity and alcohol abuse. A study comparing alcohol response level at age 20 to whether the individual developed alcoholism showed that initial sensitivity to alcohol is associated with lower abuse potential in later life, and initial resistance is associated with higher abuse potential (Schuckit 1994, Schuckit 1997).

The Self Rating Effects (SRE) of Alcohol aims to quantify whether a person has a low or high alcohol sensitivity based on how many drinks it took for them to feel different, exhibit dizziness or slurred speech, show uncoordinated movements and sleep or pass out (Schuckit 1997). The standard SRE asks participants to answer the questions based on the first five times they consumed alcohol, but can be modified to consider different drinking periods. Numerous studies have shown that the SRE is a highly robust indicator of alcohol use outcomes (Ray et. al 2011).

1e. Identification and function of genes influencing AUD risk

Historically, attempts to identify genes that contribute to the risk of developing alcoholism were in the form of linkage studies in large families consisting of affected and unaffected individuals (Lipner et. al 2018). The basis of linkage studies is to determine whether any known markers or SNPs segregate with affected individuals more than chance would predict (Lipner et. al 2018). Linkage studies rely on recombination; recombination is unlikely when the marker loci and disease loci are close together, so they are more likely to segregate together than if they were located further away (Ott et. al 2015). Linkage studies are important as they can give

important information about inheritance patterns, penetrance, gene-gene interactions and co-segregation (Lipner et. al 2018). They are highly useful when researchers do not have preliminary insight into what genes may be involved in the pathogenesis of a specific condition, as it is unbiased and takes the entire genome into consideration (Ott et. al 2015, Lipner et. al 2018). However, linkage studies are limited in their scope. They are most informative when diseases follow a Mendelian inheritance pattern and the chromosomal regions implicated are often too large to be functionally meaningful without considerable follow-up (Lipner et. al 2018).

In more recent years, GWAS have become a leading method for identifying genetic polymorphisms contributing to a phenotype. The basic principle of GWAS is to scan the entire genome of unaffected and affected individuals to assess whether any polymorphisms are observed more often in the affected population than the unaffected (Lipner et. al 2018). GWAS has several strengths, including the ability to identify common variants that may account for a percentage of overall disease risk. Additionally, due to their case-control structure, collecting data from families is unnecessary and statistical analysis is relatively straightforward (Ducci & Goldman 2008, Lipner et. al 2018). However, very large sample sizes are needed to provide informative results (Ducci & Goldman 2008, Lipner et. al 2018). In complex phenotypes like alcohol use, individual genes tend to have small effects. This, coupled with a high proportion of studies being underpowered make it difficult to unambiguously identify involved genes (Ducci & Goldman 2018). Another issue is phenotype heterogeneity. It is important to note that in order to be diagnosed with some level of AUD, individuals must experience a minimum of two of the criteria laid out by the DSM-V. Given the number criteria and the different classifications of severity, there are hypothetically 2036 ways for an individual to meet the criteria (Edenberg et. al 2018). Also, DSM-V criteria are very different from previous DSM-IV criteria; therefore, it can be difficult to reproduce older data under new definitions (Edenberg et. al 2013, Edenberg et. al 2018). Additionally, different studies use different inclusion and exclusion criteria and often use

different alcohol phenotypes (response level, alcohol dependency, AUD, etc.) making reproducibility difficult (Edenberg et. al 2013, Edenberg et. al 2018)

Given these challenges, there are a limited number of genes with consistent evidence linking them to AUD. There is very reproducible evidence of the involvement of ADH (alcohol dehydrogenase) and ALDH (aldehyde dehydrogenase) genes in risk of developing alcohol dependence (Edenberg et. al 2018). These genes are involved in the metabolism of ethanol to aldehyde and aldehyde to acetate, respectively (Edenberg et. al 2018). There are seven ADH genes; proteins corresponding to six have been isolated *in vivo* (Edenberg et. al 2018). Class I ADHs consist of ADH1A, ADH1B and ADH1C and have been shown to have significant roles in ethanol metabolism and together, can affect risk for development of alcohol dependence (Edenberg et, al 2018). These genes are 90% identical and arose from gene duplication events (Edenberg et. al 2018). Given their similarity, the gene products can heterodimerize (Edenberg et. al 2018). Variants in these chromosomal regions are generally in linkage disequilibrium and inherited together. ADH1B is highly expressed in the liver and is believed to be most involved in the oxidation of alcohol to aldehyde (Edenberg et. al 2018). ADH1B has three isoforms: ADH1B*1, ADH1B*2 and ADH1B*3. Each isoform is associated with specific rates of alcohol oxidation, and frequencies of these alleles are population specific (Edenberg et. al 2018). For example, ADH1B*1 metabolizes alcohol the most slowly and is most commonly observed globally (Edenberg et. al 2018). It is associated with a 3-fold increase in alcohol dependence risk compared to ADH1B*2 (Edenberg et. al 2018). ADH1B*2 is associated with rates of metabolism 11x that of the ADH1B*1 isoform and candidate gene studies strongly show that it is associated with a protective effect against alcohol dependence in Asian populations (Edenberg et. al 2018). However, there is heterogeneity in allele frequencies and amount of protection, even within Asian populations (Edenberg et. al 2018). It is difficult to assess the impact of this isoform on other populations, as its frequency is very low in European and African populations. GWAS has also shown that ADH1B*2 is associated with Alcohol Use Disorder Identification

Test (AUDIT) scores (Sanchez-Roige et. al 2019). AUDIT is a screening tool made up of ten items spanning three dimensions (consumption (AUDIT-C), dependence and problematic alcohol use (AUDIT-P). GWAS of UK BioBank and 23and Me for AUDIT-C and AUDIT-P identified that ADH1B*2 was strongly associate with AUDIT (Sanchez-Roige et. al 2018; Edenberg et. al 2018). The third isoform, ADH1B*3 is primarily found in African populations and is also associated with a protective effect against alcohol misuse (Edenberg et. al 2018). ADH1B*2 and ADH1B*3 are both coding variants that increase kinetic activity of ADH1B (Edenberg et. al 2018). ADH1C, which is expressed in the liver at about 30% the expression of ADH1B has two isoforms: ADH1C*1 shows protective effect against alcoholism, has high metabolic activity and is more highly expressed in various population groups than ADH1C*2 (Edenberg et. al 2018).

ALDHs consist of 19 genes, although ALDH2, ALDH1A and ALDH1B are the primary genes that are involved in the irreversible oxidation of acetylaldehyde to acetate (Edenberg et. al 2018). Gene products from all three form homotetramers. ALDH1A1 and ALDH1B1 have minor effects on risk of developing alcohol dependence (Edenberg et. al 2018). ALDH2 is ubiquitously expressed with high expression levels in the liver (Edenberg et. al 2018). It has two isoforms: ALDH2*1 and ALDH2*2. ALDH2*1 is most commonly expressed in individuals globally (Edenberg et. al 2018). ALDH2*2 is primarily found in some Asian populations and is protective against alcoholism (Edenberg et. al 2018). It makes the ALDH2 homotetramer inactive, causing individuals with the mutation to have higher blood aldehyde levels. Consequently, when individuals with the ALDH2*2 variant consume alcohol, there is a toxic buildup of acetylaldehyde, leading to symptoms such as severe flushing, nausea and increased skin temperatures (Edenberg et. al 2018). These negative side effects cause people with the mutation to generally consume less alcohol and therefore, decrease the risk of developing alcohol dependence.

Recent GWAS (genome wide association study) performed with large sample sizes have identified genes of strong interest for further molecular study. Sanchez-Roige et. al performed a GWAS on AUDIT-C and AUDIT-P scores from a UK BioBank and 23andMe population and identified several associated risk loci, some of which map to the following genes: *LINC01833*, *GCKR*, *KLB*, *METAP1*, *JCAD*, and the alcohol metabolism gene *ADH1C* (Sanchez-Roige et. al 2019). Another GWAS studying the genetic etiology of alcohol and tobacco use had a population cohort of 1.2 million individuals (Liu et. al 2019). Liu et. al identified several genes associated with drinks consumed per week, including: *ADH1B*, *GCKR*, *SLC39A8*, *SERPINA1*, *ACTR1B*, *TNFSF12-13* and *HGFAC* (Liu et. al 2019). Genes identified by GWAS, and especially those that are tagged by multiple GWAS could be meaningful candidates for further study in model organisms.

2. *Drosophila melanogaster* as a model organism for alcohol-related behaviors

2a. Conservation between D. melanogaster and humans

Given the complexity and potential confounding social factors of studying AUD in humans, model organism studies have become a key platform to identify and investigate the contribution of genes thought to be involved in alcohol-related behaviors and molecular mechanisms. The fruit fly, *Drosophila melanogaster* is a powerful model organism to functionally and mechanistically characterize the function of various genes for several reasons. Flies possess orthologs for 75% of human disease-causing genes (Yamaguchi & Yoshida 2018; Engel et. al 2019) and have 80% conserved functional protein domains (Yamaguchi & Yoshida 2018) *D. melanogaster* are low cost, are easy to maintain, take up little space and have fast generation times (Engel et. al 2019). It is possible to perform many types of high throughput genetic analyses and easily manipulate individual genes of interest. Additionally, a breadth of bioinformatics resources to perform various analyses with fly-derived data exist (Engel et. al 2019).

Flies are a useful model organism to study alcohol-related behaviors due to shared nervous system molecular machinery (reviewed in Grotewiel & Bettinger 2015, Engel et. al 2019). Additionally, flies exhibit conserved behavioral responses to alcohol exposure as humans: when exposed to low doses, both species show increased psychomotor and locomotor activity. High doses lead to sedation in both species (Engel et. al 2019). Like humans, flies also develop tolerance and withdrawal to alcohol as well as reward learning and memory behaviors (Engel et. al 2019). Though the fly is extremely short lived compared to humans, it is important to note that relatively, the time to develop tolerance and withdrawal symptoms is similar to humans (Adkins et. al 2017).

2b. Approaches used to investigate genes contributing to alcohol behaviors in flies

Specific phenotypes associated with AUD, including alcohol sensitivity, tolerance and preference can be studied in flies through the use of a diverse range of behavioral assays. Initial sensitivity to alcohol is one of the most compelling predictors of whether a person later develops alcohol use disorder (Schuckit 1997, Engel et. al 2019). Thus, measuring the amount of time it takes flies of different genotypes to become sedated upon ethanol exposure (which is essentially a measure of ethanol sensitivity) can identify genes that contribute to the development of AUD. Several methods exist to measure ethanol sedation times. One example is the inebriometer. An inebriometer device is essentially a glass column with mesh slats that flies can attach to (Sass et. al 2020; Berger et. al 2004). Once flies become sedated, they drop to the bottom of the column, and the number of sedated flies is measured every minute (Sass et. al 2020; Berger et. al 2004). This method has historically been prevalent; however, it is low throughput and inefficient for studying individual flies (Sass et. al 2020; Berger et. al 2004). Sedation can also be measured by exposing groups of flies to ethanol vapor and recording the number of flies that become sedated at predetermined time intervals. This data can be used to calculate the Sedation Time 50 (ST50), or the amount of time required for 50% of flies in a

group to become sedated. Increased ST50 times indicate increased resistance to ethanol, and vice versa (Sandhu et. al 2015). This method is higher throughput, allows multiple genotypes to be assessed at once and is group based.

Like humans, flies can develop tolerance to alcohol. Generally, tolerance can be thought of as needing an increased amount of the drug to obtain the same response after chronic or multiple exposures, or the same amount of drug eliciting a lower response level (Engel et. al 2019, Berger et. al 2004,). Humans and flies can develop three types of tolerance: acute tolerance, rapid tolerance and chronic tolerance. Acute tolerance is the development of tolerance within a single session of consumption (for example, binge drinking), rapid tolerance is a decrease in the intensity of response after recovery from a first alcohol exposure and chronic tolerance is developed from multiple repeated exposures (Engel et. al 2019). In flies, developing tolerance can look like exhibiting lower levels of locomotor impairment or taking longer to become sedated. Rapid tolerance can be measured using sedation assays. Flies undergo a first exposure to alcohol and are then allowed to recover in the absence of alcohol. They are then re-exposed. ST50's from the first and second exposure are compared to determine whether they have developed rapid tolerance. ST50 values from the second exposure being significantly higher than the first is evidence of rapid tolerance (Engel et. al 2019, Berger et. al 2004). Flies can also be raised on alcohol containing food to model chronic tolerance (Engel et. al 2019).

Preference assays build on the idea that the rewarding properties of alcohol contribute to the development of AUD (Engel et. al 2019). An example of an assay to assess fly preference is the proboscis extension response (PER) (Kaun et. al 2011). This assay consists of fixing individual starved flies to a plate. Flies are offered small amounts of alcohol and nonalcoholic food and the rate at which they extend their proboscii is measured. Increased rate of extension is associated with more appetitive substances (Kaun et. al 2011, Shiraiwa 2007). Another example is the two-choice Capillary Feeder (CAFÉ) assay. Flies are able to choose which of two capillaries containing ethanol and non-ethanol containing liquid food they would like to feed

on (Devineni et. al 2009, Engel et. al 2019). Results are measured by calculating the preference index (PI); a value that quantifies the ratio of alcoholic food consumption relative to total consumption (Devineni et. al 2009). Positive and negative PI scores indicate ethanol preference and repulsion respectively. CAFÉ assays have shown that flies increase their preference for alcohol over time: a study conducted by Devnineni et. al found that flies' PI value for alcohol consumption increased steadily over the five days they were exposed to alcohol containing food in the experiment and self-administered alcohol to pharmacologically relevant levels (Devineni et. al 2009). Flies were offered a choice between alcohol containing food laced with quinine, an aversive substance to flies, and regular food. Flies developed a preference for quinine laced alcoholic food compared to non-alcoholic food, showing that they overcame negative stimuli to access alcohol (Devineni et. al 2009, Engel et. al 2019).

Humans and flies both experience hyperactive nervous systems and seizures as a potential consequence of alcohol withdrawal. Alcohol is a depressant of the central nervous system (CNS). Chronic alcohol exposure results in long term excitatory neural adaptations to maintain homeostatic balance (Robinson et. al 2013). Upon withdrawal however, these adaptations result in an overactive nervous system, consequently causing seizures. Robinson et. al found that *D. melanogaster* larvae raised on alcohol containing food show hyper-excitable nervous systems when they are removed from alcohol containing food (Robinson et. al 2013). Seizure susceptibility can be measured in flies by using electrodes to shock the fly brain and induce seizures. Flies undergoing withdrawal were shown to need a lower stimulus voltage to induce seizures (Ghezzi et. al 2014). These experiments show that flies can be powerful models to study withdrawal.

Relapse is commonly observed in individuals attempting to cease problematic alcohol consumption (Melemis 2015). Devenini et. al showed that flies can be used to study relapse behavior by conducting a CAFÉ assay where flies were given the choice between alcoholic and nonalcoholic food (Devineni et. al 2009). Once flies demonstrated a preference for alcohol

containing food, they were deprived of it. During this period, PI values neared 0, since all food tubes available to flies were nonalcoholic. However, when alcohol was reintroduced PI values quickly returned to peak values indicating that flies are able to maintain strong memories of ethanol (Devineni et. al 2009).

2c. Genetic manipulation of the fly genome

There are various approaches to manipulate the fly genome to study underlying genetic mechanisms that may be involved in behavioral alcohol response. A fundamental strategy is the GAL4-UAS system, a bipartite strategy that allows for ectopic expression or knockdown of essentially any transgene or gene in various tissues (Southall et. al 2008, Caygill et. al 2016, Duffy et. al 2002). *GAL4* is endogenously expressed and regulates galactose metabolism in yeast (Caygill et. al 2016). The GAL4 protein consists of 881 amino acids and contains an N terminal DNA binding domain and C terminal transcription activation domain (Duffy et. al 2002). GAL4 dimerizes and binds the Upstream Activation Sequence (UAS), a specific 17-nucleotide sequence. Mediator and essential transcription machinery are recruited, leading to transcription. This activity is retained when GAL4 is expressed in *D. melanogaster* (Southall et. al 2008, Caygill et. al 2016).

In flies, the system requires two transgenic parental lines: the gene of interest is expressed under the control of the UAS, while the other parental line contains the GAL4 driver. When these lines are crossed, the gene of interest, or an RNA interference (RNAi) is expressed in the progeny (Caygill et. al 2016). The UAS-GAL4 system has several advantages that make it an attractive method. Since it is a bipartite system and GAL4 and UAS are in distinct parental lines, it is possible to use a particular UAS with various GAL4 drivers to study the effect of that specific gene in multiple tissue types (Southall et. al 2008). On the other hand, it is possible to use a singular GAL4 driver to study the effect of various genes in a particular tissue type. In the

absence of GAL4, UAS is generally silent or expressed at low levels. Therefore, this system can be used to study toxic or apoptotic proteins (Southall et. al 2008).

Apart from expressing genes of interest, this system can also be used to express RNAi. RNAi is an endogenous cellular mechanism that results in the degradation of RNA molecules. When double stranded RNA (dsRNA) is detected, a protein complex containing DICER is formed (Heigwer et. al 2018). This complex degrades the dsRNA into 21 bp fragments. These dsRNA fragments bind argonaute that subsequently binds to other proteins, forming the RNAi-induced silencer complex (RISC) (Heigwer et. al 2018). RISC is able to recognize mRNA complementary to the dsRNA fragment bound by it, and degrade it, effectively silencing the gene of interest (Heigwer et. al 2018). This system can be applied in coordination with the UAS-GAL4 system. When RNAi is under the control of UAS, the gene of interest will essentially be knocked down when GAL4 is present.

There are several extensions of the UAS-GAL4 system. For example, GeneSwitch (GS) uses a modified GAL4 driver to induce gene expression at specific times (Osterwalder et. al 2001). In this inducible system, GAL4 is in an inactive conformation until exposure to a steroid, mifepristone (RU486). Upon mifepristone treatment, GAL4 conforms to its active state, binds UAS to activate the gene or RNAi of interest (Osterwalder et. al 2001). There are several advantages of this system compared to traditional UAS-GAL4 in regards to flexibility in the timing of gene expression/knockdown. Genes can be variably expressed in different life stages, allowing for the study of specific genes at particular developmental stages. By using this system, it is possible to conduct experiments involving genes that are essential to developmental processes, and for which constitutive knockdown would be lethal (Osterwalder et. al 2001).

2d. Genes involved in Drosophila melanogaster alcohol behaviors

Upwards of 150 fly genes with roles in alcohol-related behaviors have been identified (reviewed in Grotewiel & Bettinger 2015). Some of those genes that have been well characterized will be discussed below. *Mef2* (myocyte enhancer factor 2), the central gene for my thesis project, is known to modulate ethanol sensitivity and is discussed in the subsequent sections.

Autism susceptibility gene (*AUTS2*) is a human gene that has been implicated in several neurological disorders including autism spectrum disorder and developmental delay (Schumann et. al 2011). *AUTS2* has been identified as relevant to alcohol-related behaviors in multiple species. A GWAS studying alcohol consumption identified an intronic SNP in *AUTS2* as significantly associated with alcohol consumption in humans (Engel et. al 2019, Schumann et. al 2011). In mice, the gene was identified between high and low alcohol preferring (HAP1 and LAP1, respectively) lines (Engel et. al 2019). *Tay*, the fly ortholog of *AUTS2* is a negative regulator of the epidermal growth factor receptor (EGFR) pathway (Engel et. al 2019, Schumann et. al 2011, Morozova et. al 2015). Mutations or neuronal RNAi leading to decreased expression of *tay* is associated with a decreased sensitivity to alcohol. Thus, *AUTS2* may play a role in AUD via affecting alcohol sensitivity (Engel et. al 2019, Morozova et. al 2015).

Other examples are the ALDH and ADH genes. Like humans, naturally occurring variation has been observed in these genes and they are involved in ethanol metabolism. Adh^F has higher enzymatic activity than Adh^S and is associated with flies that are more resistant to alcohol (Edenberg et. al 2018). This is similar to naturally occurring isoforms of human ADH1B/A that are associated with higher or lower enzymatic activity and consequently, altered risk of developing alcohol dependence.

Many genes of interest have been identified via GWAS. While GWAS is a powerful method of identification, it has some limitations (Engel et. al 2019). For example, GWAS cannot capture genes because they exhibit changes in expression due to chronic alcohol exposure

(Engel et. al 2019). Transcriptomics level experiments like microarray and RNA-seq are well suited for this. Microarrays have advanced quite a bit since their inception; however, they are still limited by the number of probes available (Engel et. al 2019, Hitzemann et. al 2013). Unlike microarray, RNA-seq is not limited by probe quantity, and well suited to identify transcriptional complexities such as alternative splicing and non-coding RNA (Hitzemann et. al 2013). *Slowpoke (slo)* is a BK channel and ethanol exposure induces expression of the gene in the nervous system. Alternative splicing in the *slo* gene has been found to mediate tolerance in flies (Cowmeadow et. al 2015).

These types of genomic and transcriptomic level studies have not only been crucial to identifying individual genes involved in the phenotype, but have also provided insight into the functional networks within which these genes operate. Networks regulating metabolic activity, stress pathways, chromatin remodeling and immune response, among others, have all been implicated in alcohol behavior (Morozova et. al 2015). For example, Ghezzi et. al and Krishnan et. al used Chromatin immunoprecipitation sequencing (ChIP-seq) to elucidate histone acetylation changes related to mutations in the *slo* genes (Ghezzi et. al 2014, Krishnan et. al 2016). 6b is an element of *slo* that is important for behavioral tolerance. ChIP-seq showed that flies that had been sedated with ethanol vapors and allowed to recover showed a spike in 6b acetylation from the 6-hour mark until the 48-hour time point (Krishnan et. al 2016). The antibody used tags all acetylated histone H4, but spikes were only observed in element 6b. Normal flies show effects of rapid tolerance up to 14 days post initial sedation, but flies with mutated element 6b (*slo*Δ6b) showed tolerance lasting at least 21 days (Krishnan et. al 2016). Therefore, *slo* may play a role in developing ethanol tolerance.

Other genes, such as *icarus (ics)* have been shown to affect ethanol sensitivity. *Ics* encodes Rsu1, the fly ortholog of human *RSU1* (Ras suppressor 1) (Ojelade et. al 2015). *RSU1* is associated with AUD and lifetime frequency of drinking in adults and adolescent populations, respectively (Ojelade et. al 2015). In flies, interrupting *ics* by a P-element leads to decreased

ethanol sensitivity and expression of UAS-Rsu1 led to rescue of the phenotype (Ojelade et. al 2015). This could potentially be a therapeutic avenue to explore further.

Overall, given the ability to model specific genetic questions and the number of alcohol related genes identified in flies, they are a powerful method to investigate the molecular genetic mechanisms underlying the pathogenesis of AUD in humans.

3. Myocyte enhancer factor 2 (*Mef2*)

3a. Mef2 in vertebrates and Drosophila melanogaster

In vertebrates, the myocyte enhancer factor-2 (MEF2) family encodes a transcription factor family with myriad functions including key roles in myogenesis and morphogenesis of skeletal, cardiac and smooth muscle cells (Pon & Marra 2016). Evolutionarily, these proteins belong to the ancient MADS (MCM1, agamous, deficiens, serum response factor) box family (Black & Olson 1998, Pothoff & Olson 2007, Sivachenko et. al 2013). Mammals have four MEF2 genes: MEF2A, MEF2B, MEF2C and MEF2D. Each gene encodes a protein with a highly conserved MADS-box and MEF2 domains at the N terminus (Pon & Marra 2016). These domains are crucial to dimerization, DNA binding ability and cofactor binding. Between species, these areas are more than 80% conserved (Crittenden et. al 2018). C-terminal transcriptional activation domains are less conserved and can undergo complex splicing patterns. In mammals, MEF2's role is highly dependent on which cofactors are present (Crittenden et. al 2018). For example, in culture, MEF2 and Nkx2.5, MEF2 and MASH1 and MEF2 and myogenin coactivate to induce cardiac muscle formation, neuronal phenotypes and skeletal muscle differentiation, respectively (Crittenden et. al 2018)..

Given this complexity, studying MEF2 in *Drosophila,* which possess a single copy of the gene (*Mef2)* can help understand MEF2's conserved roles in a streamlined way. Like mammalian species, *Drosophila Mef2* has been shown to be essential for the differentiation of numerous cell lines (Crittenden et. al 2018). Neuronally, *Mef2* is localized in Kenyon neurons,

which make up the mushroom bodies in the fly brain. *Mef2* is crucial to mushroom body formation in the embryonic fly brain as null or hypomorphic *Mef2* mutants exhibit significantly fewer differentiated mushroom body neurons (Crittenden et. al 2018). Additionally, adult flies with *Mef2* mutations were found to have abnormal wings showing that the gene is essential to normal wing development (Crittenden et. al 2018). *Mef2* has also been shown to be required for the daily fasciculation/defasciculation cycle in small ventral lateral neurons (s-LNv) and contribute clock information to neuronal remodeling machinery (Sivachenko et. al 2013).

As previously detailed, *Mef2* is a transcription factor (Black & Olson 1998). Sivachenko et. al conducted a ChIP-seq using fly brains that revealed that 342 fly genes were bound by *Mef2.* (Sivachenko et. al 2013) These fly genes are orthologous to over 500 human genes. *Mef2* and the genes it binds are the main focus of this thesis project.

3b. Mef2 influences ethanol sedation in Drosophila melanogaster

Mef2 has been shown to influence ethanol sedation in *D. melanogaster*. As previously outlined, initial resistance to ethanol is a key predictor of later alcohol abuse in humans. The SRE is a questionnaire designed to elucidate whether a person has a high or low initial sensitivity to alcohol (Schuckit 1997). A meta-analysis of two population based GWAS studies of SRE interrogating upwards of 18,000 genes showed that 37 were nominally significant ($p < 0.001$) for SRE (Schmitt et. al 2019). Schmitt et. al found that 29 of those 37 genes had appropriate *Drosophila* orthologs and ultimately selected nine human genes (APP, BORC8, MEF2B, GPD2, ISL1, PCDH15 and SFSWAP) to follow up on based on reports suggesting their potential involvement in behavior or neurological disease (Schmitt et. al 2019).

The nine chosen genes are orthologous to 12 fly genes (Schmitt et. al 2019). RNAi transgenes against each gene of interest was expressed in fly neurons using the GAL4-UAS system with a neuron specific GAL4 driver (*elaV*-GAL4) and UAS-RNAi construct (Schmitt et. al 2019). Sedation experiments were performed for each RNAi. Three separate RNAi transgenes

against and mutations in *Mef2* were shown to increase ST50 values compared to controls (Schmitt et. al 2019). Additionally, another study reported that pan-neuronal expression of *Mef2* RNAi increases ethanol sensitivity and dominant negative *Mef2* in all neurons or mushroom bodies alone decreases tolerance (Adhikari et. al 2018). This study also reported that *Hr38*, a gene downstream of *Mef2*, influences ethanol tolerance and preference behaviors (Adhikari et. al 2018)

4. Significance

As previously discussed, AUD is a significant public health concern and a leading cause of injury, development and progression of various physical and mental diseases, loss of productivity and preventable death in the US and abroad (NIAAA 2016, WHO 2018). Treatment options are limited. Only three medicines are approved to treat AUD: disulfiram, naltrexone and acamprosate (Kranzler & Soyka 2018). These drugs have distinct mechanisms. Disulfiram for example, causes negative symptoms like nausea and skin flushing when an individual consumes alcohol. If a person knows that drinking will cause these side effects, they may be deterred from consumption. Naltrexone and Acamprosate are aimed at curbing cravings by targeting neural reward systems (Kranzler & Soyka 2018). However, these options are limited and rates of relapse are high. Behavioral counseling is also a viable treatment option; however, many rates of relapse are high (Kranzler & Soyka 2018, Melemis 2015).

Given the level of conservation between humans and *D. melanogaster*, experiments in flies can be powerful avenues to identify candidate genes relevant to alcohol related behaviors. AUD is a complex disease, and better understanding the molecular genetic mechanisms underlying pathogenesis of the disease is a springboard for developing more robust therapeutic interventions to help those experiencing AUD or other forms of alcohol dependence in the future.

CHAPTER 2: RNAi SCREEN OF CANDIDATE GENES DOWNSTREAM OF *MEF2*

1. Introduction & rationale

Studies in the Grotewiel laboratory demonstrated that mutations in *Mef2* or neuronal expression of RNAi transgenes against *Mef2* decrease ethanol sensitivity in flies (Schmitt et. al 2019). Additionally, Wolf and co-workers (Adhikari et. al 2018) reported that pan-neuronal expression of Mef2 RNAi increases ethanol sensitivity in flies and that dominant negative Mef2 in all neurons or mushroom body neurons alone decreases ethanol tolerance. This study also showed that genes regulated by *Mef2* play a role in ethanol related behaviors. *Hr38* is a *Drosophila* homolog of the mammalian Nr4a1/Nr4a2/Nr4a3 gene family (Adhikari et. al 2018). In humans, these genes are transcriptionally activated by MEF2 (Adhikari et. al 2018). When drug naive flies were exposed to ethanol, *Hr38* was the only surveyed gene with induced expression (Adhikari et. al 2018). Flies heterozygous for *Hr38* were observed to display significantly altered ethanol tolerance and preference (Adhikari et. al 2018). Similarly to mammals, *Hr38* was transcriptionally induced by *Mef2,* in the presence of alcohol (Adhikari et. al 2018). In other words, alcohol activates *Mef2,* allowing it to induce *Hr38.* Increased levels of Hr38 were associated with higher tolerance and increased preferences for alcohol, but did not have an effect on ethanol sensitivity. Another group performed ChIP-seq on DNA isolated from fly heads and identified 342 Mef2-bound genes (Sivachenko et. al 2013). Exploring the function of genes downstream of *Mef2* in ethanol sedation is a primary goal of this thesis, and particularly, of this chapter.

Given that *Mef2* regulates ethanol sedation (Schmitt et. al 2019, Adhikari et. al 2018), is a transcription factor (Black & Olson 1998; Taylor & Hughes 2017), is known to bind 342 genes in flies (Sivachenko et. al 2013) and at least one gene downstream of *Mef2* (*Hr38*) influences ethanol tolerance and preference (Adhikari et. al 2018), we hypothesized that other genes downstream of *Mef2* play a role in ethanol sedation. We began testing this hypothesis by identifying genes of interest by determining the intersection of (i.e. overlap between) genes

bound by *Mef2,* genes associated with human externalizing behavior (Dick, personal communication 2020, Linnér et. al 2020), genes known to influence fly and worm alcohol behavior (reviewed in Grotewiel & Bettinger 2015), human genes associated with expression changes related to alcohol consumption (Silviu Bacanu, personal communication) and genes related to human/mouse disease linked genes and gene ontology (Michael Miles, personal communication). This approach identified the genes *spin*, *unc79*, *Bx*, *CtBP*, *Fas2* and *For* as being bound by Mef2, associated with externalizing behavior in humans, and (for *spin*) associated with alcohol consumption in humans. These six genes of interest were therefore high priority candidates that I investigated for roles in fly ethanol sedation.

1a. Functions of genes of interest

The functions of the genes of interest are quite varied. *spin* is the fly ortholog of the human *SPNS1* gene. In flies, the *spin* gene encodes a multi-pass transmembrane late endosomal/lysosomal protein (Kim 2017, Sweeney & Davis 2002). *spin* is known to be involved in nervous system development, cell death control and is required for TGF-β signaling (Kim 2017). Additionally, *spin* is involved in eye development via control of glial cell migration in flies (Yuva-Aydemir et. al 2011). Loss of function *spin* mutants displayed patterns of early endosome recycling. This led to accumulation of autophagosomes and autolysosomes, leading to abnormal head growth (Kim 2017). *spin* mutant females are also known to avoid mating (Kim 2017). Previous work by Katlyn Myers in the Grotewiel laboratory found that neuronal expression of RNAi transgenes against *spin* and transposon insertions in or near the *spin* locus made flies resistant to ethanol sedation, strongly suggesting the gene is involved in ethanol sedation sensitivity (Myers 2020).

unc79 is the fly ortholog of the human *UNC79* gene. In flies, it is known to be involved in sleep homeostasis and locomotor activity (Joiner et. al 2013). The fly Narrow Abdomen (NA) ion channel is orthologous to the mammalian NALCN sodium leak channel (Joiner et. al 2013). In

both humans and flies, NALCN/NA function in clock neurons to support behavioral rhythmicity. Loss of function mutations in *unc79* lead to serious defects in circadian locomotor rhythmicity (Lear et. al 2013). Immunoprecipitation and tissue specific RNAi experiments show that *unc79* operates in pacemaker neurons (Lear et. al 2013). *unc79* is also involved in behavioral responses to anesthetics. Flies with loss of function mutations in *unc79* display an increased response to anesthesia (Joiner et. al 2013). *unc79* is also known to be involved in alcohol behaviors in worms (reviewed in Grotewiel & Bettinger 2015). Worms with mutations in *unc-79*, the worm ortholog of the human *UNC79* gene were found to be dramatically hypersensitive to ethanol exposure in regards to swimming behavior (Speca et. al 2010). Additionally, the same study found that mice with a point mutation in the mouse homolog of unc-79 were similarly sensitive to alcohol exposure, as well as exhibiting a high preference for alcohol (Speca et. al 2010), suggesting a conserved function of the gene.

Beadex is the fly ortholog of the human gene *LMO1* (Milan et. al 1998). *Bx* encodes the dLMO transcription factor, which contains two conserved LIM homeodomains (Milan et. al 1998). This gene is known to be involved in cytoskeletal organization, cell specification and differentiation and organ development (Milan et. al 1998). Misexpression of dLMO is associated with disturbed dorsal-ventral boundaries and wing patterning and extraneous wing growth (Zeng et. al 1998). Additionally, flies with a loss of function point mutation in *Bx* showed enlarged abdomens due to decreased gastric emptying (Ren et. al 2014). *Bx* is also involved in blood cell development. Knockdown mutants of *Bx* are associated with an increased crystal cell count (Chatterjee et. al 2019). In addition, *Bx* is involved in follicle cell development. To that end, it has been shown to be essential to female reproduction (Karaimkonda & Nongthomba 2018). Eggs of female flies with decreased levels of *Bx* have multiple defects due to its involvement in follicle cell development (Kairamkonda & Nongthomba 2018). *Bx* is known to be involved in alcohol behaviors in flies (Grotewiel & Bettinger 2015). Flies with loss of function (LOF) mutations in *Bx* display increased resistance to ethanol, whereas flies with an overexpression construct of the

gene exhibit the opposite phenotype (Lasek et. al 2011). This relationship has also been observed in mice, suggesting a conserved effect of the gene (Lasek et. al 2011)

CtBP is the fly ortholog of the human gene of the same name. It is a transcriptional co-regulator and plays a role in regulating gene expression (Hoang et. al 2010). Hoang et. al found that *Drosophila CtBP* interacts with other proteins to mediate gene expression at multiple stages of eye development to prevent over proliferation of eye precursors (Poortinga et. al 1998). Additionally, *CtBP* has been reported to inhibit the Wnt signaling pathway by blocking β-catenin from binding T-cell factors (Poortinga et. al 1998). Interestingly, Fang et. al found that *CtBP* is involved in activating targets of the Wnt signalling pathway and blocks expression of Wnt targets (Fang et. al 2006). *CtBP* is known to be involved in alcohol behaviors in worms (Grotewiel & Bettinger 2015). Worms with mutations in *ctbp-1*, the worm ortholog of the human *CTBP* gene displayed increased time to develop acute functional tolerance, or the normalization of neural function in spite of the presence of alcohol, compared to wild type worms (Bettinger et. al 2012). Additionally, worms with increased expression of *ctbp-1* displayed faster development of AFT and resistance to alcohol (Bettinger et. al 2012).

Fas2 is the fly ortholog of the human neural cell adhesion molecule (*NCAM*). Both are part of the Ig (immunoglobulin) domain superfamily and possess homophilic cell-cell adhesion mediator activity (Neuert et. al 2020). Sivachenko et. al found that *Mef2* negatively regulates *Fas2*, and both genes are involved in circadian rhythm control (Sivachenko et. al 2013) in flies. *Fas2* is a player in neuron-glia signaling and has several isoforms that display different membrane attachment patterns and cytoplasmic domains (Neuert et. al 2020). Additionally, *Fas2* is involved in organ morphogenesis, nervous system development and synapse organization (Neuert et. al 2020). Mao et. al showed that *Fas2* inhibits EGFR (epidermal growth factor receptor) signaling during development of the eyes, wings and notum, as decreased levels of *Fas2* are associated with EGFR hyperactivity phenotypes (Mao & Freeman 2009). *Fas2* is known to be involved in alcohol behaviors in flies (reviewed in Grotewiel & Bettinger

2015). Loss of function *Fas2* mutants displayed increased sensitivity to ethanol vapors when tested using an inebriometer (Cheng et. al 2001). Expression of *Fas2* transgenes did not rescue this phenotype, suggesting that the role the gene plays in ethanol sensitivity is complex (Cheng et. al 2001).

For is the fly ortholog of the human gene *PRKG1*. *For* encodes a cGMP (cyclic guanosine-3', 5'-monophosphate)-dependent protein kinase and is known to be expressed in ellipsoid-body ring neurons in flies (Kent et. al 2009). Previous studies have shown that flies and larvae harboring *for* mutations display altered learning patterns and associative olfactory learning (Kunz et. al 2012). Additionally, *for* is known to be involved in visual pattern memory and visual orientation memory. *for* has naturally occurring isoforms: for^{S}, or "sitter" and for^{R}, or "rover" (Wang et. al 2008). Flies with the sitter variant have lower cGMP-dependent protein kinase levels and display both deficient visual pattern and visual orientation memory (Wang et. al 2008). Rover and sitter variants have altered movement and socialization patterns. As their name implies, rover larvae move further while foraging for food, whereas sitters tend to group together collectively (Wang et. al 2008). Rovers also display longer short-term memory and are more sensitive to heat, hypoxia and starvation, but can resist starvation stress (Wang et. al 2008). On the other hand, sitters show better learning patterns while in groups, showing that *for* is quite important in fly behavior (Wang et. al 2008). It is known to be involved in worm alcohol behaviors (reviewed in Grotewiel & Bettinger 2015). Worms with mutations in egl-4, the worm ortholog of the human PRKG1 gene display normal responses to ethanol when they are drug naive, but do not display symptoms of withdrawal (Mitchell et. al 2010).

2. Materials & methods

2a. Fly husbandry and stocks

Flies were raised in an environmental chamber operating on a 12-hour light/dark cycle at 25°C and 60-65% relative humidity. *Drosophila* food medium consisted of 10% sugar, 3.3% cornmeal, 2% yeast, 1% agar, 0.1 g/L ampicillin, 0.125 g/L chloramphenicol, 2 g/L tegosept, 0.02 g/L tetracycline and live yeast (Schmitt et. al 2019).

Twenty-one RNAi transgenes against the six genes of interest were obtained either from the Bloomington *Drosophila* Stock Center (BDSC, Bloomington, IN) or the Vienna *Drosophila* RNAi Center (VDRC, Vienna, Austria) (Table 1). Whenever possible, multiple RNAi transgenes were ordered per gene of interest. The VDRC online interface was used to assess the number of predicted off-target effects of each RNAi from this stock center. VDRC defines an "on-target" as any gene that is a perfect match to at least 80% of a construct's 19-mers, or 19 nucleotide length sequences (Dietzl et. al 2007). Off-targets are defined as any genes that matches to less than 80% of a construct's 19-mers, but possesses at least one match (Dietzl et. al 2007). Each RNAi is assigned a specificity score (s_{19}), calculated as $s_{19} = \sum$(on target matches) / $\sum$(on target matches + off target matches) (Dietzl 2007). A specificity score of 0.5 indicates a gene that matches perfectly to a construct 50% of the time, or a gene that is considered on-target (Dieztl et. al 2007). Only VDRC RNAi transgenes with less than two predicted off target effects were used.

2b. Identification of fly-human gene orthologs

Fly genes bound by *Mef2* (Sivachenko et. al 2013) were converted to their human orthologs by calculating DIOPT scores using the DIOPT – DRSC Integrative Ortholog Prediction Tool version 8 (Hu et. al 2011, DRSC). DIOPT allows for rapid identification of orthologous genes by integrating several approaches to allow users to identify the most appropriate orthologs for further analysis (Hu et. al 2011). The tool reports several metrics, including a

DIOPT score which is determined by the number of predictive tools that have paired the two orthologs (Hu et. al 2011). Higher DIOPT scores indicate increased tools predicting the two genes as an ortholog pair, and is a metric by which to prioritize output gene lists (Hu et. al 2011). In our work, each gene of interest was entered into the website using the appropriate input and output species and only orthologs with a DIOPT score greater than or equal to 5 were considered further.

2c. Ethanol sedation

Ethanol sedation was assessed as described previously (Sandhu et. al 2015, Chan et. al 2014). One to three-day-old female flies were collected in groups of 11 under brief CO_2 anesthesia. All collected flies visually appeared healthy, were of similar size, had normal wings and were not virgins. Eight vials of flies were collected per genotype for a total of 24 vials of flies per experiment to assess an individual RNAi transgene. Collected flies were allowed to recover overnight in inverted, non-yeasted food vials in the environmental chamber (25°C, 60-65% humidity). Eighty-five % ethanol was made a maximum of 1 week before each sedation experiment. All sedations were conducted in the same behavioral room at 20-23°C, relative humidity at 55-65% and with standard laboratory lighting. Flies were allowed to adjust to the conditions of the room for approximately 60 minutes prior to beginning each experiment.

Prior to beginning each experiment, each vial was randomly assigned an alphanumeric code such that they are in 6 sets of 4 vials each. The experimenter (AT) was blind to genotypes. Flies were transferred from the food collection tubes they had been housed in overnight to correspondingly labelled empty vials. Vials were sealed with a cotton flug. Flugs were pressed down to a uniform height and the number of immobile flies before any ethanol application was recorded. At 0 minutes, ethanol was added to each vial in the first set at five second intervals and sealed with a silicone plug. At 30 seconds, the number of sedated flies in vial 1 was counted, and each subsequent vial in set 1 was counted at 5 second intervals. Ethanol was

added into the second set of vials every five seconds, and so on. Each vial was tapped and counted every six minutes as described by Sandhu et. al (Sandhu et. al 2015).

The ST50, or amount of time it takes for 50% of the flies in a vial to become sedated was ascertained from a sigmoidal curve fits of the ethanol sedation time-course data (Sandhu et. al 2015). Vials were decoded and sorted by genotype. ST50 values were compiled for each genotype.

2d. Testing genes of interest in ethanol sedation

The effect of each RNAi transgene on ethanol sedation was assessed in individual experiments with three genotypes, each derived from three crosses performed in parallel. The *elav*-GAL4 driver stock was used to express each RNAi transgene pan-neuronally. Additionally, two other genotypes, an *elav*-GAL4 control and an RNAi control were also tested for ST50 using the ethanol sedation assay outlined in the following section to determine whether the RNAi-expressing group displayed significant ethanol sedation resistance or sensitivity compared to controls. All crosses are outlined in Table 2. All flies tested were in an F1 hybrid genetic background consisting of 50% w[A] (Grotewiel laboratory stock) and 50% w[VDRC] (the background used to generate most RNAi transgenes).

2e. Statistical analyses

Fisher's exact test was performed in R Studio (R Studio Version 1.4.1717). The exact script used was provided by Dr. Michael Miles' lab and is in the appendix. $p \leq 0.05$ was considered significant. One-way analysis of variance (ANOVA) and Bonferroni's multiple comparisons (BMC) were performed using GraphPad Prism version 9 for each sedation experiment. $P \leq 0.05$ was considered the threshold for statistical significance.

Genotype	Stock #	Description	Source
w1118; P{GD2579}v8392	8392	*Fas2* RNAi	VDRC
w1118; P{GD2579}v8393	8393	*Fas2* RNAi	VDRC
w1118; P{GD14486}v36350	36350	*Fas2* RNAi	VDRC
w1118; P{GD14486}v36351	36351	*Fas2* RNAi	VDRC
P{KK100888}VIE-260B	103807	*Fas2* RNAi	VDRC
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02918}attP2 e[*]	28990	*Fas2* RNAi	BDSC
y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01098}attP2	34084	*Fas2* RNAi	BDSC
P{KK108401}VIE-260B	107313	*CtBP* RNAi	VDRC
y1 v1; P{TRiP.JF01291}attP2/TM3, Ser1	31334	*CtBP* RNAi	BDSC
y1 sc* v1 sev21; P{TRiP.HMS00677}attP2	32889	*CtBP* RNAi	BDSC
w1118; P{GD6843}v38319	38319	*For* RNAi	VDRC
P{KK101298}VIE-260B	101298	*For* RNAi	VDRC
y1 v1; P{TRiP.JF01449}attP2/TM3, Ser1	31698	*For* RNAi	BDSC
y1 sc* v1 sev21; P{TRiP.GL00026}attP2	35158	*For* RNAi	BDSC
y1 sc* v1 sev21; P{TRiP.HMS04486}attP40	57041	*For* RNAi	BDSC
y1 v1; P{TRiP.JF03390}attP2	29454	*Bx* RNAi	BDSC
y1 sc* v1 sev21; P{TRiP.GL00484}attP2/TM3, Sb1	35637	*Bx* RNAi	BDSC
y1 sc* v1 sev21; P{TRiP.HMC04776}attP40	57465	*Bx* RNAi	BDSC
P{KK102682}VIE-260B	108132	*unc79* RNAi	VDRC
w1118; P{GD11587}v45780	45780	*unc79* RNAi	VDRC
y1 sc* v1 sev21; P{TRiP.HMC03213}attP2	51471	*unc79* RNAi	BDSC
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02782}attP2	27702	*spin* RNAi	BDSC
w1118; P{GD5039}v15550	15550	*Mef2* RNAi	VDRC
w[VDRC]	N/A	VDRC control strain, w^{1118}	VDRC
w[A]	N/A	lab control strain, w^{1118}	BDSC

Table 1. Genotypes, stock numbers and ordering source of all stocks used in subsequent experiments.

Virgin ♀	♂	**Group**
elav-GAL4	w[VDRC]	*elav*-GAL4 control
w[A]	RNAi	RNAi control
elav-GAL4	RNAi	Presumed knockdown

Table 2. Three crosses set up and resulting groups. These crosses were set up for each RNAi used.

3. Results & Discussion

3a. Identifying genes of interest

Considering that Mef2 is a transcription factor (Black & Olson 1998; Taylor & Hughes 2017) that influences ethanol behaviors in flies and possibly humans (Schmitt 2019, Adhikari 2019), we hypothesized that genes bound (and presumably regulated) by Mef2 would be good candidates for functioning downstream of Mef2 to regulate ethanol sedation. We also hypothesized that additional human genetic information could be used to prioritize genes bound by Mef2 as downstream mediators of Mef2. Toward testing these two interrelated hypotheses, we used DIOPT (DRSC, Hu et. al 2011) to convert the 342 genes bound by *Mef2* (Sivachenko et. al 2013) to 581 human orthologs. Dr. Danielle Dick's laboratory had identified 928 human genes associated with externalizing behavior (defined as a group of behaviors that are directed outwardly and pertain to self-regulation). Examples include substance abuse, antisocial disorder and poor impulse control (Dick et. al 2020); Danielle Dick, personal communication 2020; Linnér et. al 2020), and further found that 39 of the 581 human orthologs of genes bound by Mef2 were associated with externalizing behaviors (Table 3, first column; Danielle Dick, personal communication 2020). Although Fisher's exact test indicates that the overlap of 39 genes between the human orthologs of Mef2 bound genes and 928 human externalizing behavior genes is not significant (Table 4), these genes might still ultimately prove to be high-priority candidates for roles in ethanol behavior given their connections to *Mef2* and/or externalizing behavior.

I compared the 36 unique fly genes that are bound by Mef2 (Table 3) and are orthologous to 39 human genes that were associated with externalizing behavior (Table 3, Linnér et. al 2020) to a comprehensive list of 91 and genes involved in at least one aspect of fly alcohol behavioral responses (reviewed by Grotewiel & Bettinger 2015). I then converted these genes to their worm orthologs and compared them to a list of genes involved in worm alcohol responses (reviewed by Grotewiel & Bettinger 2015).

Human Gene	*Mef2* bound fly gene	DIOPT Score	p-value
NCAM1	*Fas2*	11	1.59E-48
SEMA6D	*Sema1a*	8	1.00E-39
SDK1	*sdk*	10	6.98E-36
PDE4B	*dnc*	12	2.03E-25
CELF2	*aret*	14	6.65E-20
PRKG1	*for*	13	3.15E-15
PHC2	*ph-p*	6	8.54e-14
BIRC6	*Bruce*	15	2.60E-13
NFAT5	*Nfat*	7	2.67E-11
NFIA	*nfl*	12	5.83E-11
CHD3	*mi-2*	12	25.96E-11
CALB1	*Cbp53e*	10	8.32E-11
FXR1	*Fmr1*	12	3.37E-10
UNC79	*unc79*	14	5.18E-10
CALB2	*Cbp53e*	11	1.08E-09
LONRF2	*CG32369*	12	1.20E-09
FMNL2	*Frl*	11	1.68E-09
PTPRN2	*IA-2*	10	1.75E-09
CYP3A43	*Cyp9f2*	8	2.69E-09
MLLT10	*Alh*	6	3.43E-09
MEF2C	*Mef2*	11	4.09E-09
OAZ3	*Oda*	7	4.54E-09
SCL17A3	*CG3649*	6	5.84E-09
ASPG	*CG6428*	13	6.00E-09
ARIH2	*ari-2*	15	1.20E-08
IGF1R	*lnR*	12	3.34E-08

MAPT	*tau*	7	5.67E-08
MAP4	*tau*	6	8.14E-08
SLC22A12	*CG8654*	6	8.42E-08
ISYNA1	*Inos*	13	1.12E-07
EXT1	*ttv*	13	1.25E-07
PFKFB2	*Pfrx*	11	2.06E-07
SPNS1	*spin*	13	2.09E-07
REEP1	*ReepA*	8	9.14E-07
LMO3	*Bx*	7	1.02E-06
CYP3A5	*Cyp9c1*	8	1.23E-06
CYP3A4	*Cyp6wi*	9	1.49E-06
PDE4D	*dnc*	10	1.71E-06
CTBP1	*CtBP*	12	2.45E-06

Table 3. Thirty-nine human genes implicated in externalizing behaviors, their respective fly orthologs, DIOPT scores and whether the gene is present in the list of the 342 *Mef2* bound fly genes. P-value refers to the Bonferroni corrected p-value of the association of the gene with human externalizing behavior. $P < 2.74E\text{-}6$ is considered significant.

Gene Set 1	Gene Set 2	Genome	Overlap	Expected	Odds Ratio	p-value
581 human orthologs of *Mef2* bound genes	928 human externalizing behavior genes	# genes interrogated (18,235)	39 total	29.5678	1.2070	0.2729

Table 4. Fisher's exact test to determine whether the number of genes overlapping between the human orthologs of Mef2 bound genes and externalizing behavior genes.

This identified six high priority candidate genes (*Bx*, *CtBP, Fas2, For, spin, unc79)* that are all bound by *Mef2* (Sivachenko et. al 2013), have orthologs associated with human externalizing behavior (Dick, personal communication 2020, Linner et. al 2020) and have been previously reported to be involved in or have orthologs that are involved in fly or worm alcohol behaviors (Grotewiel & Bettinger 2015). The steps taken to identify these six candidate genes are outlined in Figure 1.

The gene *spin* was previously examined as a locus involved in ethanol sedation by Katlyn Myers, a student in the Grotewiel lab as part of her Master's thesis. Within human orthologs of the 342 Mef2-bound genes (Sivachenko et. al 2013), a GSCAN (GWAS & Sequencing Consortium of Alcohol and Nicotine Use, Dr. Silviu Bacanu, personal communication) identified 15 genes including *spin* as being nominally associated with gene expression changes associated with alcohol consumption (Bacanu, unpublished). Katlyn found that flies expressing an RNAi transgene targeting *spin* pan-neuronally and flies with transposon insertions near or in the *spin* locus had increased ST50 values (Myers and Grotewiel, unpublished). We also searched for our other five candidate genes (*Bx, CtBP, Fas2, For, unc79*) within the GSCAN data. Additionally, we searched for our six candidate genes within data from a compilation of gene ontology and human/mouse disease linked genes from Dr. Miles' lab (Michael Miles, personal communication). These additional analyses via Drs. Bacau and Miles did not further implicate the candidate genes in ethanol behavior. In summary, we identified six genes or orthologs of genes for further study that are bound by Mef2, known to be involved in fly or worm alcohol behavior, and associated with externalizing behavior (Table 5). One of these six genes, *spin*, is also associated with gene expression changes related to alcohol consumption (Table 5).

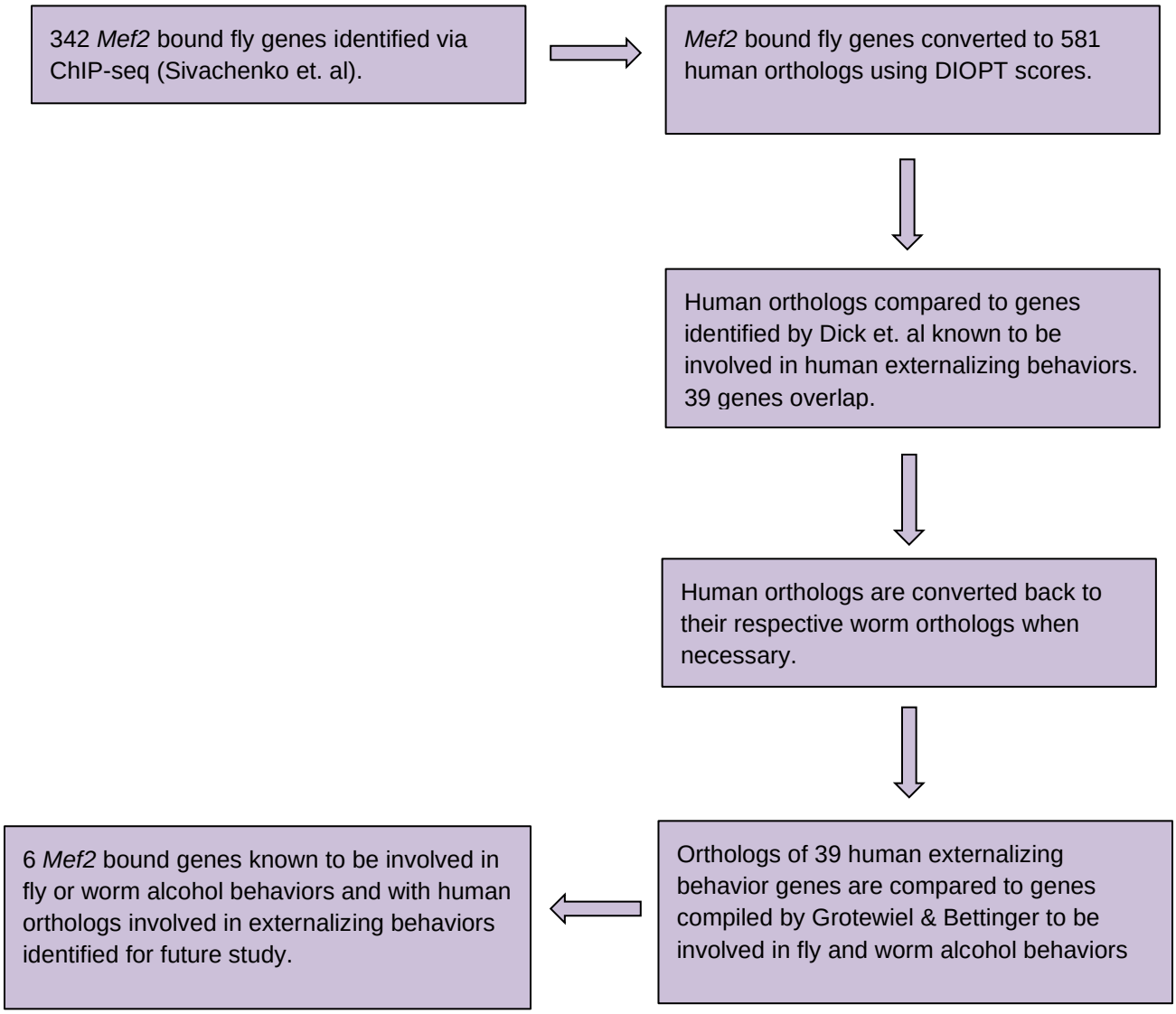


Figure 1. Outline of the steps taken to identify high priority candidate genes for possible roles in ethanol sedation. Initial *Mef2* bound genes identified by Sivachenko et. al (Sivachenko 2013) were converted to human orthologs, filtered by human genes involved in human externalizing behavior, converted back to fly and worm orthologs and then filtered against lists of genes known to be involved in fly and worm alcohol behaviors.

Fly gene	**Human gene**	**Mef2 bound?**	**Previously implicated in fly/worm alcohol behavior?**	**Dick** (rank, p value)	**Bacanu** (rank, p value)	**Miles** (rank, p value)
Fas2	*NCAM1*	yes	yes – fly	2 (1.44E-48)	50 (0.082)	not present
CtBP	*CTBP*	yes	yes – worm	798 (5.36E-07)	65 (0.139)	149 (0.008)
Bx	*LMO3*	yes	yes – fly	773 (4.27E-07)	61 (0.112)	254 (0.026)
For	*PRKG1*	yes	yes – worm	73 (8.44E-15)	403 (1)	not present
unc79	*UNC79*	yes	yes – worm	358 (1.20E-09)	440 (1)	427 (0.116)
spin	*SPNS1*	yes	yes – fly	575 (5.11E-08)	4 (6.78E-08)	293 (0.038)

Table 5. Six candidate genes, their human orthologs, their past implication in fly or worm alcohol behavior, ranking and adjusted p-values from Dr. Dick's analyses of human externalizing behaviors and Dr. Bacanu's work on gene expression changes related to alcohol consumption. Rank and p-value represent the represent the relative order of the gene on gene of interest, and all p-values indicate that each gene is significantly associated with the phenotype being studied.

3b. Confirmation and testing of roles for genes in ethanol sedation

To confirm that neuronal expression of RNAi targeting Mef2 makes flies resistant to ethanol sedation as the Grotewiel laboratory previously reported (Schmitt et. al 2019), I assessed ethanol sedation in flies with the *elav*-Gal4 (pan-neuronal driver) expressing a validated RNAi transgene against *Mef2* (v15550). Expression of v15550 in neurons (*elav*-Gal4/+;v15550/+) increased ST50 values compared to both controls (elav-Gal4/+ and v15550/+; Figure 2A). Knockdown of *Mef2* in neurons therefore made flies resistant to ethanol sedation in my studies as previously reported by the Grotewiel laboratory (Schmitt et. al 2019).

To test the role of the 6 candidate genes (Table 5) in ethanol sedation, we expressed RNAi transgenes targeting each of the genes in neurons via *elav*-Gal4 with the same overall approach as used for *Mef2* (Figure 2A). Expression of *spin* RNAi (*elav*-GAL4/+;JF02782/+) increased ST50 compared to both the *elav*-GAL4/+ and JF02782/+ controls (Figure 2B). My data on *spin* JF02782 RNAi confirm those of Katlyn Myer (Myers and Grotewiel, unpublished). Furthermore, Katlyn also found that transposon insertions in the *spin* locus increased ST50 values. Together, data from my and Katlyn's studies strongly suggest that *spin* influences ethanol sedation in flies.

Regarding the other 5 candidate genes (Table 5), expression of the *unc79* v45780 RNAi transgene (*elav*-GAL4+/v45780/+) significantly increased ST50 compared to the *elav*-GAL4 control, but not the v45780/+ control (Figure 3A). Neuronal expression of the KK102682 RNAi transgene (*elav*-GAL4+/KK102682/+) increased ST50 compared to both the *elav*-GAL4/+ and KK102682/+ controls (Figure 3B). The standard ethanol sedation protocol used for the studies in Figures 3A and 3B in the Grotewiel laboratory exposes flies to vapor from 85% ethanol. To explore the possibility that the effect of expressing *unc79* RNAi might depend on the concentration of ethanol used, these experiments were repeated using vapor from 65% ethanol. At this lower concentration of ethanol, expression of both RNAi transgenes (v45780 and

KK102682) significantly increased resistance compared to both controls (Figures 3C and 3D). Expression of another *unc79* RNAi transgene (*elav*-GAL4/+;HMC03213/+) was lethal. Taken together, these experiments show that *unc79* may influence ethanol sedation, as multiple RNAi transgenes against the gene result in significant increases in ST50.

Experiments with RNAi transgenes targeting the remaining four genes were less informative. Expression of a *Bx* RNAi transgene (*elav*-GAL4/+;JF03390/+) and a *CtBP* RNAi transgene (*elav*-GAL4/+;JF01291/+) did not significantly increase resistance or sensitivity to ethanol compared to the respective controls (Figures 4 and 5). Expression of other *Bx* (*elav*-GAL4/+;KK108513/+, *elav*-GAL4/+;GL00484/+) and *CtBP* RNAi transgenes (*elav*-GAL4/+;KK108401/+, *elav*-GAL4/+;HMS00677/+) were lethal. Expression of the v8393 *Fas2* RNAi (*elav*-GAL4/+;v8393/+) significantly increased resistance compared to both controls (Figure 6A). Expression of four other *Fas2* RNAi's, (*elav*-GAL4/+;v36350/+; *elav*-GAL4/+;v8392/+; *elav*-GAL4/+;HMS01098/+ and *elav*-GAL4/+;JF02918/+) significantly changed ST50 values compared to one control (Figures 6B, 6C, 6D, 6E), but the direction of the change (resistance vs. sensitivity) was not consistent. Expression of another *Fas2* RNAi transgene, v36351 (*elav*-GAL4/+;v36351/+) did not produce any significant changes (Figure 6F) and the KK100888 transgene (*elav*-GAL4/+;KK100888/+) was lethal. Expression of the *for* RNAi transgenes v38319 and GL00026 (*elav*-GAL4/+;v38319/+ and *elav*-GAL4/+;GL00026/+) significantly increased resistance compared to one control group (Figure 7A, D), whereas expression of two other *for* RNAi transgenes JF01449 and KK101298 (*elav*-GAL4/+;JF01449/+ and *elav*-GAL4/+;KK101298/+) did not produce significant changes (Figure 7B, C). Expression of another *for* RNAi transgene, HMS04486 (*elav*-GAL4/+;HMS04486/+) was lethal. Overall, my data on *Bx, CtBP, Fas2* and *for* did not support the hypothesis that these genes function in neurons to regulate ethanol sedation.

3c. Discussion

spin, unc79, Bx, CtBP, Fas2 and *for* are all *Mef2*-bound genes that were previously shown to influence ethanol behaviors in flies or (via orthologs) worms and have human orthologs that were associated with externalizing behavior. They were therefore lead candidates for influencing ethanol sedation by functioning downstream of *Mef2*. Expression of RNAi transgenes targeting *spin* and *unc79* (Figure 2 and Figure 3) consistently changed ST50 values compared to RNAi and *elav*-Gal4 controls. All viable RNAi transgenes against *unc79* produced an effect in the same direction and at two concentrations, though the v45780 (*elav*-GAL4/+;v45780+) was significantly different from both control genotypes only when using vapor from 65% ethanol. These results implicate *spin* and *unc79* in ethanol sedation and also raise the possibility that ethanol sedation (assessed in *Drosophila*) and externalizing behavior (assessed in humans) might be driven or influenced by shared genetic mechanisms as similarly suggested by studies in worms and humans (Mathies et. al 2017).

Expression of neuronal RNAi targeting the other candidate genes did not consistently impact ST50. This lack of effect could be explained by several possibilities including (i) the genes do not affect ethanol sedation by functioning in neurons, (ii) the RNAi transgenes did not sufficiently knockdown expression of their target genes or (iii), that lethality associated with constitutive, pan-neuronal expression RNAi transgenes is preventing us from observing the phenotype. To better understand the potential role of these genes in ethanol sedation, future studies could include expression of the transgene only in adulthood or with a different GAL4 driver that would express the transgene in select neurons, rather than pan-neuronally, thereby circumventing lethality associated with elav-Gal4 pan-neuronal expression. Additionally, genetic manipulation of the genes via mutations, overexpression or expression of dominant negatives could also provide insight into the role of these genes in ethanol sedation.

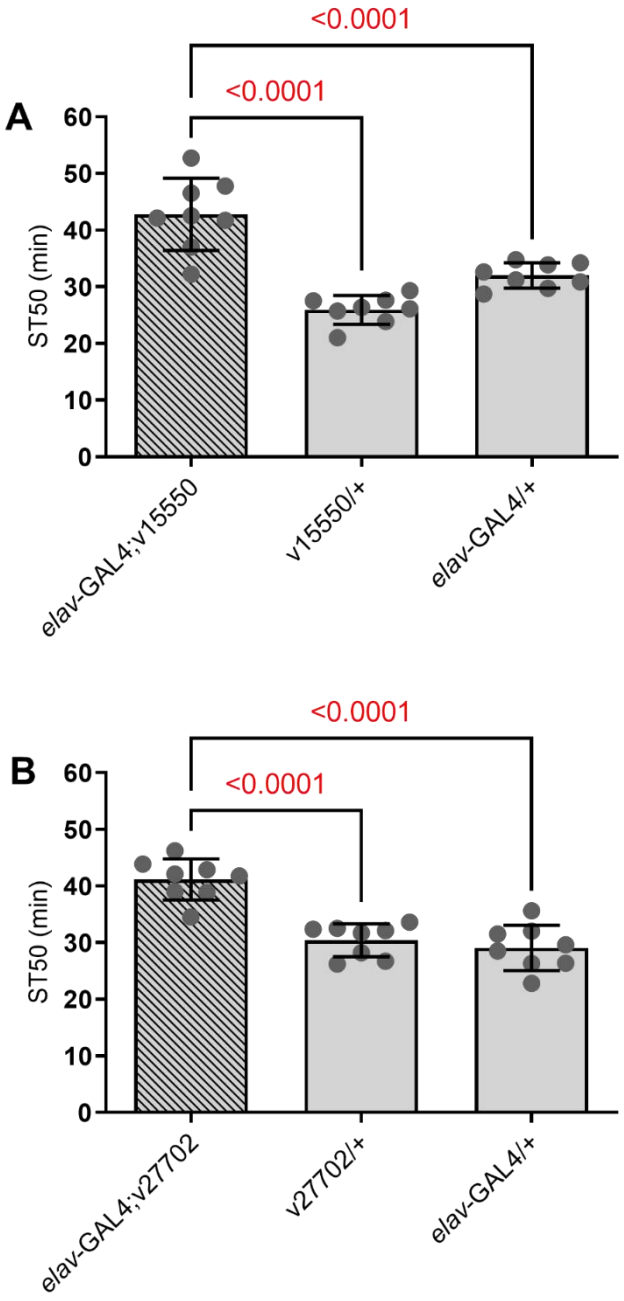


Figure 2. Confirmation that expression of *Mef2* and *spin* RNAi in neurons increases ST50 values. Data presented are ST50 values of flies with pan-neuronal expression of (A) *Mef2* (*elaV*-Gal4;v15550) and (B) *spin* (*elaV*-Gal4;JF02782) RNAi along with their respective controls (v15550/+, JF02782/+ and *elaV*-Gal4/+) exposed to vapor from 85% ethanol. Overall, genotype significantly affected ST50 (individual one-way ANOVAs, $p<0.0001$ for both panels, $n=8$). Expression of the *Mef2* (*elaV*-Gal4;v15550) and *spin* (*elaV*-Gal4;JF02782) RNAi transgenes increased ST50 compared to their respective controls. (Bonferroni's multiple comparisons, lines indicate pairwise comparisons with resulting p values above each line).

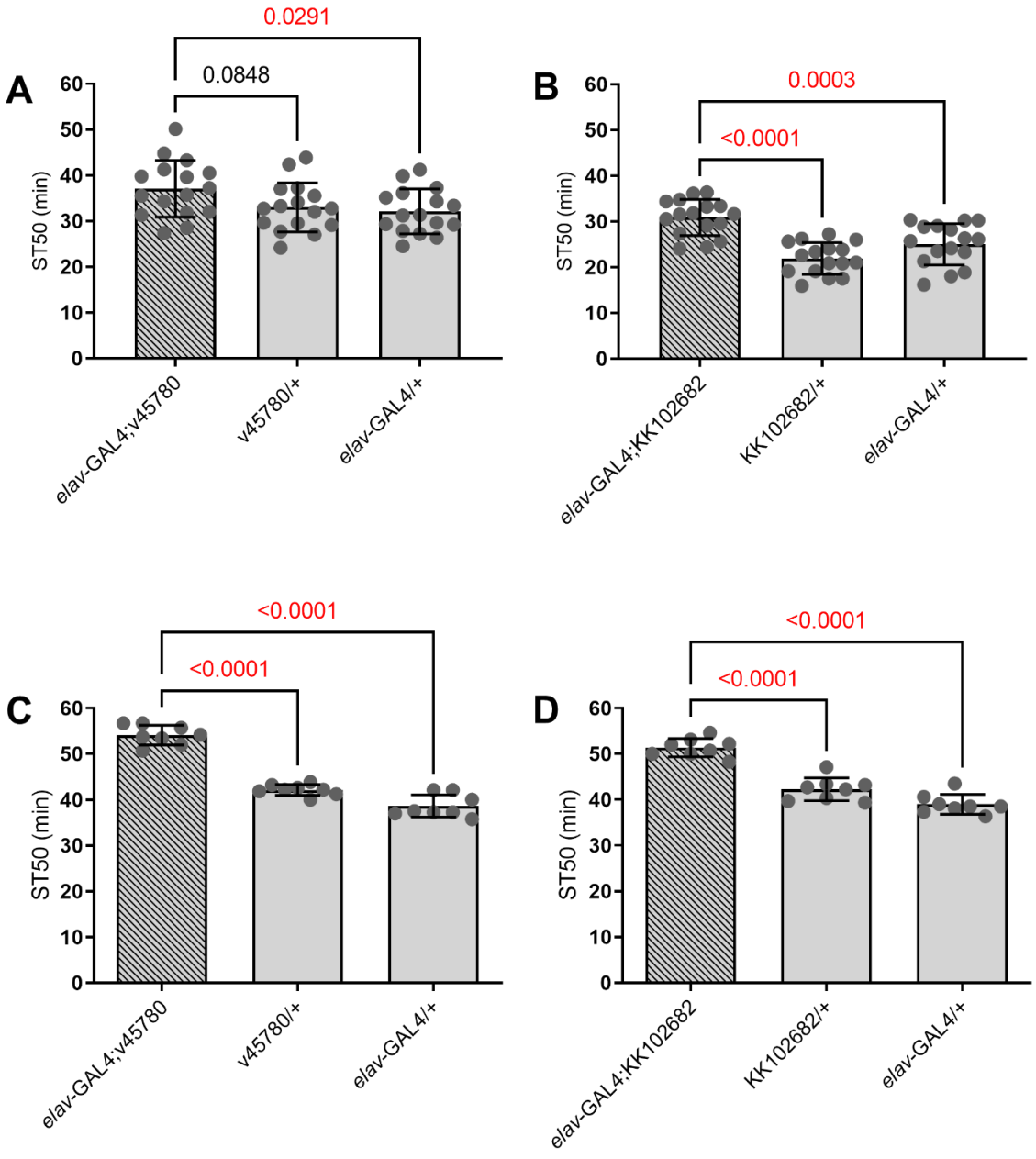


Figure 3. ST50 values of flies with pan-neuronal expression of *unc79* RNAi exposed to vapor from 85% (A, B) or 65% (C, D) ethanol. Data presented show RNAi expressing (A, C: *elaV*-Gal4;v45780; B, D: *elaV*-Gal4;KK102682) flies and their respective controls (v45780/+, KK102682/+ and *elaV*-Gal4/+). Genotype significantly affected ST50 values in all studies (individual one-way ANOVAs: A, $p=0.0332$, $n=16$; B, $p<0.0001$, $n=16$; C, $p<0.0001$, $n=8$; D, $p<0.0001$, $n=8$). ST50 values in RNAi expressing groups were significantly higher than both controls for all planned comparisons except for *elaV*-Gal4;v45780 vs 45780/+ in panel A. (Bonferroni's multiple comparisons, lines indicate pairwise comparisons with resulting p values above each line).

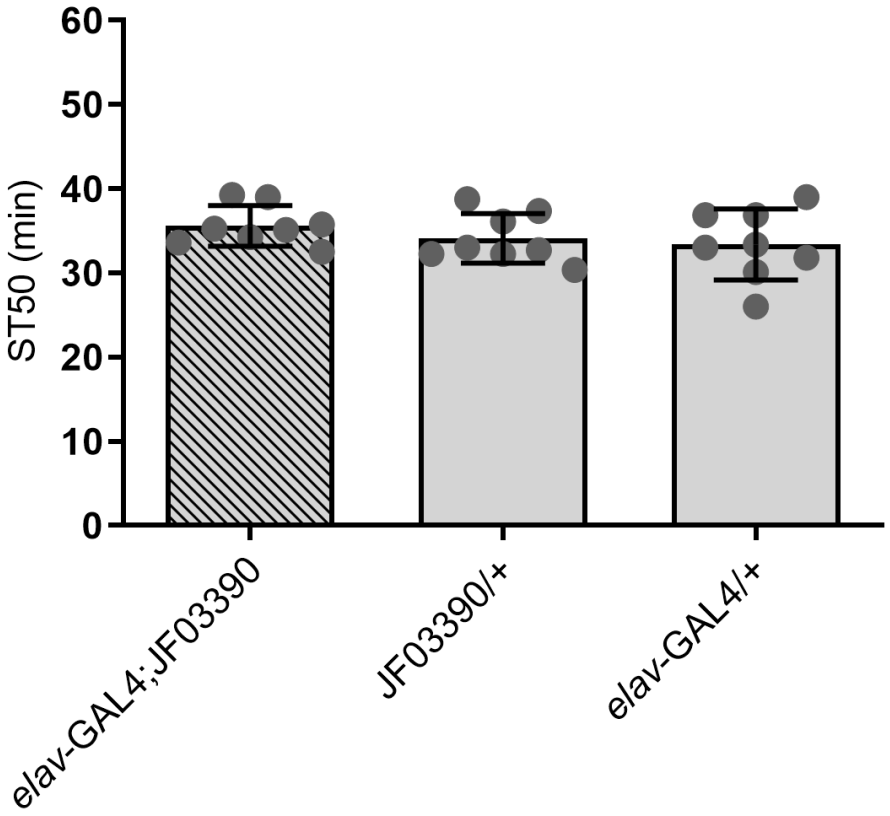


Figure 4. ST50 values of flies with pan-neuronal *Bx* RNAi (*elaV*-Gal4;JF03390) and controls (JF03390/+, *elaV*-Gal4/+) exposed to vapor from 85% ethanol. Genotype did not significantly affect ST50 (one-way ANOVA p=0.4001, n=8).

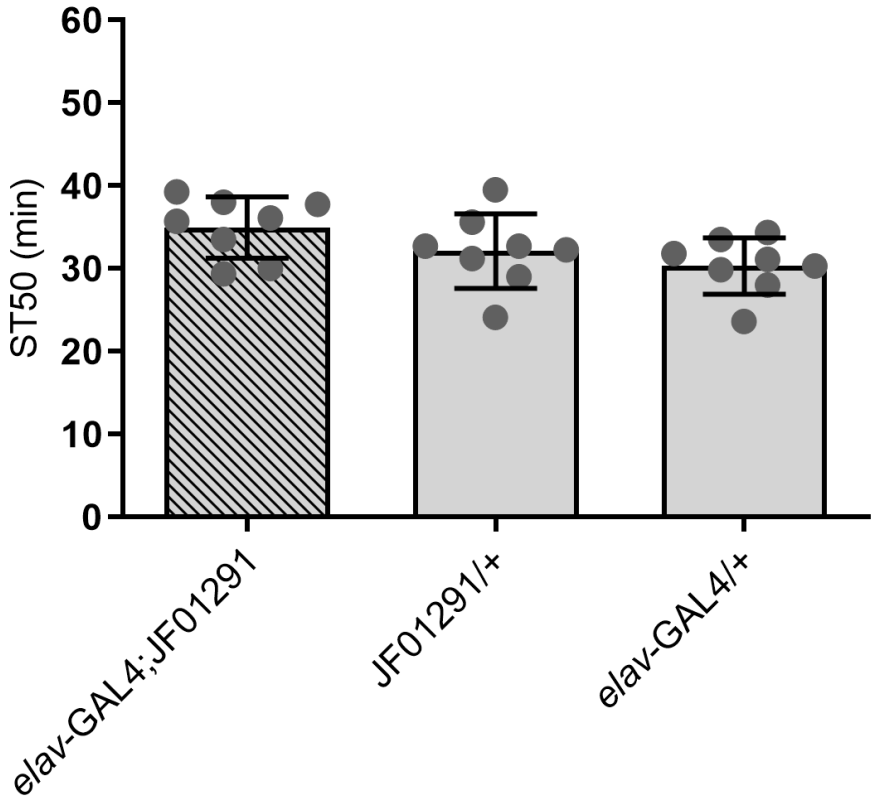


Figure 5. ST50 values of flies with pan-neuronal expression of *CtBP* RNAi (*elaV*-Gal4;JF01291) and controls (JF01291/+, *elaV*-Gal4/+) exposed to vapor from 85% ethanol. Genotype did not significantly affect ST50 (one-way ANOVA p=0.0772, n=8).

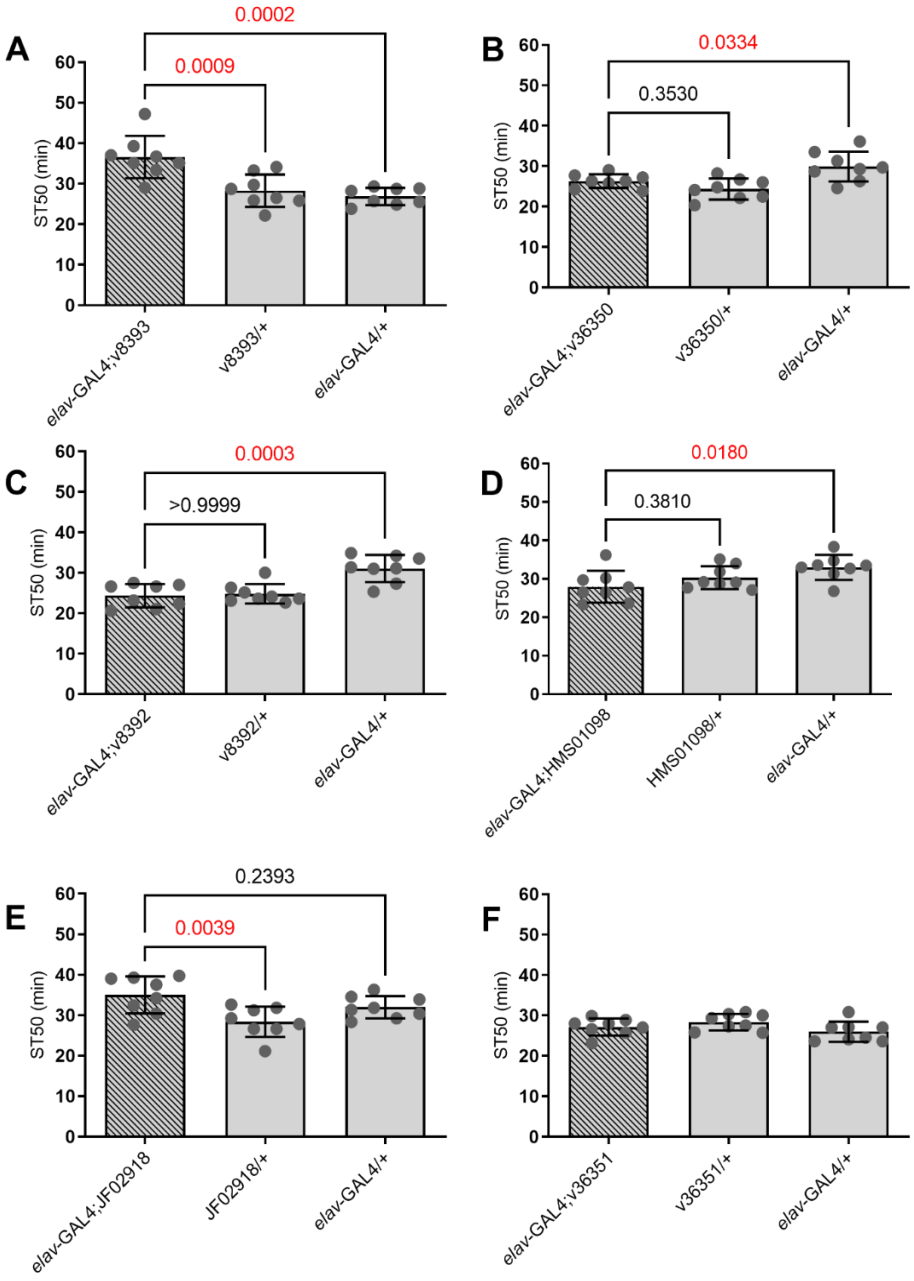


Figure 6. ST50 values of flies with pan-neuronal expression of *Fas2* RNAi (A, *elav*-Gal4;v8393; B, *elav*-Gal4;v36350; C, *elav*-Gal4;8392; D, *elav*-Gal4;HMS01098; E, *elav*-Gal4;JF02918; F, *elav*-Gal4;v36351) and their respective controls (v8393/+, v36350/+, v8392/+, HMS01098/+, JF02918/+, v36351/+ and *elav*-Gal4/+) exposed to vapor from 85% ethanol. Overall, genotype significantly affected ST50 in all studies except those in panel F (individual one-way ANOVAs; A, $p<0.0001$; B, $p=0.0023$; C, $p=0.0002$; D, $p=0.0304$; E, $p=0.0073$; F, $p=0.1256$; n=8). ST50 in flies expressing RNAi were different than both controls in panel A, and different than one control in panels B, C, D and E (Bonferroni's multiple comparisons, lines indicate pairwise comparisons with resulting p values above each line).

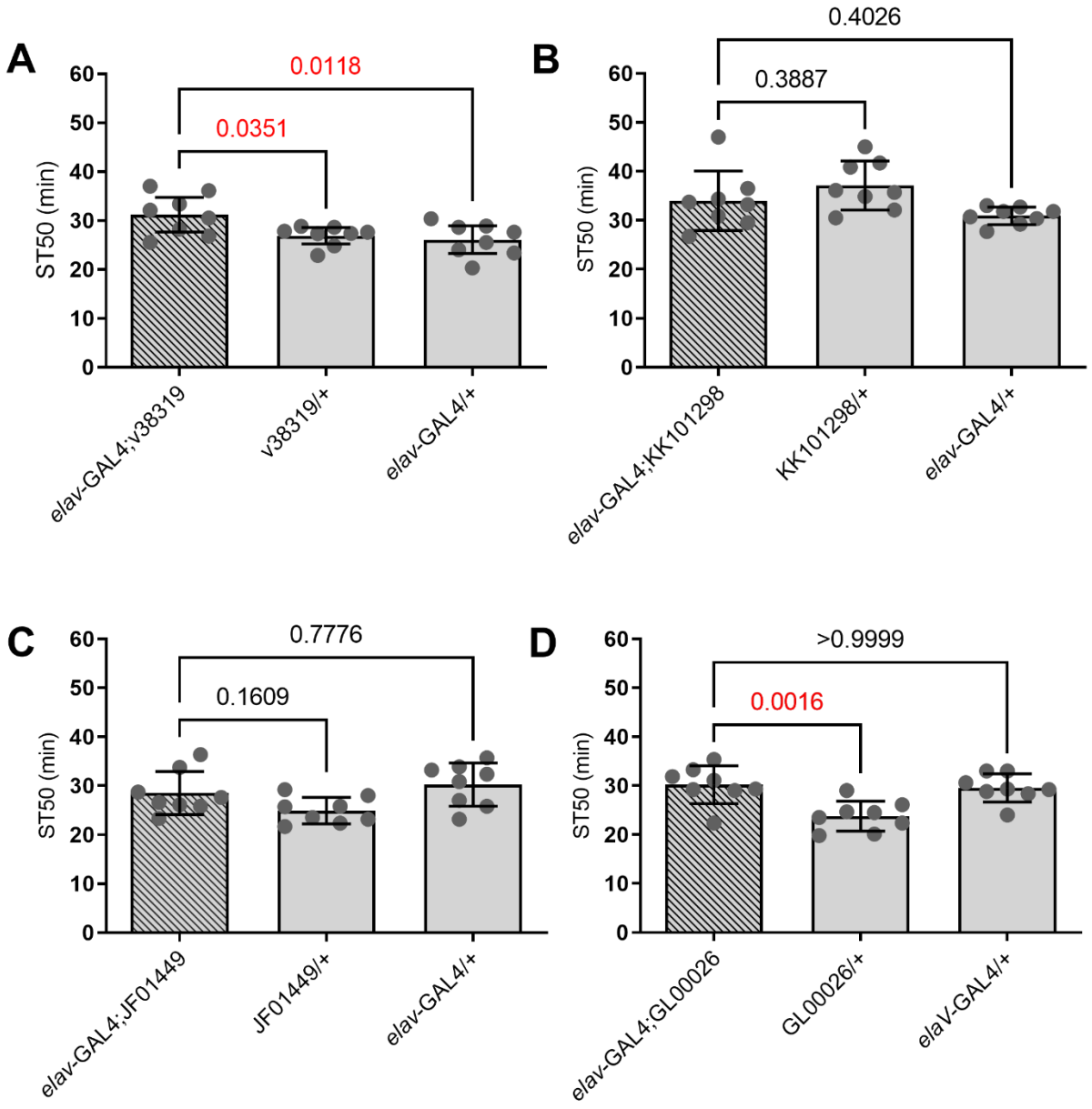


Figure 7. ST50 values of flies with pan-neuronal expression of *For* RNAi (A, *elaV*-Gal4;v38319; B, *elaV*-Gal4;KK101298; C, *elaV*-Gal4;JF01449; D, *elaV*-Gal4;GL00026) and their respective controls (v38319/+, KK101298/+, JF01449/+, GL00026/+ and *elaV*-Gal4/+) exposed to vapor from 85% ethanol. Overall, genotype significantly affected ST50 in all studies (individual one-way ANOVAs; A, p=0.0126; B, p=0.0474; C, p=0.0378; D, p=0.0013; n=8). Pan-neuronal expression of (A) v38319 (*elaV*-Gal4;v38319) increased ST50 compared to both controls, whereas expression of (D) GL00026 (*elaV*-Gal4;GL00026) RNAi increased ST50 compared to only a single control. Expression of the KK010298 (C) and JF01449 (D) transgenes did not significantly change ST50 relative to either controls (Bonferroni's multiple comparisons, lines indicate pairwise comparisons with resulting p values above each line).

CHAPTER 3: IDENTIFICATION OF MEF2-DEPENDENT GENE EXPRESSION CHANGES

1. Introduction and Rationale

Identifying gene expression changes in response to *Mef2* knockdown can help us begin to explore and understand mechanisms pertaining to ethanol sedation. To get a picture of global gene expression changes that result from *Mef2* knockdown, we performed an RNA-seq experiment comparing gene expression levels in three groups: *elav*-gal4/v15550/+ *Mef2* knockdown (hereafter KD), an *elav*-GAL4/+ control (hereafter Gal4) and a v15550/+ RNAi control (hereafter RNAi). Specific questions we addressed are:

- Are any of our six previously identified genes of interest (detailed in Chapter 2) regulated by *Mef2*?
- Are any of the 342 genes that bind *Mef2* (Sivachenko et. al 2013) regulated by *Mef2*?
- Are any known fly or worm alcohol genes (reviewed in Grotewiel & Bettinger 2015) regulated by *Mef2*?
- Are fly orthologs of previously identified genes implicated in human SRE (Schmitt 2019) regulated by *Mef2*?
- Are fly orthologs of genes implicated in human externalizing behaviors (Dick 2020, personal communication; Linnér et. al 2020) regulated by *Mef2?*
- What gene ontology terms are over-represented in the Mef2-dependent differentially expressed genes?

We hypothesized that knockdown of *Mef2* would result in expression changes in genes represented by each of these lists. Additionally, we hypothesized that gene ontology analysis would reveal biological processes in which the Mef2-dependent differentially expressed genes function. Together, we predicted that this study would mechanistically connect *Mef2* with genes previously implicated in alcohol-related behavior and relevant biological processes.

2. Materials and Methods

2a. Drosophila husbandry

Flies were grown under the same conditions described in Chapter 2. Each week, three crosses were set up in parallel: *elav*-GAL4 virgin females were crossed to v15550 (*Mef2* RNAi) males to express the *Mef2* RNAi pan-neuronally (hereafter KD for knock-down), w[A] virgin females were crossed to v15550 RNAi males to generate the RNAi controls (hereafter RNAi), and *elav*-GAL4 virgin females were crossed to w[VDRC] males to generate the *elav-Gal4* controls (hereafter Gal4). All progeny were in a uniform F1 hybrid genetic background of 50% w[A] and 50% w[VDRC].

At one to three days of adulthood, approximately 350 female flies of each genotype were collected under brief CO_2 anesthesia. Flies were collected in a rotating manner in regard to genotype. As an illustration, in collection one, KD flies were collected first, then RNAi flies, followed by Gal4 flies. In collection two, RNAi flies were collected first, then Gal4 flies and subsequently, KD flies. The order of genotypes collected from each subsequent round of crosses was similarly rotated to avoid batch effects. Flies were collected under CO_2 and immediately transferred to a 50 mL conical tube on water ice. The tube was labelled with the genotype and number of flies collected, and quickly transferred to a -80˚C freezer. Flies were on water ice for a maximum of two minutes before being transferred to the freezer. The process was repeated with the next genotype of that week's collection, and so on. After one bottle of females was collected of all three genotypes, the cycle was repeated, beginning with the first genotype and working through that week's order of collections until approximately 350 flies of each genotype were collected per cross. This was repeated weekly until six collections were obtained.

2b. Isolation of fly heads

Heads were isolated in batches from approximately 250 frozen flies of each genotype. Liquid nitrogen and dry ice were collected in a dewar and styrofoam box, respectively, from the

Sanger Hall supply center. Head preps were performed in the same rotating order as the fly collections outlined in the previous section. All head collections were performed at 4˚C.

Holes were punched through the lid of a 50 mL conical tube using a heated 18-gauge needle. The tubes of flies of the first genotype were removed from the -80˚C freezer, kept on dry ice and filled with approximately 35 mL of liquid nitrogen. The cap with holes was screwed on and the tube was vortexed for approximately one minute or until all the liquid nitrogen dissipated. This was repeated, filling the tube with about 25 mL of liquid nitrogen for a second round of vortexing. Vortexing flies in liquid nitrogen causes heads, wings, legs, abdomens and thoraxes to break apart.

A sieve was used to separate the bodies of the flies from the heads. Prior to adding flies to the sieve, it was confirmed that the layers of the sieve were in the correct order and liquid nitrogen was slowly poured into the sieve to ensure it was cold enough to prevent flies from sticking to it. Flies were added to the top layer of the sieve and the sieve was repeatedly struck laterally with forceps for at least three minutes to help move the fly body parts to various levels of the sieve. Heads, representing the smallest body parts, were collected in the bottom vessel of the sieve.

Collected heads were transferred into labelled 1.7 mL snap-cap tubes and kept on dry ice, then quickly transferred to a -80˚C freezer. The bodies of the flies were quickly examined under the microscope to ensure that heads were actually separated from other body parts. The sieve was cleaned and dried and the process was repeated with the next two genotypes in the order determined for that week. Throughout the head isolations, protective gear such as safety glasses, a lab coat and cryo-gloves were worn. This procedure was repeated for all six collections. The full protocol is in the Appendix.

2c. Preparation of RNA

Prior to beginning RNA preps, the lab bench, pipettes, pipette boxes and anything else on the lab bench were wiped down thoroughly with 100% ethanol. Plastic pestles were placed into a 50 mL conical tube and covered with chloroform (to inactivate RNAses) under the fume hood. After soaking for 20 minutes, the pestles were transferred to a new tube and allowed to dry for another 20 minutes.

Total RNA was isolated using a combination of a previously published protocol (Weston et. al 2021; Lee et. al 2021) and Qiagen reagents. Tubes of fly heads were retrieved from the 80˚C freezer and immediately placed on ice. For the rest of the RNA extraction protocol, samples were processed in a rotating order in regard to genotype, as described in previous sections. A chloroform-soaked pestle was properly secured to an electric drill. 50 µL of Trizol (ThermoFisher Scientific) was very slowly added to the first tube of heads and homogenized with the drill for 30 s. Another 200 µL of Trizol was added to the same vial, and the flies were again homogenized for 90 s. This was repeated with all remaining genotypes. 100 µL of chloroform was added to each vial. Vials were then vortexed for 15 seconds and incubated at room temperature for 3 minutes. Samples were taken to the cold room and centrifuged at 14,000 x g for 15 minutes. After centrifugation, new vials were labelled for each genotype and the upper aqueous layer was transferred to the appropriate new vial, taking care not to pipette any fly parts or other layers. Qiagen RNeasy Mini Kit and on-column DNase digestion protocols were performed per the manufacturer's instructions (Qiagen), with a final elution in 60 µl of RNase free water. Eluted RNA samples were stored at -80˚C until being sent for sequencing. The full protocols are in the Appendix.

2d. Initial RNA quality assessments

Concentration and initial quality assessments were performed by taking absorbance readings of 1:20 dilutions of each sample at 260 nm and 280 nm using a Pharmacia Biotech Ultraspec 2000 spectrophotometer. From these values, RNA concentration and the A260/A80

ratio were calculated. Generally, an A260/280 ratio of between ~1.8 - 2.1 is indicative of purified RNA (ThermoFisher Scientific 2012).

1:20 dilutions of all samples were also run on the Agilent Bioanalyzer as an additional measure of concentration and to assess RNA quality (performed by Sati Afshari in Dr. Babette Fuss' laboratory, Virginia Commonwealth University). The Bioanalyzer calculates an RNA integrity (RIN score) from peaks at 28s and 18s ribosomal RNA (rRNA) for mammalian samples; however, fly RNA has a double peak at 18s, precluding the calculation of a reliable RNA integrity number (RIN). Therefore, I assigned grades (A, B, C, etc.) to each trace from the Bioanalyzer based on visual inspection noting subjective interpretations of peaks beyond the 28s and 18s peaks (which if present, would indicate degradation). Samples that appeared to have higher degradation were assigned lower scores, and those with decreased degradation received higher grades. The five sets of three samples each that overall received the best grades were ultimately sent for sequencing by GeneWiz. The remaining three RNA samples were not analyzed further.

2e. RNA-sequencing and related analyses performed by GeneWiz

GeneWiz guidelines recommended sending at least 2 µg of RNA at a minimum concentration of 50 ng/µl per sample. The amount of RNA to send was calculated per sample. For all samples, at least 1.5 times, and whenever possible, double the amount of sample required was sent. Each sample vial was labelled, sealed and packaged in dry ice in a Styrofoam box. The package was overnighted to GeneWiz' facilities in South Plainfield, NJ for standard RNA-seq analysis.

GeneWiz performed their standard bioinformatic analysis as part of the RNA sequencing package. Their overall workflow consisted of assessing RNA integrity, generating cDNA libraries from my RNA samples, sequencing the cDNAs on a single lane of an Illumina Hi-Seq machine with 150 base paired-end reads, evaluating sequence quality, trimming reads, mapping reads to

the BDGP6 version of the fly genome, generating hit counts for genes and exons, comparing counts to assess read depth and differentially expressed genes and analyzing initial gene ontology as outlined in Figure 1. GeneWiz also performed initial differential gene expression analyses with a two-fold cut-off, but we did not use these analyses and instead performed our own differentially expressed gene analysis using iDEP as described below.

2f. Identification of differentially expressed genes and related analyses using iDEP

iDEP9.2 (integrated Differential Expression and Pathway analysis version 9.2) is a web-based tool for exploratory data analysis that integrates R/Bioconductor packages frequently used for RNA-seq analysis, such as DESeq2, to analyze RNA-seq data. It is available at http://ge-lab.org/idep/. Users are able to upload gene expression data and the tool allows detection of differentially expressed genes (DEGs), creation of Venn diagrams to visualize the overlap of DEGs, perform principal component analyses (PCA), identify gene ontology terms and perform pathway analysis.

We uploaded raw read counts received from GeneWiz into iDEP and set the species to *Drosophila melanogaster*. In the pre-processing stage, iDEP allows users to filter out genes that have extremely low read counts (less than 0.5 counts per million). iDEP transforms the data for downstream clustering and PCA analysis. We selected to transform the data using the edgeR package. Using the “plot gene” function of iDEP’s pre-processing stage, we plotted levels of specific genes of interest, such as the *white* marker gene and *Mef2*.

Next, we used iDEP to perform principal component analyses. We assessed all possible pairwise combinations of principal components 1-5. We also used iDEP to view differentially expressed genes using the website’s “DEG1” tab. We set the method to identify differentially expressed genes to the R/Bioconductor DESeq2 package, the false discovery rate (FDR) to 0.1 and the fold change to 1. This identifies all differentially expressed genes with an adjusted FDR less than or equal to 0.1 regardless of the fold-expression level. We also generated Venn

diagrams of all DEGs, and then specifically of up- and down-regulated DEGs by using the “Venn diagram” tab. On this tab, we also downloaded gene lists of the differentially expressed genes to later filter in MS Excel to understand overlaps between the DEGs and other biologically relevant gene sets.

2g. Analysis of overlapping genes

We used the conditional formatting tool in Microsoft Excel (version 1808, Redmond, WA, USA) to determine the overlap between DEGs identified by our analyses in iDEP and other biologically relevant gene sets. Fisher’s exact test (R Studio Version 1.4.1717) was used to assess whether the amount of overlap we observed between our DEGs and other gene lists was significantly different than that expected by chance. The R script used was provided by Dr. Michael Miles and Maren Smith (provided in the appendix). P values ≤ 0.05 were considered significant.

2h. Gene Ontology (GO)

DAVID (Database for Annotation, Visualization and Integrated Discovery version 6.8, https://david.ncifcrf.gov/) was used for gene ontology (GO) analyses. Using the functional annotation tool, we uploaded our list of DEGs into DAVID and set the species to *Drosophila melanogaster*. In the annotation summary results, we looked at the Biological Process (GOTERM_BP_DIRECT), Cellular Component (GOTERM_CC_DIRECT) and Molecular function (GOTERM_MF_DIRECT) tabs to understand the various gene ontology terms tagged and the genes associated with each one. Additionally, we used the “functional annotation clustering” selection for a combined view. Each GO term, the category it was associated with, the genes involved, the percentage of genes involved in the GO term compared to the full list of DEGs, p-values and adjusted p-values were all compiled. P values ≤ 0.05 were considered significant.

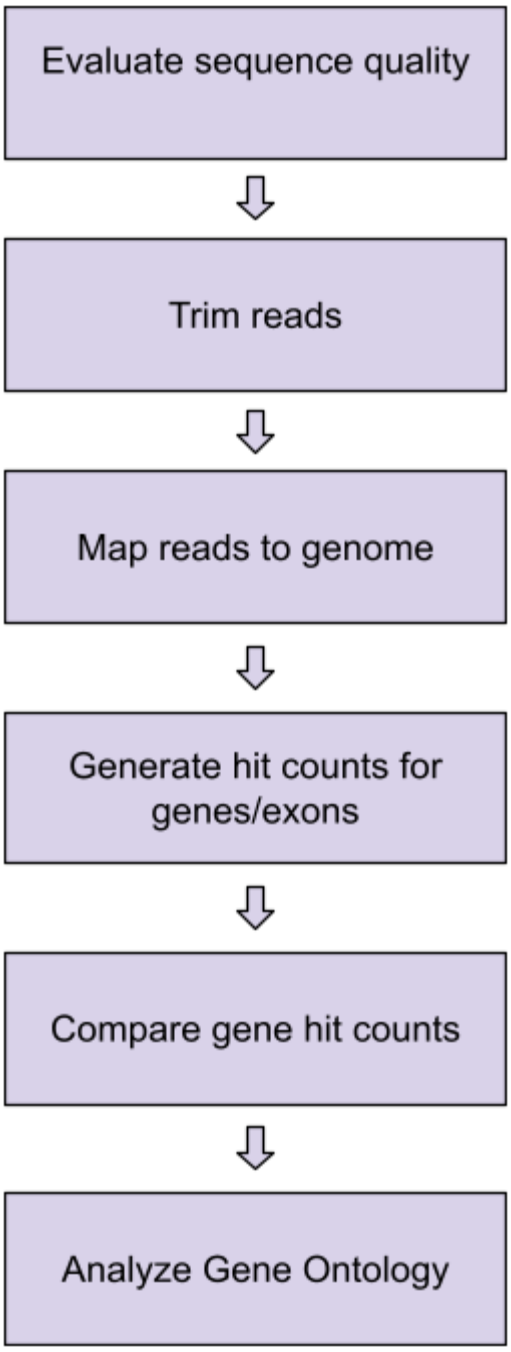


Figure 1. Overall workflow performed by GeneWiz as part of their standard RNA-seq analysis package.

3. Results & Discussion

3a. Rationale and overall experimental design

Mef2 is a transcription factor (Black & Olson 1998; Taylor & Hughes 2017) that regulates ethanol sedation behaviors (Schmitt et. al 2019; Adhikari et. al 2018). Mef2 binds 342 genes in flies (Sivachenko et. al 2013) and *Hr38*, a gene downstream of *Mef2*, influences ethanol preference and tolerance (Adhikari et. al 2018). Thus, we were interested in identifying other genes downstream of *Mef2*, as these genes may also play a role in ethanol-related behaviors, specifically sedation. Additionally, we hoped to understand the relationship between *Mef2* and individual, previously identified genes connected to ethanol-related behaviors. We hypothesized that an RNA-seq study would both identify genes that were differentially expressed upon *Mef2* knockdown and elaborate on the direction by which *Mef2* regulates its downstream genes.

We had three groups of flies in this experiment, as described in the Methods section of this chapter. The KD group had an RNAi transgene against *Mef2* expressed pan-neuronally, and the elav-Gal4 (Gal4) and RNAi groups served as our two controls. As *Mef2* was knocked down pan-neuronally, we used fly heads as the starting material for RNA isolations. After preparing RNA, it was sent for sequencing and the RNA-seq data we received are the basis of the analyses described in this chapter.

3b. RNA & RNA-seq Quality Control Assessments

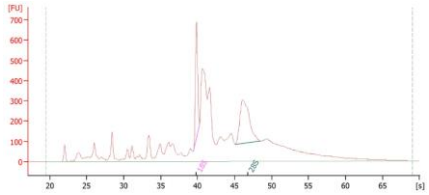
RNA preps were performed in six sets of three samples each. Each set consisted of one sample of each genotype, and absorbance measurements of 1:20 dilutions of each sample were taken at 260 and 280 nm (A260, A280) to estimate concentration and purity (Koetsier et. al 2019). Pure RNA has an A260/A280 ratio of 2.1; a decreased ratio corresponds to protein contamination (Koetsier et. al 2019). My RNA samples had A260/A280 ratios of 1.85-2.29 and concentrations of 48-733 µg/ml (Table 1, columns 3-6), making them suitable for further analysis.

1:10 dilutions of each sample were analyzed on Dr. Babette Fuss' lab's Agilent Bioanalyzer machine. The Agilent 2100 Bioanalyzer system is an automated, microfluidic chip-based electrophoresis machine that has utility in measuring RNA quality for downstream applications (Agilent 2020), in our case, RNA-seq. The machine works by measuring the quality of ribosomal RNA to generate an RNA integrity (RIN) score. RIN scores are an indicator of sample integrity, which has significant implications on downstream applications (Agilent 2020). However, the 18S and 28S eukaryotic RNA ribosomal peaks the machine uses to calculate the RIN score are not consistent with insect RNA, which has a double peak at 18s, rendering it impossible for the BioAnalyzer to generate reliable RIN scores. Therefore, I visually inspected each Bioanalyzer sample trace and assigned each a relative grade (A, B, C, etc.) based on the perceived amount of degradation as indicated by peaks beyond the 18s and 28s rRNA peaks. Figure 2 shows each trace, while Table 1, column 7 shows the grades assigned. Ultimately, the five sets of three samples with the best grades (outlined in black boxes in Table 1) were sent for sequencing. All of the samples sequenced had grades ≥ B.

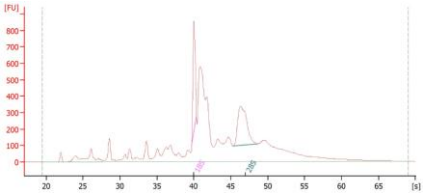
Sample	Genotype	Abs_{260}	Abs_{280}	Abs Ratio	Concentration (µg/mL)	Grade	RIN	DV200	Quality score
1	KD	0.325	0.165	1.97	260.0	C	n/a	n/a	n/a
2	RNAi	0.351	0.172	2.04	280.8	B	n/a	n/a	n/a
3	Gal4	0.378	0.182	2.08	302.4	B	n/a	n/a	n/a
4	KD	0.625	0.292	2.14	500.0	B+	9.7	94.69	35.71
5	RNAi	0.162	0.075	2.16	129.6	B+	9.8	93.27	35.69
6	Gal4	0.595	0.280	2.13	476.0	A	10.0	94.27	35.75
7	KD	0.403	0.198	2.04	322.4	A	10.0	93.01	35.73
8	RNAi	0.134	0.068	1.97	107.2	A	10.0	92.19	35.70
9	Gal4	0.072	0.039	1.85	57.6	A	9.6	92.36	35.70
10	KD	0.917	0.421	2.18	733.6	A-	10.0	94.67	35.73
11	RNAi	0.344	0.177	1.94	275.2	A-	10.0	94.07	35.75
12	Gal4	0.664	0.314	2.11	531.2	B+	9.9	92.92	35.76
13	KD	0.169	0.089	1.90	135.2	B	9.4	92.45	35.77
14	RNAi	0.060	0.032	1.88	48.0	B	9.4	83.36	35.75
15	Gal4	0.160	0.060	2.67	128.0	A	10.0	93.37	35.78
16	KD	0.204	0.089	2.29	163.2	B+	9.8	91.04	35.76
17	RNAi	0.102	0.047	2.17	81.6	A	10.0	94.03	35.76
18	Gal4	0.196	0.090	2.18	156.8	B+	10.0	94.94	35.75

Table 1. RNA sample number and sample genotypes are shown in columns 1 and 2, respectively. Absorbance values at 260 nm, 280 nm, and the ratio between these values (for 1:20 dilutions) are shown in columns 3-5. RNA concentrations based on the A260 values are in column 6. Quality grades based on visual assessments of chromatograms in Figure 2 are shown in t column 7. RIN and DV200 values from GeneWiz are shown in the final two columns.

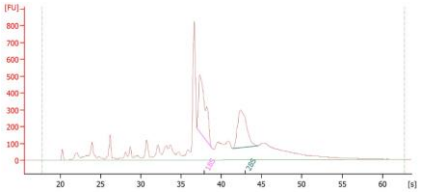
SAMPLE 1



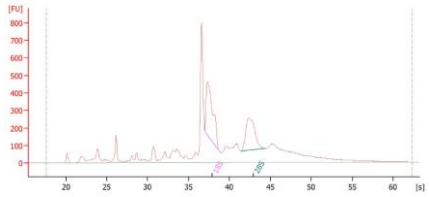
SAMPLE 2



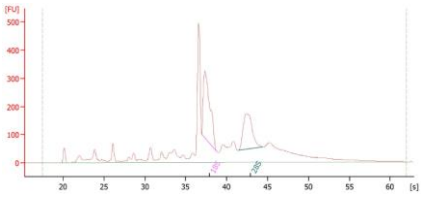
SAMPLE 3



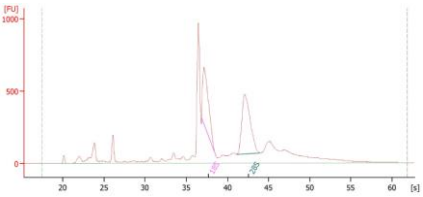
SAMPLE 4



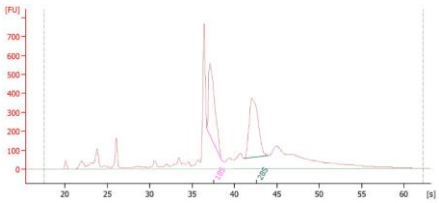
SAMPLE 5



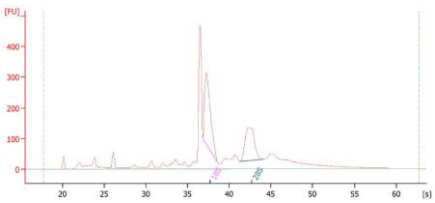
SAMPLE 6



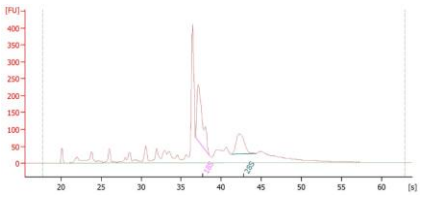
SAMPLE 7



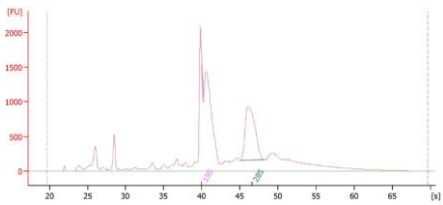
SAMPLE 8



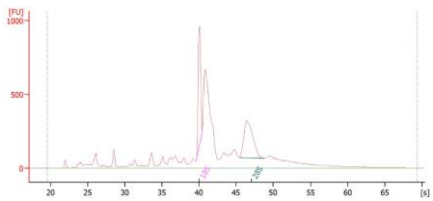
SAMPLE 9



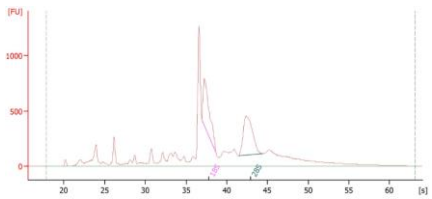
SAMPLE 10



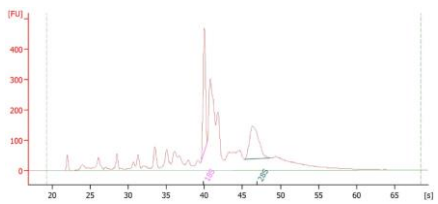
SAMPLE 11



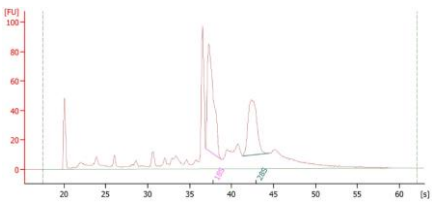
SAMPLE 12



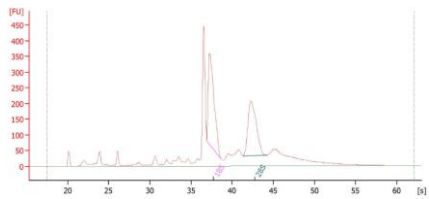
SAMPLE 13



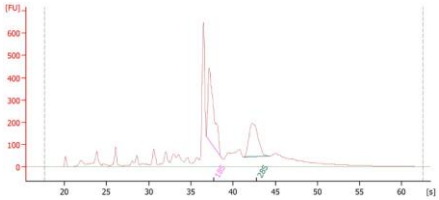
SAMPLE 14



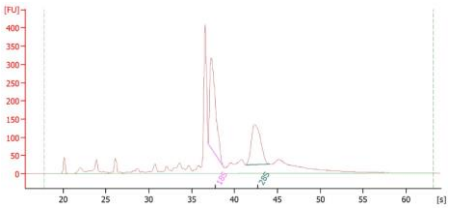
SAMPLE 15



SAMPLE 16



SAMPLE 17



SAMPLE 18

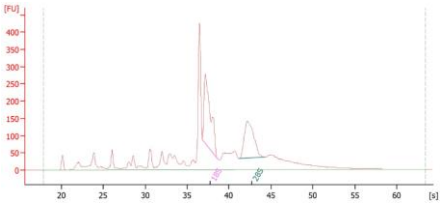


Figure 2. Bioanalyzer traces for each RNA sample. Qualitative grades were assigned to each sample (Table 1, column 7) based on the amount of degradation evidenced by peaks to the left of the 18S doublet. Data generated by Fatemah S. Afshari and Dr. Babette Fuss, Virginia Commonwealth University.

RNA quality control data (RIN and DV200 values) were also provided by GeneWiz (Table 1, columns 8-9). These were determined through the Agilent TapeStation system. Like the Bioanalyzer, this is an automated electrophoresis machine that is used for determining nucleic acid size, concentration and integrity (Graf 2017). GeneWiz provided chromatogram outputs generated via TapeStation and corresponding RIN scores for each sample (Figure 3). Chromatograms are generally consistent, with the largest peaks at the lower end and 18S. Visually, samples with increased noise in the areas between the lower and 18S peaks (such as Figure 3, samples 1 and 6) or smaller 18S peaks (such as Figure 3, samples 10 and 11) tended to have relatively lower RIN scores. Sample 7 is noticeably different from the rest of the samples in that there was an additional peak at 28S and more degradation than other samples. However, the GeneWiz RIN is 10.0 for this sample, and representatives from GeneWiz confirmed the accuracy of the score. Overall, all RNA samples had RINs (Table 1, column 8) that were well above GeneWiz's recommended minimum suggestion of 7 for samples to be sequenced.

To address the accuracy of my grading system, I plotted my grades against the respective RIN values provided by GeneWiz. My grades generally corresponded to the RIN values, and were very strongly correlated with the RIN values (Figure 4). We therefore feel confident that visual grading can be a useful method of analyzing Bioanalyzer traces for *Drosophila* RNA (Figure 2) and (of the 18 RNA samples prepared) that we selected the 15 samples of the highest quality (Table 1).

Another indicator of RNA sample quality is DV200, or percentage of RNA fragments longer than 200 nucleotides. Like sample integrity, fragment size is an important contributor to good library yield and sequencing downstream (Graf 2017). Samples with a DV200 less than 30% are not recommended for downstream applications (Illumina 2016). All of our samples had DV200 values much higher than this minimum threshold, with our lowest DV200 value being 83.36 (Table 1, column 9). Taken together, my subjective grades coupled with the more formal

DV200 and RIN scores from GeneWiz indicate our RNA samples were appropriate for library preparation and sequencing.

Differences in sequencing depth can affect downstream analyses. Data provided by GeneWiz (Figure 5) show that across all samples and genotypes, RNA-seq read counts from our RNA samples were consistent. Quality scores from sequencing ranged from 35.69 - 35.77. Quality score is defined by the formula: $Q = -10\log_{10}(e)$, where e is the estimated probability of the base call being incorrect (Illumina 2016). Higher scores are indicative of high quality, whereas lower scores are associated with increased rates of false positives and potentially unusable data. Quality scores above 30 indicate a 1 in 1000 chance of a base being called incorrectly and correspondingly, a base call accuracy of 99.9% (Illumina 2016). As all our samples had Q scores well above 30 (Table 1, column 10), we are confident that the sequencing is accurate. Additionally, representatives from GeneWiz confirmed that libraries were generated together and all samples were run on the same lane of the same sequencing machine at the same time to avoid batch effects.

Together, the data in Table 1 and Figures 2-5 indicate that the RNA samples sent for sequencing were of high quality and that the sequencing itself was of high quality, supporting our confidence in the results of this study.

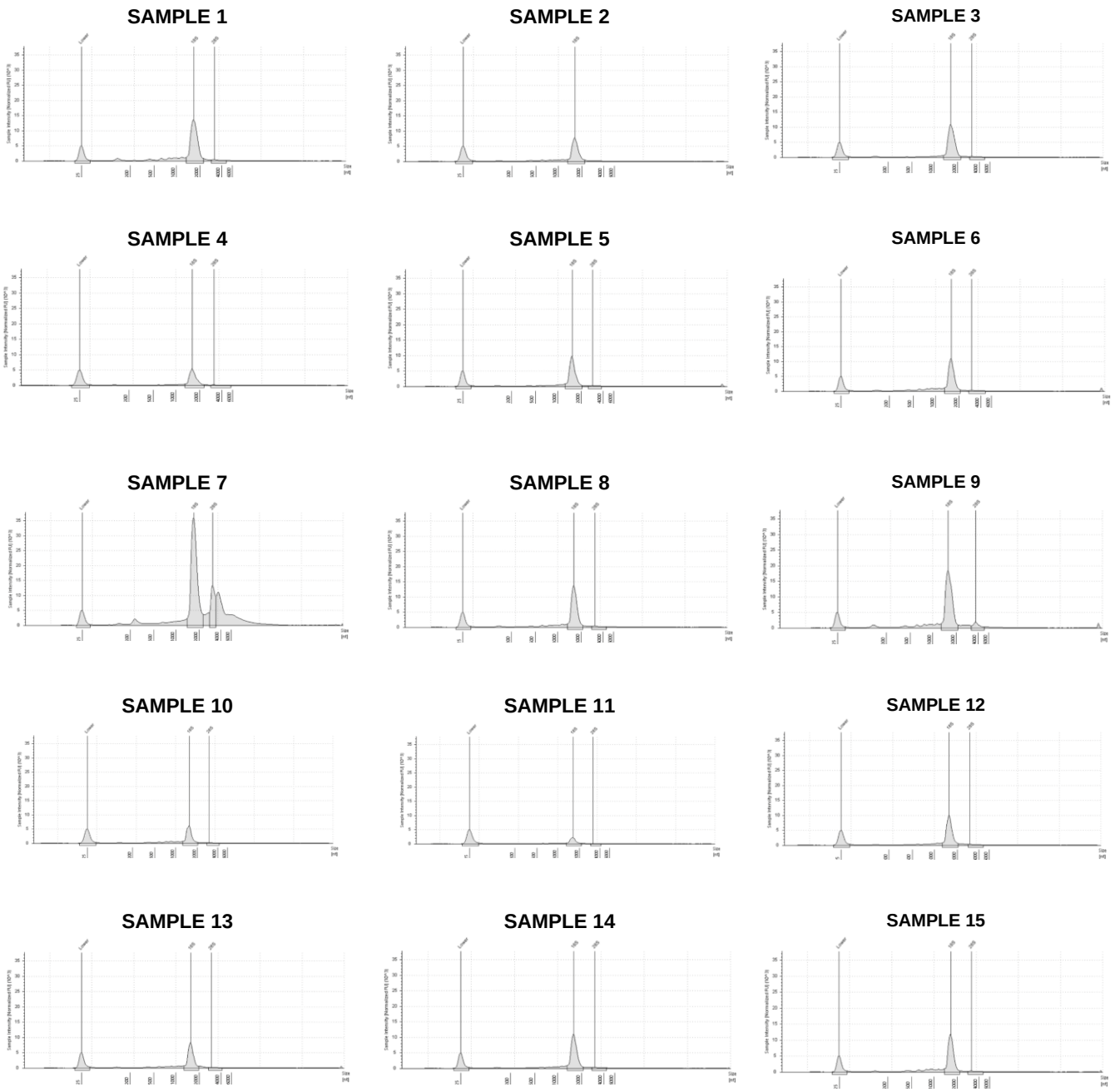


Figure 3. TapeStation chromatogram traces and RIN scores provided by GeneWiz. Generally, samples with increased noise between lower and 18S peaks, or relatively smaller 18S peaks have lower RIN scores, though all samples have very high RIN scores. Data generated by GeneWiz.

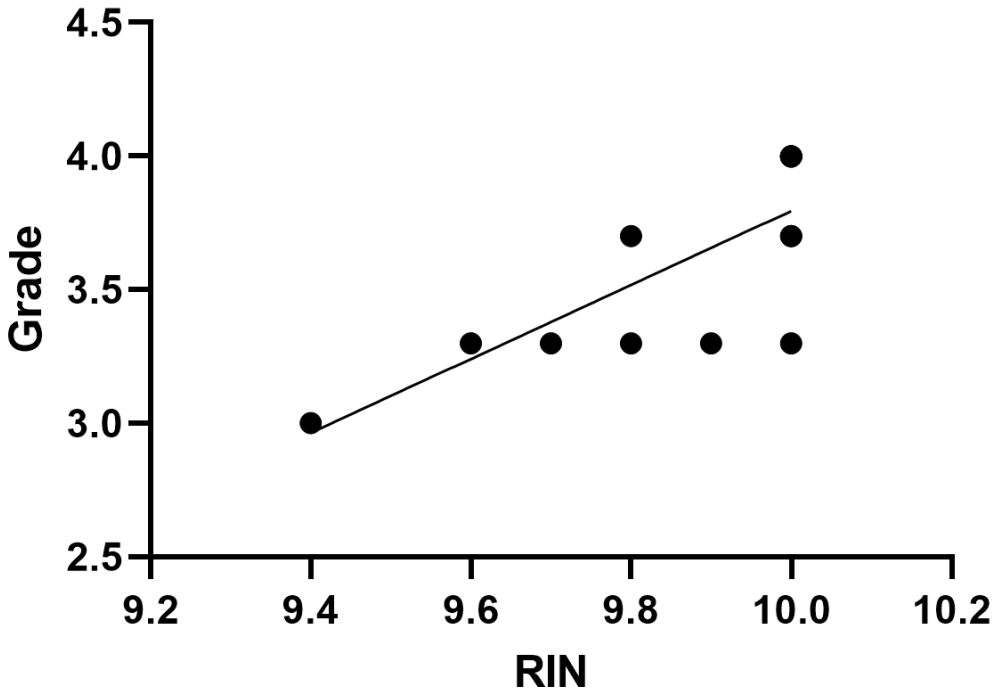


Figure 4. Relationship between the grades assigned to each sample and their RIN scores. RIN values correlated with grades (Spearman correlation, r=0.8238, p=0.005, n=15). Some X-Y pairs overlap others, giving the appearance of less than 15 samples. Line is best-fit linear regression (R^2=0.6515).

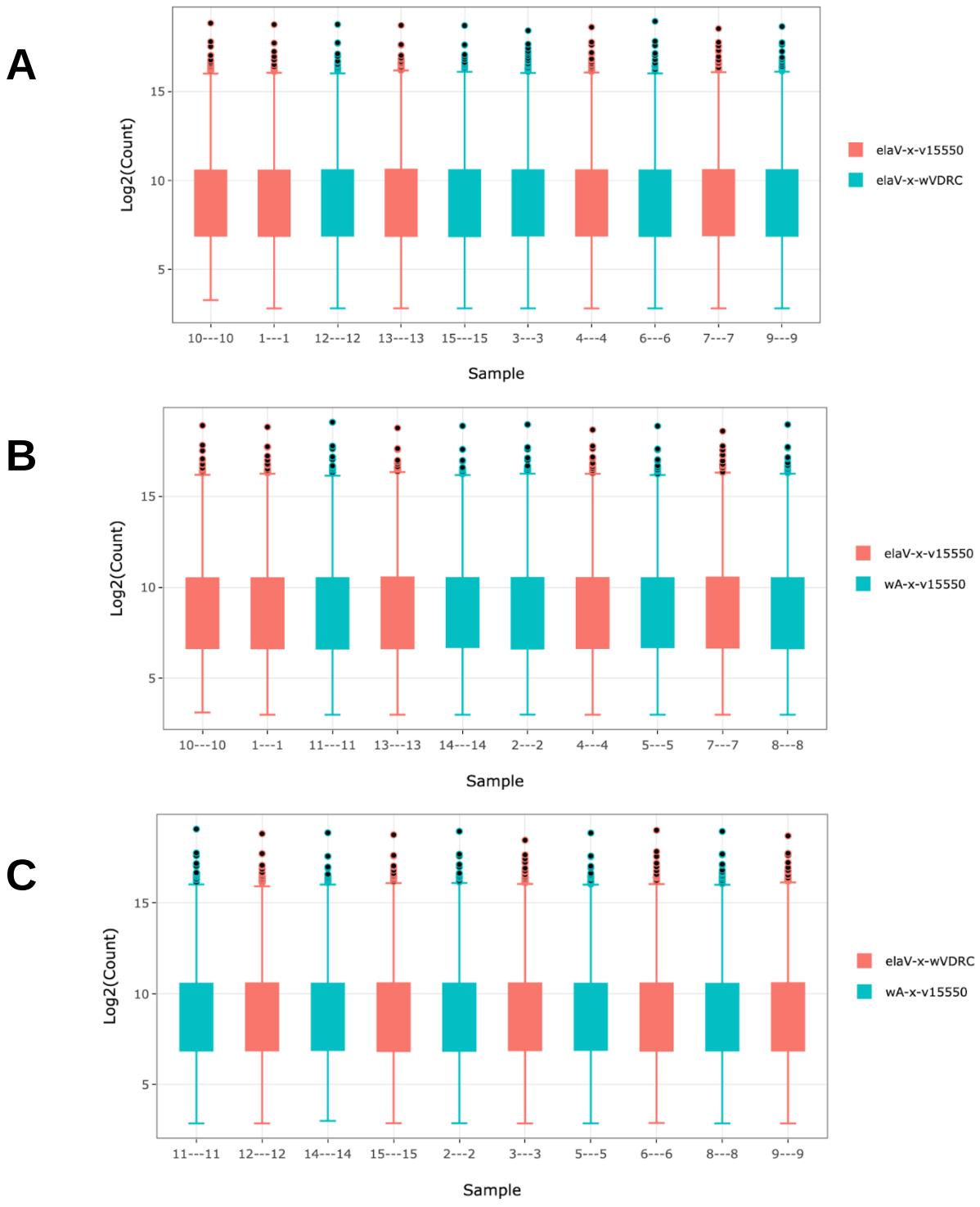


Figure 5. Distribution of normalized read counts for all samples, generated by GeneWiz. Panels A, B and C show pairwise comparisons of read counts for KD//Gal4, KD//RNAi and Gal4//RNAi respectively. Each comparison shows that read counts are consistent across all samples.

3c. Analysis of Mef2-dependent differentially expressed genes

Principal component analysis (PCA) linearly transforms data (RNA-seq data for my project) to allow for visualization on a two-dimensional plane (GeneWiz). Each principal component accounts for an amount of variance, with principal component 1 accounting for the largest single factor for the variance, and each subsequent component accounting for less (GeneWiz). A priori, we expected to see each genotype cluster separately from the other two. The PCA analysis of the first two components of the RNA-seq data did not show a clean clustering or segregation of the three genotypes (Figure 6A). To determine whether additional principal components might provide cleaner clustering patterns, we generated PCA plots with all possible combinations of principal components 1-5 using iDEP (Figure 6). No combination of principal components 1-5 produced obvious clustering of genotypes (Figure 6), suggesting there might not be large-scale, consistent patterns of differential gene expression across the three genotypes. As part of their analyses, GeneWiz provided a heatmap of the top 30 DEGs (identified by p-value) between the KD and Gal4 groups (Figure 7). The expression levels of these genes in each genotype were very similar, suggesting that these two samples did not have large scale differences in the most significantly differentially expressed genes.

Given that the PCA (Figure 6) and heatmap data (Figure 7) were not consistent with large-scale, substantive changes in *Mef2* expression in knockdown animals, we explored whether sample mishandling at some point in the workflow could have occurred. We addressed this possibility by examining the expression of the *white* gene. *white* is a marker for the RNAi and Gal4 transgenes that we expected to be expressed at a molecular level corresponding to the observable eye pigmentation phenotype regulated by the *white* gene. Of the three genotypes in our study, RNAi flies have the least eye pigmentation, Gal4 flies have an intermediate level of eye pigmentation, and KD flies (because they harbor both the RNAi and Gal4 transgenes) have the strongest eye pigmentation. Reassuringly, levels of *white* were lowest in RNAi flies, intermediate in Gal4 flies, and highest in KD flies (Figure 8). This pattern of

expression level is evident in the mean raw counts as well as the raw counts from individual samples (Figure 8), fully consistent with the recorded genotypes used to generate RNA samples and the RNA-seq data. Additionally, *Mef2* expression was slightly decreased in the KD compared to the two control groups, although this was significantly different only relative to the RNAi control (Figure 8B). Given that expression of this same *Mef2* RNAi transgene in neurons using the same elav-Gal4 driver decreases expression of Mef2 protein in the central fly brain (Schmitt et. al 2019), we believe that the marginal knockdown of *Mef2* observed in our RNA-seq study is likely due to endogenous expression of *Mef2* in other head tissues including head muscle (Velasco et. al 2006) contributing to the overall *Mef2* read counts. Importantly, the *white* and *Mef2* expression levels from the RNA-seq data strongly indicate that no sample mishandling or mislabeling occurred that could explain the somewhat unexpected similarities in gene expression in our three genotypes.

In many transcriptomics studies, including those from flies using Gal4 to express RNAi transgenes (e.g. Nitta 2017, Picchio 2013, Zeng 2018), differentially expressed genes (DEGs) are defined relative to a single Gal4 control. The standard design of experiments using Gal4 to drive transgene expression (like our study), however, has two control genotypes, one for the Gal4 driver alone and another for the RNAi transgene alone. We therefore identified and analyzed *Mef2*-dependent DEGs using two approaches. In the first approach, we defined DEGs relative to the Gal4 control alone. In the second approach, we defined DEGs as being consistently changed relative to both the Gal4 alone and the RNAi alone. Additionally, we tabulated DEGs between the Gal4 and RNAi controls themselves.

Using a false discovery rate of 0.1 (Ge et. al 2018) and allowing for any level of fold-change, we found 172 DEGs in KD vs Gal4 fly head RNA samples (Figure 9, Table 2). This KD//Gal4 set of DEGs was of principal interest and was further analyzed below. Unexpectedly, we found that the number of DEGs (1,063) was considerably greater in KD//RNAi fly heads, and was even greater (2,238) in Gal4//RNAi (Figure 9). The overlap between these three sets of

DEGs followed a pattern generally consistent with their sizes (Figure 10). There were 51 genes common to the 172 KD//Gal4 and 1,063 KD//RNAi DEGs. These 51 KD//Gal4//RNAi DEGs are differentially expressed in KD compared to both controls, also warranting their further analysis below.

Although we do not understand why the greatest number of DEGs was found between the two control genotypes, this relatively large number of DEGs is not related to differences in genetic background (both controls are in the same F1 hybrid genetic background), RNA sample or sequence quality (Table 1, Figures 2 and 3), batch effects due to the order of sample preparation or sequencing (see earlier descriptions of experimental design), or sample misidentification (Figures 8A, 8B). Additionally, this relatively large number of DEGs does not appear to have a major effect on ethanol sedation (Chapter 2, Figure 2A). It is possible that the large gene expression differences between controls is driven by the insertion of the v15550 RNAi transposon. However, if that were the case, we would also expect to observe these DEGs in the KD//Gal4 group, which we largely do not. Given that we frankly do not have a satisfactory explanation for this large set of DEGs at this time and that this set of DEGs does not seem to be a major driver of ethanol sedation, we did not further analyze them.

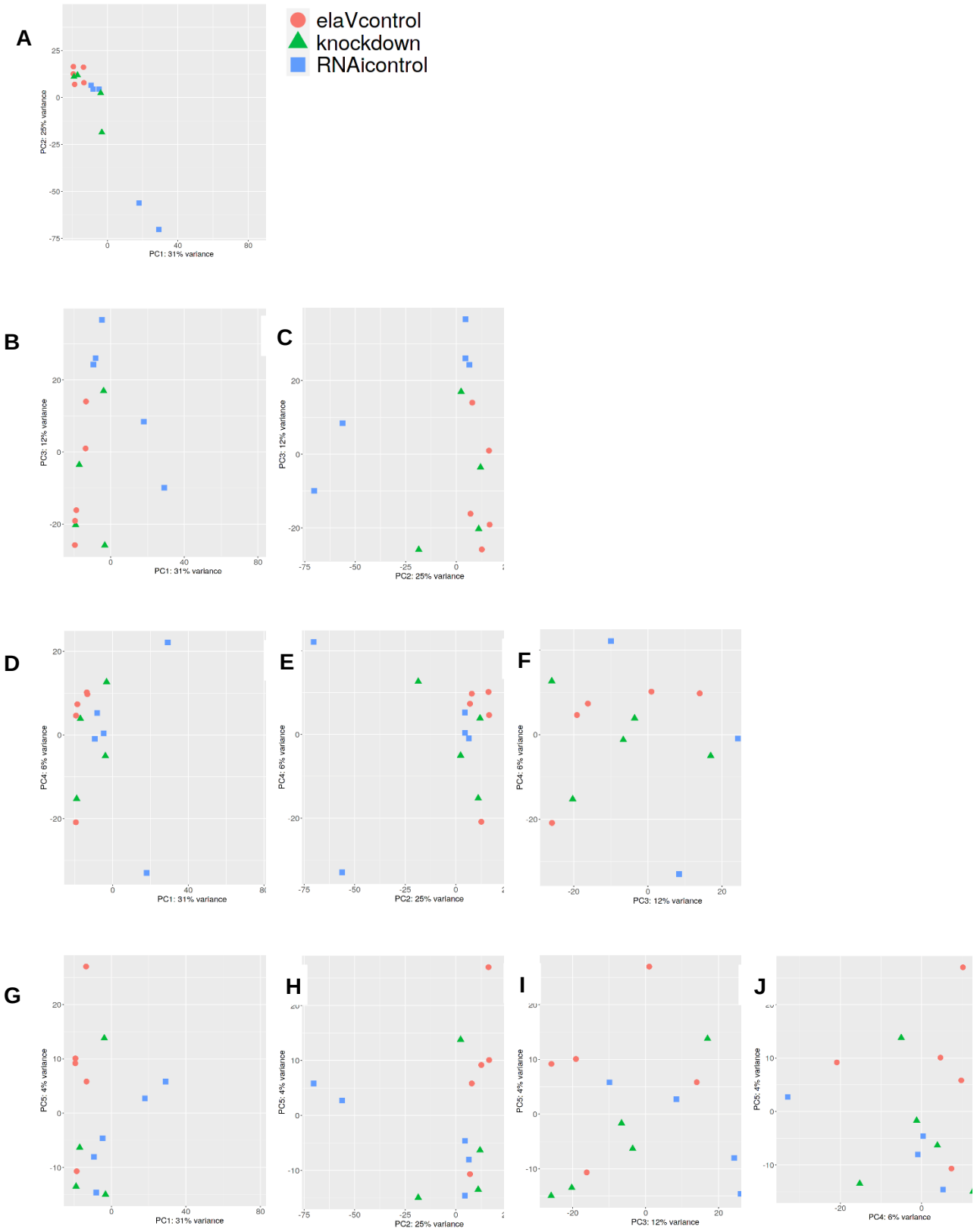


Figure 6. PCA plots showing all combinations of two PCAs for the first 5 components.

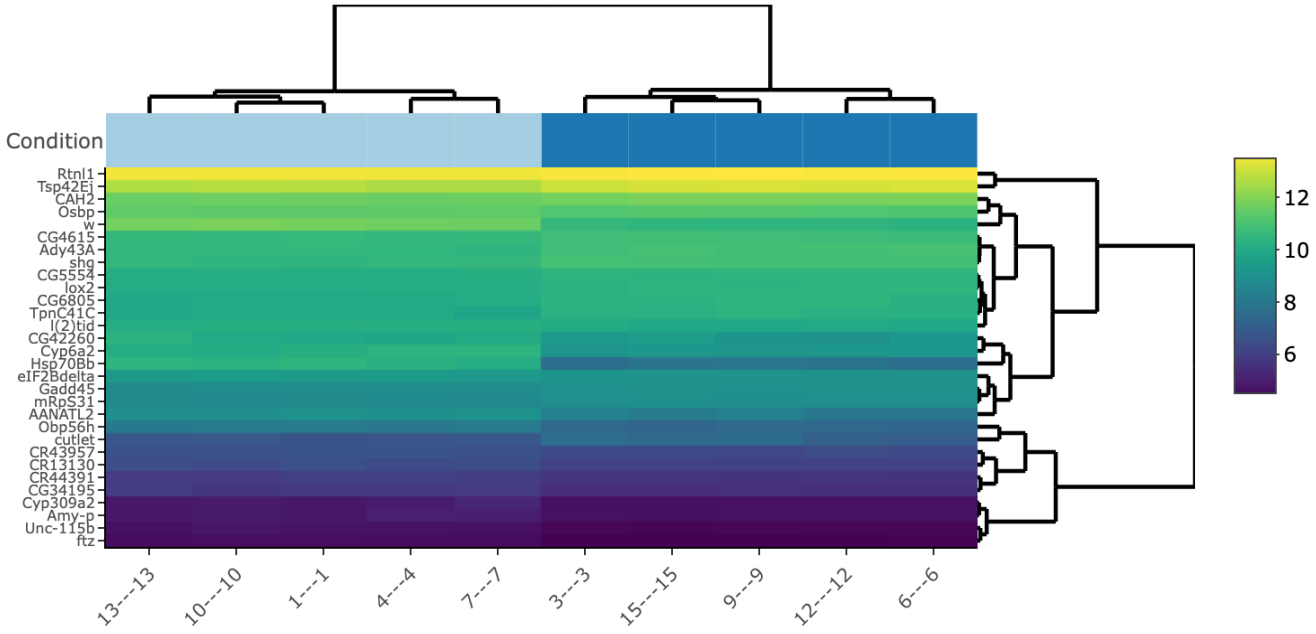


Figure 7. Clustering heatmap provided by GeneWiz to show the similarity between the top 30 DEGs (sorted by p value) between the KD//Gal4. Samples 1, 4, 7, 10 and 13 (light blue cluster) are the KD group and samples 3, 6, 9, 12 and 15 (dark blue cluster) are the Gal4 group. The scale represents the log2 transformed expression values of each DEG.

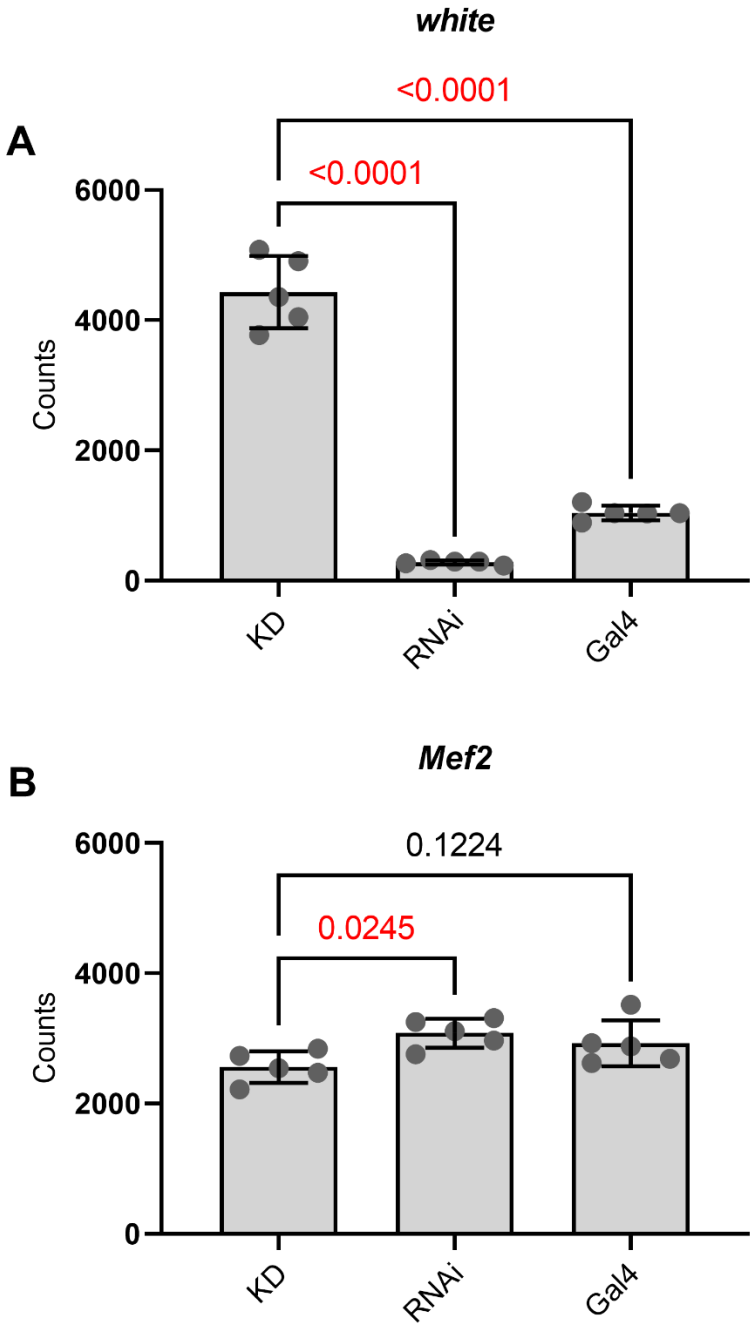


Figure 8. Raw read counts of the *white* (Panel A), *Mef2* (Panel B) genes across three genotypes. As expected, for the *white* gene (A), the KD group has highest counts, followed by the Gal4 group, and lastly, the RNAi group (one-way ANOVA $p<0.0001$) For *Mef2* (B), the KD group did display decreased levels of gene expression, but was only statistically significant when compared to the Gal4 group (one-way ANOVA $p=0.0335$). (n=5 per genotype; Bonferroni's multiple comparisons, lines indicate pairwise comparisons with resulting p values above each line).

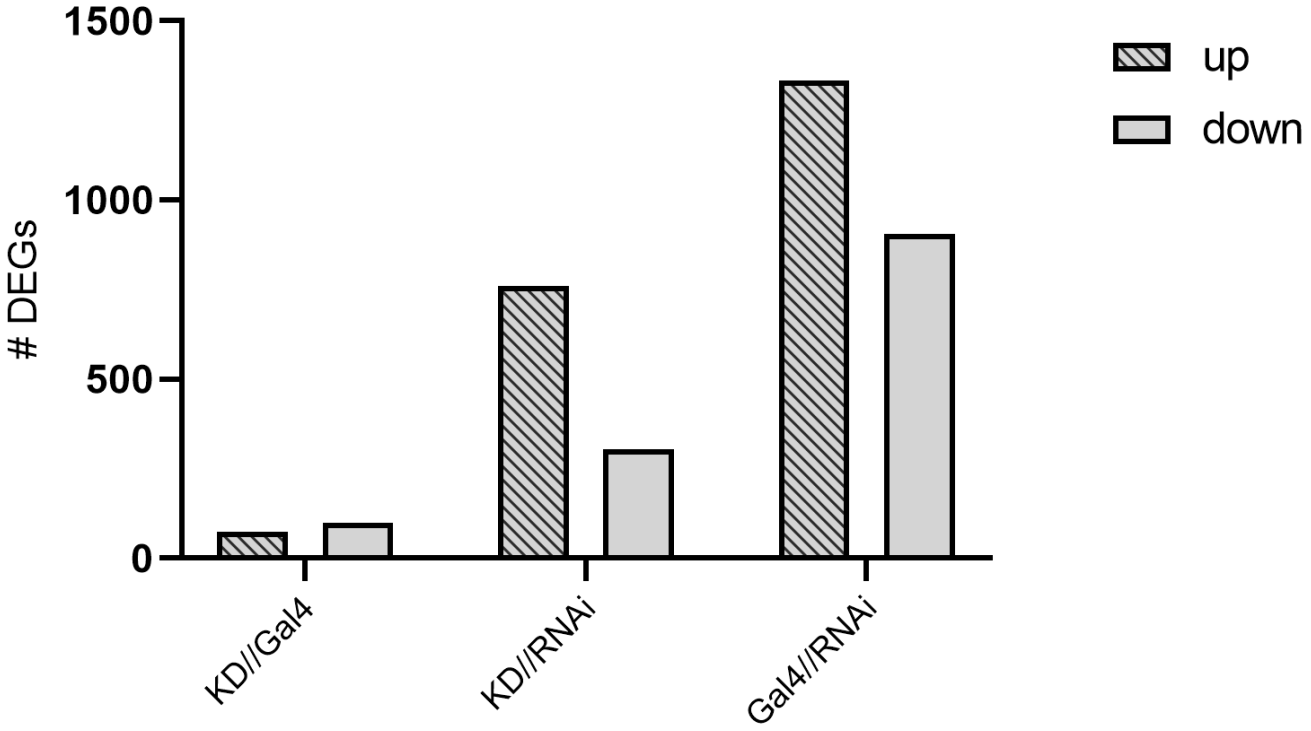


Figure 9. Number of DEGs for each comparison with FDR of 0.1 and any fold change.

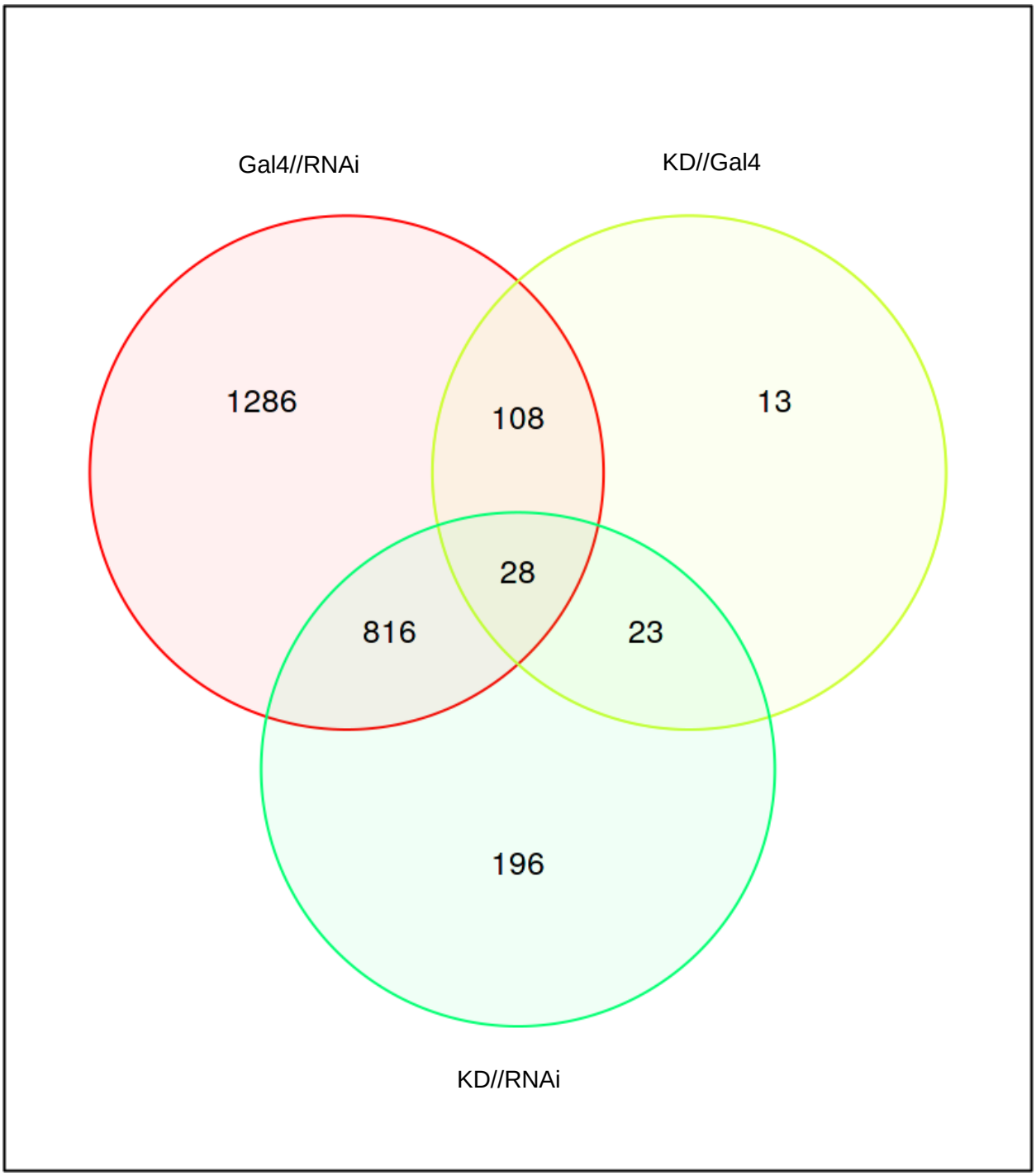


Figure 10. Venn Diagram showing the overlap between all sets of DEGs with FDR of 0.1 and any fold change. Figure generated from iDEP

3d. Analysis of KD//Gal4 DEGs

Of 172 KD vs Gal4 DEGs, 148 genes had detectable expression in high throughput studies of adult fly brains (FlyAtlas-RNA.adult, FB2021_03 and Table 2). Of these 148 genes, 51 had moderate to high to very high expression and an additional 52 had lower, but detectable, expression in fly brains (Table 2). Eight-seven percent of the set of 172 DEGs is therefore known to be expressed in the fly brain, increasing our confidence that our overall experimental design identified potential molecular determinants of brain function and behavior. The 172 DEGs were examined for enrichment for categories of the gene ontology terms Biological Processes, Cellular Components and Molecular Function (Table 3). However, among all categories, only the cellular adhesion term was significantly enriched, based on the Benjamini adjusted p-value (Table 3). The Benjamini-Hochberg Procedure is a form of multiple testing corrections to minimize false discovery (Huang et. al 2009). Additionally, the 172 DEGs represented four annotation clusters (Table 4). (Huang et. al 2009) The largest of these clusters contains terms for oxido-reduction processes, while the other clusters involve GTPase activity, extracellular matrix components, and splicesomes; however, the only significant term based on the Benjamini-Hochberg adjusted p-value is endoplasmic reticulum membrane (Table 4). We also performed functional annotation clustering, a measure of the extent of shared genes between annotations (Huang et. al 2009) All enriched GO terms and functional annotation clusters are included in Tables 3 and 4 for completeness. The set of 172 DEGs might be involved in a very wide range of biological, cellular and molecular processes.

Gene	Function	Up- or down-regulated	FDR	Expressed in neurons?
Actin 87E	involved in cell motility and muscle contraction	down	0.044	low (25.6)
Amylase proximal	dietary enzyme required to hydrolyze starch	down	0.004	no data
Cytochrome p450 6a2	involved in breakdown of synthetic insecticides and insect hormone metabolism	up	5.50E-55	low (47.7)
Pka-C3	encodes a cAMP-dependent protein kinase	down	0.000194	moderate (276.9)
engrailed	encodes a homeodomain-containing transcription factor needed to posterior compartment identity	down	0.072957	low (14.5)
Fas3	encodes a cell adhesion molecule	down	0.006449	moderate (284.5)
FER tyrosine kinase	encodes a tyrosine-protein kinase involved in formation of actin cables during embryogenesis	down	0.000409	low (98.9)
fushi tarazu	predicted to play a role in specifying neuronal identity and known to be required in embryogenesis	up	1.36E-12	no expression (1.3)
Heat shock protein 68	encodes a protein involved in determination of lifespan and response to heat shock and starvation	down	0.004797	low (27.6)
Ecdysone-inducible gene L2	involved in embryogenesis and normal nervous system development	down	0.004864	moderate (421.7)
knirps-like	encodes a nuclear hormone receptor with C4 zinc finger motif and no ligand-binding domain	down	0.054153	no expression (4.1)
l(2)tid	predicted to be a tumor suppressor in larval imaginal disks	up	2.11E-11	moderate (200.6)
lethal (3) malignant blood neoplasm	tumor suppressor involved in normal blood cell proliferation	down	0.045032	no expression (4.6)
CG2150	predicted to be a structural component of chitin-based cuticle and localize to the extracellular matrix	down	0.041810	no expression (6.9)
Punch	isoforms are involved in eye pigment production and normal embryonic segmentation	down	0.009554	moderate (119.6)
pawn	encodes a protein involved in chaeta morphogenesis	down	0.047142	no expression (3.7)
Ras oncogene at 85D	encodes a protein that regulates normal tissue growth and development	down	0.035028	very high (1042.5)
raw	encodes a membrane protein involved in dendritic patterning and localization of JNK signaling components	down	0.006410	moderate (292.9)
shotgun	calcium-dependent cell adhesion protein and has roles in cell sorting, oogenesis and body asymmetry	down	2.46E-31	no expression (9)
scarlet	encodes a protein component involved in transport of pigment precursor to pigment cells	down	0.030551	no expression (5.1)

white	ABC-type guanine transporter involved in transporting cyclic GMP, various amines and pigments	up	4.06E-119	low (13)
Cysteine string protein	encodes a protein that is essential to maintaining synaptic function	down	0.046875	very high (1093.9)
Recombination repair protein 1	involved in cellular response to oxidative stress via DNA repair mechanisms	down	0.032122	moderate (107)
Juvenile hormone esterase	involved in ethanol response and catabolism of juvenile hormone	up	0.043431	no expression (3.7)
Signal recognition particle 54k	thought to be involved in GTPase activity	down	0.027455	moderate (195.5)
ALG3	predicted to be involved in regulating the tumor necrosis factor-mediated signalling pathway	up	0.002876	low (66.6)
Heat shock protein 70Bb	encodes a protein involved in heat shock and hypoxia response	up	2.98E-45	no data
Troponin C at 41C	encodes a protein that binds calcium and regulates muscle contraction	down	0.000239	no expression (3.3)
Rho-like	involved in hemocyte maturation	down	0.000138	low (72)
Histone H3.3A	variant histone that is enriched in active chromatin	up	0.071674	high (636)
cutlet	thought to aid in DNA clamp loader activity and be a positive regulator of DNA polymerase activity	down	2.49E-07	low (10.4)
α-Esterase-8	has carboxylesterase activity	down	0.081458	low (84.3)
Rab7	encodes a GTPase that is involved in regulation of vesicle trafficking	down	0.006512	very high (1056.9)
late bloomer	facilitates synapse formation	up	0.013701	low (30)
Phm	encodes a hydroxylating monooxygenase that is essential to neuropeptide biosynthesis	down	0.012614	no data
drongo	thought to have GTPase activator activity and involvement in regulating GTPase activity	down	0.001262	moderate (453.5)
piopio	encodes a zona pellucida domain protein that is involved in apical epithelial adhesion	down	0.005852	no data
Oxysterol binding protein	encodes a protein involved in sperm development	up	1.37E-07	moderate (299.8)
Splicing regulatory protein 54	encodes a protein involved in regulating alternative splicing, selecting a transcriptional start site and processing the 3’ end of mRNA	down	0.093683	moderate (240.2)
Bicoid interacting protein 1	recruits Sin3A-HDAC1 by interacting with transcription factor	up	0.006512	moderate (172.1)
PRL-1 phosphatase	encodes a growth inhibitor	down	0.042707	high (608.8)
Rab27	encodes a Rab GTPase that is involved in exosomal secretion	down	0.001413	moderate (270.5)

Adenosine kinase 3	has adenosine kinase activity	down	6.90E-30	moderate (435.7)
Prolyl-tRNA synthetase, mitochondrial	has proline-tRNA ligase and prolyl-tRNA aminoacylation activity	up	0.077856	low (62.8)
lilipod	encodes a transmembrane protein that promotes ovarian stem cell maintenance	up	0.003229	moderate (177.3)
Carbonic anhydrase 2	encodes a dehydratase that is involved in catalyzing the hydration and dehydration of carbon dioxide	down	2.25E-08	low (22.9)
Sulfonylurea receptor	predicted to encode an ATP-sensitive K[+] channel	up	0.010093	low (24.2)
Replication factor C 38kD subunit	predicted to be involved in DNA clamp loading and DNA repair	up	0.000214	low (79.7)
CG3823	thought to have phosphatidylinositol biphosphate binding activity	up	0.074056	low (7.3)
CG4615	predicted to be involved in cytolysis	down	2.11E-11	moderate (301.9)
CG1636	known to be expressed in the adult head	down	0.015781	moderate (165.5)
CG8034	predicted to be involved in monocarboxylic acid transport and have monocarboxylic acid transmembrane transporter activity	down	0.098764	low (28.8)
Odorant-binding protein 19b	thought be involved in smell perception and have odorant binding activity	down	0.062071	no expression (6.1)
galectin	encodes a galactoside binding protein with involvement in target recognition	up	0.050380	moderate (853.7)
CG3345	expressed in spermatozoa	up	0.045032	low (14.9)
CG4213	known to expressed in the embryonic brain	down	0.042814	no expression (2.4)
CG14340	thought to be involved in vesicle-mediated transport	down	0.006512	low (57.1)
CG5397	thought to localize to the extracellular space	down	0.002797	low (29.3)
calcutta-cup	thought to localize to the membrane and endomembrane system	up	4.66E-05	no expression (1.2)
papi	involved in the piwi-interacting RNA metabolic process	down	0.011749	moderate (192.6)
CG17224	predicted to have uridine phosphorylase activity and have nucleoside metabolic and catabolic processes	up	0.019329	low (37.3)
CG17264	no information	up	0.000822	low (16.3)
Cytochrome p450 28d2	could be involved in metabolizing insect hormones and breakdown of insecticides	down	0.043661	low (27.5)
Cytochrome p450 6a16	thought to have heme and iron ion binding activity and monooxygenase activity	down	0.088621	no expression (5.8)

CG9107	predicted to be involved in rRNA procession, ribosomal subunit assembly and nucleic acid binding activity	up	0.001391	no data
CG9147	thought to have catalytic activity	down	0.002013	low (96.6)
Arylalkylamine-N-acetyltransferase-like 2	has acyl-coA-N-acyltransferase and aralkylamine N-acetyltransferase activity	up	3.72E-19	low (42.9)
farjavit	encodes lysophospholipid acyltransferase	up	0.003083	moderate (342.7)
CG11319	is involved in proteolysis and dipeptidyl-peptidase activity	down	0.007751	high (729.2)
CG5973	has phosphatidylinositol biphosphate binding activity	up	0.042707	moderate (103.8)
CG13133	thought to be involved in protein folding	up	0.044090	no expression (1.2)
CG4839	thought to have cGMP-dependent protein kinase activity and be involved in protein phosphorylation	up	0.006983	no expression (2.2)
CG10383	encodes a hydrolase that regulates glycosylphosphatidylinositol metabolism	up	0.071674	low (87.3)
Tetraspanin 42Ej	encodes a lysosomal protein essential to degrading endocytosed rhodopsin	down	6.28E-13	high (991.3)
Growth arrest and DNA damage-inducible 45	interacts with JNK pathway and affects egg asymmetry	down	1.80E-10	low (43.8)
CG12164	no information	down	1.52E-05	no expression (1.6)
CG11123	has RNA binding activity and involvement in ribosome biogenesis	up	0.001391	low (50.4)
Death resistor Adh domain containing target	encodes an alcohol dehydrogenase enzyme that is involved in ethanol-induced apoptosis	down	0.015354	very high (1364.2)
CG2064	has NADP-retinol dehydrogenase activity	up	0.000313	low (47.1)
Transmembrane protein 63	has calcium activated cation channel activity	down	0.054427	moderate (150)
CG12910	thought to have UDP-galactosyltransferase and N-glycan processing activity	up	0.044090	low (43.7)
Cuticular protein 47Ef	thought to be a structural component of chitin-based cuticle	down	0.006449	no data
CG10257	involved in I-kappaB kinase/NF-kappaB signalling	down	0.050380	moderate (113.2)
CG8157	no information	down	0.071674	no expression (3.8)
CG15706	localizes to the membrane	down	0.028913	low (69.1)
CG15617	expressed in the adult head	up	0.072325	no expression

				(3.3)
resilin	encodes a component of the extracellular matrix	down	0.006449	no data
CG6472	thought to have lipase activity and be involved in lipid catabolism	down	0.025441	no expression (3.6)
CG6805	has phosphatidylinositol-4,5-biphosphate 5-phosphatase activity and be involved in inositol phosphate dephosphorylation	down	4.95E-21	moderate (111)
CG15611	has guanyl-nucleotides exchange factor activity and is involved in endoplasmic reticulum stress	up	0.006247	low (64.6)
Lysyl oxidase-like 2	predicted to have protein-lysine 6 oxidase activity and be involved in cell adhesion	down	5.96E-09	no expression (5.4)
Odorant-binding protein 56a	predicted to aid in the transport of hydrophobic odorants	down	0.021457	no expression (3.4)
Odorant-binding protein 56h	predicted to aid in the transport of hydrophobic odorants	up	1.51E-32	low (13.1)
CG11099	no information	down	0.013701	high (863.1)
CG13869	no information	up	0.000578	no expression (8.3)
EF-hand domain containing 1.2	thought to have alpha-tubulin binding activity and be involved in cilium assembly	down	0.032580	low (11.8)
dim γ-tubulin 3	plays a role in centrosome-independent spindle microtubule generation and is a part of the augmin complex	down	0.001022	no expression (6.3)
Lysyl oxidase-like 2	predicted to have protein-lysine 6 oxidase activity and be involved in cell adhesion	down	5.96E-09	no expression (5.4)
CG11475	thought to be involved in regulating response to DNA damage, protein methylation and have enzyme binding an phosphatase activity	up	0.002458	no expression (6)
Golgi matrix protein 130 kD	encodes a protein involved in Golgi organization	up	8.59E-08	moderate (130.8)
Cytochrome p450 6d2	involved in detoxifying toxic compounds and encodes a cytochrome p460 oxidase protein	up	0.016460	no expression (7.6)
nahoda	known to be expressed in the adult head	down	0.069003	low (26.2)
CG3500	thought to be involved in endoplasmic reticulum to Golgi vesicle mediated transport	up	0.071674	moderate (132.5)
eukaryotic translation initiation factor 2B subunit delta	thought to have translation initiation involvement	up	4.06E-10	moderate (113.1)
CG5554	predicted be involved in cell redox and homeostasis and be involved in disulfide oxidoreductase activity	down	1.39E-19	moderate (385.4)
Kahuli	encodes a transcription factor	down	0.0503809 2274	no expression (1.3)

spatzle 5	acts as a ligand in promoting motor axon targeting and neuronal survival in the CNS	down	0.006983	no expression (2.1)
CG1136	thought to be a structural constituent of chitin-based larval cuticle	down	0.070749	low (29.4)
CG13705	known to be expressed in the embryonic dorsal epidermis	down	0.054153	no expression (8.3)
CG9953	has dipeptidyl-peptidase activity and thought to play a role in proteolysis	down	1.00E-08	moderate (416)
CG13670	thought to be a structural constituent of chitin-based larval cuticle	up	3.37E-05	no expression (1)
Valyl-tRNA synthetase, mitochondrial	thought to have valine-tRNA ligase activity and to be involved in valyl-tRNA aminoacylation	down	0.050380	low (55.9)
CG5653	thought to have polyamine oxidase activity	down	3.45E-05	no expression (5.1)
CG4461	involved in heat response	down	0.081458	no expression (6)
Aps	encodes a diphosphoinositol-polyphosphate diphosphatase with hydrolysis functions	down	0.082274	moderate (386.9)
Corazonin receptor	involved in ethanol-induced behavior and ethanol metabolism	up	0.051862	low (55)
CG13458	thought to localize to the mitochondrial inner membrane	down	0.045032	low (48.1)
CG17027	has inositol monophosphate 1-phosphatase activity and be involved in inositol metabolic processes	down	0.005325	low (24.6)
mitochondrial ribosomal protein S31	thought to be a structural component of the ribosome and be involved in translation	down	3.13E-07	moderate (237.1)
artichoke	encodes a leucine-rich extracellular matrix protein that contributes to cilium assembly and integrity	up	0.090325	low (31.7)
CG11370	expressed in the dorsal trunk primordium	down	0.050380	no expression (0.7)
CG18473	has hydrolase activity and is involved in catabolic processes	up	4.08E-05	no expression (7.1)
Pinin	involved in mRNA splicing	down	7.47E-05	moderate (123)
CG14715	known to localize to the endomembrane system	down	3.49E-05	moderate (114.5)
CG9813	known to be expressed in the adult head	down	0.046834	very high (2028.2)
CG9759	no information	down	0.032216	no data
Adenosine deaminase-related growth factor C	thought to have adenosine deaminase activity and be involved in the adenosine catabolic process	down	0.077174	low (78.4)
CG14880	has chitin binding activity and involved in chitin metabolism	down	0.053265	no expression

				(1.7)
CG11951	thought to have metalloaminopeptidase activity and zinc ion binding activity	down	0.076151	no expression (.5)
Regulatory particle triple-A ATPase 6-related	predicted to have ATPase and TBP protein binding activity	down	0.087649	no expression (5.5)
PIP4K	encodes an enzyme that that is involved in mTOR signaling and cell size control	down	0.041810	no data
UDP-glycosyltransferase family 317 member A1	thought to have UDP-glycosyltransferase activity	down	0.098038	low (53.7)
lncRNA:CR13130	no information	up	4.11E-14	moderate (114.6)
Cytochrome P450 309a2	predicted to be involved in insect hormone and synthetic hormone metabolism	up	2.34E-11	no expression (5.6)
Apollo	involved in protein transport to the nucleus	up	0.001535	no data
Glutamate receptor IIC	encodes a subunit of muscle glutamate receptor	up	2.10E-07	no expression (6.8)
CG30033	expressed in adult fat bodies	down	0.000393	no expression (2.3)
CG30082	thought to have serine-type endopeptidase activity	down	0.071674	no expression (3.4)
Ipk1	has a role in protein phosphorylation	up	0.004797	no expression (7.6)
Maltase A5	has alpha-glucosidase activity and is involved in metabolizing carbohydrates	up	0.047306	moderate (453.6)
CG30472	no information	up	0.004797	no expression (0.7)
CG31643	thought to have protein kinase activity	up	0.032216	low (54.5)
CG31664	no information	up	0.008963	low (9.4)
CG31997	known to localize to cytosol	down	0.019329	moderate (290.5)
CG32428	known to be expressed in the extended germ band embryo	down	0.027758	moderate (213.4)
Syndapin	encodes a protein with roles in regulating cellularization, endocytosis and membrane tubulation	down	0.071674	moderate (340.8)
p24-related-2	encodes a protein involved in oviposition post-mating	up	0.000122	low (17.9)
Reticulon-like 1	encodes a protein that promotes ER tubule formation	down	1.87E-10	very high (1399.7)
Ventrally-expressed protein D	encodes a protein that is expressed in the mesoderm and has roles in development	up	0.050380	moderate (181.5)

CG7600	no information	up	9.30E-06	moderate (218.9)
debra	encodes a transcriptional coactivator	up	4.17E-06	moderate (157.7)
CG34150	expressed in the embryonic midgut, endoderm and midgut primordium	up	0.088378	low (21.1)
CG34195	thought to be involved in tRNA methylation and has predicted tRNA methyltransferase activity	up	1.56E-12	low (20.3)
CG34423	thought to have ATPase binding and ATPase inhibitor activity	up	2.30E-07	no data
Methenyltetrahydrofolate synthetase	though to be involved in tetrahydrofolate interconversion	up	0.066414	no expression (12.2)
palisade	encodes a protein essential to coordinating assembly during oogenesis	down	1.87E-07	no expression (6.1)
CG42260	encodes a protein that makes up a subunit of a nucleotide-gated ion channel	up	2.43E-10	no data
spaghetti-squash activator	encodes a myosin light chain kinase-like protein that is required for starvation-induced autophagy	down	0.006512	no data
flare	encodes a protein that is involved in F-actin disassembly	down	0.002745	moderate (106.4)
Uncoordinated 115b	thought to have actin filament binding activity and be involved in lamellipodium assembly	up	2.65E-19	low (17.9)
CG42615	no information	up	0.029931	no data
Sterol regulatory element binding protein	encodes a membrane protein that has roles of regulating lipogenesis and activating the transcription of lipogenic genes	up	0.004320	moderate (314)
CR43086 pseudogene	no information	down	0.084479	no data
Nucleoporin 160kD	encodes a component of the nuclear pore complex and is involved in polyA+ RNA transport	up	0.004320	no data
Myc	encodes a transcription factor that is involved in cell growth and proliferation	up	0.071674	no data
antisense RNA:CR43609	no information	up	0.042707	no data
lncRNA:CR43785	no information	up	0.083834	no data
antisense RNA:CR43957	no information	up	1.49E-09	no data
CR44391 pseudogene	no information	up	1.95E-08	no data
lncRNA:CR44525	no information	up	0.050380	no data

lncRNA:CR44608	no information	down	0.003224	no data
lncRNA:CR45267	no information	up	0.074123	no data

Table 2. Description of function of 172 KD//Gal4 DEGs, their direction of regulation, expression pattern in the brain. All gene descriptions were found via FlyBase (FB2021_03) and expression data was accessed via the FlyAtlas Anatomical Microarray Expression Data (FlyAtlas-RNA.adult, FB2021_03).

Category		Genes	% (involved genes/total genes)	p-value	Benjamini adj. p-value
Biological Process	cell adhesion	ImpL2, His3.3A, PIP4K, Pnn, RhoL, lox2, shg	4.10%	7.20E-04	3.10E-01
	transport	CG2823, CG5973, Obp56a, Obp56h, p24-2, st	3.50%	9.60E-03	1.00E+00
	sleep	Amy-p, CG10383, Rab27, CG5973, Drat, l(2)tid	3.50%	1.30E-02	1.00E+00
	hemocyte migration	Ras85D, RhoL, shg	1.80%	2.00E-02	1.00E+00
	insecticide catabolic process	Cyp6d2, Cyp28d2, Cyp6a2	1.80%	2.80E-02	1.00E+00
	regulation of cell shape	raw, His3.3A, PIP4K, RhoL, sqa	2.90%	3.20E-02	1.00E+00
	protein folding	CG14715, CG5554, Csp, Hsp68, l(2)tid	2.90%	3.50E-02	1.00E+00
	response to DDT	Cyp6d2, Cyp28d2, Cyp6a2	1.80%	4.10E-02	1.00E+00
	cell proliferation	Adgf-C, Myc, cutlet, Ras85D	2.30%	4.10E-02	1.00E+00
	positive regulation of cell growth	Myc, PIP4K, Ras85D	1.80%	4.30E-02	1.00E+00
	cellular response to ethanol	CrzR, Drat	1.20%	4.50E-02	1.00E+00
	gonad development	raw, en, shg	1.80%	4.60E-02	1.00E+00
	negative regulation of ATPase activity	CG34423, CG34424	1.20%	5.60E-02	1.00E+00
	eye pigment precursor transport	st, w	1.20%	5.60E-02	1.00E+00
	response to hypoxia	Drat, Hsp70Bb, Sur	1.80%	5.80E-02	1.00E+00
	border follicle cell migration	Rab7, Ras85D, RhoL, flr, shg	2.90%	6.40E-02	1.00E+00

	gonad morphogenesis	raw, shg	1.20%	6.70E-02	1.00E+00
	exosomal secretion	Rab27, Rab7	1.20%	6.70E-02	1.00E+00
	regulation of apoptotic process	Myc, Ras85D, l(2)tid	1.80%	7.70E-02	1.00E+00
	gonadal mesoderm development	ftz, shg	1.20%	8.80E-02	1.00E+00
Cellular Component	endoplasmic reticulum membrane	CG14715, Cyp309a2, Cyp6d2, Cyp28d2, Cyp6a2, Rtnl1, SREBP, l(2)not, p24-2	5.30%	5.80E-03	4.70E-01
	synapse	Rab27, CG5397, Rab7, GluRIIC, Jhe, lbm	3.50%	8.60E-03	4.70E-01
	membrane	CG11210, CG14340, drongo, Phm, Ras85D, SREBP, c-cup, lox2, st, w	6.40%	2.10E-02	7.60E-01
	plasma membrane	CG11951, raw, CG4839, CG5807, Rab7, Csp, FER, Fas3, PRL-1, PIP4K, Ras85D, RhoL, Synd, l(3)mbn, st, shg, w	9.90%	3.00E-02	8.40E-01
	nuclear membrane	CG32165, Fas3, SREBP	1.80%	6.50E-02	1.00E+00
	organelle membrane	Cyp309a2, Cyp6d2, Cyp28d2, Cyp6a2	2.30%	7.30E-02	1.00E+00
	extracellular matrix	CG13670, resilin, Cpr47Ef, atk, l(3)mbn	2.90%	9.70E-02	1.00E+00
Molecular Function	ATP binding	Act87E, Ipk1, cutlet, CG34424, CG4839, CG5660, FER, Hsp68, Hsp70Bb, Aats-pro, Pka-C3. Rpt6R, Sur, l(2)tid, st, sqa, w	9.90%	2.20E-02	1.00E+00
	monooxygenase activity	Cyp309a2, Cyp6d2, Cyp28d2, Cyp6a16, Cyp6a2	2.90%	2.40E-02	1.00E+00

	heme binding	Cyp309a2, Cyp6d2, Cyp28d2, Cyp6a16, Cyp6a2, c-cup	3.50%	2.40E-02	1.00E+00
	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	Cyp309a2, Cyp6d2, Cyp28d2, Cyp6a16, Cyp6a2	2.90%	2.60E-02	1.00E+00
	oxidoreductase activity	Cyp309a2, CG2064, Cyp6d2, CG5653, Cyp28d2, Cyp6a2, Drat	4.10%	3.30E-02	1.00E+00
	ATPase inhibitor activity	CG34423, CG34424	1.20%	3.70E-02	1.00E+00
	structural constituent of cuticle	CG1136, CG13670 resilin, Cpr47Ef, l(3)mbn	2.90%	5.00E-02	1.00E+00
	pigment binding	st, w	1.20%	6.00E-02	1.00E+00
	receptor activity	CG5397, Fas3, Jhe, shg	2.30%	6.20E-02	1.00E+00
	GTP binding	Rab27, Rab7, Pu, Ras85D, RhoL, Srp54k	3.50%	6.70E-02	1.00E+00
	GTPase activity	Rab27, Rab7, Ras85D, RhoL, Srp54k	2.90%	8.30E-02	1.00E+00
	ATPase binding	CG34423, CG34424	1.20%	9.50E-02	1.00E+00
	ATPase activity, coupled to transmembrane movement of substances	Sur, st, w	1.80%	9.90E-02	1.00E+00

Table 3. Summary of the GO terms tagged for each category (biological process, cellular component and molecular function), genes involved in each term, the percentage of total DEGs that the involved genes make up, their p-values and Benjamini adjusted p-values.

Annotation Cluster 1	Enrichment Score 1.46	Count	p-value	Benjamini adj. p-value
	endoplasmic reticulum membrane	9	5.80E-03	4.70E-01
	monooxygenase activity	5	2.40E-02	1.00E+00
	heme binding	6	2.40E-02	1.00E+00
	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	5	2.60E-02	1.00E+00
	insecticide catabolic process	3	2.80E-02	1.00E+00
	oxidoreductase activity	7	3.30E-02	1.00E+00
	response to DDT	3	4.10E-02	1.00E+00
	organelle membrane	4	7.30E-02	1.00E+00
	oxidation-reduction process	9	1.00E-01	1.00E+00
	iron ion binding	5	1.10E-01	1.00E+00
Annotation Cluster 2	**Enrichment Score 1.15**	**Count**	**p-value**	**Benjamini adj. p-value**
	border follicle cell migration	5	6.40E-02	1.00E+00
	GTP binding	6	6.70E-02	1.00E+00
	GTPase activity	5	8.30E-02	1.00E+00
Annotation Cluster 3	**Enrichment Score 0.93**	**Count**	**p-value**	**Benjamini adj. p-value**
	structural constituent of cuticle	5	5.00E-02	1.00E+00
	extracellular matrix	5	9.70E-02	1.00E+00
	structural constituent of chitin-based larval cuticle	4	1.40E-01	1.00E+00
	chitin-based cuticle development	4	2.70E-01	1.00E+00
Annotation Cluster 4	**Enrichment Score 0.24**	**Count**	**p-value**	**Benjamini adj. p-value**
	catalytic step 2 spliceosome	3	4.80E-01	1.00E+00
	precatalytic spliceosome	3	5.50E-01	1.00E+00
	mRNA splicing, via spliceosome	3	7.40E-01	1.00E+00

Table 4. Summary of functional annotation clusters, their enrichment scores, number of genes involved, p values and Benjamini adjusted p-values.

An initial goal for our analysis was to determine whether our 172 KD vs Gal4 DEGs were represented or overrepresented within:

- our six candidate genes of interest (*spin*, *unc79*, *Bx*, *CtBP*, *Fas2* or *For,* Chapter 2, Table 5)
- the 342 genes bound by Mef2 (Sivachenko et. al 2013)
- known fly alcohol genes and orthologs of worm alcohol genes (Grotewiel & Bettinger 2015)
- fly orthologs of genes nominally associated with SRE (Schmitt et. al 2019)
- human genes associated with externalizing behavior (Danielle Dick, personal communication)

As necessary, we used fly genes or fly orthologs of human and worm genes (defined by DIOPT scores ≥5) for these analyses. We found that eight fly genes overlapped between the 172 DEGs and 342 *Mef2* bound genes (Sivachenko et. al 2013), no fly genes overlapped between the DEGs and our 6 candidate genes (*unc79*, *spin*, *Bx*, *CtBP*, *Fas2*, *for*), two fly genes overlapped between the DEGs and known fly alcohol genes (reviewed by Grotewiel & Bettinger 2015), one fly gene overlapped between the DEGs and fly orthologs of worm alcohol genes (reviewed by Grotewiel & Bettinger 2015), no fly genes overlapped between the DEGs and fly orthologs of SRE-related genes (Schmitt et. al 2019), and nine genes overlapped between the DEGs and fly orthologs of human genes associated with externalizing behavior (Dick, personal communication 2020, Linnér et. al 2020) (Table 5). To determine whether the amount of overlap between the DEGs and each other gene lists was significantly more than expected by chance, we performed Fisher's exact test using a custom R script (generously provided by Mike Miles and Maren Smith, Virginia Commonwealth University, detailed in the Appendix). The number of overlapping genes was not significantly greater than expected by chance for any of the six gene sets (Table 6), suggesting that the 172 DEGs and the other gene sets of interest might be involved in biologically distinct processes. The inclusion of the *white* gene is of somewhat

suspect biological value given that it is a marker for the Gal4 and RNAi transgenes used in our study, but is included here for completeness.

The lack of overlap between the 342 *Mef2* bound genes and 172 KD//Gal4 DEGs could be due to several reasons. First, if Mef2 regulates the same or similar sets of genes in muscle and neurons, it is possible that the expression of *Mef2* regulated genes in the muscle could mask the effect *Mef2* knockdown on gene expression in neurons. If this is the case, then it is possible this study primarily identified neuron-specific genes downstream of the 342 Mef2 bound genes identified by Sivachenko et. al (Sivachenko et. al 2013). Another possibility is that *Mef2* knockdown during development and into adulthood could have led to gene expression changes that are not represented by the 342 Mef2 bound genes, as these genes are experimentally defined as being Mef2 bound in adulthood. Additionally, because Sivachenko et. al identified the initial Mef2 bound genes using head tissue, it may be possible that their study largely identified genes bound by Mef2 in head muscle. If that is the case, and the genes *Mef2* is bound by or regulates do not substantially overlap with those bound or regulated in neurons, we would not necessarily expect to observe a large overlap between our 172 KD//Gal4 DEGs and the 342 Mef2 bound genes identified by Sivachenko et. al (Sivachenko et. al 2013).

DEGs vs. Mef2 bound genes	DEGs vs. 6 candidate genes	DEGs vs. fly alcohol genes	DEGs vs. fly orthologs of worm alcohol genes	DEGs vs. fly orthologs of human SRE genes	DEGs vs. fly orthologs of human externalizing factor genes
- raw *- white* *- drongo* *- PRL-1 phosphatase* *- lilipod* *- Death resistor Adh domain containing target* *- CG9813* *- Reticulon-like 1*	none	*- white* *- Corazonin receptor*	*- Sterol regulatory binding protein*	none	*- Cytochrome P450 6a2* *- FER tyrosine kinase* *- Rab7* *- papi* *- CG11319* *- CG4839* *- Valyl-tRNA synthetase, mitochondrial* *- Syndapin* *- Sterol regulatory element binding protein*

Table 5. Summary of each of the genes that overlap between the 172 genes identified as differentially expressed between the KD and Gal4 groups and various gene lists of interest.

Gene Set 1	Gene Set 2	Genome	Overlap	Expected	Odds Ratio	p-value
172 DEGs	342 *Mef2* bound	Fly (13,968 genes)	8	4.2113	1.8286	0.0936
172 DEGs	6 candidate genes	Fly (13,968 genes)	0	0.0739	0.0	1.0
172 DEGs	91 fly alcohol genes	Fly (13,968 genes)	2	1.1206	1.7509	0.3229
172 DEGs	51 fly orthologs of worm alcohol genes	Fly (13,968 genes)	1	0.6280	1.5667	0.4776
172 DEGs	22 fly orthologs of human SRE genes	Fly (13,968 genes)	0	0.2709	0.0	1.0
172 DEGs	817 fly orthologs of human externalizing genes	Fly (13,968 genes)	9	10.0604	0.8307	0.7503

Table 6. Fisher's exact test data to determine whether the number of overlapping genes between our 172 identified DEGs and each gene list of interest is significant. The amount of overlap between our DEGs and no other gene set is significant.

3e. Analysis of KD//Gal4//RNAi DEGs

Our initial analysis indicated that 51 genes were differentially expressed in both the KD vs Gal4 comparison and the KD vs RNAi comparison (Figure 10). We surmised that the most relevant genes would be consistently up-regulated or consistently down-regulated in these two sets of DEGs. Among these 51 DEGs, we therefore identified those genes that were up-regulated or down-regulated within both the KD//Gal4 and KD//RNAi. We found 13 genes that were up-regulated in both the KD//Gal4 and KD//RNAi sets (Figure 11A, Table 7), and similarly found 18 genes that were consistently down-regulated in these same two gene sets (Figure 11B, Table 7). Of these 31 genes, 27 have brain expression data available for them (FlyAtlas-RNA.adult, FB2021_03). Twelve of these 27 genes are moderately, highly or very highly expressed in the brain, with 23 total having detectable expression in this tissue (Table 7), again strongly suggesting that our workflow led to genes with potential roles in nervous system function and behavior. Interestingly, four of the genes in this set of 31 DEGs were also identified as being bound by Mef2 (Sivachenko et. al 2013) which was greater than expected by chance (Tables 8 and 9). Although this set of 31 DEGs did not significantly overlap with the other five gene sets of interest (Table 8), three of the four genes common to the KD//Gal4//RNAi DEGs (*raw, PRL-1 phosphatase* and *CG9813*) and the 342 genes bound by Mef2 (Table 9) could be high priority candidates for future studies on the role of genes directly downstream of Mef2 in ethanol sedation. The inclusion of the fourth overlapping gene, *white*, is of somewhat suspect biological value, given that it is a marker for the Gal4 and RNAi transgenes used in our study, but is included here for completeness.

PRL-1 phosphatase has three human orthologs with DIOPT scores ≥ 8 each: *PTP4A2*, *PTP4A1* and *PTP4A3* (DIOPT scores are 12, 11, 8, respectively). Neither *PTP-1 phosphatase* nor its orthologs have been implicated in fly or human ethanol responses. *Raw* and *CG9813* do not have any human orthologs.

Ontology analysis of the 31 genes (Table 10) indicates that none of the enriched GO terms are significant based on their Benjamini-Hochberg adjusted p-value. The relatively small size of the gene set (31) and lack of significantly terms limits interpretation of these results.

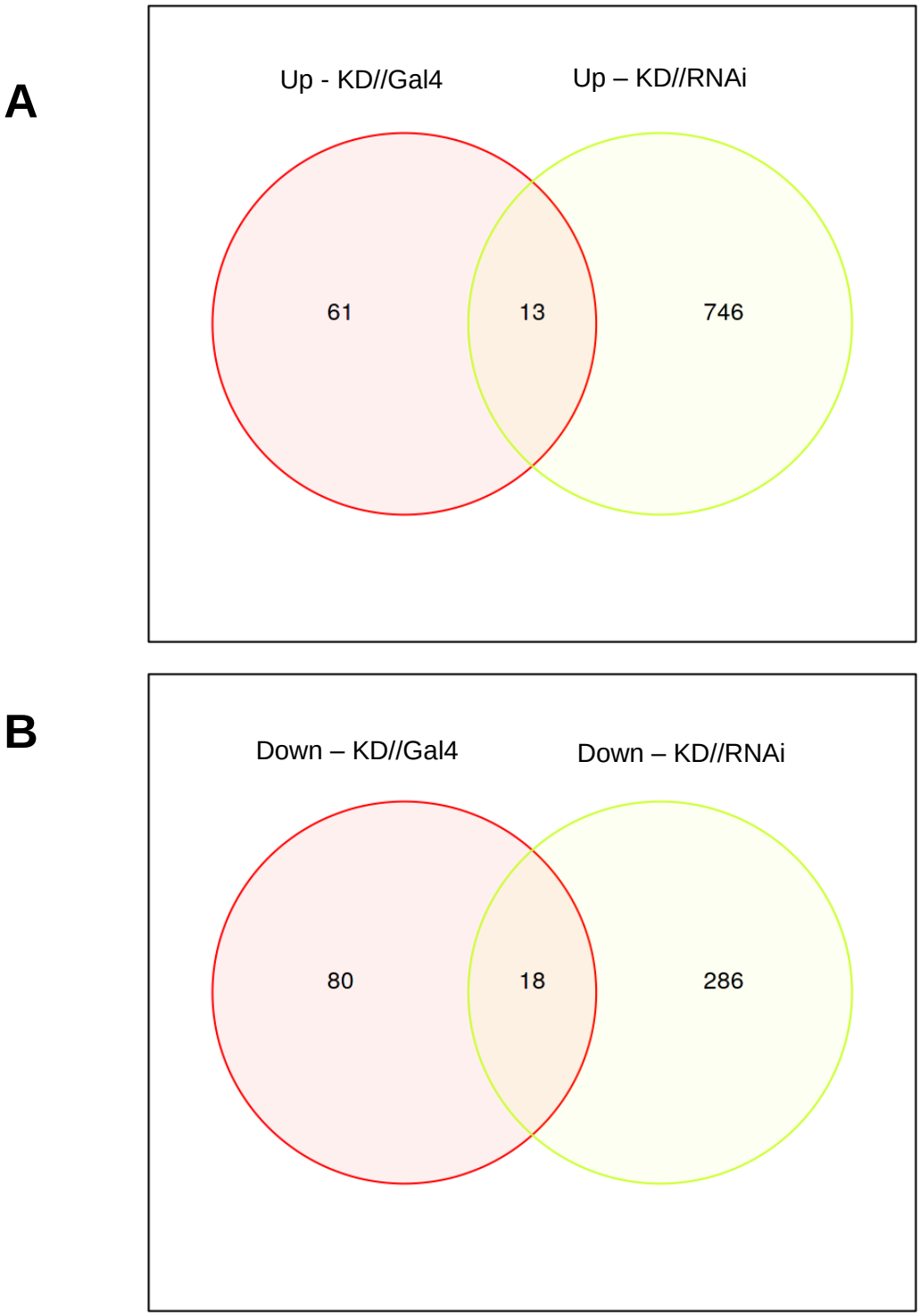


Figure 11. Venn diagrams showing the number genes found to be up or down regulated both in comparison of the knockdown group vs. Gal4 control and the RNAi control. Panel A shows overlap of upregulated genes and panel B shows overlap of downregulated genes.

Gene	Function	Up- or down-regulated	Expressed in neurons?
Cytochrome p450 6a2	involved in breakdown of synthetic insecticides and insect hormone metabolism	up	low (47.7)
Fushi tarazu	predicted to play a role in specifying neuronal identity and known to be required in embryogenesis	up	no expression (1.3)
Juvenile hormone esterase	involved in ethanol response and catabolism of juvenile hormone	up	no expression (3.7)
Bicoid interacting protein 1	recruits Sin3A-HDAC1 by interacting with transcription factor	up	moderate (172.1)
CG3823	thought to have phosphatidylinositol biphosphate binding activity	up	low (7.3)
Arylalkylamine N-acetyltransferase-like 2	has acyl-coA-N-acyltransferase and aralkylamine N-acetyltransferase activity	up	low (42.9)
CG12910	thought to have UDP-galactosyltransferase and N-glycan processing activity	up	low (43.7)
artichoke	encodes a leucine-rich extracellular matrix protein that contributes to cilium assembly and integrity	up	low (31.7)
Cytochrome p450 309a	predicted to be involved in insect hormone and synthetic hormone metabolism	up	no expression (5.6)
Apollo	involved in protein transport to the nucleus	up	no data
Methenyltetrahydrofolate synthetase	though to be involved in tetrahydrofolate interconversion	up	no expression (12.2)
white	ABC-type guanine transporter involved in transporting cyclic GMP, various amines and pigments	up	low (13)
Heat-shock protein-70Bb	encodes a protein involved in heat shock and hypoxia response	up	no data
Pka-C3	encodes a cAMP-dependent protein kinase	down	moderate (276.9)
Ecdysone-inducible gene L2	involved in embryogenesis and normal nervous system development	down	moderate (421.7)
raw	encodes a membrane protein involved in dendritic patterning and	down	moderate

	localization of JNK signaling components		(292.9)
shotgun	calcium-dependent cell adhesion protein and has roles in cell sorting, oogenesis and body asymmetry	down	no expression (9)
Recombination repair protein 1	involved in cellular response to oxidative stress via DNA repair mechanisms	down	moderate (107)
Troponin C at 41C	encodes a protein that binds calcium and regulates muscle contraction	down	no expression
PRL-1 phosphatase	encodes a growth inhibitor	down	high (608.8)
CG4615	predicted to be involved in cytolysis	down	moderate (301.9)
CG1636	known to be expressed in the adult head	down	moderate (165.5)
CG14340	thought to be involved in vesicle-mediated transport	down	low (57.1)
CG10257	involved in I-kappaB kinase/NF-kappaB signaling	down	moderate (113.2)
CG6805	has phosphatidylinositol-4,5-biphosphate 5-phosphatase activity and be involved in inositol phosphate dephosphorylation	down	moderate (111)
CG4461	involved in heat response	down	no expression (6)
mitochondrial ribosomal protein S31	thought to be a structural component of the ribosome and be involved in translation	down	moderate (237.1)
CG9813	known to be expressed in the adult head	down	very high (2028.1)
CG9759	no information	down	no data
CG30033	expressed in adult fat bodies	down	no expression (2.3)
spaghetti-squash activator	encodes a myosin light chain kinase-like protein that is required for starvation-induced autophagy	down	no data

Table 7. Description of function of 31 DEGs, their direction of regulation, expression pattern in the brain. All gene descriptions were found via FlyBase (FB2021_03) and expression data was accessed via the FlyAtlas Anatomical Microarray Expression Data (FlyAtlas-RNA.adult, FB2021_03).

Gene Set 1	Gene Set 2	Genome	Overlap	Expected	Odds Ratio	p-value
31 DEGs	342 *Mef2* bound	Fly (13,968 genes)	4	0.3893	5.1263	0.0106
31 DEGs	6 candidate genes	Fly (13,968 genes)	0	0.0133	0.0	1.0
31 DEGs	91 fly alcohol	Fly (13,968 genes)	1	0.2020	4.9060	0.1908
31 DEGs	51 fly orthologs of worm alcohol	Fly (13,968 genes)	1	0.1132	0.0	1.0
31 DEGs	22 fly orthologs of human SRE	Fly (13,968 genes)	0	0.0488	0.0	1.0
31 DEGs	817 fly orthologs of human externalizing	Fly (13,968 genes)	1	1.8132	0.5180	1.0

Table 8. Fisher's exact test contingency tests to determine whether the number of overlapping genes between our 31 DEGs differentially expressed in the same direction and each gene list of interest is significant. The amount of overlap between our DEGs and 342 *Mef2* bound genes is significant; the amount between our previously identified 6 genes of interest, known fly alcohol genes, fly orthologs of known worm alcohol genes, fly orthologs of previously identified genes implicated in SRE and fly orthologs of human genes involved in externalizing behavior is not.

DEGs vs. Mef2 bound genes	DEGs vs. 6 candidate genes	DEGs vs. fly alcohol genes	DEGs vs. fly orthologs of worm alcohol genes	DEGs vs. fly orthologs of human SRE genes	DEGs vs. fly orthologs of human externalizing factor genes
- *white* - *raw* - *PRL-1 phosphatase* - *CG9813*	none	- *white*	none	none	- *Cytochrome p450 6a2*

Table 9. Summary of each of the genes that overlap between the 31 genes differentially expressed between the KD and Gal4 groups and various gene sets of interest.

Category	Term	Genes	% (involved genes/total genes)	P-value	Benjamini adj. p value
Biological Process	gonad morphogenesis	raw, shg	6.5	1.3E-2	1.00E+00
	gonadal mesoderm development	ftz, shg	6.5	1.7E-2	1.00E+00
	morphogenesis of an epithelium	raw, shg	6.5	5.5E-2	1.00E+00
	gonad development	raw, shg	6.5	6.4E-2	1.00E+00
	germ cell migration	ftz, shg	6.5	8.2E-2	1.00E+00
Cellular Process	none				
Molecular Function	ATP binding	CG34424, Hsp70Bb, Pka-C3, sqa, w	16.1	6.2E-2	1.00E+00

Table 10. Summary of the GO terms tagged for each category (biological process, cellular component and molecular function), genes involved in each term, the percentage of total DEGs that the involved genes make up, their p-values and Benjamini adjusted p-values.

4. Discussion

We performed an RNA-seq study to better understand the role of *Mef2* in ethanol sedation. We confirmed the quality of our RNA samples collected for this project in several ways: by measuring A260/A280 ratios, assessing data from the Agilent Bioanalyzer and interpreting RIN numbers and DV200 data provided by GeneWiz. Additionally, we confirmed that sequencing quality was reliable and accurate via the mean quality scores provided by GeneWiz, that sequencing read depths were comparable across samples, and that the genotypes of the flies collected and the identities of the resulting samples were accurate. We also designed the study to prevent batch effects, and normalized the genetic background of all flies used. We are therefore confident in the overall design of our study and the RNA-seq data derived from it.

We were expecting to see larger-scale differences in gene expression in the KD//Gal4 comparison and clear distinctions between the three genotypes in PCA analyses. The low number of DEGs could potentially indicate that *Mef2* does not regulate many genes; however, given that it is a transcription factor (Black & Olson 1998; Taylor & Hughes 2017), is known to bind 342 genes (Sivachenko et. al 2013) and has a reproducible impact on ethanol sedation patterns (Schmitt et. al 2019, Myers 2020, Chapter 2), we do not think this is likely. A reasonable explanation might be that our experimental design might be masking the true impact of pan-neuronal *Mef2* knockdown. *Mef2* is highly expressed in head muscle (Velasco et. al 2006), so sequencing RNA isolated from fly brains alone might provide increased sensitivity and allow us to capture the full scope of global gene expression changes caused by neuronal knockdown of *Mef2*.

We used our RNA-seq data to address the biological functions of two differentially expressed gene sets (KD//Gal4 and KD//Gal4//RNAi) and to address whether these two sets of differentially expressed genes significantly overlapped with our previously identified candidate genes (Chapter 2), 342 *Mef2* bound genes (Sivachenko et. al 2013), known fly alcohol genes

(Grotewiel & Bettinger 2015), fly orthologs of known worm alcohol genes (Grotewiel & Bettinger 2015), fly orthologs of human genes previously implicated in SRE and fly orthologs of genes implicated in human externalizing behaviors. Although the set of 172 KD//Gal4 DEGs did not significantly overlap with any other gene list of interest, there are still some meaningful conclusions we may be able to draw. For one, approximately 70% of the DEGs were expressed in the brain at some level (FlyAtlas-RNA.adult, FB2021_03), indicating that our experimental design is able to capture expression changes occurring in the brain. Though the amount of overlap is not significant, the genes that are shared between the KD//Gal4 DEGs and each other set may still be worthwhile candidates to pursue in future studies. GO analysis highlighted a few terms of interest: sleep and ethanol response. Previous studies have shown that *Mef2* is required for the daily fasciculation/defasciculation cycle, and that its transcription is regulated by the master circadian rhythm transcription complex (Sivachenko et. al 2013). *Mef2*'s role in ethanol response is a key focus of this thesis, as well as previous literature (Schmitt et. al 2019, Adhikari et. al 2018). The enrichment of these GO terms, which are consistent with *Mef2*'s previously identified functions, could serve as evidence of the validity of our experiment. However, only two genes are associated with the ethanol response GO term: *Death resistor Adh domain containing target (Drat)* and *Corazonin receptor*. While this is quite a low number of genes, it is interesting to note that both overlap with other gene lists of interest - *Drat* is known to bind Mef2, and *Corazonin receptor* is a known fly alcohol gene. This indicates that these genes could be potential candidate genes for future behavioral studies.

The amount of overlap (four genes) between the 31 KD//Gal4//RNAi genes and the 342 *Mef2* bound genes is significant, and each of the four are expressed in the brain at least some extent, indicating that these genes may also be high priority candidates for future experiments. Additionally, two of the four (*white* and *PRL-1 phosphatase*) have multiple strong orthologs (defined as DIOPT≥5), indicating that studies done on these genes could eventually impact understanding of AUD in humans. *white* is also a known fly alcohol gene. PRL-1 phosphatase

and the other two overlapping DEGs (*raw* and *CG9813*) are not known to be involved in any aspect of ethanol response, experiments pertaining to them could still be meaningful in that they may provide a better understanding of the role of other genes downstream of *Mef2* in ethanol sedation, as well as a better understanding of *Mef2*'s interactions with its downstream genes. In the larger set of 31 KD//Gal4//RNAi genes, *Juvenile hormone esterase* (*Jhe*), is known to be involved in ethanol response and is known to be upregulated in both KD//Gal4 and KD//RNAi comparisons, indicating that it may also be a promising candidate gene.

Future iterations of this experiment might be successful in identifying more DEGs if fly brains were used as the starting material. Additionally, we could perform qPCR of fly heads with a knockdown of *Mef2* to see whether we get the same results in terms of decrease of *Mef2* expression. If we do, this could be further evidence to extract RNA from fly brains in the future. We could also perform qPCR with reagents specifically for the six candidate genes identified in Chapter 2, or the potential new candidate genes laid out in this chapter to understand how knockdown of *Mef2* affects those genes specifically.

CHAPTER 4: DISCUSSION

GWAS performed by Schmitt et. al (Schmitt et. al 2019) identified human MEF2B as a gene implicated in SRE. *Mef2*, the fly ortholog of MEF2B has been shown to have an impact on ethanol sedation behaviors in *Drosophila* (Schmitt et. al 2019, Myers 2020); flies with pan-neuronal expression RNAi transgenes against *Mef2* and mutations in the gene display significantly decreased ethanol sedation sensitivity. *Mef2* is a transcription factor (Black & Olson 1998, Taylor & Hughes 2017), and a ChIP-seq was performed to identify any genes bound by *Mef2* (Sivachenko et. al 2013). Additionally, work by other groups has shown that a gene downstream *of Mef2* influences ethanol tolerance and preference (Adhikari et. al 2018) The goal of our first aim of this study was to begin testing the hypothesis that one or more genes of interest downstream of *Mef2* might influence ethanol sedation.

Through a collaboration with Danielle Dick, we determined that 38 human orthologs of the genes bound by Mef2 were associated with human externalizing behavior (Dick 2020, personal communication, Lanier et. al 2020). We found that six of these 38 genes had been implicated in behavioral responses to alcohol in flies and worms (Grotewiel & Bettinger 2015) and confirmed that 1 of them (*spin*) was associated with gene expression changes related to alcohol consumption in humans (Bacanu, personal communication). These six genes (*spin*, *unc79*, *Bx*, *CtBP*, *Fas2* and *for*) became our high priority candidate genes for analysis in Chapter 2. Using available RNAi reagents, these genes were tested to determine whether RNAi-mediated knockdown of any of the genes influenced ethanol sedation.

Pan-neuronal expression of RNAi targeting two genes, *spin* and *unc79,* produced consistent changes in ethanol sedation. Expression of all viable RNAi transgenes targeting these genes decreased ethanol sedation sensitivity. Expression of RNAi targeting other candidate genes either did not produce an effect on ethanol sedation or did not produce a consistent effect when multiple RNAi transgenes against one gene were tested. This could be

due to the genes potentially not functioning in neurons, or the specific RNAi transgenes not adequately knocking down the target gene. To address this, future iterations of this study could focus on manipulating the candidate genes via mutations rather than RNAi transgenes. Additionally, assessing the impact of overexpression or expression of dominant negatives could also provide insight into the potential role each candidate gene plays in ethanol sedation. Several of the RNAi reagents we tested were lethal. Expressing these genes only in adulthood or only in specific neurons to bypass lethality might also aid in better understanding of the role these genes play in ethanol sedation.

The second component of this project was an RNA-seq to identify the genes regulated by *Mef2*. We isolated RNA from fly heads of three genotypes, the *elav*-GAL4/+;v15550/+ knockdown genotype (KD), an *elav*-Gal4/+ control (Gal4) and a v15550/+ control (RNAi) and sent them for sequencing. We established confidence in the RNA quality, sequencing quality and that there was no disordering of samples.

In total, we identified 172 DEGs in the KD//Gal4 comparison, 1,063 DEGs in the KD//RNAi comparison and 2,238 DEGs in the Gal4//RNAi comparison. Though we do not understand and were not expecting to observe such large expression differences between our two controls, we know that the differences do not stem from differences in genetic background, mix-up of RNA samples, sample or sequence quality or batch effects, due to our experimental design. As the two controls consistently do not show behavioral differences in sedation experiments, we focused the first part of our analyses on the 172 KD//Gal4 DEGs.

We found that 70% of these DEGs were expressed in the fly brain to some extent, indicating that our experiment was able to identify gene expression changes at the neuronal level. Gene ontology terms representing an expansive spectrum of biological, cellular and molecular processes were found to be enriched, indicating that these DEGs may have a far-reaching scope. Specific GO terms such as sleep and ethanol response were also enriched – these terms are consistent with previously identified functions of *Mef2* (Sivachenko et. al 2013;

Schmitt et. al 2019; Adhikari et. al 2018), further highlighting the validity of our experiment. We conducted overlap analyses to understand whether genes differentially expressed between the KD and Gal4 groups overlapped significantly with other gene lists of interest: our six candidate genes (Chapter 2), known Mef2 bound genes (Sivachenko et. al 2013), known fly alcohol genes and fly orthologs of known worm alcohol genes (reviewed in Grotewiel & Bettinger 2015), fly orthologs of human SRE genes (Schmitt et. al 2019) and fly orthologs of human externalizing factor genes (Dick, personal communication; Linnér et. al 2020). While we did not find that overlap between the DEGs and any other list of interest was significant, the two genes associated with the ethanol response GO term (*Drat* and *corazonin receptor*) are known to be Mef2 bound, and a known fly alcohol gene, respectively, and may be promising candidates for future experiments.

54 genes were shown to be differentially expressed between both the KD//Gal4 and KD//RNAi comparisons. We found that 31 of these genes were regulated in the same direction (18 are downregulated in both comparisons, and 13 are upregulated in both comparisons). 76.7% of these genes (23 out of 31) are expressed in the brain at some level, again demonstrating that our experimental design is able to identify gene expression changes in neurons and identify genes with neuronal roles. Interpretation of GO results is limited by the small number of DEGs; however, these genes may be involved in regulating development and GTP-dependent processes. Overlap analysis showed that a significant number of genes *(PTP-1 phosphatase*, *raw*, *white* and *CG9813*) are shared between the 31 DEGs and known Mef2 bound genes. *PTP-1 phosphatase* and *white* both have three strong human orthologs (DIOPT ≥ 5) each, and while neither they nor their human orthologs have known function in ethanol behaviors or response, they may be promising candidate genes. Additionally, while *raw* and *CG9813* do not have human orthologs, studying them further may still provide insight into *Mef2*'s regulation of its downstream genes and the role of *Mef2* regulated genes in ethanol sedation.

Overall, the results from the RNA-seq were not what we expected. There could be several explanations for this, (i) that *Mef2* doesn't actually regulate many genes. However, this is unlikely, due to the multiple studies that have implicated *Mef2* as a transcription factor (Black & Olson 1998, Taylor & Hughes 2017), the large number of genes that *Mef2* is known to bind to (Sivachenko et. al 2013) and the behavioral differences that we observe in ethanol sedation when the gene is knocked down. Another possible reason is batch effects, though this is also not likely as we took great care to collect flies, separate fly heads and perform RNA preps in a rotating manner and representatives from GeneWiz confirmed that the possibility of large-scale batch effects was highly unlikely from their end as well. Therefore, a more viable explanation may be that our starting material for the RNA-seq was not optimal. *Mef2* is expressed in neurons (Schmitt et. al 2019, Crittenden et. al 2018), but it is also expressed in head muscle (Valasco et. al 2006). Our knockdown of *Mef2* was pan-neuronal, not in the entire fly head. Therefore, in the future, sequencing RNA from fly brains might provide increased sensitivity to capture changes resulting from gene expression changes *Mef2* in future repetitions of this project. Targeting qPCR can also be done to understand individual relationships between *Mef2* and its downstream genes.

Overall, these types of experiments can provide a better understanding of the molecular mechanisms underlying AUD in humans, especially if implicated genes have orthologs in other model organisms or humans, as they could then be studied across species to get a better sense of their potentially conserved role in ethanol behaviors. *Drosophila* are powerful tools to study alcohol use, as their behavioral responses to ethanol, including locomotor and sedation behaviors, withdrawal and tolerance are quite conserved to human responses to alcohol (Grotewiel & Bettinger 2015). Therefore, identifying genes that are relevant to ethanol behaviors in fruit flies could be a springboard to examine the impact of those genes in human systems.

References

Adhikari, P., Orozco, D., & Wolf, F. W. (2017). *Mef2 induction of the immediate early gene Hr38/Nr4a is terminated by Sirt1 to promote ethanol tolerance*. Cold Spring Harbor Laboratory. http://dx.doi.org/10.1101/205666

Adkins, A., Hack, L., Bigdeli, T., Williamson, V., McMichael, G., Mamdani, M., Edwards, A., Aliev, F., Chan, R., Bhandari, P., Raabe, R., Alaimo, J., Blackwell, G., Moscati, A., Poland, R., Rood, B., Patterson, D., Walsh, D., Whitfield, J., . . . Montgomery, G. (2017). Genomewide Association Study of Alcohol Dependence Identifies Risk Loci Altering Ethanol Response Behaviors in Model Organisms. *Alcoholism: Clinical and Experimental Research*. Published. https://doi.org/10.1111/acer.13362

Alcohol use disorder: A comparison between DSM–IV and DSM–5. (n.d.). Retrieved July 30, 2021, from https://www.niaaa.nih.gov/publications/brochures-and-fact-sheets/alcohol-use-disorder-comparison-between-dsm

Aliev, F., Wetherill, L., Bierut, L., Bucholz, K. K., Edenberg, H., Foroud, T., & Dick, D. M. (2015). Genes associated with alcohol outcomes show enrichment of effects with broad externalizing and impulsivity phenotypes in an independent sample. *Journal of Studies on Alcohol and Drugs*, *76*(1), 38–46. https://doi.org/10.15288/jsad.2015.76.38

Babor, T., Holder, H., Caetano, R., Homel, R., Casswell, S., Livingston, M., Edwards, G., Österberg, E., Giesbrecht, N., Rehm, J., Graham, K., Room, R., Grube, J., Rossow, I., & Hill, L. (2010). Alcohol: No ordinary commodity. In *Alcohol: No Ordinary Commodity* (pp. 11–22). Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199551149.003.002

Bagnardi, V., Rota, M., Botteri, E., Tramacere, I., Islami, F., Fedirko, V., Scotti, L., Jenab, M., Turati, F., Pasquali, E., Pelucchi, C., Galeone, C., Bellocco, R., Negri, E., Corrao, G., Boffetta, P., & La Vecchia, C. (2014). Alcohol consumption and site-specific cancer risk:

A comprehensive dose–response meta-analysis. *British Journal of Cancer*, *112*(3), 580–593. https://doi.org/10.1038/bjc.2014.579

Barr, P. B., & Dick, D. M. (2019). The genetics of externalizing problems. In *Recent Advances in Research on Impulsivity and Impulsive Behaviors* (pp. 93–112). Springer International Publishing. http://dx.doi.org/10.1007/7854_2019_120

Barr, P. B., Salvatore, J. E., Wetherill, L., Anokhin, A., Chan, G., Edenberg, H. J., Kuperman, S., Meyers, J., Nurnberger, J., Porjesz, B., Schuckit, M., & Dick, D. M. (2020). A family-based genome wide association study of externalizing behaviors. *Behavior Genetics*, *50*(3), 175–183. https://doi.org/10.1007/s10519-020-09999-3

Berger, K. H., Heberlein, U., & Moore, M. S. (2004). Rapid and chronic: Two distinct forms of ethanol tolerance in drosophila. *Alcoholism: Clinical & Experimental Research*, *28*(10), 1469–1480. https://doi.org/10.1097/01.alc.0000141817.15993.98

Bettinger, J. C., Leung, K., Bolling, M. H., Goldsmith, A. D., & Davies, A. G. (2012). Lipid Environment Modulates the Development of Acute Tolerance to Ethanol in Caenorhabditis elegans. *PLoS ONE*, *7*(5), e35192. https://doi.org/10.1371/journal.pone.0035192

Black, B. L., & Olson, E. N. (1998). TRANSCRIPTIONAL CONTROL OF MUSCLE DEVELOPMENT BY MYOCYTE ENHANCER FACTOR-2 (MEF2) PROTEINS. *Annual Review of Cell and Developmental Biology*, *14*(1), 167–196. https://doi.org/10.1146/annurev.cellbio.14.1.167

Briasoulis, A., Agarwal, V., & Messerli, F. H. (2012). Alcohol consumption and the risk of hypertension in men and women: A systematic review and meta-analysis. *The Journal of Clinical Hypertension*, *14*(11), 792–798. https://doi.org/10.1111/jch.12008

Cao, Y., Willett, W. C., Rimm, E. B., Stampfer, M. J., & Giovannucci, E. L. (2015). Light to moderate intake of alcohol, drinking patterns, and risk of cancer: Results from two prospective US cohort studies. *BMJ*, h4238. https://doi.org/10.1136/bmj.h4238

Caygill, E. E., & Brand, A. H. (2016). The GAL4 system: A versatile system for the manipulation and analysis of gene expression. In *Methods in Molecular Biology* (pp. 33–52). Springer New York. http://dx.doi.org/10.1007/978-1-4939-6371-3_2

Cederbaum, A. I. (2020). Alcohol metabolism. In *Encyclopedia of Gastroenterology* (pp. 47–55). Elsevier. http://dx.doi.org/10.1016/b978-0-12-801238-3.65618-0

Chan, R. F., Lewellyn, L., DeLoyht, J. M., Sennett, K., Coffman, S., Hewitt, M., Bettinger, J. C., Warrick, J. M., & Grotewiel, M. (2014). Contrasting Influences ofDrosophila white/mini-whiteon Ethanol Sensitivity in Two Different Behavioral Assays. *Alcoholism: Clinical and Experimental Research*, *38*(6), 1582–1593. https://doi.org/10.1111/acer.12421

Chatterjee, A., Aavula, K., & Nongthomba, U. (2019). Beadex, a homologue of the vertebrate LIM domain only protein, is a novel regulator of crystal cell development in Drosophila melanogaster. *Journal of Genetics*, *98*(5). https://doi.org/10.1007/s12041-019-1154-6

Correction for Schumann et al., Genome-wide association and genetic functional studies identify autism susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. (2011). *Proceedings of the National Academy of Sciences*, *108*(22), 9316–9316. https://doi.org/10.1073/pnas.1106917108

Cowmeadow, Roshani. B., Krishnan, Harish. R., Ghezzi, A., Al'Hasan, Y. M., Wang, Yan. Z., & Atkinson, N. S. (2006). Ethanol Tolerance Caused by slowpoke Induction in Drosophila. *Alcoholism: Clinical and Experimental Research*, *30*(5), 745–753. https://doi.org/10.1111/j.1530-0277.2006.00087.x

Crittenden, J. R., Skoulakis, E. M. C., Goldstein, Elliott. S., & Davis, R. L. (2018). *Drosophila mef2 is essential for normal mushroom body and wing development*. Cold Spring Harbor Laboratory. http://dx.doi.org/10.1101/311845

DAVID functional annotation bioinformatics microarray analysis. (n.d.). Retrieved July 30, 2021, from https://david.ncifcrf.gov/

de Velasco, B., Mandal, L., Mkrtchyan, M., & Hartenstein, V. (2005). Subdivision and developmental fate of the head mesoderm in Drosophila melanogaster. *Development Genes and Evolution*, *216*(1), 39–51. https://doi.org/10.1007/s00427-005-0029-4

Devineni, A. V., & Heberlein, U. (2009). Preferential ethanol consumption in drosophila models features of addiction. *Current Biology*, *19*(24), 2126–2132. https://doi.org/10.1016/j.cub.2009.10.070

Dick, D. M., & Foroud, T. (2003). Candidate genes for alcohol dependence: A review of genetic evidence from human studies. *Alcoholism: Clinical & Experimental Research*, *27*(5), 868–879. https://doi.org/10.1097/01.alc.0000065436.24221.63

Dietzl, G., Chen, D., Schnorrer, F., Su, K.-C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S., Couto, A., Marra, V., Keleman, K., & Dickson, B. J. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature*, *448*(7150), 151–156. https://doi.org/10.1038/nature05954

DRSC - DRSC integrative ortholog prediction tool. (n.d.). Retrieved July 30, 2021, from https://www.flyrnai.org/cgi-bin/DRSC_orthologs.pl

Ducci, F., & Goldman, D. (2008a). Genetic approaches to addiction: Genes and alcohol. *Addiction*, *103*(9), 1414–1428. https://doi.org/10.1111/j.1360-0443.2008.02203.x

Ducci, F., & Goldman, D. (2008b). Genetic approaches to addiction: Genes and alcohol. *Addiction*, *103*(9), 1414–1428. https://doi.org/10.1111/j.1360-0443.2008.02203.x

Duffy, J. B. (2002). GAL4 system indrosophila: A fly geneticist's swiss army knife. *Genesis*, *34*(1–2), 1–15. https://doi.org/10.1002/gene.10150

Edenberg, H. J., & Foroud, T. (2013). Genetics and alcoholism. *Nature Reviews Gastroenterology & Hepatology*, *10*(8), 487–494. https://doi.org/10.1038/nrgastro.2013.86

Edenberg, H. J., & McClintick, J. N. (2018). Alcohol dehydrogenases, aldehyde dehydrogenases, and alcohol use disorders: A critical review. *Alcoholism: Clinical and Experimental Research*, *42*(12), 2281–2297. https://doi.org/10.1111/acer.13904

Engel, G. L., Taber, K., Vinton, E., & Crocker, A. J. (2019). Studying alcohol use disorder using Drosophila melanogaster in the era of 'Big Data.' *Behavioral and Brain Functions*, *15*(1). https://doi.org/10.1186/s12993-019-0159-x

Fang, M., Li, J., Blauwkamp, T., Bhambhani, C., Campbell, N., & Cadigan, K. M. (2006). C-terminal-binding protein directly activates and represses Wnt transcriptional targets in Drosophila. *The EMBO Journal*, *25*(12), 2735–2745. https://doi.org/10.1038/sj.emboj.7601153

Gao, B., & Bataller, R. (2011). Alcoholic liver disease: Pathogenesis and new therapeutic targets. *Gastroenterology*, *141*(5), 1572–1585. https://doi.org/10.1053/j.gastro.2011.09.002

Ge, S. X., Son, E. W., & Yao, R. (2018). iDEP: An integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics*, *19*(1). https://doi.org/10.1186/s12859-018-2486-6

Ghezzi, A., Krishnan, H. R., & Atkinson, N. S. (2012). Susceptibility to ethanol withdrawal seizures is produced by BK channel gene expression. *Addiction Biology*, *19*(3), 332–337. https://doi.org/10.1111/j.1369-1600.2012.00465.x

Goldman, D., Oroszi, G., & Ducci, F. (2006). The Genetics of addictions: Uncovering the genes. *FOCUS*, *4*(3), 401–415. https://doi.org/10.1176/foc.4.3.401

Grotewiel, M., & Bettinger, J. C. (2015). DrosophilaandCaenorhabditis elegansas Discovery Platforms for Genes Involved in Human Alcohol Use Disorder. *Alcoholism: Clinical and Experimental Research*, *39*(8), 1292–1311. https://doi.org/10.1111/acer.12785

Heigwer, F., Port, F., & Boutros, M. (2018). RNA interference (rnai) screening in drosophila. *Genetics*, *208*(3), 853–874. https://doi.org/10.1534/genetics.117.300077

Hitzemann, R., Bottomly, D., Darakjian, P., Walter, N., Iancu, O., Searles, R., Wilmot, B., & McWeeney, S. (2012). Genes, behavior and next-generation RNA sequencing. *Genes, Brain and Behavior*, *12*(1), 1–12. https://doi.org/10.1111/gbb.12007

Hoang, C. Q., Burnett, M. E., & Curtiss, J. (2010). Drosophila CtBP regulates proliferation and differentiation of eye precursors and complexes with Eyeless, Dachshund, Dan, and Danr during eye and antennal development. *Developmental Dynamics*, *239*(9), 2367–2385. https://doi.org/10.1002/dvdy.22380

Hu, Y., Flockhart, I., Vinayagam, A., Bergwitz, C., Berger, B., Perrimon, N., & Mohr, S. E. (2011). An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics*, *12*(1), 357. https://doi.org/10.1186/1471-2105-12-357

iDEP.92. (n.d.). Retrieved July 30, 2021, from http://bioinformatics.sdstate.edu/idep/

Joiner, W. J., Friedman, E. B., Hung, H.-T., Koh, K., Sowcik, M., Sehgal, A., & Kelz, M. B. (2013). Genetic and anatomical basis of the barrier separating wakefulness and anesthetic-induced unresponsiveness. *PLoS Genetics*, *9*(9), e1003605. https://doi.org/10.1371/journal.pgen.1003605

Kairamkonda, S., & Nongthomba, U. (2018). Beadex , a Drosophila LIM domain only protein, function in follicle cells is essential for egg development and fertility. *Experimental Cell Research*, *367*(1), 97–103. https://doi.org/10.1016/j.yexcr.2018.03.029

Kaun, K. R., Azanchi, R., Maung, Z., Hirsh, J., & Heberlein, U. (2011). A Drosophila model for alcohol reward. *Nature Neuroscience*, *14*(5), 612–619. https://doi.org/10.1038/nn.2805

Kendler, K. S. (1992). A population-based twin study of alcoholism in women. *JAMA: The Journal of the American Medical Association*, *268*(14), 1877. https://doi.org/10.1001/jama.1992.03490140085040

Kent, C. F., Daskalchuk, T., Cook, L., Sokolowski, M. B., & Greenspan, R. J. (2009). The Drosophila foraging Gene Mediates Adult Plasticity and Gene–Environment Interactions

in Behaviour, Metabolites, and Gene Expression in Response to Food Deprivation. *PLoS Genetics*, *5*(8), e1000609. https://doi.org/10.1371/journal.pgen.1000609

Kim, J.-S., Park, K.-S., Park, O., Kim, S. H., & Jeon, S.-H. (2017). Novel phenotypes of Drosophila spinster locus on the head formation during embryogenesis. *Genes & Genomics*, *39*(2), 237–242. https://doi.org/10.1007/s13258-016-0513-4

Kranzler, H. R., & Soyka, M. (2018). Diagnosis and pharmacotherapy of alcohol use disorder. *JAMA*, *320*(8), 815. https://doi.org/10.1001/jama.2018.11406

Krishnan, H. R., Li, X., Ghezzi, A., & Atkinson, N. S. (2016). A DNA element in the slo gene modulates ethanol tolerance. *Alcohol*, *51*, 37–42. https://doi.org/10.1016/j.alcohol.2015.12.003

Kuntz, S., Poeck, B., Sokolowski, M. B., & Strauss, R. (2012). The visual orientation memory of Drosophila requires Foraging (PKG) upstream of Ignorant (RSK2) in ring neurons of the central complex. *Learning & Memory*, *19*(8), 337–340. https://doi.org/10.1101/lm.026369.112

Lasek, A. W., Giorgetti, F., Berger, K. H., Tayor, S., & Heberlein, U. (2011). Lmo Genes Regulate Behavioral Responses to Ethanol in Drosophila melanogaster and the Mouse. *Alcoholism: Clinical and Experimental Research*, no-no. https://doi.org/10.1111/j.1530-0277.2011.01506.x

Larkin A, Marygold SJ, Antonazzo G, Attrill H, dos Santos G, Garapati PV, Goodman JL, Gramates LS, Millburn G, Strelets VB, Tabone CJ, Thurmond J and the FlyBase Consortium (2021) FlyBase: updates to the Drosophila melanogaster knowledge base. Nucleic Acids Res. 49(D1) D899–D907

Lear, B. C., Darrah, E. J., Aldrich, B. T., Gebre, S., Scott, R. L., Nash, H. A., & Allada, R. (2013). UNC79 and UNC80, putative auxiliary subunits of the NARROW ABDOMEN ion channel, are indispensable for robust circadian locomotor rhythms in drosophila. *PLoS ONE*, *8*(11), e78147. https://doi.org/10.1371/journal.pone.0078147

Lee, K. M., Talikoti, A., Shelton, K., & Grotewiel, M. (2021). Tyramine synthesis, vesicular packaging, and the SNARE complex function coordinately in astrocytes to regulate Drosophila alcohol sedation. *Addiction Biology*, *26*(4). https://doi.org/10.1111/adb.13019

Karlsson Linnér, R., Mallard, T.T., Barr, P.B., Sanchez-Roige, S., Madole, J.W., Driver, M.N., Poore, H.E., de Vlaming, R., Grotzinger, A.D., Tielbeek, J.J., Johnson, E.C., Liu, M., Rosenthal, S.B., Ideker, T., Zhou, H., Kember, R.L., Pasman, J.A., Verweij, K.J.H., Liu, D.J., Vrieze, S., COGA Collaborators, Kranzler, H.R., Gelernter, J., Harris, K.M., Tucker-Drob, E.M., Waldman, I.D., Palmer, A.A., Harden, K.P., Koellinger, P.D., & Dick, D.M. (in press). Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nature Neuroscience*.

Lipner, E. M., & Greenberg, D. A. (2018). The rise and fall and rise of linkage analysis as a technique for finding and characterizing inherited influences on disease expression. In *Methods in Molecular Biology* (pp. 381–397). Springer New York. http://dx.doi.org/10.1007/978-1-4939-7471-9_21

Mao, Y., & Freeman, M. (2009). Fasciclin 2, the Drosophila orthologue of neural cell-adhesion molecule, inhibits EGF receptor signalling. *Development*, *136*(3), 473–481. https://doi.org/10.1242/dev.026054

Mathies, L. D., Lindsay, J. H., Handal, A. P., Blackwell, G. G., Davies, A. G., & Bettinger, J. C. (2020). SWI/SNF complexes act through CBP-1 histone acetyltransferase to regulate acute functional tolerance to alcohol. *BMC Genomics*, *21*(1). https://doi.org/10.1186/s12864-020-07059-y

Melemis, S. (2015). Relapse Prevention and the Five Rules of Recovery. *Yale Journal of Biology and Medicine*. Published.

Milan, M., Diaz-Benjumea, F. J., & Cohen, S. M. (1998). Beadex encodes an LMO protein that regulates Apterous LIM-homeodomain activity in Drosophila wing development: A model

for LMO oncogene function. *Genes & Development*, *12*(18), 2912–2920. https://doi.org/10.1101/gad.12.18.2912

Morozova, T. V., Huang, W., Pray, V. A., Whitham, T., Anholt, R. R. H., & Mackay, T. F. C. (2015). Polymorphisms in early neurodevelopmental genes affect natural variation in alcohol sensitivity in adult drosophila. *BMC Genomics*, *16*(1). https://doi.org/10.1186/s12864-015-2064-5

Neuert, H., Deing, P., Krukkert, K., Naffin, E., Steffes, G., Risse, B., Silies, M., & Klämbt, C. (2019). The Drosophila NCAM homolog Fas2 signals independent of adhesion. *Development*. https://doi.org/10.1242/dev.181479

Nitta, Y., Matsui, S., Kato, Y., Kaga, Y., Sugimoto, K., & Sugie, A. (2019). Analysing the evolutional and functional differentiation of four types of Daphnia magna cryptochrome in Drosophila circadian clock. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-45410-w

Ojelade, S., Jia, T., Rodan, A., Chenyang, T., Kadrmas, J., Cattrell, A., Ruggeri, B., Charoen, P., Lemaitre, H., Banaschewski, T., Buchel, C., Bokde, A., Carvalho, R., Conrod, P., & Flor, H. (2015). Rsu1 regulates ethanol consumption in Drosophila and humans. *Proceedings of the National Academy of Sciences of the United States of America*. Published. https://doi.org/10.1073/pnas.1417222112

Osterwalder, T., Yoon, K. S., White, B. H., & Keshishian, H. (2001). A conditional tissue-specific transgene expression system using inducible GAL4. *Proceedings of the National Academy of Sciences*, *98*(22), 12596–12601. https://doi.org/10.1073/pnas.221303298

Ott, J., Wang, J., & Leal, S. M. (2015). Genetic linkage analysis in the age of whole-genome sequencing. *Nature Reviews Genetics*, *16*(5), 275–284. https://doi.org/10.1038/nrg3908

Picchio, L., Plantie, E., Renaud, Y., Poovthumkadavil, P., & Jagla, K. (2013). Novel Drosophila model of myotonic dystrophy type 1: Phenotypic characterization and genome-wide view

of altered gene expression. *Human Molecular Genetics*, *22*(14), 2795–2810. https://doi.org/10.1093/hmg/ddt127

Pon, J. R., & Marra, M. A. (2015). MEF2 transcription factors: Developmental regulators and emerging cancer genes. *Oncotarget*, *7*(3), 2297–2312. https://doi.org/10.18632/oncotarget.6223

Poortinga, G., Watanabe, M., & Parkhurst, S. M. (1998). Drosophila CtBP: A Hairy-interacting protein required for embryonic segmentation and Hairy-mediated transcriptional repression. *The EMBO Journal*, *17*(7), 2067–2078. https://doi.org/10.1093/emboj/17.7.2067

Potthoff, M. J., & Olson, E. N. (2007). MEF2: A central regulator of diverse developmental programs. *Development*, *134*(23), 4131–4140. https://doi.org/10.1242/dev.008367

Ray, L. A., Hart, E. J., & Chin, P. F. (2011). Self-Rating of the Effects of Alcohol (SRE): Predictive utility and reliability across interview and self-report administrations. *Addictive Behaviors*, *36*(3), 241–243. https://doi.org/10.1016/j.addbeh.2010.10.009

Ren, J., Zhu, H., Chi, C., Mehrmohamadi, M., Deng, K., Wu, X., & Xu, T. (2014). Beadex affects gastric emptying in Drosophila. *Cell Research*, *24*(5), 636–639. https://doi.org/10.1038/cr.2014.24

Robinson, B. G., Khurana, S., & Atkinson, N. S. (2013). Drosophilalarvae as a model to study physiological alcohol dependence. *Communicative & Integrative Biology*, *6*(2), e23501. https://doi.org/10.4161/cib.23501

Sanchez-Roige, S., Fontanillas, P., Elson, S. L., Gray, J. C., de Wit, H., Davis, L. K., MacKillop, J., & Palmer, A. A. (2017). Genome-wide association study of alcohol use disorder identification test (AUDIT) scores in 20 328 research participants of European ancestry. *Addiction Biology*, *24*(1), 121–131. https://doi.org/10.1111/adb.12574

Sanchez-Roige, S., Palmer, A. A., Fontanillas, P., Elson, S. L., Adams, M. J., Howard, D. M., Edenberg, H. J., Davies, G., Crist, R. C., Deary, I. J., McIntosh, A. M., & Clarke, T.-K.

(2019). Genome-Wide association study meta-analysis of the alcohol use disorders identification test (AUDIT) in two population-based cohorts. *American Journal of Psychiatry*, *176*(2), 107–118. https://doi.org/10.1176/appi.ajp.2018.18040369

Sandhu, S., Kollah, A. P., Lewellyn, L., Chan, R. F., & Grotewiel, M. (2015). An inexpensive, scalable behavioral assay for measuring ethanol sedation sensitivity and rapid tolerance in drosophila. *Journal of Visualized Experiments*, *98*. https://doi.org/10.3791/52676

Sass, T. N., MacPherson, R. A., Mackay, T. F. C., & Anholt, R. R. H. (2020). High-Throughput Method for Measuring Alcohol Sedation Time of Individual Drosophila melanogaster. *Journal of Visualized Experiments*, *158*. https://doi.org/10.3791/61108

Schmitt, R. E., Shell, B. C., Lee, K. M., Shelton, K. L., Mathies, L. D., Edwards, A. C., & Grotewiel, M. (2019). Convergent Evidence From Humans and Drosophila melanogaster Implicates the Transcription Factor MEF2B/Mef2 in Alcohol Sensitivity. *Alcoholism: Clinical and Experimental Research*, *43*(9), 1872–1886. https://doi.org/10.1111/acer.14138

Schuckit, M. A. (1994). Low level of response to alcohol as a predictor of future alcoholism. *American Journal of Psychiatry*, *151*(2), 184–189. https://doi.org/10.1176/ajp.151.2.184

Schuckit, M. A., Tipp, J. E., Smith, T. L., Wiesbeck, G. A., & Kalmijn, J. (1997). Self-Rating of the effects of alcohol. *PsycTESTS Dataset*. https://doi.org/10.1037/t04178-000

Section 5 PE tables – results from the 2019 national survey on drug use and health: Detailed tables, SAMHSA, CBHSQ. (n.d.). Retrieved July 29, 2021, from https://www.samhsa.gov/data/sites/default/files/reports/rpt29394/NSDUHDetailedTabs2019/NSDUHDetTabsSect5pe2019.htm?s=5.4&#tab5-4a

Shiraiwa, T., & Carlson, J. R. (2007a). Proboscis extension response (PER) assay in drosophila. *Journal of Visualized Experiments*, *3*. https://doi.org/10.3791/193

Shiraiwa, T., & Carlson, J. R. (2007b). Proboscis extension response (PER) assay in drosophila. *Journal of Visualized Experiments*, *3*. https://doi.org/10.3791/193

Sivachenko, A., Li, Y., Abruzzi, K. C., & Rosbash, M. (2013). The transcription factor mef2 links the drosophila core clock to fas2, neuronal morphology, and circadian behavior. *Neuron*, *79*(2), 281–292. https://doi.org/10.1016/j.neuron.2013.05.015

Southall, T. D., Elliott, D. A., & Brand, A. H. (2008). The GAL4 System: A versatile toolkit for gene expression in drosophila. *Cold Spring Harbor Protocols*, *2008*(8), pdb.top49-pdb.top49. https://doi.org/10.1101/pdb.top49

Speca, D. J., Chihara, D., Ashique, A. M., Bowers, M. S., Pierce-Shimomura, J. T., Lee, J., Rabbee, N., Speed, T. P., Gularte, R. J., Chitwood, J., Medrano, J. F., Liao, M., Sonner, J. M., Eger, E. I., Peterson, A. S., & McIntire, S. L. (2010). Conserved Role of unc-79 in Ethanol Responses in Lightweight Mutant Mice. *PLoS Genetics*, *6*(8), e1001057. https://doi.org/10.1371/journal.pgen.1001057

Sweeney, S. T., & Davis, G. W. (2002). Unrestricted Synaptic Growth in spinster—a Late Endosomal Protein Implicated in TGF-β-Mediated Synaptic Growth Regulation. *Neuron*, *36*(3), 403–416. https://doi.org/10.1016/s0896-6273(02)01014-0

Taylor, M. V., & Hughes, S. M. (2017). Mef2 and the skeletal muscle differentiation program. *Seminars in Cell & Developmental Biology*, *72*, 33–44. https://doi.org/10.1016/j.semcdb.2017.11.020

Understanding Alcohol Use Disorder. (n.d.). Retrieved July 29, 2021, from https://www.niaaa.nih.gov/publications/brochures-and-fact-sheets/understanding-alcohol-use-disorder

VDRC Stock Center: Main page. (n.d.). Retrieved July 30, 2021, from http://stockcenter.vdrc.at/control/main

Verhulst, B., Neale, M. C., & Kendler, K. S. (2014). The heritability of alcohol use disorders: A meta-analysis of twin and adoption studies. *Psychological Medicine*, *45*(5), 1061–1072. https://doi.org/10.1017/s0033291714002165

Wang, Z., Pan, Y., Li, W., Jiang, H., Chatzimanolis, L., Chang, J., Gong, Z., & Liu, L. (2008). Visual pattern memory requires foraging function in the central complex of Drosophila. *Learning & Memory*, *15*(3), 133–142. https://doi.org/10.1101/lm.873008

Weston, R. M., Schmitt, R. E., Grotewiel, M., & Miles, M. F. (2021). *Transcriptome analysis of chloride intracellular channel knockdown in drosophila identifies oxidation-reduction function as possible mechanism of altered sensitivity to ethanol sedation*. Cold Spring Harbor Laboratory. http://dx.doi.org/10.1101/2021.01.20.427413

WHO. (2018). Global status report on alcohol and health. *Geneva: World Health Organization*.

Yamaguchi, M., & Yoshida, H. (2018). Drosophila as a model organism. In *Advances in Experimental Medicine and Biology* (pp. 1–10). Springer Singapore. http://dx.doi.org/10.1007/978-981-13-0529-0_1

Yuva-Aydemir, Y., Bauke, A.-C., & Klambt, C. (2011). Spinster controls dpp signaling during glial migration in the drosophila eye. *Journal of Neuroscience*, *31*(19), 7005–7015. https://doi.org/10.1523/jneurosci.0459-11.2011

Zeng, C., Justice, N. J., Abdelilah, S., Chan, Y.-M., Jan, L. Y., & Jan, Y. N. (1998). The Drosophila LIM-only gene, dLMO, is mutated in Beadex alleles and might represent an evolutionarily conserved function in appendage development. *Proceedings of the National Academy of Sciences*, *95*(18), 10637–10642. https://doi.org/10.1073/pnas.95.18.10637

APPENDICES

Basic Fly Handling and Husbandry

A. Standard Fly Lab Lingo:

1. Stock or strain: a culture of flies with a particular genotype. Balanced stocks have a special chromosome called a balancer that is marked with a dominant phenotype and suppresses recombination on the corresponding sister chromosome. Balanced stocks are often weak (i.e. grow poorly).

2. Seeding: putting adult flies into a new bottle or vial. Also called 'setting-up'.

3. Transfer: moving flies without anesthesia from one vial or bottle to another. One-to-one transfer means moving flies from one bottle/vial to one new bottle/vial. Two-to-one transfer means moving flies from 2 vials/bottles to 1 new vial/bottle. Also called 'flipping'.

4. Clearing: removing all of the adults from a bottle or vial. Can be done with or without anesthesia.

5. Anesthesia: CO_2 used to temporarily immobilize flies.

6. Brood: refers to the number of times a set of adults has been used to seed bottles. Using flies for 2 broods is common, with 3 broods being possible in some cases.

7. white plus (w^+): indicates eye color. white minus (w^-) flies have white eyes. w^+ flies have eyes that can vary from light peach to deep red.

8. Food: All of our fly food currently has antibiotics on it (ampicillin, tetracycline and chloramphenicol; i.e. ATC). Yeasted (Y) food vials and bottles have live yeast on added. Yeasted food should be used for seeding new vials and bottles for *growing* flies. Non-yeasted (NY) food has no yeast on it and should be used to *house* flies prior to behavioral studies and for *storing* virgin females and males prior to setting-up crosses.

B. Standard Fly Husbandry

1. Remove necessary number of yeasted bottles or vials from the cold room. Use bottles to grow lots of flies for behavioral or other large experiments. Use vials for smaller numbers of flies in limited scale crosses or other small scale experiments.

2. Before putting in new flies, bottles and vials must be dried 2 hours to overnight in the environmental chamber so that all condensation on the walls evaporates. The food will pull away from the wall of the bottle or vial if they are over-dried. It is poor practice to use over-dried food.

3. Turn on the CO_2. Clean microscope, CO_2 pad and counter with ethanol. Clean before starting, between each genotype and after you are finished. Be sure the CO_2 is on before putting ethanol on the pad.

4. Open CO_2 to pipette, invert bottle or vial, insert pipette along cotton plug and tap bottle/vial gently. Flies will become anesthetized quickly and should fall onto the plug and/or the neck of the bottle/vial.

5. Click off CO_2 to pipette, remove CO_2 pipette from vial/bottle. Hold inverted bottle/vial over CO_2 pad. Remove plug and gently shake/tap flies onto pad into a pile. Return plug to bottle/vial and set aside.

6. Use brush or spatula to place anesthetized flies in a row and sort flies according to needs. Short CO_2 times are important. For collecting flies that will be used in behavioral studies, goals are (1) all genotypes experience the same CO_2 exposure and (2) all flies are anesthetized for less than 5 minutes.

7. Set-up new bottles/vials by putting sorted flies from step 6 into dried bottles/vials. Anesthetized flies should be kept on the wall of the bottle/vial. If they fall into the food, many of them will stick there and die. Robust strains such as w[A] will do well with 10 females (♀, see below) per bottle or 3 females per vial. It is good practice to include a comparable number of males (♂, see below). Weaker stocks will need more females, up to as many as 50 per bottle and 15 per vial. When working with a stock that is new to you it is good practice to seed bottles or vials with a range of females (10-25/bottle for example) and then use an optimum number thereafter based on how the various bottles/vials grow.

8. Insert cotton plug, invert new bottle/vial and tap anesthetized flies onto the plug. Lay the bottle/vial on its side, label with genotype and date. First broods (i.e. bottles or vials in which the flies are new parents) are marked with a single slash.

9. Wait for flies to regain locomotor activity. Turn bottles/vials upright and place in environmental chamber to grow.

10. Beginning at around 4 days after seeding, check bottles/vials daily for larval activity (darkish band on top of food). When larval activity is obvious, either discard the adults or—if a second brood is needed—transfer adults to new bottles/vials (dried appropriately). Label second brood with genotype, date and two slashes.

11. Beginning at around 4 days after seeding the second brood, check bottles/vials daily for larval activity. Discard adults when larval activity is obvious. If necessary, a third brood is possible in some cases.

12. You should expect to see obvious larval activity 4 to 7 days after seeding and obvious pupae 5-10 days after seeding. New adults should begin emerging ~10 days after seeding.

Some strains, especially balanced strains, can take up to 4 additional days to generate adults. Perfectly seeded bottle/vials will have robust larval activity followed by large numbers of pupae that populate the bottom three-fourths of the wall of the vial or bottle. Pupae will not typically be in the food or on the plug in these bottles. Large numbers of healthy adults suitable for experiments will emerge from perfectly seeded bottles/vials.

13. Common Problems: If your bottles/vials are too dry or wet (as described below), the resulting adults should not be used for behavioral, stress or gene expression studies. The resulting adults are fine genotype-wise and reproduction-wise, though, and can be used to set-up new bottles/vials as necessary.

a. Food too dry after 4-7 days of new adults in bottle/vial: The food should not be so dry that it detaches from the wall of the bottle of vial and the pupae are in the food. In cases like this, the food was either over-dried, there were not enough females placed in the bottle/vial, or possibly both. If this occurs across several strains that have grown well in the past, it is likely due to over-drying. If it occurs with a subset of strains, it is more likely due to insufficient numbers of females being used for those specific strains. The appropriate fixes are to decrease drying time, add more females next time, or both.

When you transfer flies from the first to second brood or when clearing the second brood, note the quality of the culture and food. If the food in some bottle/vials is detached from the wall after 7 days, go ahead and transfer/clear the adults and then add ddH_2O (NOT ETHANOL!) to the bottle/vial until the gap between the food and the wall is filled. In many cases this will help the larvae quite a lot and you still might get a decent yield of adults, although they might be delayed a few days due to lack of water.

b. Food too wet after 4-7 days of new adults in bottle/vial: The food should not be so wet that it runs down the wall of the bottle/vial when it is inverted and the pupae are on the plug. If this happens, the food was not dried sufficiently before adults were added, too many adults were added, or possibly both. If this occurs across several strains that have not had this problem in the past, it is likely due to under-drying the food. If it occurs with only a subset of strains, it is more likely due to too many females being added in those specific strains. The fixes are to increase the drying time for bottles/vials, decrease the number of females used, or both.

If you notice that your bottles are too wet when transferring from the first to second brood or when clearing the second brood, you can put a folded Kim wipe in the bottle/vial so that it touches both the food and the plug. This will not result in a miraculous drying of the bottle/vial, but it can convert a bottle/vial that is far too wet into one that can be managed with some care.

C. The Basics of Setting-Up Crosses

1. You will need males (♂, mated or unmated) and virgin females (♀ with a ‘v’ on top) for your crosses. Grow bottles or vials as above for strains required to generate males and virgin females. For planning purposes, you can comfortably collect 50-100 males and/or 25-50 virgin females from a robust bottle. Likewise, you can probably count on collecting 15-20 males and 5-10 virgin females from each well-seeded vial.

2. Around day 10 after seeding, begin to collect virgin females, identified by their light body pigmentation and female genitalia (see below). Typically, one collects virgin females first thing in the morning, again around noon, and again last thing before leaving for the day.

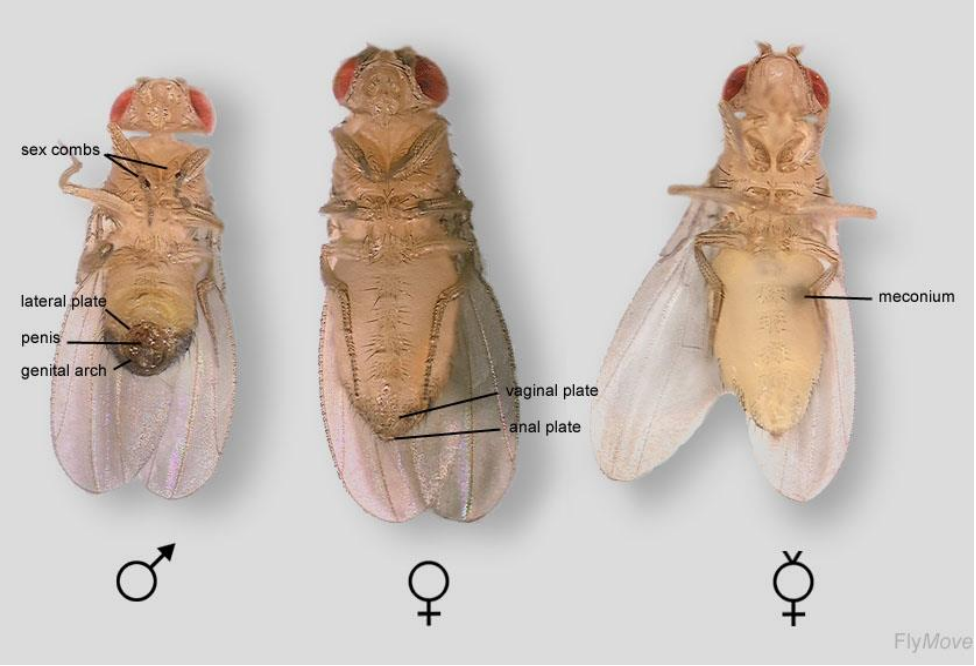
3. Keep virgin females in non-yeasted vials with no more than 25 females/vial. Label each vial with genotype, date and number of virgins collected. Keep collected virgins in environmental chamber until ready to use. One will often collect virgin females over several days until a sufficient number of virgin females has been collected. Also, it is convenient to store virgin females in upside-down vials.

4. When sufficient numbers of virgin females have been collected (~10% more than you plan to use) or when it is obvious that you will be able to collect all the virgin females you will need, collect all males into non-yeasted vials needed for your crosses. Males are identified by their male genitalia (see below).

5. Set-out yeasted bottles or vials to warm and dry as described above. On the day of the cross, check all virgin female vials for larvae using the microscope. Any vials with larvae MUST be discarded because at least one of the females has mated. Use only virgin females from vials with no larvae.

6. To set-up a cross, anesthetize the males and check them, anesthetize the virgin females on the same plate and check them, and put appropriate numbers of males and females into yeasted bottles/vials as described in steps B7-B9 above. Handle them thereafter as described in B10-B12 above.

7. Make sure that you know what progeny to expect from your crosses before you set them up.



Ethanol Sedation Assay

A. Day before assay

1. Collect flies (reared for behavioral assays) in groups of 11 (single sex) under brief CO_2 (~5 minutes) following standard procedures for behavioral assays. Collect only those flies that look healthy, are relatively the same size, have normal wings, and appear dry. Flies should be transferred from the CO_2 plate into an Eppendorf tube using a funnel and then dumped from the Eppendorf tube into a non-yeasted vial.

2. Allow flies to recover overnight in upside-down non-yeasted food vials in the environmental chamber. It is possible to test a maximum of 24 vials of flies in a single experiment.

3. Dilute ethanol solution as necessary (85% is our standard concentration). ~250 ml of ethanol solution can be stored in a sealed 500ml bottle or other sealed container for a week without a problem. Make ethanol fresh weekly. Diluted ethanol is exothermic and should be stored overnight at room temperature before use.

B. Day of assay

1. For each vial of flies to be tested, you will need (a) a clean, empty food vial; i.e. testing vial, (b) a new Flug, (c) a silicone #4 plug and (d) 1.0 ml of ethanol solution (85% ethanol is our standard concentration).

2. Turn on humidifier and allow relative humidity in testing room to rise to 55-65%. Temperature should be 20-23°C. Record humidity and temperature on test log.

3. Have someone else in the lab assign a unique code to each group of vials for each genotype and—IMPORTANTLY—record the code for later. Place coded vials with flies in testing room to acclimate.

4. Label empty testing vials to match codes on fly vials from B.3.

5. Construct a testing log by entering the code for each vial into the Test Log E or Test Log EE sheet within the Excel Sedation file SA E EE 6 min SIGMOIDAL 2015.10.05. Use a random or cycling order. Add other pertinent information (% ethanol, sex, etc.) to the Test Log worksheet and print for use during testing.

6. Using the Test Log as a guide, arrange coded food vials with flies and empty testing vials into matching arrays with 4 vials in each row. The maximum possible number of vials that can be tested in a single experiment is 24 vials (i.e. 6 rows of 4 vials each).

7. Transfer flies from food vials into matched/labeled testing vials one at a time and immediately insert Flugs into testing vials until Flugs are a uniform distance below the vial tops. Use the Fluginator to push Flugs down into vials.

8. Time 0 assessment: Grasp each vial individually with thumb and forefinger, tap gently on the table three times to knock flies to the bottom of the vial, wait 30 seconds and then count the number of flies that are immobile. Typically, this is 0 or 1 at time 0. Record the number of immobile flies for each vial at time 0 in the printed Testing Log.

9. Hereafter, each row of four vials will be handled as a set at staggered one-minute intervals.

Start timer counting up at time 0 and immediately begin adding 1 ml of ethanol to the Flug in the vials for the first row/set of 4 vials. Add ethanol to the vials at 5 second intervals in the order they will be tested. Add ethanol to the Flugs in a circular motion so that all ethanol is absorbed as uniformly as possible. When ethanol has been added to all four testing vials in the set, insert a silicone #4 plug in each vial to seal it.

At times 1, 2, 3, 4 and 5 minutes on the timer, add 1 ml of ethanol to the second, third, fourth and fifth sets of 4 vials, respectively. Continue inserting #4 plugs after adding ethanol to each set of 4 vials.

10. At time 6 minutes, test the first set of 4 vials by grasping the first vial with thumb and forefinger and then tapping gently on the table three times to knock flies to the bottom of the vial. Tap the other 3 vials in the set the same way at 5 second intervals. 30 seconds after tapping the first vial, count and record the total number of flies that are sedated. Count and record the number of sedated flies in the other 3 vials at 5 second intervals. Flies are scored as sedated if they do not appear to have productive locomotion.

The specific schedule is:

Vial	Tap	Assess
1	6 min 0 s	6 min 30 s
2	6 min 5 s	6 min 35 s
3	6 min 10 s	6 min 40 s
4	6 min 15 s	6 min 45 s

At times 7, 8, 9, 10 and 11 minutes, test the second, third, fourth, fifth and sixth sets of vials, respectively, as done for the first set.

11. At time 12 minutes, test the first set of 4 vials again as described in B10 and continue testing the second, third, fourth, fifth and sixth sets of vials at 13, 14, 15, 16 and 17 minutes, respectively.

Continue testing flies as described in B10 and B11 until all flies are sedated (typically 60-90 min).

12. Record the total number of flies in each vial.

13. Clean-up is (a) turn off humidifier, (b) remove #4 plugs for washing and reuse, (c) discard Flugs/vials/flies, (d) remove any trash from and straighten up testing room and (e) turn off light in testing room.

14. Enter the total number of flies in each vial and the number of flies sedated at each time point in the Test Log within the Excel worksheet. Percent Active flies will be automatically calculated and graphed below the Test Log. Press ‘Ctrl + s’ to calculate ST50s for each vial and sort the data by group in the Sorted Data worksheet.

15. Note any flagged data in Sorted Data worksheet. Consider excluding data that looks qualitatively poor.

M Grotewiel, R Schmitt, K Lee: 7/2014, 3/2015, 7/2016

List of lethal RNAi's tested in Chapter 2

Gene	RNAi
unc79	HMC03213
Bx	KK108513
Bx	GL00484
CtBP	KK108401
CtBP	HMS00677
Fas2	KK100888
For	HMS004486

RNA Prep

Part A: Fly collection

Whole body fly collection

1. Collect 25 flies of desired age, genotype and gender in a 1.5mL snap cap tube. 1 tube = 1 n. Place tubes on ice immediately after flies enter tube.

2. Once done collecting, place flies in -80. Once flies are frozen and dead, you can proceed to Part B

Head Preps

1. Collect whole body flies of desired age, genotype and gender in a conical tube. 1 tube = 1 n. For head preps, about 250 flies should be in each tube (absolute minimum = 150 flies, but this is not recommended). Store flies on ice at all times. After each collection, place flies immediately back in the -80 freezer.

2. Once collected, bring the following equipment to the cold room
   - large plastic box
   - styrofoam box with sieve and tube holder in it
   - large metal forceps
   - a vortex
   - funnel(s)
   - large orange-capped conical tubes → ONLY USE CAPS WITH HOLES
   - labeled 1.7 snap-cap tubes for the number of preps you are doing
   - Cryogloves → WEAR AT ALL TIMES

3. Obtain liquid N_2 (in dewar) and dry ice (in styrofoam container) from 6th floor supply center. 3/4th full liquid N_2 and 1/2 full dry ice is sufficient for ~12 preps.

4. Store samples on dry ice before and after prep.

5. Fill conical tube 3/4 full (in styrofoam container) with liquid N_2.

6. Add flies to tube, screw on cap **WITH HOLES** (or it will violently explode), and vortex (stopping as little as possible) until the liquid N_2 is almost gone (~1 min).

7. Repeat—filling tube with flies 1/2 full with N_2 and vortexing again.

8. Pour N_2 into sieve so it is sufficiently cold (otherwise the flies will stick and you'll get nothing). Dump flies into sieve and beat laterally with heavy forceps for several minutes.

9. Using funnel, quickly collect heads (middle layer) or bodies (top layer) using funnel and labeled 1.7 snap-cap tube.

9. Between genotypes: take a break to clean and completely dry sieve (or it will freeze together and form an ice layer over the holes). Get a new funnel and conical tube.

10. Store at -80°C until use.

**Wear safety glasses, lab coat, 2 pairs of latex gloves and cryo gloves (plus warm clothes and long pants).

Part B: RNA extraction

** All water used is DEPC water

** Samples must be kept on ice at all times, unless otherwise stated

1. Wipe down bench and all pipettes, pipette boxes, ect with 100% ETOH. Place clean plastic pestles in 50mL conical tube, cover pestles with chloroform, and let them soak for 20 minutes. Transfer pestles to new clean empty 50mL conical tube and allow to air dry for 20 minutes.

** all chloroform is stored under the hood, and any procedures involving chloroform should always be done under the hood

2. While under the fume hood, add 250µL Trizol (pink, stored in fridge) to each tube of flies. Homogenize for 1 minute with drill and pestle

3. Add 100µL of chloroform to each tube. Vortex for 15 seconds. Incubate for 3 minutes at room temperature

4. Centrifuge samples at maximum speed (14,000 x g) for 15 minutes in the cold room

5. Label new 1.5mL tubes appropriately. Remove roughly 200µL of the upper aqueous phase and place in new tube. If you accidently pipette any fly parts or other liquid, centrifuge that sample again (i.e. repeat step 4) and then attempt this step. Discard tubes with fly parts and the pink liquid.

6. Add 250µL isopropanol (labeled ISO in RNA reagents station) to each tube containing the upper aqueous sample. Invert the tube 10 times. Incubate samples for 10 minutes at room temperature. After, centrifuge samples at maximum speed (14,000 x g) for 10 minutes in cold room

7. Continue with Qiagen RNeasy MiniKit protocol

Fisher's Test R Script from Michael Miles

The contingency table cells are as follows: n_A_B = overlapping genes between A and B
n_A = number of genes in group A
n_B = number of genes in group B
n_C = number of genes in genome (mouse)

```
matrix=matrix(c(n_A_B,n_A,n_B,n_C-(n_B+n_A+n_A_B)),2,2)
matrix
sum(matrix)
fisher.test(matrix)
```

342 Mef2 bound genes and 581 human orthologs described in Chapter 2

Mef2 Bound gene	Human Ortholog	DIOPT Score
jar	MYO6	14
tlk	TLK2	11
	TLK1	8
CG11405	ATF3	5
	JDP2	4
	FOS	3
	FOSL1	3
	FOSL2	3
	ATF7	2
	ATF2	2
	CREB5	2
l(1)G0007	DHX38	13
Gga	GGA1	13
	GGA2	12
	GGA3	13
HDAC4	HDAC4	11
	HDAC5	10
	HDAC9	9
	HDAC7	8
Ptp61F	PTPN2	11
	PTPN1	11
CG14186	none	
tara	none	
CG34299	none	
l(2)k16918	none	
CG18317	SLC25A36	14
	SLC25A33	10
bun	TSC22D1	10
	TSC22D2	7
	TSC22D4	5
	TSC22D3	5

CG11873	none	
att-ORFA	SLC25A42	14
	SLC25A16	5
for	PRKG1	13
	PRKG2	5
Ntl	SLC6A7	5
CG4577	none	
CG31216	none	
fray	STK39	14
	OXSR1	13
bnl	FGF20	7
	FGF16	6
	FGF9	5
	FGF10	5
	FGF22	5
	FGF13	5
	FGF4	5
	FGF6	5
	FGF5	5
	FGF11	5
raw	none	
CG17034	ATP8A1	14
	ATP8A2	12
CG32521	none	
stv	BAG3	9
	BAG4	6
CG8036	TKTL2	14
	TKT	14
	TKTL1	
cbt	KLF11	6
CG5385	none	
c11.1	MROH1	13
	MROH2B	9

	MROH2A	8
Gpdh	GPD1	14
	GPD1L	12
dpp	BMP2	9
	BMP4	8
d	none	
Hr38	NR4A2	13
	NR4A3	11
	NR4A1	10
Pdp1	HLF	11
	DBP	8
	TEF	8
jdp	DNAJC12	11
cwo	HES1	2
	HES2	2
	HES4	2
Kr-h1	none	
mrj	DNAJB2	11
	DNAJB6	11
	DNAJB7	10
	DNAJB8	9
Atf-2	NPDC1	7
	ATF7	6
	CREB5	6
	ATF2	6
crp	TFAP4	11
CG18472	SPAG1	9
	TOMM34	6
mt:ATPase8	none	
Rim	RIMS1	10
	RIMS2	10
	RIMS3	7
	RIMS4	6

sr	EGR2	6
	EGR3	5
	EGR1	5
hts	ADD1	14
	ADD2	11
	ADD3	11
CG14322	none	
dnc	PDE4B	12
	PDE4D	10
	PDE4C	9
	PDE4A	8
CG7337	WDR62	10
	MAPKBP1	10
spen	SPEN	11
CG9775	none	
Cyp6w1	CYP3A4	7
	CYP3A43	7
CG4662	MICU3	14
	MICU2	5
Eip75B	none	
rdx	SPOP	10
	SPOPL	9
l(2)k01209	UCKL1	13
cnc	NFE2L2	8
	NFE2L1	8
	NFE2	7
	NFE2L3	5
Mnt	MNT	6
CG14497	none	
Nfl	NFIA12	
	NFIC	11
	NFIB	11
	NFIX	10

mt:Cyt-b	CYTB	10
CG9005	FAM214A	8
	FAM214B	5
miple	none	
CG32647	DGCR2	6
CG4966	HPS4	9
CG14509	none	
CG10365	CHAC1	12
	CHAC2	5
vri	NFIL3	7
CG9813	none	
PRL-1	PTP4A2	12
	PTP4A1	11
	PTP4A3	8
milt	TRAK1	13
	TRAK2	11
sif	TIAM1	11
	TIAM2	8
CG15926	none	
CG31183	NPR2	13
	NPR1	13
IA-2	PTPRN	11
	PTPRN2	10
EDTP	MTMR14	12
ph-p	PHC3	9
	PHC2	6
	PHC1	5
CG1869	CHIT1	11
	CHIA	10
	CHI3L1	8
	CHI3L2	
CG34360	ZNF704	10
	ZNF395	9

	SLC2A4RG	6
Hex-A	GCK	13
	HK1	10
	HKDC1	9
	HK3	9
	HK2	9
nmo	NLK	14
cpo	RBPMS	9
	RBPMS2	8
NK7.1	none	
CdGAPr	ARHGAP32	9
	ARHGAP32	9
	ARHGAP31	5
Oscillin	GNPDA2	14
	GNPDA1	13
kirre	KIRREL3	11
	KIRREL1	11
	KIRREL2	10
CG30389	MACO1	13
mib	none	
CG14515	none	
CG9339	TBC1D24	14
CG10737	C2CD2	6
lola	none	
Fit1	FERMT2	14
	FERMT1	13
	FERMT3	10
CG7272	SLC50A1	5
CG4599	DNAJC7	13
CG32626	AMPD2	13
	AMPD3	5
	AMPD1	5
CG14207	HSPB1	5

	CRYAA	5
fru	none	
CG11347	none	
CG17836	none	
14-3-3zeta	YWHAZ	13
	YWHAB	11
	YWHAH	6
	YWHAG	6
	YWHAQ	5
	SFN	5
l(3)82Fd	OXR1	12
	NCAO7	12
	TLDC2	5
nrv1	ATP1B1	14
	ATP1B2	12
	ATP1B3	10
	ATP1B4	9
	ATP4B	7
Dad	SMAD7	10
	SMAD6	9
Mi-2	CHD5	12
	CHD3	12
	CHD4	11
CG32432	none	
CG9413	SLC7A9	14
	SLC7A13	5
	SLC7A6	5
	SLC7A5	5
	SLC7A8	5
	SLC7A7	4
capt	CAP1	14
	CAP2	12
CG31183	NPR2	13

	NPR1	13
CG10947	CAMKMT	12
CG32486	CYHR1	15
Argk	CKM	11
	CKMT2	10
	CKB	10
	CKMT1A	9
	CKMT1B	5
exba	BZW1	14
	BZW2	12
CG33523	MOSPD2	14
CG3161	HES1	7
	ATP6V0C	13
	HES4	7
	HES2	5
EcR	NR1H3	10
	NR1H2	9
CG4068	SINHCAF	9
CG34417	SMTNL1	5
CG8321	ARL6IP6	5
CG9990	none	
W	ABCG2	8
	ABCG1	6
	ABCG4	5
ttv	EXT1	13
	EXTL1	6
Sdc	SDC4	6
	SDC1	6
	SDC2	6
	SDC3	5
kst	SPTBN5	13
	SPTBN2	3
	SPTBN1	3

	SPTBN4	3
chic	PFN4	9
stv	BAG4	6
CG9005	FAM214A	5
SelR	MSRB3	14
	MSRB2	7
CG31915	COLGALT2	13
	COLGALT1	12
	CERCAM	11
bocksbeutel	none	
Cbp53E	CALB2	11
	CALB1	10
	SCGN	9
Sema-1a	SEMA6A	8
	SEMA6B	6
	SEMA6C	6
	SEMA3D	5
	SEMA3A	5
CG30421	USP31	11
	USP43	10
CG4766	MAB21L2	14
	MAB21L1	12
CG32813	none	
CG7458	SLC22A1	5
CG7458	SLC22A13	5
	SLC22A3	5
	SLC22A12	5
	SLC22A11	5
	SCL22A14	5
	SLC22AG	5
	SLC22A4	5
	SCL22A5	5
	SLC22A13	5

	SLC22A3	5
	SLC22A12	5
	SLC22A11	5
	SLC22A14	5
CG7990	PGAP2	5
Klp98A	KIF16B	11
CG3600	NR6A1	8
wun	PLPP1	14
	PLPP3	12
	PLPP2	9
CG18812	GDAP2	14
CG10311	none	
DopEcR	none	
cdc14	CDC14B	10
	CDC14A	9
	CDC14	5
ari-2	ARIH2	15
CG8745	ETNPPL	14
	PHYKPL	11
drongo	AGFG1	10
	AGFG2	5
Slob	none	
wdb	PPP2R5E	13
	PPP2R5A	12
	PPP2R5B	9
	PPP2R5C	5
kek4	none	
Sin3A	SIN3A	13
	SIN3B	13
sty	SPRY2	11
	SPRY4	9
	SPRY3	9
	SPRY1	9

Pdk	PDK3	14
	PDK2	13
	PDK1	13
	PDK4	13
Top1	TOP1	
	TOP1MT	
CG31140	DGKQ	14
drp	none	
Or45b/?		
CG11033	KDM2A	13
	KDM2B	11
CG12433	none	
sdk	SDK2	11
	SDK1	10
RhoGAP19D	ARHGAP23	8
	ARHGAP21	7
tai	NCOA3	5
	NCOA1	5
	NCOA2	5
CG12355	MPV17L	
5-HT7	HTR7	6
	HTR1B	4
CG14234	TMEM198	13
comm3	none	
cn	KMO	14
Mef2	MEF2C	11
	MEF2A	11
	MEF2D	10
	MEF2B	5
CG33724	none	
Blimp-1	PRDM1	12
	ZNF683	5
CG10419	GEMIN2	13

cdi	TESK2	8
	TESK1	6
Rtnl1	RTN4	9
	RTN3	9
	RTN1	9
	RTN2	7
CrebB-17A	CREB1	13
	CREM	11
	ATF1	10
kay	FOSL2	6
Bx	LMO1	8
	LMO3	7
CG3703	RUNDC1	14
CG6398	none	
stnA	none	
CG18135	GPCPD1	8
Cyp6v1	CYP3A43	5
	CYP3A7-CYP3A51P	5
	TBXAS1	4
foi	SLC39A10	
	SLC39A6	
	SLC39A4	
Pkn	PKN2	13
	PKN1	12
	PKN3	10
CG7378	DUPD1	8
	DUSP13	8
	DUSP26	7
	DUSP3	5
	DUSP10	5
Hr38	NR4A2	13
	NR4A3	11
	NR4A1	10

pum	PUM2	12
	PUM1	11
Idh	IDH1	14
	IDH2	5
DopEcR	none	
Fas2	NCAM1	11
	NCAM2	11
Hn	PAH	14
CG30190	REEP2	9
	REEP4	8
	REEP1	8
	REEP3	7
NFAT	NFAT5	7
	NFATC3	6
	NFATC2	5
Pino	none	
CG30377	none	
CtBP	CTBP2	14
	CTBP1	12
CG2162	R3HCC1L	10
	R3HCC1L	6
l(2)08717	SLC17A2	5
	SLC17A3	5
	SLC17A4	5
	SLC17A5	5
	SLC17A7	5
	SLC17A8	5
CG3249	AKAP1	11
CG6934	FRMPD4	9
	FRMPD3	5
	FRMPD1	5
CG34045	none	
CG14619	USP2	9

	SUP21	6
	MPRIP	
osp	MPRIP	9
	TRIOBP	6
Lasp	LASP1	9
CG4238	AREL1	13
CG6191	CABLES1	10
	CABLES2	8
ph-d	PHC3	7
	PHC2	6
	PHC1	4
Ptr	PTCHD3	11
	PTCHD1	7
	PTCHD4	7
Bsg	NPTN	8
	BSG	7
CG32369	LONRF3	12
	LONRF2	12
	LONRF1	11
Thd1	TDG	10
Fmr1	FXR1	12
	FXR2	11
	FMR1	11
CG15630	NCAM2	4
	NCAM1	4
bin3	MEPCE	10
spir	SPIRE1	13
	SPIRE2	11
cg	none	
tamo	SPATA2	7
	SPATA2L	5
CG16944	SLC25A4	12
	SLC25A5	10

	SLC25A6	9
	SLC25A31	8
CG9007	KMT2E	8
	SETD5	8
CG8032	PAOX	13
	SMOX	10
CG5122	CRAT	7
bip1	none	
h	none	
CHES-1-like	FOXn3	6
Smr	NCOR	10
	NCSO2	9
CG1600	none	
shakB	none	
CG1888	none	
sr	none	
NK7.1	none	
CG8490	none	
CG6770	NUPR1	8
	NUPR5	5
Sema-1a	none	
CG11489	SRPK3	9
	SRPK1	8
	SRPK2	8
jim	none	
CG14764	none	
tau	MAPT	7
	MAP4	6
	MPA2	5
Trp1	SEC6S	13
elB	ZNF703	9
	ANF503	8
CG30492	ZFAND6	12

	ZFAND5	11
CG5237	UNC79	14
Act5C	ACTB	11
	ACTG1	10
	POTEJ	5
	POTEI	5
	POTEF	5
Mbs	PPP1R12B	13
	PPP1R12A	10
	PPP1R12C	5
CG12214	TBCEL	14
CG31150	none	
Nmdmc	MTHFD2	14
	MTHFD2	12
CG13330	none	
sick	NAV2	9
	NAV3	7
CG1090	none	
th	DBH	15
	TH	15
	TXNRD2	10
	XIAP	7
	BIRC3	6
	BIRC2	6
	TXNRD1	5
glob1	CYGB	9
glob1	HBG2	5
	HBE1	5
	HBG1	5
d	none	
oa2	none	
Eip75B	none	
CG5830	CTDSPI1	15

	CTDSPL	12
	CTDSP2	12
spin	SPNS1	13
	SPNS3	9
	SPNS2	9
CG10543	none	
Lis-1	PAFAH1B1	14
ctp	DYNLL2	13
	DYNLL1	9
CG8489	none	
Sp7	none	
CG9335	none	
stck	LIMS1	14
	LIMS2	13
	LIMS4	5
retn	ARID3A	11
	ARID3C	9
	ARID3B	8
aret	CELF2	14
	CELF1	14
CG5807	LMBR1L	14
	LMBR1	13
CG7982	AGAP1	13
	AGAP3	11
	AGAP2	9
	AGAP4	7
	AGAP5	6
	AGAP6	6
CG7378	none	
Atg4	ATG4A	14
	ATG4D	13
	ATG4B	13
	ATG4C	12

dpp	none	
CG34113	none	
CG5346	none	
CG32138	FMNL3	13
	FMNL1	13
	FMNL2	11
emc	ID4	10
	ID1	10
	ID2	9
	ID3	8
CG10737	none	
Inos	ISYNA1	13
Alh	MLLT6	7
	MLLT10	6
CG11791	none	
TepIII	CD109	11
	PZP	6
	A2M	6
	A2ML1	5
	C5	5
dl	REL	9
	RELB	7
	RELA	6
CG13624	CREBRF	8
CG33967	WWC2	13
	WWC1	13
	WWC3	11
mys	ITGB1	14
	ITGB2	10
	ITGB7	10
	ITGB3	8
	ITGB6	6
	ITGB5	5

	ITGB4	5
Pk61C	PDPK1	13
	PDPK2P	5
CG3927	KHDRBS3	7
	KHDRBS1	7
	KHDRBS2	6
Pect	PCYT2	14
CG3847	none	
alt	none	
InR	IGF1R	12
	INSR	10
	INSRR	9
Oda	OAZ2	9
	OAZ1	8
	OAZ3	7
lin-52	LIN52	12
tud	TDRD6	10
	TDRD15	9
unc-104	KIF1A	11
	KIF1B	9
	KIF1C	6
puc	DUSP10	8
sgg	GSK3B	12
	GSK3A	11
CG5758	none	
CG31738	FNDC3A	12
	FNDC3B	9
CG8486	PIEZO2	13
	PIEZO1	10
hig	none	
CG31300	none	
CHES-1-like	none	
CG32352	none	

Pvf3	none	
caps	LRRN2	5
	LRRC70	2
rdgBbeta	PITPNC1	14
jeb	none	
CG12054	JAZF1	11
CG2225	none	
CG33144	RNF144A	10
	RNF144B	5
aralar1	none	
tws	PPP2R2D	15
	PPP2R2A	13
	PPP2R2B	13
	PPP2R2C	11
Hsromega	none	
Ald	ALDOC	13
	ALDOA	13
	TTK	10
	ALDOB	11
scyl	DDIT4	11
	DDIT4L	11
mbc	DOCK1	13
	DOCK2	12
	DOCK5	11
	DOCK3	5
CG31619	ADAMTSL3	10
	ADAMTSL1	9
Eip74EF	ELF2	7
betaTub56D	TUBB4B	12
	TUBB4A	8
	TUBB	8
	TUBB2B	8
	TUBB2A	6

	TUBB6	5
CG9086	UBR2	14
	UBR1	12
brk	none	
Sara	ZFYVE16	11
	ZFYVE9	11
PGRP-LC	none	
Treh	TREH	14
beat-Iib	none	
sba	MBD5	7
osp	none	
beta-Spec	SPTB	11
Ack	TBK2	12
	TBK1	8
pk	PRICKLE2	12
	PRICKLE1	11
	PRICKLE3	10
glec	none	
scrib	SCRIB	7
Dhc93AB	DNAH9	14
	DNAH17	11
	DNAH11	8
SK	KCNN1	11
	KCNN2	10
	KCNN3	10
CG14207	HSPB1	5
Rdl	none	
CG5151	none	
l(1)G0232	PTPN9	11
shep	RBMS3	11
	RBMS2	10
	RBMS1	10
LpR1	VLDLR	11

	LDLR	9
	LRP8	9
klg	none	
Msp-300	SYNE1	5
Pfrx	PFKFB1	13
	PFKFB3	12
	PFKFB4	12
	PFKFB2	11

928 human externalizing behavior genes and respective fly orthologs described in Chapter 2

gene #	928 genes	fly ortholog	DIOPT	FBGN
1	SEMA6D	Sema1a	8	FBGN011259
		Sema1b	7	FBGN016059
2	CADM2	none >5		
3	LSAMP	Ama	8	FBGN0000071
		DIP-epsilon	8	FBGN0259714
		DIP-kappa	8	FBGN0051814
		DIP-alpha	7	FBGN0052791
		DIP-eta	7	FBGN0031725
		DIP-iota	7	FBGN0031837
		DIP-beta	7	FBGN0259245
		DIP-theta	6	FBGN0051646
		DIP-zeta	6	FBGN0051708
		DIP-gamma	6	FBGN0039617
		DIP-delta	6	FBGN0085420
		Lac	5	FBGN0010238
4	TENM2	Ten-m	13	FBGN0004449
		Ten-a	11	FBGN0267001
5	SDK1	sdk	10	FBGN0021764
6	DCC	fra	12	FBGN0011592
7	MAGI2	Magi	10	FBGN0034590
8	THSD7B	none >5		
9	SORCS3	none >5		
10	MAML3	none >5		
11	MSRA	Eip72CD	11	FBGN0000565
12	DAB1	Dab	8	FBGN0000414
13	PDE4B	dnc	12	FBGN0000479
14	ROBO2	robo1	11	FBGN0005631
		robo3	7	FBGN0041097
		robo2	5	FBGN0002543
15	XKR6	CG32579	10	FBGN0053579

		CG18635	5	FBGN0034279
16	HIST1H3I	His3:CG33866	5	FBGN0053866
		His3: CG33863	5	FBGN0053863
17	CAMTA1	Camta	12	FBGN0259234
18	LRRC4C	kek3	5	FBGN0028370
19	NCAM1	Fas2	11	FBGN0000635
20	WDPCP	frtz	13	FBGN0086698
21	HS6ST3	Hs6st	13	FBGN0038755
22	LAMA2	wb	8	FBGN0261563
23	EFNA5	none >5		
24	NTM	DIP-tehta	8	FBGN0051646
		DIP-zeta	8	FBGN0051708
		DIP-epsilon	8	FBGN0259714
		DIP-kappa	8	FBGN0051814
		DIP-alpha	7	FBGN0052791
		DIP-delta	7	FBGN0085420
		DIP-iota	7	FBGN0031837
		DIP-beta	7	FBGN0259245
		DIP-gamma	7	FBGN0039617
		DIP-eta	7	FBGN0031725
		Ama	6	FBGN0000071
		Lac	5	FBGN0010238
25	SNTG1	Syn2	14	FBGN0034135
26	CTNNA2	alpha-Cat	13	FBGN0010215
27	CELF2	bru1	14	FBGN0000114
		bru2	10	FBGN0262475
28	ST3GAL3	none >5		
29	HYAL1	none >5		
30	HIST1H4L	His4r	5	FBGN0013981
31	FOXP2	FoxP	8	FBGN0262477
32	BDNF	none >5		
33	CHD13	none >5		
34	PTPRF	Lar	11	FBGN0000464

35	C10orf32	CG18065	9	FBGN0034519
36	BTN3A2	none >5		
37	OTX1	oc	7	FBGN0004102
38	ZSCAN12	none >5		
39	BTN2A1	none >5		
40	EXOC4	Sec8	14	FBGN0266672
41	LAMB2	LanB1	13	FBGN0261800
42	ZKSCAN3	none >5		
43	ZKSCAN5	none >5		
44	ERI1	none >5		
45	NEGR1	Lac	9	FBGN0010238
		DIP-zeta	8	FBGN0051708
		DIP-epsilon	7	FBGN0259714
		DIP-eta	7	FBGN0031725
		DIP-beta	7	FBGN0259245
		DIP-kappa	6	FBGN0051814
		DIP-iota	6	FBGN0031837
		DIP-alpha	6	FBGN0052791
		DIP-gamma	6	FBGN0039617
		DIP-delta	6	FBGN0085420
		Ama	5	FBGN0000071
		DIP-theta	5	FBGN0051646
46	ELAVL4	fne	13	FBGN0086675
		Rbp9	12	FBGN0010263
		elav	7	FBGN0260400
47	OPCML	DIP-iota	9	FBGN0031837
		DIP-theta	8	FBGN0051646
		DIP-epsilon	8	FBGN0259714
		DIP-gamma	8	FBGN0039617
		DIP-eta	8	FBGN0031725
		DIP-alpha	7	FBGN0052791
		Ama	7	FBGN0000071
		DIP-zeta	7	FBGN0051708

		DIP-beta	7	FBGN0259245
		DIP-kappa	7	FBGN0051814
		DIP-delta	6	FBGN0085420
		Lac	5	FBGN0010238
48	MDH1	Mdh1	14	FBGN0262782
49	INPP4B	CG42271	13	FBGN0262782
50	CAMKV	none >5		
51	PGBD1	none >5		
52	MAD1L1	Mad1	12	FBGN0026326
53	ZFHX3	zfh2	12	FBGN0004607
54	ZIC4	opa	9	FBGN0003002
55	HIST1H2BN	His2B: CG33884	8	FBGN0053884
56	FBXL16	CG32085	12	FBGN0052085
57	CYP17A1	none >5		
58	ZSCAN16	none >5		
59	ZSCAN4	none >5		
60	FURIN	Fur1	10	FBGN004509
61	TCF4	da	12	FBGN0267821
62	ZSCAN8	none >5		
63	TMEM161B	CG7638	14	FBGN0036133
64	STK32C	CG32944	11	FBGN0052944
65	ZSCAN9	none >5		
66	MFHAS1	none >5		
67	HIST1H3J	His3: CG33830	5	FBGN0053830
		His3: CG33863	5	FBGN0053863
68	UTRN	Dys	9	FBGN0260003
69	ZNF789	none >5		
70	WDR24	Wdr24	14	FBGN0025718
71	NT5C2	CG32549	14	FBGN0052549
72	PINX1	CG11180	7	FBGN0034528
73	HCN1	Ih	5	FBGN0263397
74	FAM002A	none >5		
75	RBFOX1	Rbfox1	9	FBGN0052062

76	ENO4	none >5		
77	ASCC3	obe	15	FBGN0038344
78	RP1L1	none >5		
79	GSI-259H13.10	none >5		
80	TCF20	CG5098	6	FBGN0034300
81	PPP6C	PpV	15	FBGN0003139
82	ESRRG	ERR	14	FBGN0035849
83	C10orf32-ASMT	none >5		
84	ARID5B	none >5		
85	CNNM2	uex	12	FBGN0262124
86	RP11-159G9.5	none >5		
87	TCTA	none >5		
88	CGGBP1	none >5		
89	WBP1L	none >5		
90	ARL3	dnd	12	FBGN0038916
91	C3orf38	CG13876	14	FBGN0035109
92	ZNF423	Oaz	12	FBGN0284250
93	GABRB1	Lcch3	13	FBGN0010240
94	RABEPK	none >5		
95	A3GALT2	none >5		
96	PRKG1	for	13	FBGN0000721
		Pkg21D	7	FBGN0000442
		CG4839	5	FBGN0032187
97	BPTF	E(bx)	13	FBGN0000541
98	NDUFAF2	CG43346	6	FBGN0263051
99	REV3L	mus205	12	FBGN0002891
100	HLA-G	none >5		
101	KLHL29	none >5		
102	OR5V1	none >5		
103	BRINP1	none >5		
104	ZNF165	none >5		
105	NAALADL2	none >5		
106	NQO1	none >5		

107	ZNF322	none >5		
108	ISL1	tup	14	FBGN0003896
109	TNKS	Tnks	12	FBGN0027508
110	SFXN2	Sfxn2	14	FBGN0036843
111	ARPC1B	Arpc1	13	FBGN0001961
112	LRRC4	none >5		
113	TRIM8	none >5		
114	ARPC1A	Arpc1	14	FBGN0001961
115	ZIC1	opa	9	FBGN0003002
116	LRPPRC	bsf	13	FBGN0284256
		Lrpprc2	8	FBGN0027794
117	ERCC8	none >5		
118	SOX7	Sox15	5	FBGN0005613
119	HIC1	none >5		
120	AGBL4	CG31019	13	FBGN0051019
121	SCAI	CG13293	14	FBGN0035677
122	ZNF655	none >5		
123	BIRC6	Bruce	15	FBGN0266717
124	DAG1	Dg	10	FBGN0034072
125	ZNF536	none >5		
126	RANBP17	Ranbp16	12	FBGN0053180
127	MCTP1	Mctp	13	FBGN0034389
128	PCCA	Mccc1	5	FBGN0039877
129	AKT3	Akt1	12	FBGN0010379
130	MON1A	Mon1	13	FBGN0031640
131	CLU	none >5		
132	YIPF4	none >5		
133	RASSF1	none >5		
134	CA10	CARPB	13	FBGN0052698
		CARPA	10	FBGN0029962
135	PCDH7	none >5		
136	PCDH15	Cad99C	11	FBGN0039709
137	CDH8	none >5		

138	RUNX1T1	nvy	11	FBGN0005636
139	ICK	CG42366	9	FBGN0259712
140	ATP5J2	ATPsynF	12	FBGN0035032
141	MLTK	none >5		
142	GATA4	pnr	10	FBGN0003117
		GATAe	6	FBGN0038391
143	WWP2	Su(dx)	11	FBGN0003557
		Nedd4	5	FBGN0259174
144	NOB1	CG2972	14	FBGN0030177
145	ELAVL2	fne	14	FBGN0086675
		Rbp9	12	FBGN0010263
		elav	7	FBGN0260400
146	ICA1L	ICA69	13	FBGN0037050
147	CACNA1D	Ca-alpha1D	13	FBGN0001991
		cac	5	FBGN0263111
148	UBE2E3	Ubc2	13	FBGN0015320
		CG5440	5	FBGN0031331
149	IP6K1	CG10082	12	FBGN0034644
150	PDAP1	CG11444	13	FBGN0029715
		CG4438	13	FBGN0032115
151	SOX7	Sox15	5	FBGN0005613
152	TRAIP	nopo	10	FBGN0034314
153	DHODH	Dhod	14	FBGN0000447
154	HP	none >5		
155	NECAB1	none >5		
156	BTN1A1	none >5		
157	FAM167A	none >5		
158	ACTN2	Actn	13	FBGN0000667
159	KIA1919	none >5		
160	KCNJ6	Irk2	6	FBGN0039081
		Irk1	5	FBGN0265042
161	C15orf59	none >5		
162	NFAT5	NFAT	7	FBGN0030505

163	TRIM39	none >5		
164	EMB	none >5		
165	AFF3	lilli	8	FBGN0041111
166	GALNTL6	Pgant6	14	FBGN0035375
		Pgant4	11	FBGN0051956
		Pgant8	7	FBGN0036529
		CG31776	6	FBGN0051776
		CG7579	5	FBGN0036528
		CG7304	5	FBGN0036527
167	RP11-180C1.1	none >5		
168	WDR12	CG6724	14	FBGN0032298
169	CCDC36	none >5		
170	AUTS2	none >5		
171	ARID4A	htk	11	FBGN0085451
172	ELOVL7	ELOVL	15	FBGN0037534
		CG31522	14	FBGN0051522
		CG31523	10	FBGN0051523
		CG6660	7	FBGN0039030
		eloF	6	FBGN0037762
		CG18609	6	FBGN0034382
		sit	6	FBGN0038986
		CG33110	6	FBGN0053110
		CG30008	5	FBGN0050008
		CG8534	5	FBGN0037761
		CG9458	5	FBGN0037765
		CG9459	5	FBGN0037764
		CG31141	5	FBGN0051141
		CG5326	5	FBGN0038983
		CG17821	5	FBGN0034383
		bond	5	FBGN0260942
		CG16904	5	FBGN0037763
		Elo68beta	5	FBGN0036128
173	CALB1	Cbp53E	10	FBGN0004580

174	OR2J2	none >5		
175	EYS	eys	5	FBGN0031414
176	CARF	none >5		
177	C3orf84	none >5		
178	ACTR10	Arp10	14	FBGN0031050
179	TRIM27	none >5		
180	HPR	none >5		
181	C8orf12	none >5		
182	GOLGA1	cbs	12	FBGN0086757
183	C17orf58	none >5		
184	TBC1D5	TBC1D5	9	FBGN0038129
185	NFKB2	Rel	6	FBGN0014018
186	EHBP1	Ehbp1	12	FBGN0034180
187	AC025287.1	none >5		
188	NFIA	NfI	12	FBGN0042696
189	C20orf112	CG46301	7	FBGN0283651
190	MST1R	none >5		
191	MAPKAP1	Sin1	15	FBGN0033935
192	NRXN3	Nrx-1	11	FBGN0038975
193	HIST1H1B	His1:CG33825	8	FBGN0053825
		His1: CG33807	7	FBGN0053807
		His1: CG33801	7	FBGN0053801
		His1: CG33834	6	FBGN0053834
		His1: CG33861	6	FBGN0053861
194	OR2B2	none >5		
195	ANO4	CG6938	10	FBGN0036235
		CG10353	7	FBGN0030349
		subdued	7	FBGN0038721
196	GBF1	garz	15	FBGN0264560
197	ZNF23	none >5		
198	NKAIN2	NKAIN	9	FBGN0085442
199	WFIKKN1	none >5		
200	ERBB3	Egfr	9	FBGN0003731

201	MTMR9	CG5026	15	FBGN0035945
202	LRRC27	none >5		
203	FBXL17	none >5		
204	FHIT	none >5		
205	TOX	CG12104	5	FBGN0035238
206	GRID1	none >5		
207	FOXP1	FoxP	8	FBGN0262477
208	GRM8	none >5		
209	BNC1	disco-r	9	FBGN0285879
		disco	9	FBGN0000459
210	CNTNAP5	Nrx-IV	10	FBGN0013997
211	CBX8	Pc	9	FBGN0003042
212	KIAAI522	none >5		
213	BRWD1	BRWD3	9	FBGN0011785
214	CUEDC2	CG9636	14	FBGN0037556
215	ADAT1	Adat1	9	FBGN0028658
216	KPNA2	Pen	12	FBGN0267727
217	TFAP2B	TfAP-2	9	FBGN0261953
218	NPAS3	trh	9	FBGN0262139
219	SUFU	Su(fu)	14	FBGN0005355
220	TNRC6A	gw	8	FBGN0051992
221	NCOA5	Neos	10	FBGN0024542
222	WDR38	none >5		
223	PHACTR1	CG32264	8	FBGN0052264
224	SND1	Tudor-SN	13	FBGN0035121
225	PLXNA4	PlexA	12	FBGN0025741
226	USP4	none >5		
227	CPNE4	none >5		
228	LNPEP	CG3502	5	FBGN0046253
		CG31445	5	FBGN0051445
		SP1029	5	FBGN0263236
229	NRXN1	Nrx-1	12	FBGN0038975
230	ACTR1A	Arp1	14	FBGN0011745

231	PRDX5	Prx5	13	FBGN0038570
232	NRAP	none >5		
233	ARID1B	osa	12	FBGN0261885
234	SMIM19	none >5		
235	HIVEP1	shn	7	FBGN0003396
236	IQCJ-SCHIP1	Schip1	9	FBGN0032221
237	TRAF3	none >5		
238	PIGQ	PIG-Q	8	FBGN0086448
239	TMEM163	none >5		
240	LRFN2	none >5		
241	PSD	Efa6	10	FBGN0051158
242	CADPS2	Cadps	11	FBGN0053653
243	GGACT	Tina-1	14	FBGN0035083
		CG2811	14	FBGN0035082
244	CABP1	none >5		
245	SGCD	Scgdelta	13	FBGN0025391
246	PSMA3	Prosalpha7	15	FBGN0023175
247	PMFBP1	none >5		
248	USP28	none >5		
249	CDHR4	none >5		
250	TSR1	Tsr1	13	FBGN0037073
251	RBMS	CG4896	14	FBGN0031319
		CG4887	11	FBGN0031318
252	GAPVD1	Gapvd1	14	FBGN0030286
253	BTBD1	none >5		
254	RNF123	CG6752	14	FBGN0038296
255	SYNGAP1	CG42684	9	FBGN0261570
256	ABT1	CG32708	13	FBGN0052708
		CG10993	12	FBGN0030524
		CG41562	11	FBGN0085693
		CG40813	11	FBGN0085521
		CG32706	10	FBGN0052706
		CG6999	10	FBGN0030085

257	CTNNA3	alpha-Cat	7	FBGN0010215
258	HMGN4	none >5		
259	TM6SF1	none >5		
260	KDM4B	Kdm4A	11	FBGN0033233
		Kdm4B	9	FBGN0053182
261	FAM186B	none >5		
262	HOOK1	hook	10	FBGN0001202
263	NICN1	none >5		
264	FXR1	Fmr1	12	FBGN0028734
265	RAB5B	Rab5	13	FBGN0014010
266	FHIT	none >5		
267	MCRS1	Rcd5	14	FBGN0263832
268	CYB561D2	CG10165	9	FBGN0032801
		CG13078	7	FBGN0032809
		CG13077	7	FBGN0032810
269	PHC2	ph-p	6	FBGN0004861
		pd-d	6	FBGN0004860
270	KDM4A	Kdm4A	11	FBGN0033233
		Kdm4B	10	FBGN0053182
271	RHOA	Rho1	13	FBGN0014020
272	BSN	none >5		
273	SEMA3F	Sema2a	7	FBGN0011260
		Sema2a	6	FBGN0264273
274	GNAT1	Galphai	6	FBGN0001104
275	TMEM115	CG9536	12	FBGN0031818
276	ZNF654	none >5		
277	ZSCAN31	none >5		
278	TRAF3IP2	none >5		
279	PTCD1	CG4611	10	FBGN0035591
280	ATP5J2-PTCD1	CG4611	5	FBGN0035591
281	BLK	Src64B	6	FBGN0262733
282	AS3MT	none >5		
283	KIAA1598	none >5		

284	FES	FER	11	FBGN0000723
285	WDR90	none >5		
286	RHOT2	Miro	12	FBGN0039140
287	STUB1	STUB1	15	FBGN0027052
288	MDGA1	none >5		
289	GRID1	none >5		
290	ELMO1	Ced-12	13	FBGN0032409
291	HDGFRP3	CG7946	7	FBGN0039743
292	CALB2	Cbp53E	11	FBGN0004580
293	FEZF1	erm	9	FBGN0031375
294	ESR1	none >5		
295	OR12D3	none >5		
296	C14orf37	none >5		
297	LONRF2	CG32369	12	FBGN0052369
298	GABARAPL2	Atg8a	5	FBGN0052672
299	NBEAL1	CG43367	8	FBGN0263110
300	PDE11A	Pde6	7	FBGN0038237
		Pde11	6	FBGN0085370
301	AMBRA1	none >5		
302	RAB19	Rab19	8	FBGN0015793
303	ARFGAP2	ArfGAP3	14	FBGN0037182
304	ATG13	Atg13	11	FBGN0261108
305	AC117395.1	none >5		
306	TRIM26	none >5		
307	IGSF11	none >5		
308	HIST1H3C	His3: CG33863	6	FBGN0053863
309	DGKZ	rdgA	14	FBGN0261549
310	SLC24A3	zyd	9	FBGN0265767
		CG17167	5	FBGN0039941
311	TDRKH	papi	11	FBGN0031401
312	SCN2A	para	11	FBGN0258944
313	LRRTM4	none >5		
314	PRPF40B	CG3542	13	FBGN0031492

315	REST	none >5		
316	CCDC71	none >5		
317	CPSF4	Clp	15	FBGN0015621
318	SUOX	shop	15	FBGN0030966
319	AF131215.5	none >5		
320	PSMD11	Rpn6	14	FBGN0028689
321	SPAST	spas	15	FBGN0039141
322	ARHGAP15	none >5		
323	DIAPH3	dia	12	FBGN0011202
324	CIT	sti	12	FBGN0002466
325	RAB40C	Rab40	14	FBGN0030391
326	ZNF207	BuGZ	10	FBGN0032600
327	IGF1	none >5		
328	RBM6	CG4896	6	FBGN0031319
		CG4887	5	FBGN0031318
329	PARD3B	baz	11	FBGN0000163
330	TRIM31	none >5		
331	KIAA1328	none >5		
332	HYAL3	none >5		
333	HHLA2	none >5		
334	CCDC53	CCDC53	14	FBGN0031979
335	FMNL3	Frl	13	FBGN0267795
336	PLCB3	Plc21C	12	FBGN0004611
337	PTK2	Fak	12	FBGN0020440
338	HIST1H2AJ	His2A: CG33832	6	FBGN0053832
		His2A: CG33859	5	FBGN0053859
339	DLG2	dlg1	11	FBGN0001624
340	ABDH17C	CG33096	12	FBGN0053096
341	IPO13	cdm	14	FBGN0261532
342	DEFB134	none >5		
343	AAGAB	CG32109	8	FBGN0052109
344	IMPDH2	ras	10	FBGN0003204
345	ZSCAN23	none >5		

346	CHD3	Mi-2	12	FBGN0262519
		Chd3	7	FBGN0023395
347	PPP1R3B	Gbs-70E	9	FBGN0036428
		Gbs-76A	5	FBGN0036862
348	ASPG	CG6428	13	FBGN0029689
		CG8526	13	FBGN0037759
349	POM121L2	none >5		
350	DUSP6	Mkp3	10	FBGN0036844
351	OAZ3	Oda	7	FBGN0014184
352	TMEM71	none >5		
353	RNF19B	none >5		
354	RP11-463C8.4	none >5		
355	IMMP2L	CG11110	15	FBGN0034535
356	C8orf88	none >5		
357	TERF21P	none >5		
358	IKZF2	none >5		
359	NYAP2	none >5		
360	NBN	nbs	10	FBGN0261530
361	RPS6KA4	JIL-1	8	FBGN0020412
362	CNTN4	Cont	12	FBGN0037240
363	ZNF19	none >5		
364	PARPBP	none >5		
365	ELL	Su(Tpl)	9	FBGN0014037
366	S100PBP	none >5		
367	TRAPPC1	Bet5	14	FBGN0260860
368	SLC39A5	Zip71B	8	FBGN0035461
369	BUD31	l(1)10Bb	14	FBGN0001491
370	PTPRD	Lar	12	FBGN0000464
371	KCNH3	Elk	11	FBGN0011589
372	C17orf75	none >5		
373	METTL16	CG7544	12	FBGN0033994
374	GUCA1C	none >5		
375	ITIH3	none >5		

376	ZFYVE21	none >5		
377	PFN2	none >5		
378	NLRC4	none >5		
379	TMBIM6	BI-1	14	FBGN0035871
380	RP11-481A20.11	none >5		
381	THBS4	Tsp	12	FBGN0031850
382	ZNF652	none >5		
383	UNC79	unc79	14	FBGN0038693
384	NUP37	Nup37	14	FBGN0039301
385	PXDNL	Pxn	11	FBGN0011828
386	MST1	none >5		
387	NAA25	psidin	15	FBGN0243511
388	MORC1	none >5		
389	ISYNA1	Inos	13	FBGN0025885
390	XXcos-LUCA11.5	none >5		
391	CACNB4	Ca-beta	12	FBGN0259822
392	VT11A	Vti1a	14	FBGN0260862
393	YARS	TyrRS	15	FBGN0027080
394	CYP2J2	Cyp18a1	9	FBGN0010383
		Cyp305a1	6	FBGN0036910
395	SLC30A6	none >5		
396	SFMBT1	none >5		
397	TXNL4B	none >5		
398	BCL11A	CG9650	11	FBGN0029939
399	CHDH	CG9514	8	FBGN0030592
		Gld	7	FBGN0001112
		CG12398	7	FBGN0030596
		CG9518	7	FBGN0030590
		CG9519	6	FBGN0030589
		CG12539	6	FBGN0030586
		CG9522	6	FBGN0030587
		CG9521	6	FBGN0030588
		CG6142	6	FBGN0039415

		CG9503	6	FBGN0030598
		CG9512	6	FBGN0030593
		fiz	6	FBGN0030594
400	TTC29	none >5		
401	LINGO1	none >5		
402	PLEKHA5	CG34383	5	FBGN0085412
403	ASPSCR1	CG33722	12	FBGN0064126
404	LUZP2	none >5		
405	FAF1	casp	14	FBGN0034068
406	NSF	Nsf2	14	FBGN0266464
		comt	14	FBGN0000346
		CG31495	6	FBGN0051495
407	TMEM55A	CG6707	12	FBGN0036058
408	MTX3	CG9393	11	FBGN0037710
409	BAI3	none >5		
410	PTPRN2	IA-2	10	FBGN0031294
411	XXbac-BPG2J3.20	none >5		
412	TMEM180	none >5		
413	SSBP4	Ssdp	12	FBGN0011481
414	VARS	ValRS	15	FBGN0027079
		ValRS-m	7	FBGN0035942
415	PTPRG	Ptp99A	9	FBGN0004369
416	NFATC2IP	none >5		
417	SIPA1L2	none >5		
418	CPSF6	Cpsf6	13	FBGN0035872
419	RAB7A	Rab7	13	FBGN0015795
420	C1orf87	none >5		
421	QK1	how	12	FBGN0264491
422	BTN3A3	none >5		
423	HIST1H2AL	none >5		
424	ITIH4	none >5		
425	CCDC88B	Girdin	5	FBGN0283724
426	PITX3	Ptx1	8	FBGN0020912

427	TRIM39-RPP21	none >5		
428	TGFBRAP1	none >5		
429	IQCH	none >5		
430	PPP1R13B	ASPP	14	FBGN0034606
431	USMG5	Neb-cGP	7	FBGN0083167
432	DPP10	CG17684	9	FBGN0263780
		CG11319	9	FBGN0031835
		CG11034	9	FBGN0031741
		ome	8	FBGN0259175
		CG45002	6	FBGN0266354
433	CHST10	none >5		
434	CD40	none >5		
435	ZKSCAN2	none >5		
436	OTX2	oc	6	FBGN0004102
437	PSME4	none >5		
438	CD47	none >5		
439	KCNJ3	Irk2	5	FBGN0039081
440	AGBL1	none >5		
441	ZNF408	none >5		
442	DXO	CG9125	14	FBGN0030793
		cuff	7	FBGN0260932
443	CUL3	Cul3	13	FBGN0261268
444	SLC39A13	none >5		
445	HIST1H4A	His4: CG33887	5	FBGN0053887
446	HTR1F	none >5		
447	SPOCK1	Cow	9	FBGN0039054
448	CYP3A43	Cyp9f2	8	FBGN0038037
		Cyp9c1	7	FBGN0015040
		Cyp9b1	7	FBGN0015038
		Cyp6a17	7	FBGN0015714
		Cyp9h1	6	FBGN0033775
		Cyp9b2	6	FBGN0015039
		Cyp6a8	6	FBGN0013772

		Cyp6a2	6	FBGN0000473
		Cyp6a23	6	FBGN0033978
		Cyp6a18	6	FBGN0039519
		Cyp6a22	6	FBGN0013773
		Cyp6a9	6	FBGN0013771
		Cyp6a19	6	FBGN0033979
		Cyp6a13	5	FBGN0033304
		Cyp6g1	5	FBGN0025454
		Cyp6w1	5	FBGN0033065
		Cyp6a20	5	FBGN0033980
		Cyp6v1	5	FBGN0031126
		Cyp6g2	5	FBGN0033696
		Cyp6a21	5	FBGN0033981
449	CD19	none >5		
450	DENND1B	CG18659	10	FBGN0027561
451	MARCH1	none >5		
452	HARBI1	CG12253	8	FBGN0026148
		CG43088	5	FBGN0262534
453	CLEC18A	none >5		
454	GRIK2	KaiR1D	10	FBGN0038837
		Ekar	9	FBGN0039916
		Grik	8	FBGN0038840
		GluRIID	7	FBGN0028422
		clumsy	7	FBGN0026255
		GluRIIE	7	FBGN0051201
		CG11155	7	FBGN0039927
455	BHLHE22	Oli	12	FBGN0032651
456	C11orf49	none >5		
457	PLD5	CG9248	6	FBGN0032923
		CG43345	5	FBGN0263050
458	NRXN2	Nrx-1	12	FBGN0038975
459	SLC1A5	Eaat1	7	FBGN0026439
460	TECPR1	Pex23	14	FBGN0052226

461	ERBB4	Egfr	10	FBGN0003731
462	GPR137	none >5		
463	HECTD4	none >5		
464	PWWP2B	none >5		
465	MYT1L	CG43689	7	FBGN0263772
466	RBM43	none >5		
467	SPNS1	spin	13	FBGN0086676
468	TNRC6B	gw	8	FBGN0051992
469	ALMS1	none >5		
470	JADE2	rno	7	FBGN0035106
471	FAM135B	CG32333	13	FBGN0052333
472	DBNDD1	none >5		
473	C1orf173	none >5		
474	GABRA2	CG8916	10	FBGN0030707
		Grd	6	FBGN0001134
475	ARTN	none >5		
476	EFTUD1	CG33158	15	FBGN0053158
477	ZCCHC7	Zcchc7	6	FBGN0036668
478	CTD-2330K9.3	none >5		
479	CNPY2	sel	12	FBGN0263260
480	CACYBP	CG3226	14	FBGN0029882
481	RPL6	RpL6	14	FBGN0039857
482	DCD1	none >5		
483	SGK223	none >5		
484	PHLPP2	Phllp	9	FBGN0032749
485	DNAJC11	CG8531	14	FBGN0033918
486	ARL17B	none >5		
487	FAT3	kug	13	FBGN0251574
488	KANSL1	nsl1	9	FBGN0262527
489	HACE1	none >5		
490	FDFT1	none >5		
491	STON1-GTF2A1L	TfIIA-L	5	FBGN0011289
492	CPNE7	none >5		

493	CCDC101	Sgf29	13	FBGN0050390
494	ZNF75A	none >5		
495	C6orf48	none >5		
496	SPI1	none >5		
497	MAPT	tau	7	FBGN0266579
498	QRICH1	none >5		
499	SZT2	none >5		
500	CPEB4	orb2	9	FBGN0264307
501	GABRB2	Lcch3	12	FBGN0010240
502	CRHR1	Dh44-R2	9	FBGN0033744
		Dh44-R1	9	FBGN0033932
503	ERP29	wbl	15	FBGN0004003
504	RYR2	RyR	14	FBGN0011286
505	FPGT-TNNI3K	none >5		
506	TNNI3K	none >5		
507	IGF1R	InR	12	FBGN0283499
508	NOTUM	Notum	11	FBGN0044028
509	APEH	none >5		
510	MPL	none >5		
511	STH	none >5		
512	SLC20A2	NaPi-III	12	FBGN0260795
513	STAT2	none >5		
514	WNT3	none >5		
515	RP11-977G19.10	none >5		
514	ABHD16A	CG1309	14	FBGN0035519
517	ATAT1	CG17003	10	FBGN0031082
518	CS	kdn	14	FBGN0261955
		CG14740	7	FBGN0037988
519	SMG6	Smg6	11	FBGN0039260
520	SMYD2	Smyd3	10	FBGN0011566
		SmydA-9	6	FBGN0030102
		SmydA-5	5	FBGN0033061
521	RP5-966M1.6	none >5		

522	PAN2	PAN2	14	FBGN0033352
523	ARIH2	ari-2	15	FBGN0025186
524	POSTN	mfas	10	FBGN0260745
		Fas1	7	FBGN0285925
		CG43333	5	FBGN0263038
525	SCAND3	none >5		
526	ZNF385D	none >5		
527	PIPOX	none >5		
528	PHF10	e(y)3	9	FBGN0087008
529	DALRD3	CG8097	11	FBGN0030660
530	KLHDC8B	none >5		
531	ASTN2	none >5		
532	PCGF6	none >5		
533	PSMG1	none >5		
534	AGO2	AGO1	14	FBGN0262739
535	ZNF263	none >5		
536	AP1G1	AP-1gamma	13	FBGN0030089
537	POU1F1	none >5		
538	RSRC1	none >5		
539	SLC17A3	CG3649	6	FBGN0034785
		MFS12	6	FBGN0033234
		CG30265	6	FBGN0050265
		CG9825	6	FBGN0034783
		dmGlut	6	FBGN0010497
		MFS17	5	FBGN0058263
		CG2003	5	FBGN0039886
		CG7881	5	FBGN0033048
		MFS1	5	FBGN0050272
		MFS14	5	FBGN0010651
		CG6978	5	FBGN0029727
		CG9254	5	FBGN0028513
		CG12490	5	FBGN0034782
		CG15096	5	FBGN0034394

		MFS15	5	FBGN0034392
540	USP19	none >5		
541	WDR27	none >5		
542	AC011239.1	none >5		
543	DPEP1	CG44837	8	FBGN0266100
		CG42750	6	FBGN0261804
		CG5282	6	FBGN0036986
544	GPX1	none >5		
545	TRIM38	none >5		
546	LZTS2	CG15365	6	FBGN0030077
547	DPYD	su(r)	15	FBGN0086450
548	SRR	Srr	8	FBGN0037684
549	SMIM15	none >5		
550	NLGN1	Nlg3	11	FBGN0083963
		Nlg2	10	FBGN0031866
		Nlg4	8	FBGN0083975
		Nlg1	7	FBGN0051146
551	ANKRD52	none >5		
552	AC010547.9	none >5		
553	TRAFD1	none >5		
554	SPPL2C	none >5		
555	PTCRA	none >5		
556	DEFB136	none >5		
557	BMPR2	wit	9	FBGN0024179
558	ZNF407	none >5		
559	MACROD2	none >5		
560	TMEM110	none >5		
561	BRI3	CG12012	12	FBGN0035444
562	FBXO9	CG5961	15	FBGN0038056
563	TMEM110-MUSTN1	none >5		
564	PHF1	Pcl	10	FBGN0003044
564	AC079354.1	none >5		
566	CCSER1	none >5		

567	CDH12	none >5		
568	CAMK1D	CaMKI	13	FBGN0016126
569	ARHGAP1	RhoGAP68F	15	FBGN0036257
570	RBM14	none >5		
571	QARS	GlnRS	14	FBGN0027090
572	ERAP2	CG3502	5	FBGN0046253
		CG8773	5	FBGN0038135
		CG31445	5	FBGN0051445
		SP1029	5	FBGN0263236
573	CHADL	none >5		
574	MSH5-SAPCD1	none >5		
575	MSH5	none >5		
576	HPCA	Nca	10	FBGN0013303
577	KCNAB3	Hk	10	FBGN0263220
578	GRIA4	GluRIB	12	FBGN0264000
		GluRIA	11	FBGN0004619
579	FAM13C	CG6424	6	FBGN0028494
580	WDR6	CG33172	13	FBGN0053172
581	CTTNBP2	none >5		
582	MEF2C	Mef2	11	FBGN0011656
583	ATP6V0B	VhaPPA-1	15	FBGN0028662
		VhaPPA1-2	9	FBGN0262514
584	ATP2A1	SERCA	14	FBGN0263006
585	SLC22A12	CG8654	6	FBGN0034479
		Orct	6	FBGN0019952
		Orct2	6	FBGN0086365
		CG6126	5	FBGN0038407
		Balat	5	FBGN0033778
		CG6356	5	FBGN0039178
		CG7458	5	FBGN0037144
		SLC22A	5	FBGN0037140
		CG4630	5	FBGN0033809
586	SLC25A20	colt	14	FBGN0019830

		MME1	10	FBGN0031881
587	SLC4A2	none >5		
588	AHCYL1	AhcyL1	14	FBGN0035371
		AhcyL2	10	FBGN0015011
589	TET2	none >5		
590	GSTO2	GstO2	12	FBGN0035906
		GstO3	12	FBGN0035904
		se	12	FBGN0086348
		GstO1	11	FBGN0035907
591	PSMC3	Rpt5	13	FBGN0028684
592	MLLT10	Alh	6	FBGN0261238
593	FOXO3	foxo	9	FBGN0038197
594	GPR135	none >5		
595	AP1S1	AP-1sigma	11	FBGN0039132
596	C10orf2	mtDNA-helicase	14	FBGN0032154
		Chmp1	12	FBGN0036805
597	SLC35G5	none >5		
598	MSI1	Rbp6	11	FBGN0260943
599	POLR2F	RpII18	14	FBGN0003275
600	MAPK14	p38b	15	FBGN0024846
		p38a	13	FBGN0015765
		p38c	5	FBGN0267339
601	TRHDE	none >5		
602	IGSF21	none >5		
603	PRKAR2A	Pka-R2	13	FBGN0022382
		Pka-R1	5	FBGN0259243
604	UBXN1	CG8209	15	FBGN0035830
605	GRID2	none >5		
606	RP11-894J14.5	none >5		
607	CCDC24	none >5		
608	CNTNAP4	Nrx-IV	10	FBGN001397
609	LY6H	none >5		
610	MEIS2	hth	11	FBGN0001235

611	RP11-410N8.4	none >5		
612	RFT1	CG3149	13	FBGN0027564
613	KIF26A	CG14535	7	FBGN0031955
614	SPG7	Spg7	13	FBGN0024992
615	LAT	none >5		
616	FNDC5	none >5		
617	SULT1A2	St1	8	FBGN0034887
		St3	7	FBGN0265052
		St4	7	FBGN0033887
		St2	6	FBGN0037665
618	ARPC5L	Arpc5	13	FBGN0031437
619	XRCC3	spn-B	15	FBGN0003480
620	RPL7L1	RpL7	6	FBGN0005593
621	TMEM116	none >5		
622	GPC6	dlp	12	FBGN0041604
623	PFKFB2	Pfrx	11	FBGN0027621
624	CTDP1	Fcp1	13	FBGN0035026
625	EP300	nej	12	FBGN0261617
626	L3MBTL2	Sfmbt	10	FBGN0032475
627	CCDC167	none >5		
628	IFRD2	CG31694	13	FBGN0051694
629	BSCL2	Seipin	12	FBGN0040336
630	GAS8	Gas8	13	FBGN0029667
631	ATAD2B	none >5		
632	IL23A	none >5		
633	SORCS1	none >5		
634	HOMER2	homer	15	FBGN0025777
635	ZBTB16	none >5		
636	COQ10A	CG9410	13	FBGN0033086
637	SREBF1	SREBP	12	FBGN0261283
638	DPP4	CG11034	10	FBGN0031741
		ome	9	FBGN0259175
		CG17684	6	FBGN0263780

639	ANKS1B	CG4393	10	FBGN0039075
		CG11168	10	FBGN0039249
640	MSI2	Rbp6	13	FBGN0260943
		msi	5	FBGN0011666
641	PLCL2	none >5		
642	STRIP1	Strip	14	FBGN0035437
643	SEMA4G	Sema2a	5	FBGN0011260
		Sema2b	5	FBGN0264273
644	PRR7	none >5		
645	CNPY3	CNPYb	10	FBGN0036847
646	PDE1C	Pde1c	8	FBGN0264815
647	PLD1	Pld	12	FBGN0286511
648	ARHGAP27	none >5		
649	MRPL43	mRpL43	14	FBGN0034893
650	TMPO	none >5		
651	ZAP70	Shark	7	FBGN0015295
652	TTC9C	none >5		
653	PLEKHM1	none >5		
654	DHX30	none >5		
655	ITIH1	none >5		
656	DEPDC1B	none >5		
657	BPNT1	CG7789	12	FBGN0039698
658	PACSIN3	Synd	11	FBGN0053094
659	HNRNPUL2-BSCL2	none >5		
660	TOP2B	Top2	12	FBGN0284220
661	NAT16	none >5		
662	USF2	Usf	8	FBGN0029711
663	FKRP	CG15651	14	FBGN0034567
664	KIAA1462	none >5		
665	TCAP	none >5		
666	PIK3R3	PiK21B	13	FBGN0020622
667	BTAF1	Hel89B	12	FBGN0022787
668	NDFIP2	Ndfip	13	FBGN0052177

669	AMIGO3	none >5		
670	GMPPB	CG1129	14	FBGN0037279
671	CHMP3	Vps24	6	FBGN0037231
672	PLAGL2	none >5		
673	AL645922.1	none >5		
674	P4HTM	none >5		
675	BEND7	none >5		
676	CMTR1	CG6379	12	FBGN0029693
677	RP11-831H9.11	none >5		
678	TEAD3	sd	12	FBGN0003345
679	RNF103-CHMP3	Vps24	10	FBGN0037231
680	TPM3	Tm2	11	FBGN0003721
		Tm1	11	FBGN0003721
681	NKAPL	CG6066	11	FBGN0039488
682	RABEP2	none >5		
683	CCNJ	CycJ	8	FBGN0010317
684	PHF13	none >5		
685	MKRN1	CG5347	13	FBGN0030578
		Mkrn1	11	FBGN0029152
		CG5332	10	FBGN0030577
		CG12477	7	FBGN0036809
686	TOM1L2	CG3529	13	FBGN0035995
687	DHX16	l(2)37Cb	14	FBGN0086444
688	PHYHIPL	none >5		
689	CASP7	Drice	12	FBGN0019972
		Dcp-1	12	FBGN0010501
690	RORA	Hr3	9	FBGN0000448
691	HNRNPUL2	CG30122	8	FBGN0050122
692	DLX6	Dll	7	FBGN0000157
693	LYZ	LysP	13	FBGN0004429
		LysE	12	FBGN0004428
		LysD	11	FBGN0004427
		LysS	11	FBGN0004430

		LysB	11	FBGN0004425
		CG7798	10	FBGN0034092
		LysX	10	FBGN0004431
		CG16799	8	FBGN0034538
		CG11159	8	FBGN0034539
		CG16756	7	FBGN0029765
		CG30062	6	FBGN0050062
		CG8492	5	FBGN0035813
694	ADAM15	Meltrin	9	FBGN0265140
695	SORL1	none >5		
696	MCHR2	none >5		
697	JKAMP	CG2126	14	FBGN0039876
698	HYI	Gip	14	FBGN0011770
699	LIN7C	veli	12	FBGN0039269
700	SLIT2	sli	12	FBGN0264089
701	CYP7B1	none >5		
702	C15orf40	CG14966	12	FBGN0035415
703	VPS54	scat	14	FBGN0011232
704	MAP2K5	none >5		
705	AC005544.1	none >5		
706	SFTA2	none >5		
707	C11orf48	none >5		
708	SIKE1	Fgop2	11	FBGN0031871
709	CIPC	none >5		
710	LMO3	Bx	7	FBGN0032196
711	NSUN3	none >5		
712	CUX1	ct	10	FBGN0004198
713	USP3	none >5		
714	BAD	none >5		
715	PNMA2	none >5		
716	ZBTB46	none >5		
717	LINGO4	none >5		
718	PSORS1C2	none >5		

719	DYPSL2	CRMP	10	FBGN0023023
720	HLA-B	none >5		
721	HSPA5	none >5		
722	KLHL21	none >5		
723	TMEFF2	none >5		
724	GSTO1	se	14	FBGN0086348
		GSTO1	13	FBGN0035907
		GSTO2	11	FBGN0035906
725	AZGP1	none >5		
726	CTBP1	CtBP	12	FBGN0020496
727	B4GLT2	beta4GalNAcTA	12	FBGN0027538
728	VPS33B	Vps33B	7	FBGN0039335
729	FTO	none >5		
730	KIAA1551	none >5		
731	BAIAP2L1	IRSp53	8	FBGN0052082
732	C22orf23	CG5280	12	FBGN0035952
733	SMARCC1	mor	13	FBGN0002783
734	CSNK2B-LY6G5B-1181	none >5		
735	EXT1	ttv	13	FBGN0254974
736	ADAP2	none >5		
737	SNRK	CG8485	11	FBGN0033915
738	BNIP3L	CG5059	5	FBGN0037007
739	FDPS	Fpps	15	FBGN0025373
740	ATXN2L	Atx2	10	FBGN0041188
741	ZBTB9	none >5		
742	SERPINC1	none >5		
743	CDKN2C	none >5		
744	UBXN2A	p47	8	FBGN0033179
		CG42383	7	FBGN0259729
745	MYO15A	Myo10A	13	FBGN0263705
746	ZBTB37	none >5		
747	RP11-10A14.4	none >5		
748	NOA1	CG10914	15	FBGN0034307

749	SGIP1	CG8176	8	FBGN0037702
750	HISTH1A	Hist1: CG33834	5	FBGN0053834
		Hist1: CG33807	5	FBGN0053807
		Hist1: CG33801	5	FBGN0053801
		Hist1: CG33825	5	FBGN0053825
751	C16orf11	none >5		
752	ARID2	Bap170	13	FBGN0042085
753	REV1	Rev1	13	FBGN0035150
754	C16orf3	none >5		
755	CLIC1	Clic	10	FBGN0030529
756	RBM19	CG3335	15	FBGN0036018
757	HMGA1	none >5		
758	COQ5	Coq5	13	FBGN0030460
759	PWWP2A	none >5		
760	GCAT	CG10361	15	FBGN0036208
761	GPR75-ASB3	none >5		
762	GRM7	none >5		
763	OLFM1	none >5		
764	CHMP1A	none >5		
765	MAP4	tau	6	FBGN0266579
766	SH2B1	Lnk	10	FBGN0028717
767	ZFHX4	zfh2	12	FBGN0004607
768	TMEM132B	dtn	13	FBGN0262730
769	MED17	MED17	14	FBGN0038578
770	ACYP2	CG18371	10	FBGN0033893
		CG14022	10	FBGN0031700
		Acyp2	9	FBGN0038363
		CG34161	8	FBGN0085190
		Acyp	8	FBGN0025115
		CG11052	7	FBGN0040524
771	C2orf69	CG31122	13	FBGN0051122
772	AQP8	none >5		
773	PSMA6	CG30382	12	FBGN0050382

		Prosalpha1	11	FBGN0263121
774	MAST2	dop	11	FBGN0267390
775	SPATA33	none >5		
776	TYW5	none >5		
777	NMD3	Nmd3	13	FBGN0023542
778	NOL9	CG8414	9	FBGN0034073
779	CTD-2260A17.2	CG3502	5	FBGN0046253
		CG8773	5	FBGN0038135
		CG31445	5	FBGN0051445
		SP1029	5	FBGN0263236
780	CTNNAL1	alpha-Catr	14	FBGN0029105
781	DPCR1	none >5		
782	GSTA4	none >5		
783	MED27	MED27	14	FBGN0037359
784	ALEX3	none >5		
785	FLJ27365	none >5		
786	C11orf80	none >5		
787	POLDIP2	POLDIP2	12	FBGN0037329
788	UBA7	Uba1	7	FBGN0023143
789	ANO10	Axs	12	FBGN0000152
790	CSMD1	none >5		
791	UBE2E1	Ubc2	12	FBGN0015320
792	HPSS	p	9	FBGN0086679
793	ATAD5	elg1	10	FBGN0036574
794	U2SURP	CG9346	13	FBGN0034572
795	MTX1	CG9393	14	FBGN0037710
796	BCL11B	CG9650	12	FBGN0029939
797	ASNS	AsnS	11	FBGN0270926
798	AL132989.1	none >5		
799	HFE	none >5		
800	UVSSA	none >5		
801	TMEM17	none >5		
802	TUFM	mEFTu1	13	FBGN0024556

		mEFTu2	5	FBGN0033184
803	RP11-529K1.3	none >5		
804	SEZ6	none >5		
805	PLEKHG5	CG42674	6	FBGN02161556
806	DEF8	CG11534	15	FBGN0046296
807	DYNC1I2	sw	13	FBGN0003654
		Sdic1	12	FBGN0067861
		Sdic4	11	FBGN0053499
		SdicB	7	FBGN0283433
		Sdic3	7	FBGN0052823
		Sdic2	6	FBGN0053497
		SdicC	5	FBGN0283434
808	PTPRM	none >5		
809	PRKAB1	alc	13	FBGN0260972
810	BTBD9	BTBD9	14	FBGN0030228
811	C6orf1	none >5		
812	MEGT1	none >5		
813	TMEM160	none >5		
814	CSPG5	none >5		
815	PRDM5	none >5		
816	ANKRD54	CG10809	10	FBGN0036052
817	SLC26A7	Prestin	6	FBGN0036770
818	BAIAP2	IRSp53	10	FBGN0052082
819	TM9SF4	TM9SF4	14	FBGN0028541
		TM9SF2	5	FBGN0032880
820	ATXN1L	Atx-1	7	FBGN0029907
821	SOX10	Sox100B	5	FBGN0024288
822	CHMP2B	CHMP2B	14	FBGN0035589
823	TSPAN2	none >5		
824	AKIRIN2	akirin	14	FBGN0082598
825	CASC10	none >5		
826	CLN3	Cln3	13	FBGN0036756
827	KLF7	luna	8	FBGN0040765

828	OR12D2	none >5		
829	LY6G6F	none >5		
830	COPZ2	zetaCOP	7	FBGN0040512
831	PRKD1	PKD	13	FBGN0038603
832	ALS2CL	Als2	9	FBGN0037116
833	RP11-297N6.4	none >5		
834	CHCHD3	Chchd3	11	FBGN0010808
835	MRPS14	mRpS14	13	FBGN0044030
836	SGCE	Scgalpha	14	FBGN0032013
837	FAM216B	none >5		
838	PKLR	PyK	11	FBGN0267385
		CG2964	6	FBGN0031462
		CG7069	5	FBGN0038952
		CG7362	5	FBGN0038258
839	HSD17B12	spidey	14	FBGN0029975
		CG31809	6	FBGN0051809
		CG13284	6	FBGN0032614
		CG31810	5	FBGN0051810
840	ADCK1	Adck	14	FBGN0035039
841	KCND2	Shal	13	FBGN0005564
842	IQCJ	none >5		
843	PROS1	none >5		
844	DNAJB1	CG5001	12	FBGN0031322
		DnaJ-1	12	FBGN0263106
		CG2887	6	FBGN0030207
845	CACNA1E	cac	9	FBGN0263111
846	EMR2	none >5		
847	MTFR1	none >5		
848	HEYL	Hey	7	FBGN0027788
849	GRK6	Gprk2	12	FBGN0261988
850	DRG2	CG6195	15	FBGN0038723
851	REEP1	ReepA	8	FBGN0261564
		CG5539	7	FBGN0034907

		Reepl1	5	FBGN0030313
852	PIK3C2G	Pi3K68D	8	FBGN0015278
853	LY6G5C	none >5		
854	CLN3	Cln3	13	FBGN0036756
855	SLC6A9	GlyT	7	FBGN0034911
856	PP1L	CG7115	14	FBGN0027515
857	DBN1	none >5		
858	TAS1R1	none >5		
859	KCNE1	none >5		
860	SLC6A10	kar	13	FBGN0001296
861	UNC119B	unc-119	12	FBGN0025549
862	PCLO	none >5		
863	L3HYPDH	none >5		
864	FAM206A	CG9288	13	FBGN0260464
865	DXL5	none >5		
866	OVOL1	ovo	8	FBGN0003028
867	KARS	LysRS	15	FBGN0027084
868	TNF	none >5		
869	TMEM69	none >5		
870	SLC17A1	CG2003	5	FBGN0039886
		CG7881	5	FBGN0033048
		CG12490	5	FBGN0034782
		CG3649	5	FBGN0034785
		MFS12	5	FBGN0033234
		CG30265	5	FBGN0050265
		CG9825	5	FBGN0034783
871	GNL1	Ns4	14	FBGN0032882
872	KIF2B	Klp10A	9	FBGN0030268
		Klp59C	5	FBGN0034824
		Klp59D	5	FBGN0034827
873	HELZ	CH9425	12	FBGN0036451
874	TTL	none >5		
875	RPL29	RpL29	9	FBGN0016726

876	LSM 2.00	none >5		
877	ZNF394	none >5		
878	HPCAL4	Nca	5	FBGN0013303
		CG7646	5	FBGN0036926
879	CCDC175	none >5		
880	PPP1R18	none >5		
881	ALS589765.1	none >5		
882	RIIAD1	none >5		
883	MSH6	Msh6	13	FBGN0036486
884	KCNA4	Sh	8	FBGN0003380
885	STON1	none >5		
886	TMIGD1	none >5		
887	CTSB	CtsB1	14	FBGN0030521
888	SMARCD3	Bap60	11	FBGN0025463
889	KIAA0586	none >5		
890	TEAD1	sd	13	FBGN0003345
891	AC068039.1	none >5		
892	CSNK2B	CkIIbeta	13	FBGN0000259
		CkIIbeta2	6	FBGN0026136
		Ssl	5	FBGN0015300
		Ste: CG33237	5	FBGN0053237
893	VWA7	none >5		
894	ACTR1B	Arp1	12	FBGN0011745
895	GPT	CG1640	13	FBGN0030478
896	MICALL1	MICAL-like	12	FBGN0036333
897	FAM19A5	none >5		
898	C11orf83	none >5		
899	TUBB	betaTub56D	8	FBGN0284243
		betaTub85D	7	FBGN0003889
		betaTub60D	5	FBGN0003888
900	TNXB	none >5		
901	BAG5	none >5		
902	B3GALT4	none >5		

903	SNX29	CG5439	9	FBGN0032476
904	ZBTB20	none >5		
905	CELF3	bru3	10	FBGN0264001
906	GID4	none >5		
907	GNAI2	Galphai	11	FBGN0001104
908	ASXL1	Asx	6	FBGN0261823
909	OR4C12	none >5		
910	C6orf136	CG16787	10	FBGN0034940
911	ERMARD	none >5		
912	PNMT	none >5		
913	SHMT1	Shmt	14	FBGN0029823
914	NRGN	none >5		
915	ASZ1	Gasz	6	FBGN0033273
916	TLR9	none >5		
917	TSPAN18	none >5		
918	CHEK2	lok	13	FBGN0019686
919	OR14J1	none >5		
920	DOT1L	gpp	12	FBGN0264495
921	PPP1R16A	MYPT-75D	12	FBGN0036801
922	DLGAP1	vlc	7	FBGN0259978
923	NMI	none >5		
924	MARK3	par-1	12	FBGN0260934
925	IP6K3	CG10082	11	FBGN0034644
926	TRIM40	none >5		
927	HLA-F	none >5		
928	IP6K2	CG10082	10	FBGN0034644

VITA

Ananya Talikoti was born on September 5, 1997, in Raleigh, North Carolina and is an American citizen. She attended North Carolina State University (Raleigh, NC) from 2015 to 2019 and received her Bachelor of Science in Genetics. She joined Virginia Commonwealth University's Department of Human and Molecular Genetics and Dr. Mike Grotewiel's laboratory in the same department in 2019.