2018

The Power of GWAS: Leveraging Genome Wide Association Studies to identify novel regulators of autophagy in Drosophila melanogaster

Axelle Weeger
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Genetics Commons

Recommended Citation
https://lib.dr.iastate.edu/etd/16486

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
The power of GWAS: Leveraging Genome Wide Association Studies to identify novel regulators of autophagy in *Drosophila melanogaster*

by

Axelle Weeger

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Genetics and Genomics

Program of Study Committee:
Hua Bai, Major Professor
Clark Coffman
Diane Bassham

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this thesis. The Graduate College will ensure this thesis is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2018

Copyright © Axelle Weeger, 2018. All rights reserved.
DEDICATION

I would like to first thank those closest to me who have made this possible. Mom, Dad, Solène and Ariane, your love and support are what carried me through the toughest times in my life. You taught me that if I put my mind to something, I can reach any height. This is the result of that support.

Tyler, you were there when I thought I was never going to complete this thesis. When life called other people on their own paths, you decided to stick with me and move cross country. I can never thank you enough.

Allison, Thea, Emma, Majd, Angela, Christian, Kaitlin (to name only a few), you all made sure I stayed sane. That in itself is an achievement, so good job.

Finally, Annabelle, you’re never going to read this because you are a cat, but you need to be mentioned. Your fluffiness provided much needed comfort.

And to all of those I have not mentioned, for there are too many to count, you’ve helped me become who I am today.

I look forward to the rest of my life with everyone and seeing us all reach for our best selves.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>List</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>vii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>viii</td>
</tr>
<tr>
<td>CHAPTER 1. GENERAL INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>mTOR: The Upstream Integrator</td>
<td>2</td>
</tr>
<tr>
<td>Molecular Mechanisms of Autophagy</td>
<td>3</td>
</tr>
<tr>
<td>Autophagy, Disease and Aging</td>
<td>7</td>
</tr>
<tr>
<td>CHAPTER 2. THE POWER OF GWAS: LEVERAGING GENOME WIDE ASSOCIATION STUDIES TO IDENTIFY NOVEL REGULATORS OF AUTOPHAGY IN <em>DROSOPHILA MELANOGASTER</em></td>
<td>11</td>
</tr>
<tr>
<td>Abstract</td>
<td>11</td>
</tr>
<tr>
<td>Background</td>
<td>11</td>
</tr>
<tr>
<td>Results</td>
<td>11</td>
</tr>
<tr>
<td>Conclusions</td>
<td>12</td>
</tr>
<tr>
<td>Introduction</td>
<td>12</td>
</tr>
<tr>
<td>Methods</td>
<td>13</td>
</tr>
<tr>
<td>Fly Husbandry and Tissue Dissection</td>
<td>13</td>
</tr>
<tr>
<td>Overview</td>
<td>13</td>
</tr>
<tr>
<td>Fly Husbandry</td>
<td>14</td>
</tr>
<tr>
<td>Treatment and Dissection</td>
<td>15</td>
</tr>
<tr>
<td>Imaging</td>
<td>17</td>
</tr>
<tr>
<td>Image Processing and Quantification</td>
<td>17</td>
</tr>
<tr>
<td>Genome Wide Association Study</td>
<td>18</td>
</tr>
<tr>
<td>Overview</td>
<td>18</td>
</tr>
<tr>
<td>Analysis</td>
<td>19</td>
</tr>
<tr>
<td>Candidate Selection</td>
<td>20</td>
</tr>
<tr>
<td>Functional Validation Using RNAi</td>
<td>20</td>
</tr>
<tr>
<td>Overview</td>
<td>20</td>
</tr>
<tr>
<td>Fly Husbandry</td>
<td>21</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Schematic representation of the autophagy process within the cell.</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>Dissection of the dorsal fat body in <em>D. melanogaster</em>.</td>
<td>16</td>
</tr>
<tr>
<td>2.2</td>
<td>Variation of lysosome density phenotypes in the DGRP.</td>
<td>26</td>
</tr>
<tr>
<td>2.3</td>
<td>Variation of lysosome density phenotypes under basal conditions.</td>
<td>27</td>
</tr>
<tr>
<td>2.4</td>
<td>Variation of lysosome density phenotypes after starvation stress.</td>
<td>28</td>
</tr>
<tr>
<td>2.5</td>
<td>Schematic and visual representation of phenotypic variation throughout the</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DGRP.</td>
<td>29</td>
</tr>
<tr>
<td>2.6</td>
<td>Manhattan plot of Single SNP GWAS for lysosome density phenotypes.</td>
<td>31</td>
</tr>
<tr>
<td>2.7</td>
<td>Gene sets correlated with basal and stress induced lysosome density do not</td>
<td></td>
</tr>
<tr>
<td></td>
<td>overlap.</td>
<td>32</td>
</tr>
<tr>
<td>2.8</td>
<td>Lysosome density in S106-GS&gt;UAS-RNAi males following RNAi induction.</td>
<td>33</td>
</tr>
<tr>
<td>2.9</td>
<td>Lysosome density before and after induction of RNAi in S106-GS&gt;UAS-RNAi</td>
<td></td>
</tr>
<tr>
<td></td>
<td>males.</td>
<td>34</td>
</tr>
<tr>
<td>2.10</td>
<td>kin17, elk, Myo61F and crz regulate lysosome density.</td>
<td>34</td>
</tr>
<tr>
<td>2.11</td>
<td>qRT-PCR confirms gene knockdown in kin17, corazonin and elk.</td>
<td>35</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 2.1: List of primers for qRT-PCR knockdown verification..................................................24
Table B.1: DGRP Fly Stock........................................................................................................54
Table B.2: UAS-RNAi Fly Stock .................................................................................................58
ACKNOWLEDGMENTS

I would like to thank my committee chair, Hua Bai, for his mentorship and continued advice, as well as my committee member Clark Coffman and Diane Bassham for their support throughout my research. Your comments helped me make this what it has become.

In addition, I would like to thank all of my friends and colleagues, as well as the department faculty and staff for making the time I spent at Iowa State University so rich and varied in experiences. I have learned so much and could not have done so without you. Thank you for making this possible, for the coffee breaks, the midnight emails, the long conversations, the laughter and the joy of achievement. I could not have done this alone.
ABSTRACT

Autophagy was first isolated as a marker of starvation but has since then been identified and implicated in a variety of processes both pathological or not. Lysosome density within a cell has long been used as a marker of activity of this pathway, as it represents the final location of degradation. Under normal conditions, lab strains of *Drosophila melanogaster* exhibit a low lysosome density considered to be universally observed. When challenged with starvation, this activity is expected to increase 2-fold. We do not know if these characteristics are representative of natural populations, or the regulatory factors used to control the lysosome density within tissues.

By phenotyping 178 lines of the Drosophila Genetic Reference Panel for their lysosome density under resting and starved states, we were able to identify a previously uncharacterized variance in lysosome density phenotypes. Resting variation included multiple lines displaying an unusually dense lysosome phenotype previously not observed in lab strains. Furthermore, a Genome Wide Association Study performed on this data revealed that the gene sets associated with variance in resting and starved lysosome density were non-overlapping. Finally, RNA interference studies identified 4 previously non-implicated regulators of lysosome density; Myosin61F and corazonin as negative regulators and kin17 and elk as positive regulators.

This analysis presents a more complete picture of natural variance of lysosome density, and thus autophagy activation, as well as identifies novel regulators that may be responsible for controlling this observed variation.
CHAPTER 1. GENERAL INTRODUCTION

Cells are delicate living machines, requiring a slew of different pathways to function properly. One of those pathways can be generally referred to as the autophagy pathway, although it comprises multiple types of actions. From the Greek "auto"-self and "phagy" –eating, autophagy allows cells to recycle damaged organelles and unused or malfunctioning proteins into their component parts, thus providing energy back to the cell (Codogno et al. 2011). This pathway is stimulated in cells as a housekeeping mechanism throughout their lives and can also be up-regulated in situations of stress to boost internal access to nutrients.

The inception of autophagy characterization was in rat liver tissue when Dr. Christian De Duve identified vesicles termed "dense bodies", containing a very high concentration of various proteases and hydrolases (De Duve et al. 1955). This structure later termed "lysosome" is one of the central players in the autophagic process. This organelle is responsible for the breakdown of all material brought to it and is critical for proper function of the cell. The next landmark discovery happening in mouse kidney cells, where a secondary vacuole was identified as containing protein aggregates as well as various organelles including mitochondria (Arstila & Trump 1968). The presence of this specialized double membrane-bound vacuole, later termed "autophagosome", could be induced in the cell using stress assays, at which point it was established that nutrient stress was a very strong predictor for the presence of this structure. It was observed that while glucagon could induce its presence, insulin would repress its formation in the cell, thus cementing the function of these structures as being tightly linked with the nutritional state in the cell (Deter et al. 1967). The autophagosome would then be shown to fuse with the lysosome, losing one of its membrane layers in the process and allowing its cytosolic
contents to be degraded by the enzyme-rich media of the lysosome. This stage is coined "autolysosome". Once all content has been degraded they will then be excreted back into the cytoplasm (Arstila & Trump 1968).

This general pathway is regulated by a wide array of dynamic protein interactions each tightly regulated by sets of autophagy genes. Early nomenclature was diverse, with some genes identified by the associated phenotype, or model organisms in which the discovery was made. This was later unified under the ATG (autophagy gene) umbrella, which is now used to refer to most of the genes involved in this pathway (Klionsky et al. 2003). The first of those genes was isolated in yeast, in a landmark study by Dr. Yoshinori Ohsumi, which identified 15 mutants in yeast that failed to survive long periods of starvation (Tsukada & Ohsumi 1993). This study also established the crucial role of autophagy specifically under nutritional challenge, as these mutants did not exhibit growth defects under nutrient-rich conditions (Suzuki et al. 2011). In the years since the ranks of characterized ATG genes have grown to more than 30 genes, and this pathway has been characterized in numerous other model organisms. Subsequent studies characterized both function and interaction of key proteins required for the proper function of the pathway (Ohsumi 2014).

**mTOR: The Upstream Integrator**

One of the major signaling controller pathway associated with autophagy is the mTOR pathway, a multifunctional signaling node downstream of the insulin pathway (Noda & Ohsumi 1998). The mechanistic target of rapamycin, or mTOR for short, forms two distinct complexes within the cell, each with separate functions (Helliwell et al. 1994). Here we will focus on
mTOR Complex 1 (mTORC1) which directly interacts with autophagy. This complex is composed of mTOR itself and three associated proteins (Raptor, GbetaL and Deptor) and integrates signal from amino acid concentration, energy levels, stress cues and growth factor levels to ultimately promote cellular growth (Beck & Hall 1999; Kim et al. 2002; Tokunaga et al. 2004). It is highly activated by the Rheb protein, a small GTPase, which can be found embedded in the lysosome membrane (Kundu 2011). When active, mTORC1 will repress autophagy by phosphorylating Atg13, a subunit of the Atg kinase complex, thus stopping the activation cascade (Chang & Neufeld 2009).

When nutrient stress is felt, mTOR is unable to phosphorylate Atg13 which triggers the autophagy activation cascade (Chang & Neufeld 2009). The specific mechanism of regulation is still unknown, but localization of mTORC1 provides a potential mechanism. mTORC1 has been shown to localize to the lysosome membrane when amino acid concentration is high, which allows it to complex with its co-activator Rheb (Zoncu et al. 2011). This highly activates mTORC1, thus repressing autophagy when nutrients are plentiful. Another mechanism of autophagy activation is also hypothesized to be due to colocalization of mTOR with LC3/Atg8 on the autophagosome membrane, although its specific function at this stage is unknown (Kabeya et al. 2000).

**Molecular Mechanisms of Autophagy**

Like all cellular processes, autophagy relies on complex protein interaction to carry out its functional goal within the cell. Proper regulation of each component is crucial to ensure that the pathway activates correctly and does not yield adverse effects within the cell. In addition,
different types of autophagy (macro, micro or chaperone-mediated) have different mechanisms for delivering cargo to the lysosome.

![Autophagy Process Diagram](image)

Figure 1.1: Schematic representation of the autophagy process within the cell. Cargo is engulfed and sequestered in the autophagosome, then transported to the lysosome. Fusion of the two vacuoles allows the cargo to be broken down by the enzyme-rich media of the lysosome. Products are then excreted back into the cytoplasm.

Macro-autophagy is a complex system requiring the formation of an autophagosome to engulf and transport cargo to the lysosome for degradation. Each step in the process is tightly regulated by over 30 associated genes. This type of autophagy may be specific to targets marked for degradation, or nonspecifically activated as a response to nutrient stress.

Of particular interest, this process also includes the recycling of damaged mitochondria (mitophagy) which is very important to safeguard the cell from damage (Otto et al. 2003).
While this process can occur non-specifically if the cell is experiencing extreme nutrient stress, it is primarily targeted to limit cytotoxic conditions, clear damaged organelles, and proteins and recycle them into usable components.

Micro-autophagy is used to refer to a process lacking any secondary chaperone structure, where the lysosome itself will engulf cytosolic elements for direct degradation (Sahu et al. 2011).

Chaperone-mediated autophagy does not require an autophagosome, but it does require specific chaperone proteins. These will recognize a motif similar to the KFCRQ pentapeptide and recruit those proteins for degradation in the lysosome using LAMP receptors as membrane interfaces (Bejarano & Cuervo 2010). Here, we will focus on macro-autophagy, hereafter referred to as simply autophagy.

Autophagy is mediated by a host of ATG genes and other related proteins, the first of which, ATG1 (a serine/threonine kinase) was identified in 1993 (Tsukada & Ohsumi 1993). Since that first identification, dozens of other critical Atg genes have been identified.

Autophagy is located downstream of the mTOR pathway, a key integrator of nutrient sensing signals in the organism. In a nutrient rich environment, mTOR will induce hyperphosphorylation of Atg13, thus preventing it from complexing with Atg1 (Chang & Neufeld 2009). A secondary feedback loop reinforcing the activity of S6K (a downstream effector of mTOR) also works to inhibit the formation of the Atg1/Atg13 complex (Lee et al. 2007). In its inactive, hyperphosphorylated form, Atg1 will allow S6K to be phosphorylated, thus activating it and promoting cell growth. This, in turn, will promote mTOR activation, and keep both Atg1 and Atg13 inactive (Scott et al. 2004).

In a nutrient-depleted environment, Atg13 will very quickly dephosphorylate and complex with Atg1 and Atg17, thus forming the Atg1/Atg13 initiation complex. As a result, S6K
will be inhibited, halting cell growth. While this interaction is well understood, it's hypothesized that S6K has other regulatory functions in the autophagy pathway, but these interactions are currently not well understood (Lee et al. 2007).

Once the initiation stage has been reached, and the Atg1/Atg13 complex is active within the cell, beclin1 and LC3 proteins are recruited to the phagophore membrane, and start embedding themselves (Xie et al. 2008). LC3, also referred to as Atg8, is a ubiquitin-like protein that will undergo further post-translational modifications as the autophagosome membrane expands (Pankiv et al. 2007). Multiple Atg genes act in succession to modify LC3/Atg8 and activate it, this step is commonly considered to be the most complex and potentially error-prone, as gene expression of all participating proteins needs to be tightly regulated. Atg4, a cysteine protease, is the first to interact, cleaving the C-terminal location and exposing a glycine. This activation stage is referred to as LC3-I and needs to undergo further modifications to reach the final active stage called LC3-II (Tanida et al. 2004). Concurrently, Atg 12 and Atg5 will also form a secondary complex, with the help of Atg10 and Atg7. The Atg12/Atg5 complex will require one more interaction (Atg16) before the mature complex is integrated into the growing phagophore membrane (Mizushima et al. 1999). This complex, although not essential, will then facilitate the final modifications of LC3-I by adding PE (phosphatidylethanolamine), at which point Atg8 is generally referred to as LC3-II.

The addition of PE allows LC3/Atg8 to be inserted into the autophagosome membrane as it grows, encapsulating cargo (Papisnki & Kraft 2014). The conjugation status of Atg8 is recognized as a strong marker of autophagy, with free cytosolic Atg8 being a marker of nutrient-rich conditions, and the presence of lipidated Atg8 being strongly associated with an active autophagy pathway.
Once the autophagosome is fully formed, its fusion to the lysosome is mediated by multiple agents, including LAMP-1 a lysosomal transmembrane protein (Saftig et al. 2008). At this step, we now refer to the fused structure as an autolysosome, and degradation of both cargo and original autophagosome membrane will begin. Of particular note is the cathepsin class of proteases, which possess a cysteine in their active group and thus may be susceptible to oxidative modifications and redox regulation (Cullen et al. 2009).

The resulting pool of free amino acids is then released into the cytosol to be reused by the cell in metabolic processes (Yang et al. 2006).

**Autophagy, Disease and Aging**

At its core, autophagy is a recycling process, but this basic mechanism has implications in multiple cellular fate outcomes. By lessening the effect of malfunctioning components of the cell, it promotes optimal function and prolongs the functionally healthy life of the cell. It can also, through intense and nonselective degradation of cytosolic content, trigger apoptotic conditions in the cell (Marino et al. 2014). Well regulated, autophagy allows the cell flexibility when challenged by environmental triggers and allows it to survive conditions that would otherwise destroy the organism. On the other hand, the pathway can also damage the cell, by nonspecifically destroying essential viable components. This wide array of consequences arising from a singular process contributes to the interest in autophagy associated with various disease states. Here we will review its specific known interactions within the domain of aging, and implications for other diseases that arise with age.
Because of its action as a regulator of damaging components within the cell, autophagy has been theorized to have an effect on the process of aging, namely that an optimally functional autophagy machinery could prolong lifespan (Vellai 2009; Markaki & Tavernarakis 2011; Madeo et al. 2010). Aging is a complex multi-partite phenomenon, but many of its markers can be considered to be affected by autophagy. Literature shows that in multiple tissues, aging is linked with a down-regulation of key autophagic genes although the process by which these genes are down-regulated is still unknown (Bejarano et al. 2014). This would lead to lessened clearance ability and overall increased stress burden on the cell.

In particular, it has been shown that aging cells are more likely to misfold proteins, and thus accumulate nonfunctional protein aggregates at a faster rate (Ravikumar et al. 2010). This, coupled with a downregulation of autophagy can contribute to some of the aging phenotypes observed, mainly increased cytotoxicity. Since the source of the down-regulation is still unknown, studying the triggers of autophagy and its regulatory pathway is critical to try and safeguard a strong activation as the organism ages. In addition, as multiple pathways are integrated to activate autophagy, it is also very important to consider their dysregulations as time passes, and how this might affect activation of autophagy.

In addition, as tissues age, mitochondria become less effective and result in a higher occurrence of reactive oxygen species (ROS) being produced as the mitochondrial machinery becomes deregulated. A small amount of ROS is known to be crucial to the proper functioning of a cell, and its complete elimination would result in cell death, but large quantities are highly damaging to the cell and need to be cleared away (Sena & Chandel 2012). In this case, the action of autophagy as a clearance mechanism for dysregulated mitochondria could prevent the
accumulation of ROS to damaging levels, thus once again safeguarding the cell's integrity (Twig et al. 2008).

Considering the current consensus in the literature regarding the cytoprotective nature of autophagy, and its association with common aging phenotypes, there is a drive to leverage the autophagy machinery to lengthen both health and lifespan. Specific Atg genes have been shown to function as lifespan extenders, including a milestone study in C. elegans (Hansen et al. 2008). Other studies have shown that tissue-specific overexpression of Atg7 in the liver can improve the tissue's function when challenged by obesity conditions (Zhang & Cuervo 2008). Finally, a general overexpression of Atg5 in mice was shown to increase lifespan by 17% (Pyo et al. 2013). While these studies are encouraging, they do not yet cross into the domain of human therapeutics. All of these discoveries are limited to non-human model organisms, which does not guarantee that processes would carry over to different organisms. It is also critical to better understand the regulators of autophagy if we wish to apply that knowledge to human medicine. Most currently, therapies targeting known upstream activators of autophagy such as rapamyacin treatments targeting and down-regulating mTOR are promising leads into better managing autophagy as a factor of aging, and diminishing tissue damage (Harrison et al. 2009; Bjedov et al. 2010).

There have been some other leading studies in the field of longevity showing that one of the clinical ways to increase longevity is calorie restriction. In laboratory models, mice fed a restricted diet (1/3 fewer calories than a standard diet) saw an extension of their lifespan, an increase in lean body mass, and an overall healthier lifespan as compared to mice fed a standard caloric diet (Wohlgemuth et al. 2007).
An overall decrease in calorie causes insulin signaling to drop, thus impacting mTOR dependent pathways including autophagy (Blagosklonny 2010). It is hypothesized that the calorie restriction activates autophagy, leading to a cell that more actively cleanses itself and recycles internal components. This more active pathway leads to reduced cytotoxicity as the animal ages and retards loss of tissue homeostasis at the organismal level.

While this clinical information is of great interest in the field of longevity, it becomes hard to apply in human settings due to the restrictions of the diet. In addition, while this demonstrates a strong link between insulin metabolism and aging, the specific interactions that lead to this longevity are poorly understood.

Key factors that integrate cellular information and regulate these various pathways, including autophagy, are affected in ways we do not yet fully understand. However, studies like these that show a strong correlation between mTOR activity, autophagy and longevity are the basis for searching and understanding the missing factors that regular cellular mechanisms. By elucidating which factors are currently missing, we could bring this type of therapy to the greater human population and leverage its benefits for the human race.
CHAPTER 2.  THE POWER OF GWAS: LEVERAGING GENOME WIDE ASSOCIATION STUDIES TO IDENTIFY NOVEL REGULATORS OF AUTOPHAGY IN DROSOPHILA MELANOGASTER

Axelle Weeger and Hua Bai

Iowa State University, aweeger@iastate.edu. Corresponding author.

Iowa State University, hbai@iastate.edu

Modified for a manuscript to be submitted to BMC Genomics

Abstract

Background

Autophagy has long been studied, and its activation during nutrient stress is well characterized. What is less studied, are whether the conclusions on pathway behavior in laboratory strains translate to wild populations. We propose that phenotyping a panel created from wild-caught Drosophila melanogaster will yield valuable data on the variability of autophagy both pre and post nutrient stress. Furthermore, we propose that applying a Genome Wide Association Study pipeline to this phenotypic data will allow us to identify and validate candidate SNPs associated with high variance in this population, and thus identify novel regulators of autophagy.

Results

Using 178 lines of the DGRP, we were able to demonstrate that wild type D. melanogaster populations exhibit more variability in autophagy that previously characterized. In addition, the sets of genes most associated with variability in pre and post stress datasets did not overlap significantly. Finally, RNAi validation identified 4 novel regulators, kin17 and elk being negative regulators while Myosin61F and corazonin are positive regulators.
Conclusions

The autophagy behavior in wild type laboratory strains of *D. melanogaster* does not adequately model the phenotypic variance found in natural populations. This study also indicated that regulators involved in basal autophagy regulation may be different from those that are responsible for post-stress activation. Finally, the full regulatory network of autophagy is still not fully delineated, as new regulators are still being identified and characterized. We present here an effective method to investigate complex phenotypes in natural populations and identify minor regulators responsible for this variation.

Introduction

In times of starvation, the body will do everything it can to maintain homeostasis. When nutrients cannot be brought to the cell from outside, it will instead look inwards to its own components as a resource. Autophagy, literally “self-eating”, will be triggered to start non-specifically engulfing cytoplasmic components and degrading them. This process allows the cell to maintain essential function by recycling itself [1]. Autophagy as a mechanism has already been widely studied, and the essential pathways and regulators have been characterized [2]. What has been less studied, are the subtler regulators of the process. We understand autophagy to be active at a low level in well fed organism, acting as a housekeeping system to help maintain homeostasis. This basal level is greatly increased when nutrient stress is felt, but we do not know as of yet if these basal and stress levels are representative of natural variation [3].

To investigate such potential variance, and its regulators, we used an established wild type panel, the Drosophila Genetic Reference Panel (or DGRP) and its associated Genome Wide
Association Study (GWAS) pipeline [4]. Composed of over 200 *Drosophila melanogaster* isogenic lines, the panel was established using wild caught specimens. As such, it provides us with an unparalleled resource to study naturally occurring genetic variation. By its very nature, the panel is not concerned with extreme, lethal variation, but instead can be used to explore the inherent pleiotropic nature of complex phenotypes, such as characterizing the variance of autophagy before and after nutrient stress.

We will be using 178 lines of the DGRP to investigating an end-stage phenotype associated with autophagy, the lysosome density within the cell [5]. This indirect assay has been used extensively to characterize autophagy phenotypes, as such we consider it to be a valuable reporter. The GWAS protocol will allow us to identify highly correlated SNP candidates, which will then undergo RNA interference validation of function.

This particular combination of tools allows us a previously unavailable view of autophagy in natural populations, as well as an opportunity to identify previously uncharacterized regulators that influence the regulation of autophagy.

**Methods**

**Fly Husbandry and Tissue Dissection**

**Overview**

Flies from the Drosophila Genetic Reference Panel (DGRP), created by the MacKay lab as a platform to run Genome Wide Association studies (GWAS), were used to record phenotypic variation associated with autophagy (Mackay et al. 2012).
Progenitor flies were collected by the MacKay lab at the Raleigh, NC Farmer's market. Progenitor gravid females were inbred via full sibling mating for 20 generations, to produce the viable lines currently included in the panel. Current lines included in the panel are isogenic and fully inbred.

All obtained lines were subsequently sequenced using Illumina technology, and unique Single Nucleotide Polymorphisms (SNPs) were mapped back to each line. Globally, this sequencing effort yielded over 4 million unique SNPs within the panel, which can be queried using a Genome Wide Association tool developed specifically for this panel.

The DGRP panel was obtained from Bloomington Stock Center, where the full panel is maintained.

**Fly Husbandry**

After quarantine and screening for mites, 20 females and 5 males of selected DGRP lines were seeded into vials containing 6ml of 3% yeast food. The standard recipe for fly food (1 batch) contains 850ml of water, 7.9g of Agar (USB, Cat#10654), 25g of yeast (Genesee Scientific, Cat#62-108), 52g of yellow cornmeal (Fisher Scientific, Cat#NC9349175) and 110g of granulated sugar (C&H commercial). Tegosept (Methyl 4-hydroxybenzoate, Genesee Scientific Cat#20-258) is also added to prevent molding (2.38g, dissolved in 9.2ml of ethanol).

Flies were allowed to mate and lay eggs for 48h before transfer. Three biological replicates were created.

Flies were the kept in a 25°C incubator, with 60% moisture during development. As biological replicates reached eclosion, emerging adults were collected by date of eclosion, to allow uniform aging across genotypes.
Flies were aged for five (5) days, with both sexes present to reduce instances of death by drowning in the exuded moisture.

At day five (5) of age, males and females were separated, and treatment was set up. Males were chosen for the experiment, as their starvation time is faster than females.

**Treatment and Dissection**

Ten (10) flies were in the control group and given normal food for 16 hours.

Ten (10) flies were in the treatment group were starved for 16 hours.

Starvation: 1.5mL of 1X PBS was placed on an absorbent wipe folded and placed at the bottom of a vial. This prevented desiccation and provided minimum electrolytes while restricting all caloric intake.

After the 16-hour treatment, flies were collected for dissection.

Flies were knocked out using Flynap (Carolina Biological Supply, # 173010).

Six (6) flies for each treatment were placed in Vaseline dorsal side down and dissected.

The thorax and head were removed, along with the last segments of the abdomen (Fig. 2.1, cuts 1 and 2). The carcass is then transferred to a drop (50 uL) of 1X PBS to prevent desiccation. Entrails were then gently pulled out of the carcass, allowing more freedom of movement to create two lateral cuts (Fig. 2.1, cuts 3 and 4) and remove the ventral panel. This exposes the relevant tissue, the fat body.
Figure 2.1: Dissection of the dorsal fat body in *D. melanogaster*. Cuts 1 and 2 separate the abdomen and remove internal connections of the gut and other internal organs to the cuticle. Cuts 3 and 4 are lateral and allow the removal of the ventral cuticle. This allows a clear visualization of the dorsal fat body.

The fat body was selected as a testing organ because of its physiological similarity to human adipose tissue. This tissue is responsible for energy storage in the fly, in the form of fat droplets, which can then be degraded when nutrient stress is detected. This tissue is very reactive to nutrient stress, and as such is a very good model organ to study the variability of autophagy activation in different genetic backgrounds.

Carcasses were then incubated with 100nM LysoTracker Red DND-99 (Invitrogen, #L7528) at room temperature for 3 minutes.

LysoTracker was selected as a dye because of its affinity for acidic organelles within the cell. It has been shown to target lysosome reliably and efficiently, allowing for a consistent staining across all dissections. Some staining of the endosome is also possible, but LysoTracker is considered to be most accurate at staining lysosome and autolysosome within cells [6].

During incubation, a glass bridge was created on the slide, to prevent carcasses from getting crushed by capillary action.
**Imaging**

Carcasses were mounted to a slide using 1X PBS and capillary action to maintain position.

The fat body stain was then imaged using an Olympus BX51W1 fluorescent microscope fitted with a TRITC filter and imaging data was collected using the associated Olympus cellSens software.

An overview of the whole carcass was taken at 10x magnification, and a 19-slice (.75um offset) z-stack was taken at 20x magnification for each carcass. 5 replicates were imaged for each treatment.

**Image Processing and Quantification**

After acquisition, images were processed using a constrained iterative deconvolution process. Image data was quantified using the following protocol:

Three (3) circular areas were isolated on the z-stack, each measuring 100nm in diameter. The location of the areas of sample was as random as the sample allowed, with an aim for optimal coverage of representative areas of the sample and minimizing potentially confounding areas (locations of abnormally high tissue staining, pericardial cells, cardiac muscle, etc).

The cellSens software was then utilized to count the total number of the objects present within these three measuring areas. This measure includes all 19-vertical stack for each sample, across the three regions of interest.
Genome Wide Association Study

Overview

Genome Wide Association Studies (GWAS) have established themselves as solid and competent models to pull putative information from large sequencing experiments. Through statistical these studies allow the selection of candidate SNPs most correlated to a specific phenotype of study. This is particularly powerful, as many phenotypes are incredibly complex and multigenic, and thus cannot be isolated through single mutant screens. In the context of *D. melanogaster*, much of the genome remains non-annotated, with a large portion of proteins whose function remains cryptic.

GWAS mobilizes the increasing power of sequencing as well as relatively simple statistical methods to produce vast amounts of data. These studies rely on Single Nucleotide Polymorphism (SNP) to better understand the influence certain genes have on overall phenotypes. Each genomic location can either be a major SNP, which occurs in more than 50% of a population, or a minor SNP. Generally speaking, when referring to a SNP, the literature is referring to the minor SNP. Once a phenotype has been identified and quantified within a population, mass sequencing will reveal which SNP is present throughout the entire genome. This information itself can be valuable, but it is not the end goal of a GWAS.

After the genomic information is obtained, a massive Analysis of Variance (ANOVA) test is performed. Some other methods are also used depending on the study's hypothesis, including contingency tables and logistic regression. A general model of interaction needs to be selected, in regard to the epistatic relationship between major and minor SNP. This model also needs to consider whether allelic effects are multiplicative or additive, a factor that can affect how correlated a genotype is to phenotype. Factors that are known to influence the phenotype also need to be taken into account when designing the model, such as age, weight,
sex and many others. Taking all of these parameters into consideration, the GWAS then runs a corrected ANOVA test on every single base pair of the sequenced genomes and tests it against contribution to the phenotype of choice. A significance level of 0.5 is the most commonly used threshold of significance, but due to the extremely large number of test, a statistical correction is usually employed to minimize false positive hits. 5% of 500,000 tests is still a very high level of false positive results. To that effect, a Bonferroni correction is most commonly used, which decreases the alpha value by dividing it by the number of tests performed. This usually decreases the number of false positives down to acceptable values.

The MacKay lab established a specific pipeline using Partial Lease Regression (PLS) models to most accurately call relevant SNPs, which can be found in the supplementals of their landmark paper on the establishment of the DGRP [7].

After the analysis is complete, a list of SNPs most likely to be associated with the phenotype is produced, with smaller p-values indicating a higher likelihood of an effect. It is important to note that all the GWAS provides are correlations, and that further testing is absolutely necessary to validate whether a particular gene is involved in the control of a particular phenotype.

**Analysis**

Once the autophagy phenotyping data was collected for each genotype two data sets were created.

One set, termed the "Fed" dataset, contained the raw untransformed LysoTracker count for all control treatments of each genotype. This represents the level of autophagy/lysosome activity in a well-fed, healthy fly.
The second set, termed "Foldchange" contained data representing the foldchange in LysoTracker count from the fed and the starved treatment for each genotype. This represents a measure of autophagy activation, following stress (starvation).

Each data set was independently submitted to the GWAS software developed by the McKay lab which can be found at the following web address: http://dgrp2.gnets.ncsu.edu/. Data sets were then visualized using Manhattan plots generated using the code in Appendix A.

**Candidate Selection**

Upon recovering GWAS results, individual SNPs were screened for potential candidate genes that are involved in autophagy regulation.

The screen was conducted in two parts, one purely mathematical in nature, the next with a subjective value assigned to each SNP. SNPs were chosen if either of the following criteria was met:

- Single mixed p-value was 10E-7 or lower.
- SNP was located in a location that disrupted proper protein function (Non-synonymous, start/stop codon)
- Literature evidence indicated the SNP-associated genes might be relevant to autophagy control, regardless of the two previous criteria.

**Functional Validation Using RNAi**

**Overview**

Following candidate selection, transgenic RNAi lines were ordered for the multiple candidate genes to verify their roles in autophagy control. These lines were crossed to a gene-switch fat-body driver (S106-GS) to induce knockdown of the gene of interest only in adult
fat body. The S106-GS line is a modified Gal4 driver that enables targeted temporal activation of RNA interference models using drug inducers [8]. The GAL4 gene is fused with the activation domain for p65 and a ligand-binding domain for the human progesterone receptor. In the presence of the inductor, antiprogestin drug RU486, also known as Mifepristone (Cayman Chemical, CAS #84371-65-3), the fusion protein is able to bind to the UAS domain and induce transcription. RU486 is also used in the medical field as an abortifacient drug.

In this way, the Gal4 driver associated with the fat body tissue is not constitutively active throughout the development of the organism and is instead dormant until activated at the desired time point. This is done to limit mortality in early development, as well as general organismal stress that would influence the autophagy process in the organism. Using this system, we can be better assured that the phenotype observed is a direct consequence of the removal of the target protein from the individual and not a consequence of developmental issues due to the lack of the target protein.

**Fly Husbandry**

Flies were reared in standard conditions previously detailed.

Mating pairs of ten (10) virgin females of S106-GS-gal4 line and five (5) UAS-RNAi males were established.

Mating pairs were allowed to lay eggs for 24 hours, before being transferred to a new vial to produce a replicate.

After eclosion, males were separated from females before treatment. Lines that contained a balancer gene were carefully screened for the absence of the balancer trait in the experimental progeny.
Treatment- Seven to ten (7-10) males were fed on food containing 200 µM RU486 for 6 days, with fresh food being offered every 2 days.

Control- Seven to ten (7-10) males were fed on food containing 0 µM RU486 for 6 days, with fresh food being offered every 2 days.

At day six (6) of age, males were dissected and imaged using previously established protocol.

**Functional Validation of Knockdown via qRT-PCR**

**Overview**

To ensure that all previous RNAi testing could be trusted, we additionally performed a qRT-PCR assay to verify that the UAS-RNAi lines we were using were indeed capable of inducing knockdown of the gene of interest in. A whole body driver (Da-GS-gal4) was used in this instance, to maximize potential expression and confirm the viability of the RNAi lines being used.

**Fly Husbandry**

Flies were reared in standard conditions previously detailed.

Mating pairs of ten (10) virgin females of Da-GS-gal4 line and five (5) UAS-RNAi males were established.

Mating pairs were allowed to lay eggs for 24 hours, before being transferred to a new vial to produce a biological replicate.

After eclosion, males were separated from females before treatment. Lines that contained a balancer gene were carefully screened for the absence of the balancer trait in the experimental progeny.
Treatment- Four (4) males were fed on food containing 200 µM RU486 for 4 days, with fresh food being offered every 2 days.

Control- Four (4) males were fed on food containing 0 µM RU486 for 4 days, with fresh food being offered every 2 days.

**RNA Extraction**

Protocol adapted from Life Technology Cat #15596026

RNA was extracted from the treatment and control males, using the full body. 4 males were added to a sample tube containing 500ul of Trizol reagent (Life Technology #15596026) and a 5mm stainless steel bead (Qiagen #69989). The samples were then homogenized using a Tissuelyzer II (Qiagen, Manufactured by Retsch), for 2 minutes at a frequency of 30/s. Samples were incubated at room temperature (RT) for 10 minutes (min), before adding 100 µl of chloroform per sample and resting an additional minute. Samples were then vortexed, (10 seconds, twice) and further incubated for 8 min at RT. They were then centrifuged at 12000 g, at 4°C, for 15 minutes.

The supernatant (~250 µl) was removed to a fresh RNase free tube, and we added 250 µl of isopropanol. After a gentle shaking, samples were incubated at RT for 15 min and centrifuged at 12,000g, at 4°C, for another 15min.

All samples were decanted and washed with 700 µl 75% RNase free then centrifuged at 7,500g at 4°C for 5 minutes. This step was repeated twice then the pellet was allowed to air dry for 15 minutes. We re-suspended each sample in 17µl Nuclease-free H2O.

Samples were then treated using Turbo DNA-free DNase treatment (Ambion, cat # AM1907). Each sample received 0.1 volume 10X TURBO DNase Buffer and 1 uL TURBO DNase before incubation at 37°C for 30 minutes.
Samples were then centrifuged at 10,000g at 4°C for 2 minutes to pellet DNase inactivation reagent, and the supernatant was transferred to a fresh tube. RNA concentration was assessed using Nanodrop system.

**cDNA Synthesis**

The reaction was set up as follows:

4 ul 5x qScript reaction mix, 1 ul qScript reverse transcriptase (QuantaBio, #101414), 15 ul RNA + dH2O (total volume 20ul). RNA volume was adjusted to contain 2ug of RNA.

Reaction was run in PCR machine for one cycle (5 min/22°C, 30 min/42°C, 5 min/85°C, Hold at 4°C).

The total volume (20ul) of cDNA was diluted in 80ul of RNase free H2O.

**Quantitative qRT-PCR**

Table 2.1: List of primers for qRT-PCR knockdown verification.

<table>
<thead>
<tr>
<th>DNA Sequence, Forward and Reverse</th>
<th>Primer Name</th>
<th>Associated Gene</th>
<th>Position</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCGAGGCGCTGTCTATCGCA</td>
<td>Crz-mRNA F</td>
<td>Crz</td>
<td>291-309</td>
<td>NM_079626.3</td>
</tr>
<tr>
<td>TGGGATGGGCGCTGTTTT</td>
<td>Crz-mRNA R</td>
<td>Crz</td>
<td>404-387</td>
<td></td>
</tr>
<tr>
<td>ACAAAAAAGCGGTTGTCCCTGG</td>
<td>kin17-mRNA F</td>
<td>kin17</td>
<td>975-994</td>
<td>NM_140955.4</td>
</tr>
<tr>
<td>ACCGTCTCCAAATGAGCCTTTGA</td>
<td>kin17-mRNA R</td>
<td>kin17</td>
<td>1079-1059</td>
<td></td>
</tr>
<tr>
<td>GTGATAGCGCTGCCCTTTTG</td>
<td>elk-mRNA F</td>
<td>elk</td>
<td>646-665</td>
<td>NM_001299624.1</td>
</tr>
<tr>
<td>TCCAGATTGTGCCCCTCACC</td>
<td>elk-mRNA R</td>
<td>elk</td>
<td>784-765</td>
<td></td>
</tr>
<tr>
<td>GTGGTGCGCTCAAATGGGAGG</td>
<td>Myo61F-mRNA F</td>
<td>Myo61F</td>
<td>883-902</td>
<td>NM_167872.2</td>
</tr>
<tr>
<td>GCCATTGAGCCCATCTGTCA</td>
<td>Myo61F-mRNA R</td>
<td>Myo61F</td>
<td>1023-1004</td>
<td></td>
</tr>
<tr>
<td>AAGAAGGGCACCAAGCCTTCATC</td>
<td>RPL32-mRNA F</td>
<td>RPL32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCTGTTCGATACCCCTTGGCCTT</td>
<td>RPL32-mRNA R</td>
<td>RPL32</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Primer mix (10pmol/ul working concentration) was created as follows for each set of primers used:

10ul of 0.1 nmol/ul P-forward, 10ul of 0.1 nmol/ul P-reverse, 80ul of nanopore H2O.

The reaction mixture for each sample contained 20ul final volume
Reaction master mix (1 sample):

1ul diluted primer set; 10ul 2X SYBR green mix; 7ul dH2O.

We added 2ul of cDNA sample to each designated well, checking all wells to ensure samples are in the appropriate well. We used a riboprotein (RPL32) primer as the housekeeping gene for each reaction. Plates were sealed and centrifuged in plate centrifuge for 20 seconds.

The reaction protocol was run in QuantStudio 3 (ThermoFisher, #A28136), with settings set to comparative CT (ΔΔCT) and SYBR green as our fluorescent dye. In addition, reaction volume was 20ul, and the cover temperature was set to 105.0°C.

The qRT-PCR was run at 50.0°C for 2 minutes, 95.0°C for 10 minutes, followed by 40 cycles of 95.0°C/15 sec; 60.0°C/60 sec until the final dissociation step.

For each ΔCT value, all ΔCT values were selected and individually divided by 20. We then calculated the ΔΔCT by using the formula $2^{-\Delta\Delta CT}$. The reference for each sample is the ΔΔCT value from the RpL32 equivalent well.

**Statistical Analysis**

Phenotypic data used in the GWAS was obtained by averaging the biological replicate measures for each genotype, so that each line was associated with two unique phenotypic data points (one for each dataset).

Statistical methods used were performed in Prism Software (GraphPad).

When referring to a simple t-test, the following parameters were applied. Two groups of data were compared using a standard, unpaired Welsh’s t-test comparison. For this method, significant p-value was set at 0.05.
When referring to an ANOVA test, the following parameters were applied. Data groups were compared to the control group, and a Bonferroni correction was applied to correct for the large number of comparisons being performed (>10). The significant p-value was also set at 0.05. All graphs obtained show mean ± SEM.

**Results**

**Lysosome Density Phenotypes are More Variable in Natural Populations Than in Laboratory Strains.**

![Graph showing variation of lysosome density phenotypes in the DGRP. Basal lysosome density phenotype across the panel is shown in A, while the variation lysosome density due to response to stress (quantified as a foldchange) is shown in B.](image)
When looking at the phenotypic distribution across the panel, we observed a normal distribution across all lines in both the resting and stress datasets (Fig.2.2-A & B).

Figure 2.3: Variation of lysosome density phenotypes under basal conditions. Line 177 A had the lowest basal activation, resulting in nothing but background stain. Comparatively, line 21 B contained a high density of lysosomes, as evidenced by the distinct punctae.

Notably, Lines 177 and 21 (Fig.2.3) exhibited the most extreme phenotypes (lowest and highest baseline, respectively) for basal lysosome density.

Lines 129 and 223 (Fig.2.4) exhibited extreme changes in lysosome density phenotypes after starvation (lowest and highest foldchange, respectively).
Figure 2.4: Variation of lysosome density phenotypes after starvation stress. Line 129 A&B exhibits a complete removal of lysosomes from the tissue following starvation. Line 223 C&D exhibits a strong increase in lysosome density following starvation.

When testing the two datasets for correlation between the two observed phenotypes, we found no significant correlation (Pearson correlation: r = 0.127). We cannot predict the behavior of a line based on either one of its phenotypes.
Figure 2.5: Schematic and visual representation of phenotypic variation throughout the DGRP. Five general categories were identified within the panel, with representative lines pictures above: a Line 138, b Line 142, c Line 555, d line 48, e Line 850

During dissection, lines across the panel exhibited various phenotypes, ranging from a normal activation of the autophagy pathway to a complete disruption of the pathway characterized by a diminished lysosome activity post starvation challenge.

Literature supports that lysosome presence in the cell increases after being challenged with starvation, however we did not expect to see the variety of other phenotypes recorded.
Line 138 (Fig.2.5a), amongst others, exhibited activations patterns consistent with the lab strain, with Fed values centered around 1500 observable units and a foldchange activation value centered around 1.

Line 142 (Fig.2.5b), amongst other, exhibited a "low baseline, normal activation" phenotype, with baseline values centered around 120 units and a foldchange centered on 2.

Line 555 (Fig.2.5c), amongst others, exhibited a “low baseline, low activation” phenotype, with baseline values centered around 900 units and a foldchange centered on 0.6.

Line 48 (Fig.2.5d), amongst others, exhibited a "High baseline, normal activation" phenotype, with baseline values centered around 4000 units and a foldchange centered on 1.

Line 850 (Fig.2.5e), amongst others, exhibited a "High baseline, negative activation" phenotype, with baseline values centered around 3800 units and a foldchange centered on 0.4.

**GWAS of Lysosome Density Phenotypes Within the DGRP Reveals Non-overlapping Datasets.**

Applying the MacKay lab GWAS protocol yielded, for each dataset, which SNPs were significantly correlated with the phenotype of interest. We recovered 105 SNPs associated with baseline phenotype, of which 66 were unique (Fig.2.6a). We recovered 399 SNPs associated with stress response, of which 287 were unique (Fig.2.6b) The full accounting of each SNP call list can be found in Appendix A and B respectively for each dataset.
After accounting for redundant SNPs associated with the same gene, and eliminating all intergenic SNPs, these SNPs correspond to 39, and 178 unique genes respectively, a combined pool of 217 potential regulators associated with lysosome density within the cell, and thus autophagy. When comparing the two datasets, we found an overlap of only 4 genes, showing that the gene sets most correlating with basal or stress induced lysosome density do
not overlap (Fig.2.7). From this pool, likely candidates were further selected for confirmation experiments.

![Venn Diagram](image)

**Figure 2.7:** Gene sets correlated with basal and stress induced lysosome density do not overlap.

**Functional Validation via RNAi Identified 4 Candidates as Regulators of Lysosome Density.**

After selection based on the criteria outlined in the methods, 41 candidates were selected for further analysis. The first round of RNAi testing on candidate genes yielded varied results. Lines were assessed against the genetic background of each line (Attp2 or Attp40) and statistical difference was evaluated using both an ANOVA test, and a simple t-test.

In the Attp2 background (Fig.2.8a), 6 genes were deemed significantly different from the control via one-way ANOVA testing, with a Bonferroni correction, all with a p-value less than 0.0001. Using a t-test analysis, 7 more genes were deemed significantly different from the control.
Figure 2.8: Lysosome density in S106-GS>UAS-RNAi males following RNAi induction. Each RNAi was controlled using its corresponding site specific UAS background, a in Attp2 and b in Attp40.

In the Attp40 background (Fig.2.8b), 6 genes were deemed significantly different from the control via one-way ANOVA testing, with a Bonferroni correction. The highest p-value was 0.0432 (Glu-Rib), and the lowest was 0.0002 (kin17).

Using a t-test analysis, 4 more genes were deemed significantly different from the control.

Overall, results of this first round of RNAi were encouraging, and so a secondary set of candidates was culled from the first for further testing and confirmation. The secondary set comprised of the genes considered to be significantly different from the control by both the t-test, and the ANOVA test. This resulted in 6 genes being selected in each background, for a total of 12 candidates being considered in the secondary analysis (Fig.2.9).

Of the 12 candidates, 4 showed a significant different between control and treatment, and are thus considered potential novel regulators of lysosome density and thus autophagy.
Figure 2.9: Lysosome density before and after induction of RNAi in S106-GS>UAS-RNAi males.

Figure 2.10: kin17, elk, Myo61F and crz regulate lysosome density.
qRT-PCR Verification of Gene Knockdown Confirms Efficacy of UAS-RNAi Lines.

Figure 2.11: qRT-PCR confirms gene knockdown in kin17, corazonin and elk. While the relative gene expression of elk was not significantly reduced, technical replicates support an overall reduction in protein mRNA. Myosin61F gene expression was not detectable.

Quantitative RT-PCR was performed to confirm the knockdown efficiency of RNAi lines. In the case of both kin17 and corazonin (Fig.2.11a&b), a significant decrease in protein mRNA expression (p-value 0.0022 and 0.0245 respectively) was detected, thus confirming that the RNAi line can effectively down-regulate the expression of these genes. In addition, while elk did not show a statistically significant decrease (Fig.2.11.c), multiple technical replicates indicate that the down-regulation is present, despite the extremely low overall expression level. Finally, Myosin-61F expression was not detectable using this protocol.
Discussion

The beauty of the panel created by the MacKay lab is that it relies not on man-mediated mutagenesis, but instead on freely occurring genetic variation. All strains in the DGRP panel were caught in the wild and inbred over twenty generations until isogenic [7]. This provides us with a unique window into natural genetic variation and showcases that complex behavior will vary greatly from individual to individual. Considering the low number of lines established and the close spatiotemporal location at capture, one might have expected the panel to lack diversity, when in fact the exact opposite was revealed. Not all flies collected were able to establish themselves through the inbreeding protocol; as such we cannot consider the DGRP to be a full sample of natural diversity. It is instead, a small snapshot, revealing only a fraction of the genetic diversity that can be found in the wild. This snapshot panel was then established as a database for Genome Wide Association Studies by the MacKay lab, generating a resource that other scientists may then use to ask further questions.

Multiple studies have already utilized GWAS to generate novel interactors in complex traits such as aggression, stress response, and longevity, studies that were made possible by the intensive sequencing of all lines included in the DGRP [9–11].

It was then natural to leverage this tool to attempt a similar process in the field of autophagy. The base mechanistic pathway of autophagy is well categorized, with a comparatively simple system of protein complexes achieving a singular role, that of sequestering and degrading material within the cell. What is less understood, are the triggers that activate or inactivate this mechanism, through which we can gain both a better understanding of cellular mechanisms, but also potentially more tightly regulate the action of autophagy in clinical setting.
In this study, we focused on lysosome abundance within the cell; a system of indirect reporting that has previously been used to quantify autophagy activation within tissues. While this assay does not give a direct representation of the autophagy mechanism, it does allow the quantification of one of its main functional components, the lysosome. In addition, the dye we used (LysoTracker) is primarily effective on acidic vacuoles, which would include both the lysosome and the fused autolysosome [6]. This assay does not in itself give an account of the autophagic flux, which is the rate at which cargo is engulfed, degraded and released back into the cytoplasm. The concept and monitoring of autophagic flux is important to understand where and how a pathway may be disrupted, and without it, no solid conclusions can be made about the rate of autophagy within the cells. However, quantification of the lysosome and autolysosome does showcase an important autophagy-linked phenotype and is often one of the first steps in investigative experiments.

During the phenotyping process, we observed multiple phenotypes both expected and unexpected, thus highlighting the phenotypic complexity available to us through this panel. While some lines behaved as literature would predict others seemed to still be viable despite obvious disruptions to the expected lysosome phenotype. GWAS is a wonderful method for generating large amounts of data, from which candidates must then be isolated. Since the method of GWAS itself relies on the correlation between individual SNPs and the phenotype being assayed, further testing is always necessary to identify which candidates are actually implicated.

It is worth noting that the SNPs called up by the study did not contain any known autophagy gene. This is likely due to the underlying nature of the panel, and how it was established. The panel cannot contain extreme or lethal SNPs that would disrupt crucial cellular mechanisms, because individuals with those genotypes would be expunged from the panel
during the establishing processes. Established lines had to be isogenically stable, and able to persist through generations. This means that any major pathological SNP would not be present within the full panel genotype. As such, when querying the panel about genotype-phenotype correlations, all genes known to be crucial to the proper functionality of a pathway (here autophagy, represented indirectly by lysosome count within the cell) are likely to be functional and only rarely containing significantly disruptive SNPs. This allows the analysis itself to be more sensitive to unknown of autophagy that may act in concert with other lesser regulators. In effect, this type of study allows the study not of the major effectors of a pathway, but of the small potentially additive network of regulators that have not previously been identified by standard mutagenesis assays.

In addition, the phenotype we are using as our base query is not itself solely the domain of autophagy. By using the lysosome as an indirect reporter of autophagy, we are actually assaying SNPs and genes most correlated with lysosome genesis rather than autophagy machinery. As a result, since known autophagy genes and regulators do not influence the bio mechanism of the lysosome, they are absent from the scree. This brings a strong caveat to our results, in that we can only make deductions based on gene involvement in lysosome mechanisms. It is known that lysosomes are an important part of autophagy function, and as such projection hypotheses can be generated from our data but will need to be further verified by assays directly measuring autophagy, not relying on indirect reporters.

Overall the dataset skewed towards intergenic regions, which were not selected for the study due to the difficulties associated with studying these regions. We know that these regions are far from functionally inactive and that regulatory elements and even important cellular product may be generated in these regions. Both of these functions may have an impact on
cellular pathway activity, but the tools to study these regions are few and still being developed. We instead decided to focus on better-annotated regions of the genome, specifically SNPs associated with protein-coding regions. Due to the production of a quantifiable product, we could better study the effect of a gene by removing its product from the organisms through RNA interference.

Our first trial consisted of 41 selected candidates, based on genic location, keywords and statistical correlation with the observed phenotype provided by the GWAS output. The first analysis yielded higher than expected numbers of significantly positive phenotypes. Each line was controlled against its genetic background (Attp2 and Attp40) following activation of the RNA interference.

The first analysis allowed for top candidates to be selected for a secondary analysis using different control methods. The 6 genes most highly activated in each genetic background were selected for a secondary verification. Where before, only the active RNA interference was visualized, we now tested each line against itself as a control, active interference versus inactive.

During the totality of the experiment, candidate gene names were obscured, and lines were only referred to by their Bloomington ID number. This was done to reduce bias, as genes were picked on a qualitative basis, with certain genes being picked as favored candidates. In this way, we hoped to reduce human error that would have prioritized, or inflated results associated with preferred candidates.

The implication of Myosin-61F successfully disrupting the lysosome activity with the organism opens an interesting avenue of research. While this specific protein has not been associated with the lysosome (and autophagy) functional pathway, the myosin-kinesin family of protein is known to be involved in autophagosome transport and fusion to the lysosome [12, 13].
This particular protein may be an as of yet unrecognized key player in this important step. It is also possible that Myosin-61F serves as a reporter of kind, signaling the cell that more lysosomes are needed when fusion events are successfully completed. This would explain why the removal of this protein caused the lysosome numbers in tissues to decrease. The sustaining signal was absent; therefore, the cell did not maintain large pools of lysosome.

Elk was originally selected with other candidates for its involvement in ion transport, a behavior that is crucial for the proper function of the lysosome [14]. Without ion transport, this organelle is unable to maintain its acidic nature, and perform its function. As a result, we hypothesized that a disruption of proteins associated with this function would, in turn, disrupt the proper function of the lysosome, and by extension the autophagy mechanism. Our results showed that elk did not act as a positive regulator of lysosome count within the cell, instead acting as a negative regulator. When removed from the cell, lysosome count increased. As part of the K⁺ voltage gated ion family of proteins, elk may be responsible for the cation backflow necessary for the lysosome to maintain its acidity. This behavior however would not explain why after removal of this potentially crucial protein, lysosome number increased within the cell. It is possible that elk serves a secondary, negative regulatory function activating a negative feedback loop within the cell. As the protein is removed, the cell would compensate by inducing lysosome activity. This mechanism is as of yet not associated with elk or other members of its protein family but could be possible.

In addition, we also identified kin17 as another potential negative regulator of autophagy. Little is known about this protein, apart from a putative DNA binding function due to a zinc-finger motif. Many important regulators act at the DNA level, by regulating expressions of co-factors. The binding site of kin17 is currently unknown but is hypothesized to correspond to
regulators of autophagy. In addition, the specific mechanisms of action are unknown, but further inquiries will be explored in future works. It is possible that further explorations would implicate kin17 as a transcription factor responsible for the regulation of transcription in either the ATG genes or lysosomal biogenesis pathways.

Finally, the protein Corazonin was also isolated as a potential regulator. Corazonin is a neuropeptide and has so far only been described in cerebral and neuronal tissues, not fat bodies or other adipose-type tissues. However, this protein is also implicated in ethanol metabolism, and its function as a hormone does not preclude it from being excreted, and having a secondary function is tissues other than neuronal tissue. The connection to ethanol sensing and metabolism may link this protein as a sensory integrator, responsible for reacting to toxic stress and activating the autophagy system as an attempt to clear the toxic environment. In this sense, it would act as an indirect trigger of the pathway, rather than a direct one, and may not need to be expressed at the site of autophagy.

There are also some limitations inherent to the DGRP panel, and this specific method of discovery. While recording phenotypes, some line exhibited extreme behaviors, which is the entire baseline for this analysis. However, once relevant SNPs and thus genes are isolated, the analysis is no longer conducted in the DGRP background and instead is moved to a single gene of interest model. This allows us to pinpoint genes as causative for our specific phenotype of interest but does not necessarily correlate back to the extreme phenotypes previously observed. Since the DGRP is not man-made, there was no way to control for the number of SNPs present in any one line, leading to a complex interaction of traits to create phenotypes. It is likely that extreme phenotypes observed in the panel are in fact the result of synergistic actions by multiple unique SNPs, as well as a mix of potential direct and indirect triggers.
Indeed, autophagy is a mechanism that can be triggered directly, by the currently known cascade of actions, but it can also be triggered indirectly by defects in other cellular processes causing a situation of stress. This forms one of the main difficulties in this type of study, entangling what is a direct or an indirect action. To that effect, isolated candidates need to be further vetted in a mechanistic manner, so as to support the predictive value of the panel. GWAS, in general, are shown to be strongly predictive tools but are in themselves not sufficient to establish causation.

Conclusion

GWAS is a successful and useful tool when analyzing complex phenotypes. This technique allowed the identification of 217 potential candidates involved in lysosome density regulation within the cell. Further RNAi validation confirmed that 4 candidates (kin17, elk, corazonin and Myo61F) have a functional link to lysosome density, either up or downregulating its activity within tissues. This screen solidifies the value of GWAS as a preliminary screen from which regulators can be further characterized.

References


CHAPTER 3. GENERAL CONCLUSIONS

Overall, this study shows the strength of this specific type of method in generating new candidates in the field of autophagy control and regulation. Generating these new candidates is crucial for research to move forward, as aging continues to be on the forefront of scientific minds everywhere. The direct mechanisms of autophagy have already been strongly tied to aging, but we currently lack a complete understanding of the other pathways and proteins linked to this process. It is in these secondary triggers that potential new therapeutic research can be undertaken. By better understanding how autophagy is triggered or repressed within the cell, which specific proteins play a role in integrating cellular signaling, we come closer to a complete picture of this complex trait.

The amount of data generated by this screen far exceeds the scope of this experiment and warrants further examination. On the largest scale, this method has proven useful fruitful when considering novel actors and regulators when considering complex traits, both behavioral and mechanistic. Further phenotyping assays in this diverse panel will likely yield more data to analyze and further refine our understanding of cellular mechanisms. In relation to this specific study, we queried genes associated with a lysosomal activation phenotype due to starvation, but many other stress conditions could be assayed. In the context of aging, many other phenotypes could be recorded and queried including neuronal and cardiac health as lifespan increases or following an insult consistent with a disease model (cardiac insult in particular). This may yield better insight into how the organism responds to this type of acute or chronic stress and may lead to novel ways to both manage and treat those conditions in humans.
On a narrower scope, there is much data that was generated by this one screen that was not further investigated, which could in itself lead to more novel discoveries. The GWAS protocol yields SNPs it judges to be most correlated with the query phenotype, but that doesn't preclude a weakly correlated SNP from having an important effect. Further assaying gene-associated SNPs could increase the pool of potential novel regulators.

In addition, many of the highly correlated SNPs were not located in genic locations, with a direct impact on gene product, but rather in intergenic locations. It is known that those regions are not inactive, that some will code long-noncoding RNAs and that others are involved in activator/repressor protein interactions. This means that disruptions in those regions could have a major impact on protein concentration within the organism, and cascade down into disruption phenotypes. As such, studying the consequences of coding shifts within those regions may yield valuable data and deepen our understanding of specific cellular mechanisms.

It should also be noted that while we validated the link between phenotype and protein knockdown (qPCR) for a select few candidates, this systematic screen is not sufficient to determine the exact mechanistic action of each selected candidate. Further experiments should be designed for each candidate to elucidate whether it is a direct or indirect trigger of the functional mechanism.

Large phenotyping screens have a strong scientific presence because, by their very creation, they open the doors to the creation of large quantities of hypotheses. The amount of data generated by this screen is immense, and only a very small portion of it has at this point been investigated. The DGRP panel allows us to gather massive amounts of data, and subsequently comb through it in search of scientific novelty. In this way, this study is not unique. What does make it unique, is its value as a well of potential information to be mined, and the fact
that by existing, it allows research to be conducted multiple times on the same dataset. This study isolated 4 potential novel regulators of lysosome behavior within the cell, out of 217 other potential candidate genes. This may sound like a small number, but this marks only the beginning of discoveries. More studies can now be conducted, both on the data as a whole, and the particular candidates.

In addition, research in autophagy, and by extension aging and longevity is crucial to better understand the processes we undergo as we age and prevent the damages that come with this process. As we research and understand more, we can prevent damage and repair it when it happens. The implications of this type of research are wide and far-reaching, as they can affect therapeutic strategies, pharmaceutical ones (like the major discoveries on rapamyacin did) and affect how we engage with our environment. Current strategies are focused on increasing lifespan but gaining a better understanding of machinery at its basest level will also help us increase our health span. If we understand what damages a cell or a tissue as it ages, and how that cell is mechanistically going to respond to particular types of stress, we can prevent and preempt those actions. This may seem like a far of future, but it starts with three potential new regulators isolated in a screen such as this
REFERENCES


Wohlgemuth, S.E. et al., 2007. Autophagy in the Heart and Liver During Normal Aging and Calorie Restriction. *Rejuvenation Research*. 


APPENDIX A. R CODE

#Load libraries
library("lmerTest")
library("rmarkdown")
library("gdata")
library("RColorBrewer")
library("reshape")
library("reshape2")
library(qqman)
library(data.table)

#read gwas.all.assoc file
gwas.result.fed <- read.table("gwas.all.assoc_152fed.txt", header = T, as.is = T)

#split column "ID" into two columns based on " ", place them following the last column,
#rename them "CHR" (chromosome) and "BP" (SNP location).
gwas.result.fed$CHR = as.character(lapply(strsplit(as.character(gwas.result.fed$ID),
       split=" "), "[", 1))
gwas.result.fed$BP = as.character(lapply(strsplit(as.character(gwas.result.fed$ID),
       split=" "), "[", 2))

#Rename column:SingleMixedPval, ID. (Four columns are used to make manhattan
Make sure reshape or reshape2 was loaded last if dplyr is also loaded

#Rename chromosome:
gwas.result.fed[gwas.result.fed=="X"] <- 11
gwas.result.fed[gwas.result.fed=="2L"] <- 21
gwas.result.fed[gwas.result.fed=="2R"] <- 22
gwas.result.fed[gwas.result.fed=="3L"] <- 31
gwas.result.fed[gwas.result.fed=="3R"] <- 32
gwas.result.fed[gwas.result.fed=="4"] <- 44

#Change column from factor to numeric:
gwas.result.fed$CHR <- as.numeric(as.character(gwas.result.fed$CHR))
gwas.result.fed$BP <- as.numeric(as.character(gwas.result.fed$BP))

str(gwas.result.fed)
head(gwas.result.fed)
tail(gwas.result.fed)

as.data.frame(table(gwas.result.fed$CHR))

manhattan(gwas.result.fed)
### APPENDIX B. FLY STOCKS

**RNAi Drivers**

ywR;S106-GS-gal4

w;Da-GS-gal4

**DGRP Fly Stock**

Table B.1: DGRP Fly Stock

<table>
<thead>
<tr>
<th>FBID KEY</th>
<th>CLASSIFICATION</th>
<th>COLLECTION</th>
<th>GENOTYPE</th>
<th>SPECIES</th>
<th>STOCK #</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBst0028122</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-21</td>
<td>Dmel</td>
<td>28122</td>
</tr>
<tr>
<td>FBst0028123</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-26</td>
<td>Dmel</td>
<td>28123</td>
</tr>
<tr>
<td>FBst0028124</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-28</td>
<td>Dmel</td>
<td>28124</td>
</tr>
<tr>
<td>FBst0028125</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-31</td>
<td>Dmel</td>
<td>55014</td>
</tr>
<tr>
<td>FBst0028126</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-32</td>
<td>Dmel</td>
<td>55015</td>
</tr>
<tr>
<td>FBst0028127</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-38</td>
<td>Dmel</td>
<td>28125</td>
</tr>
<tr>
<td>FBst0028128</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-40</td>
<td>Dmel</td>
<td>29651</td>
</tr>
<tr>
<td>FBst0028129</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-41</td>
<td>Dmel</td>
<td>28126</td>
</tr>
<tr>
<td>FBst0028130</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-42</td>
<td>Dmel</td>
<td>28127</td>
</tr>
<tr>
<td>FBst0028131</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-45</td>
<td>Dmel</td>
<td>28128</td>
</tr>
<tr>
<td>FBst0028132</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-48</td>
<td>Dmel</td>
<td>55016</td>
</tr>
<tr>
<td>FBst0028133</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-57</td>
<td>Dmel</td>
<td>29652</td>
</tr>
<tr>
<td>FBst0028134</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-59</td>
<td>Dmel</td>
<td>28129</td>
</tr>
<tr>
<td>FBst0028135</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-69</td>
<td>Dmel</td>
<td>28130</td>
</tr>
<tr>
<td>FBst0028136</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-73</td>
<td>Dmel</td>
<td>28131</td>
</tr>
<tr>
<td>FBst0028137</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-75</td>
<td>Dmel</td>
<td>28132</td>
</tr>
<tr>
<td>FBst0028138</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-83</td>
<td>Dmel</td>
<td>28134</td>
</tr>
<tr>
<td>FBst0028139</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-85</td>
<td>Dmel</td>
<td>28274</td>
</tr>
<tr>
<td>FBst0028140</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-88</td>
<td>Dmel</td>
<td>28135</td>
</tr>
<tr>
<td>FBst0028141</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-91</td>
<td>Dmel</td>
<td>28136</td>
</tr>
<tr>
<td>FBst0028142</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-93</td>
<td>Dmel</td>
<td>28137</td>
</tr>
<tr>
<td>FBst0028143</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-100</td>
<td>Dmel</td>
<td>55017</td>
</tr>
<tr>
<td>FBst0028144</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-105</td>
<td>Dmel</td>
<td>28138</td>
</tr>
<tr>
<td>FBst0028145</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-109</td>
<td>Dmel</td>
<td>28139</td>
</tr>
<tr>
<td>FBst0028146</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-129</td>
<td>Dmel</td>
<td>28140</td>
</tr>
<tr>
<td>FBst0028147</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-136</td>
<td>Dmel</td>
<td>28142</td>
</tr>
<tr>
<td>FBst0028148</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-138</td>
<td>Dmel</td>
<td>28143</td>
</tr>
<tr>
<td>FBst0028149</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-142</td>
<td>Dmel</td>
<td>28144</td>
</tr>
<tr>
<td>FBst0028150</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-149</td>
<td>Dmel</td>
<td>28145</td>
</tr>
<tr>
<td>FBst0028151</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-153</td>
<td>Dmel</td>
<td>28146</td>
</tr>
<tr>
<td>FBst0028152</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-158</td>
<td>Dmel</td>
<td>28147</td>
</tr>
<tr>
<td>FBst0028153</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-161</td>
<td>Dmel</td>
<td>28148</td>
</tr>
<tr>
<td>FBst0028154</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-176</td>
<td>Dmel</td>
<td>28149</td>
</tr>
<tr>
<td>FBst0028155</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-177</td>
<td>Dmel</td>
<td>28150</td>
</tr>
<tr>
<td>FBst0028156</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-181</td>
<td>Dmel</td>
<td>28151</td>
</tr>
<tr>
<td>FBst0028157</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-189</td>
<td>Dmel</td>
<td>28152</td>
</tr>
<tr>
<td>FBst0028158</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-195</td>
<td>Dmel</td>
<td>28153</td>
</tr>
<tr>
<td>FBst0028159</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-208</td>
<td>Dmel</td>
<td>25714</td>
</tr>
<tr>
<td>FBst0028160</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-217</td>
<td>Dmel</td>
<td>28154</td>
</tr>
<tr>
<td>FBst0028161</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-227</td>
<td>Dmel</td>
<td>28156</td>
</tr>
<tr>
<td>FBst0028162</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-228</td>
<td>Dmel</td>
<td>28157</td>
</tr>
<tr>
<td>FBst0028163</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-229</td>
<td>Dmel</td>
<td>29653</td>
</tr>
<tr>
<td>FBst0028164</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-235</td>
<td>Dmel</td>
<td>28275</td>
</tr>
<tr>
<td>FBst0028165</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-237</td>
<td>Dmel</td>
<td>28160</td>
</tr>
<tr>
<td>FBst0028166</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-239</td>
<td>Dmel</td>
<td>28161</td>
</tr>
<tr>
<td>FBst0028167</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-256</td>
<td>Dmel</td>
<td>28162</td>
</tr>
<tr>
<td>FBst0028168</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-280</td>
<td>Dmel</td>
<td>28164</td>
</tr>
</tbody>
</table>
Table B.1 (Continued)

<table>
<thead>
<tr>
<th>FBID KEY</th>
<th>CLASSIFICATION</th>
<th>COLLECTION</th>
<th>GENOTYPE</th>
<th>GENUS</th>
<th>STOCK #</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBst0028165</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-287</td>
<td>Dmel</td>
<td>28165</td>
</tr>
<tr>
<td>FBst0025175</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-301</td>
<td>Dmel</td>
<td>25175</td>
</tr>
<tr>
<td>FBst0025176</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-303</td>
<td>Dmel</td>
<td>25176</td>
</tr>
<tr>
<td>FBst0025177</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-304</td>
<td>Dmel</td>
<td>25177</td>
</tr>
<tr>
<td>FBst0037525</td>
<td>whole genome sequence wild type</td>
<td>BDSC</td>
<td>DGRP-306</td>
<td>Dmel</td>
<td>37525</td>
</tr>
<tr>
<td>FBst0025179</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-307</td>
<td>Dmel</td>
<td>25179</td>
</tr>
<tr>
<td>FBst0028166</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-309</td>
<td>Dmel</td>
<td>28166</td>
</tr>
<tr>
<td>FBst0028267</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-310</td>
<td>Dmel</td>
<td>28276</td>
</tr>
<tr>
<td>FBst0025180</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-313</td>
<td>Dmel</td>
<td>25180</td>
</tr>
<tr>
<td>FBst0025181</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-315</td>
<td>Dmel</td>
<td>25181</td>
</tr>
<tr>
<td>FBst0028167</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-317</td>
<td>Dmel</td>
<td>28167</td>
</tr>
<tr>
<td>FBst0028168</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-318</td>
<td>Dmel</td>
<td>28168</td>
</tr>
<tr>
<td>FBst0055018</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-319</td>
<td>Dmel</td>
<td>55018</td>
</tr>
<tr>
<td>FBst0029654</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-320</td>
<td>Dmel</td>
<td>29654</td>
</tr>
<tr>
<td>FBst0029655</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-321</td>
<td>Dmel</td>
<td>29655</td>
</tr>
<tr>
<td>FBst0025182</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-324</td>
<td>Dmel</td>
<td>25182</td>
</tr>
<tr>
<td>FBst0028171</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-332</td>
<td>Dmel</td>
<td>28171</td>
</tr>
<tr>
<td>FBst0025183</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-335</td>
<td>Dmel</td>
<td>25183</td>
</tr>
<tr>
<td>FBst0028172</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-336</td>
<td>Dmel</td>
<td>28172</td>
</tr>
<tr>
<td>FBst0028173</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-338</td>
<td>Dmel</td>
<td>28173</td>
</tr>
<tr>
<td>FBst0028174</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-340</td>
<td>Dmel</td>
<td>28174</td>
</tr>
<tr>
<td>FBst0055019</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-348</td>
<td>Dmel</td>
<td>55019</td>
</tr>
<tr>
<td>FBst0028176</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-350</td>
<td>Dmel</td>
<td>28176</td>
</tr>
<tr>
<td>FBst0028177</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-352</td>
<td>Dmel</td>
<td>28177</td>
</tr>
<tr>
<td>FBst0055020</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-354</td>
<td>Dmel</td>
<td>55020</td>
</tr>
<tr>
<td>FBst0055038</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-355</td>
<td>Dmel</td>
<td>55038</td>
</tr>
<tr>
<td>FBst0028178</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-356</td>
<td>Dmel</td>
<td>28178</td>
</tr>
<tr>
<td>FBst0025184</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-357</td>
<td>Dmel</td>
<td>25184</td>
</tr>
<tr>
<td>FBst0025185</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-358</td>
<td>Dmel</td>
<td>25185</td>
</tr>
<tr>
<td>FBst0028179</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-359</td>
<td>Dmel</td>
<td>28179</td>
</tr>
<tr>
<td>FBst0025186</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-360</td>
<td>Dmel</td>
<td>25186</td>
</tr>
<tr>
<td>FBst0028180</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-361</td>
<td>Dmel</td>
<td>28180</td>
</tr>
<tr>
<td>FBst0028178</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-362</td>
<td>Dmel</td>
<td>25187</td>
</tr>
<tr>
<td>FBst0025445</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-365</td>
<td>Dmel</td>
<td>25445</td>
</tr>
<tr>
<td>FBst0028181</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-367</td>
<td>Dmel</td>
<td>28181</td>
</tr>
<tr>
<td>FBst0028182</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-370</td>
<td>Dmel</td>
<td>28182</td>
</tr>
<tr>
<td>FBst0028183</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-371</td>
<td>Dmel</td>
<td>28183</td>
</tr>
<tr>
<td>FBst0028184</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-373</td>
<td>Dmel</td>
<td>28184</td>
</tr>
<tr>
<td>FBst0028185</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-374</td>
<td>Dmel</td>
<td>28185</td>
</tr>
<tr>
<td>FBst0025188</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-375</td>
<td>Dmel</td>
<td>25188</td>
</tr>
<tr>
<td>FBst0028186</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-377</td>
<td>Dmel</td>
<td>28186</td>
</tr>
<tr>
<td>FBst0025189</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-379</td>
<td>Dmel</td>
<td>25189</td>
</tr>
<tr>
<td>FBst0025190</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-380</td>
<td>Dmel</td>
<td>25190</td>
</tr>
<tr>
<td>FBst0028188</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-381</td>
<td>Dmel</td>
<td>28188</td>
</tr>
<tr>
<td>FBst0028189</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-382</td>
<td>Dmel</td>
<td>28189</td>
</tr>
<tr>
<td>FBst0028190</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-383</td>
<td>Dmel</td>
<td>28190</td>
</tr>
<tr>
<td>FBst0028191</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-385</td>
<td>Dmel</td>
<td>28191</td>
</tr>
<tr>
<td>FBst0028192</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-386</td>
<td>Dmel</td>
<td>28192</td>
</tr>
<tr>
<td>FBst0055021</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-390</td>
<td>Dmel</td>
<td>55021</td>
</tr>
<tr>
<td>FBst0025191</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-391</td>
<td>Dmel</td>
<td>25191</td>
</tr>
<tr>
<td>FBst0028194</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-392</td>
<td>Dmel</td>
<td>28194</td>
</tr>
<tr>
<td>FBst0055022</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-395</td>
<td>Dmel</td>
<td>55022</td>
</tr>
<tr>
<td>FBst0055023</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-397</td>
<td>Dmel</td>
<td>55023</td>
</tr>
<tr>
<td>FBst0025192</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-399</td>
<td>Dmel</td>
<td>25192</td>
</tr>
<tr>
<td>FBst0029656</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-405</td>
<td>Dmel</td>
<td>29656</td>
</tr>
<tr>
<td>FBst0029657</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-406</td>
<td>Dmel</td>
<td>29657</td>
</tr>
<tr>
<td>FBst0028278</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-409</td>
<td>Dmel</td>
<td>28278</td>
</tr>
<tr>
<td>FBst0028196</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-426</td>
<td>Dmel</td>
<td>28196</td>
</tr>
<tr>
<td>FBst0025193</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-427</td>
<td>Dmel</td>
<td>25193</td>
</tr>
<tr>
<td>FBst0025194</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-437</td>
<td>Dmel</td>
<td>25194</td>
</tr>
<tr>
<td>FBst0029658</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-439</td>
<td>Dmel</td>
<td>29658</td>
</tr>
<tr>
<td>FBst0028197</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-440</td>
<td>Dmel</td>
<td>28197</td>
</tr>
</tbody>
</table>
Table B.1 (Continued)

<table>
<thead>
<tr>
<th>FBID KEY</th>
<th>CLASSIFICATION</th>
<th>COLLECTION</th>
<th>GENOTYPE</th>
<th>GENUS</th>
<th>STOCK #</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBst0028198</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-441</td>
<td>Dmela</td>
<td>28198</td>
</tr>
<tr>
<td>FBst0028199</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-443</td>
<td>Dmela</td>
<td>28199</td>
</tr>
<tr>
<td>FBst0028200</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-461</td>
<td>Dmela</td>
<td>28200</td>
</tr>
<tr>
<td>FBst0025195</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-486</td>
<td>Dmela</td>
<td>25195</td>
</tr>
<tr>
<td>FBst0028202</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-491</td>
<td>Dmela</td>
<td>28202</td>
</tr>
<tr>
<td>FBst0028203</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-492</td>
<td>Dmela</td>
<td>28203</td>
</tr>
<tr>
<td>FBst0028204</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-502</td>
<td>Dmela</td>
<td>28204</td>
</tr>
<tr>
<td>FBst0055024</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-505</td>
<td>Dmela</td>
<td>55024</td>
</tr>
<tr>
<td>FBst0028205</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-508</td>
<td>Dmela</td>
<td>28205</td>
</tr>
<tr>
<td>FBst0028206</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-509</td>
<td>Dmela</td>
<td>28206</td>
</tr>
<tr>
<td>FBst0025197</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-513</td>
<td>Dmela</td>
<td>29659</td>
</tr>
<tr>
<td>FBst0028219</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-517</td>
<td>Dmela</td>
<td>25197</td>
</tr>
<tr>
<td>FBst0055025</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-528</td>
<td>Dmela</td>
<td>55025</td>
</tr>
<tr>
<td>FBst0029660</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-530</td>
<td>Dmela</td>
<td>29660</td>
</tr>
<tr>
<td>FBst0028207</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-531</td>
<td>Dmela</td>
<td>28207</td>
</tr>
<tr>
<td>FBst0028208</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-535</td>
<td>Dmela</td>
<td>28208</td>
</tr>
<tr>
<td>FBst0055026</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-551</td>
<td>Dmela</td>
<td>55026</td>
</tr>
<tr>
<td>FBst0025198</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-555</td>
<td>Dmela</td>
<td>25198</td>
</tr>
<tr>
<td>FBst0055027</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-559</td>
<td>Dmela</td>
<td>55027</td>
</tr>
<tr>
<td>FBst0028211</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-563</td>
<td>Dmela</td>
<td>28211</td>
</tr>
<tr>
<td>FBst0055028</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-566</td>
<td>Dmela</td>
<td>55028</td>
</tr>
<tr>
<td>FBst0028212</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-584</td>
<td>Dmela</td>
<td>28212</td>
</tr>
<tr>
<td>FBst0028213</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-589</td>
<td>Dmela</td>
<td>28213</td>
</tr>
<tr>
<td>FBst0028215</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-595</td>
<td>Dmela</td>
<td>28215</td>
</tr>
<tr>
<td>FBst0055029</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-596</td>
<td>Dmela</td>
<td>55029</td>
</tr>
<tr>
<td>FBst0055030</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-627</td>
<td>Dmela</td>
<td>55030</td>
</tr>
<tr>
<td>FBst0055031</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-630</td>
<td>Dmela</td>
<td>55031</td>
</tr>
<tr>
<td>FBst0055032</td>
<td>-</td>
<td>BDSC</td>
<td>DGRP-634</td>
<td>Dmela</td>
<td>55032</td>
</tr>
<tr>
<td>FBst0025199</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-639</td>
<td>Dmela</td>
<td>25199</td>
</tr>
<tr>
<td>FBst0028217</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-646</td>
<td>Dmela</td>
<td>28217</td>
</tr>
<tr>
<td>FBst0028218</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-703</td>
<td>Dmela</td>
<td>28218</td>
</tr>
<tr>
<td>FBst0025744</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-705</td>
<td>Dmela</td>
<td>25744</td>
</tr>
<tr>
<td>FBst0025200</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-707</td>
<td>Dmela</td>
<td>25200</td>
</tr>
<tr>
<td>FBst0025201</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-712</td>
<td>Dmela</td>
<td>25201</td>
</tr>
<tr>
<td>FBst0025745</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-714</td>
<td>Dmela</td>
<td>25745</td>
</tr>
<tr>
<td>FBst0028219</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-716</td>
<td>Dmela</td>
<td>28219</td>
</tr>
<tr>
<td>FBst0028220</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-721</td>
<td>Dmela</td>
<td>28220</td>
</tr>
<tr>
<td>FBst0028221</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-727</td>
<td>Dmela</td>
<td>28221</td>
</tr>
<tr>
<td>FBst0025202</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-730</td>
<td>Dmela</td>
<td>25202</td>
</tr>
<tr>
<td>FBst0025203</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-732</td>
<td>Dmela</td>
<td>25203</td>
</tr>
<tr>
<td>FBst0028222</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-737</td>
<td>Dmela</td>
<td>28222</td>
</tr>
<tr>
<td>FBst0028223</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-738</td>
<td>Dmela</td>
<td>28223</td>
</tr>
<tr>
<td>FBst0028224</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-748</td>
<td>Dmela</td>
<td>28224</td>
</tr>
<tr>
<td>FBst0028226</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-757</td>
<td>Dmela</td>
<td>28226</td>
</tr>
<tr>
<td>FBst0028227</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-761</td>
<td>Dmela</td>
<td>28227</td>
</tr>
<tr>
<td>FBst0025204</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-765</td>
<td>Dmela</td>
<td>25204</td>
</tr>
<tr>
<td>FBst0025205</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-774</td>
<td>Dmela</td>
<td>25205</td>
</tr>
<tr>
<td>FBst0028229</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-776</td>
<td>Dmela</td>
<td>28229</td>
</tr>
<tr>
<td>FBst0028230</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-783</td>
<td>Dmela</td>
<td>28230</td>
</tr>
<tr>
<td>FBst0025206</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-786</td>
<td>Dmela</td>
<td>25206</td>
</tr>
<tr>
<td>FBst0028231</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-787</td>
<td>Dmela</td>
<td>28231</td>
</tr>
<tr>
<td>FBst0028232</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-790</td>
<td>Dmela</td>
<td>28232</td>
</tr>
<tr>
<td>FBst0028233</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-796</td>
<td>Dmela</td>
<td>28233</td>
</tr>
<tr>
<td>FBst0025207</td>
<td>inbred line &lt;newline&gt; whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-799</td>
<td>Dmela</td>
<td>25207</td>
</tr>
<tr>
<td>FBst0028234</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-801</td>
<td>Dmela</td>
<td>28234</td>
</tr>
<tr>
<td>FBst0028235</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-802</td>
<td>Dmela</td>
<td>28235</td>
</tr>
<tr>
<td>FBst0028236</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-804</td>
<td>Dmela</td>
<td>28236</td>
</tr>
<tr>
<td>FBst0028237</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-805</td>
<td>Dmela</td>
<td>28237</td>
</tr>
<tr>
<td>FBst0028238</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-808</td>
<td>Dmela</td>
<td>28238</td>
</tr>
<tr>
<td>FBst0028239</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-810</td>
<td>Dmela</td>
<td>28239</td>
</tr>
<tr>
<td>FBst0028240</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-812</td>
<td>Dmela</td>
<td>28240</td>
</tr>
<tr>
<td>FBst0028241</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-818</td>
<td>Dmela</td>
<td>28241</td>
</tr>
</tbody>
</table>
Table B.1 (Continued)

<table>
<thead>
<tr>
<th>FBID KEY</th>
<th>CLASSIFICATION</th>
<th>COLLECTION</th>
<th>GENOTYPE</th>
<th>GENUS</th>
<th>STOCK #</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBst0028242</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-819</td>
<td>Dmel</td>
<td>28242</td>
</tr>
<tr>
<td>FBst0025208</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-820</td>
<td>Dmel</td>
<td>25208</td>
</tr>
<tr>
<td>FBst0028243</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-821</td>
<td>Dmel</td>
<td>28243</td>
</tr>
<tr>
<td>FBst0028244</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-822</td>
<td>Dmel</td>
<td>28244</td>
</tr>
<tr>
<td>FBst0028245</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-832</td>
<td>Dmel</td>
<td>28245</td>
</tr>
<tr>
<td>FBst0028246</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-837</td>
<td>Dmel</td>
<td>28246</td>
</tr>
<tr>
<td>FBst0028247</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-843</td>
<td>Dmel</td>
<td>28247</td>
</tr>
<tr>
<td>FBst0028248</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-849</td>
<td>Dmel</td>
<td>28248</td>
</tr>
<tr>
<td>FBst0028249</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-850</td>
<td>Dmel</td>
<td>28249</td>
</tr>
<tr>
<td>FBst0025209</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-852</td>
<td>Dmel</td>
<td>25209</td>
</tr>
<tr>
<td>FBst0028250</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-853</td>
<td>Dmel</td>
<td>28250</td>
</tr>
<tr>
<td>FBst0028251</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-855</td>
<td>Dmel</td>
<td>28251</td>
</tr>
<tr>
<td>FBst0028252</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-857</td>
<td>Dmel</td>
<td>28252</td>
</tr>
<tr>
<td>FBst0025210</td>
<td>inbred line whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-859</td>
<td>Dmel</td>
<td>25210</td>
</tr>
<tr>
<td>FBst0028253</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-861</td>
<td>Dmel</td>
<td>28253</td>
</tr>
<tr>
<td>FBst0028254</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-879</td>
<td>Dmel</td>
<td>28254</td>
</tr>
<tr>
<td>FBst0028255</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-882</td>
<td>Dmel</td>
<td>28255</td>
</tr>
<tr>
<td>FBst0028256</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-884</td>
<td>Dmel</td>
<td>28256</td>
</tr>
<tr>
<td>FBst0028279</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-887</td>
<td>Dmel</td>
<td>28279</td>
</tr>
<tr>
<td>FBst0028257</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-890</td>
<td>Dmel</td>
<td>28257</td>
</tr>
<tr>
<td>FBst0028258</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-892</td>
<td>Dmel</td>
<td>28258</td>
</tr>
<tr>
<td>FBst0028259</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-894</td>
<td>Dmel</td>
<td>28259</td>
</tr>
<tr>
<td>FBst0028260</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-897</td>
<td>Dmel</td>
<td>28260</td>
</tr>
<tr>
<td>FBst0028261</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-900</td>
<td>Dmel</td>
<td>28261</td>
</tr>
<tr>
<td>FBst0028262</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-907</td>
<td>Dmel</td>
<td>28262</td>
</tr>
<tr>
<td>FBst0028263</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-908</td>
<td>Dmel</td>
<td>28263</td>
</tr>
<tr>
<td>FBst0028264</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-911</td>
<td>Dmel</td>
<td>28264</td>
</tr>
<tr>
<td>FBst0028265</td>
<td>whole genome sequence</td>
<td>BDSC</td>
<td>DGRP-913</td>
<td>Dmel</td>
<td>28265</td>
</tr>
</tbody>
</table>
### UAS-RNAi Fly Stock

**Table B.2: UAS-RNAi Fly Stock**

<table>
<thead>
<tr>
<th>STOCK</th>
<th>CG#</th>
<th>FBGN</th>
<th>GENE SYMBOL</th>
<th>GENOTYPE</th>
<th>GENE LOC</th>
<th>VECTOR</th>
<th>TRIP#</th>
<th>HAIRPIN LOCATION</th>
<th>BALANCER</th>
</tr>
</thead>
<tbody>
<tr>
<td>31766</td>
<td>CG1044</td>
<td>FBgn0016794</td>
<td>dos</td>
<td>y1 v1; P(TRiP)attP2</td>
<td>3L</td>
<td>VALIUM1</td>
<td>HM04078</td>
<td></td>
<td>attP2</td>
</tr>
<tr>
<td>31336</td>
<td>CG10645</td>
<td>FBgn0016031</td>
<td>lama</td>
<td>y1 v1; P(TRiP)attP2</td>
<td>3L</td>
<td>VALIUM1</td>
<td>JF01294</td>
<td></td>
<td>attP2</td>
</tr>
<tr>
<td>57711</td>
<td>CG11030</td>
<td>FBgn0031736</td>
<td>CG11030</td>
<td>y1 sc v1; P(TRiP)attP2</td>
<td>2L</td>
<td>VALIUM20</td>
<td>HM04900</td>
<td></td>
<td>attP2</td>
</tr>
<tr>
<td>28796</td>
<td>CG11111</td>
<td>FBgn003218</td>
<td>rdgB</td>
<td>y1 v1; P(TRiP)attP2</td>
<td>X</td>
<td>VALIUM10</td>
<td>JF03224</td>
<td></td>
<td>attP2</td>
</tr>
<tr>
<td>57400</td>
<td>CG11883</td>
<td>FBgn0033538</td>
<td>CG11883</td>
<td>y1 sc v1; P(TRiP)attP40</td>
<td>2R</td>
<td>VALIUM20</td>
<td>HM04704</td>
<td></td>
<td>attP40</td>
</tr>
<tr>
<td>60469</td>
<td>CG12296</td>
<td>FBgn0013469</td>
<td>kdu</td>
<td>y1 v1; P(TRiP)attP40/CyO</td>
<td></td>
<td>VALIUM20</td>
<td>HMJ22839</td>
<td></td>
<td>attP40</td>
</tr>
<tr>
<td>64588</td>
<td>CG13431</td>
<td>FBgn0034521</td>
<td>Mgt1</td>
<td>y1 sc v1; P(TRiP)attP40</td>
<td>2R</td>
<td>VALIUM20</td>
<td>HM05607</td>
<td></td>
<td>attP40</td>
</tr>
<tr>
<td>28762</td>
<td>CG1416</td>
<td>FBgn0032961</td>
<td>CG1416</td>
<td>y1 v1; P(TRiP)attP2</td>
<td>2L</td>
<td>VALIUM10</td>
<td>JF03190</td>
<td></td>
<td>attP2</td>
</tr>
<tr>
<td>30489</td>
<td>CG15532</td>
<td>FBgn0010113</td>
<td>hdc</td>
<td>y1 sc v1; P(TRiP)attP2</td>
<td>3R</td>
<td>VALIUM10</td>
<td>HM05231</td>
<td></td>
<td>attP2</td>
</tr>
<tr>
<td>35038</td>
<td>CG17077</td>
<td>FBgn003118</td>
<td>pnt</td>
<td>y1 sc v1; P(TRiP)attP2</td>
<td>3R</td>
<td>VALIUM20</td>
<td>HMS01452</td>
<td></td>
<td>attP2</td>
</tr>
<tr>
<td>60881</td>
<td>CG17490</td>
<td>FBgn0040009</td>
<td>CG17490</td>
<td>y1 v1; P(TRiP)attP40/CyO</td>
<td></td>
<td>VALIUM20</td>
<td>HMJ22618</td>
<td></td>
<td>attP40</td>
</tr>
<tr>
<td>56932</td>
<td>CG1977</td>
<td>FBgn0250789</td>
<td>alpha-Spec</td>
<td>y1 sc v1; P(TRiP)attP40</td>
<td>3L</td>
<td>VALIUM20</td>
<td>HM04371</td>
<td></td>
<td>attP40</td>
</tr>
<tr>
<td>28693</td>
<td>CG2368</td>
<td>FBgn0263102</td>
<td>psq</td>
<td>y1 v1; P(TRiP)attP2</td>
<td>2R</td>
<td>VALIUM10</td>
<td>JF03109</td>
<td></td>
<td>attP2</td>
</tr>
<tr>
<td>28785</td>
<td>CG2679</td>
<td>FBgn004919</td>
<td>gOl</td>
<td>y1 v1; P(TRiP)attP2</td>
<td>2R</td>
<td>VALIUM10</td>
<td>JF03213</td>
<td></td>
<td>attP2</td>
</tr>
<tr>
<td>28623</td>
<td>CG2682</td>
<td>FBgn0033015</td>
<td>d4</td>
<td>y1 v1; P(TRiP)attP2</td>
<td>2R</td>
<td>VALIUM10</td>
<td>JF03038</td>
<td></td>
<td>attP2</td>
</tr>
<tr>
<td>58341</td>
<td>CG2790</td>
<td>FBgn0027599</td>
<td>CG2790</td>
<td>y1 v1; P(TRiP)attP40/CyO</td>
<td>2R</td>
<td>VALIUM20</td>
<td>HMJ22472</td>
<td></td>
<td>attP40</td>
</tr>
<tr>
<td>58276</td>
<td>CG30268</td>
<td>FBgn0050268</td>
<td>CG30268</td>
<td>y1 v1; P(TRiP)attP40/CyO</td>
<td>2R</td>
<td>VALIUM20</td>
<td>HMJ22359</td>
<td></td>
<td>attP40</td>
</tr>
<tr>
<td>58271</td>
<td>CG30427</td>
<td>FBgn0043792</td>
<td>CG30427</td>
<td>y1 v1; P(TRiP)attP40/CyO</td>
<td>2R</td>
<td>VALIUM20</td>
<td>HMJ22354</td>
<td></td>
<td>attP40</td>
</tr>
<tr>
<td>40923</td>
<td>CG3161</td>
<td>FBgn0262736</td>
<td>Vha16-1</td>
<td>y1 sc v1; P(TRiP)attP40/CyO</td>
<td>2R</td>
<td>VALIUM20</td>
<td>HMS02171</td>
<td></td>
<td>attP40</td>
</tr>
<tr>
<td>25943</td>
<td>CG32975</td>
<td>FBgn0028875</td>
<td>nAcRalpha-34E</td>
<td>y1 v1; P(TRiP)attP2</td>
<td>2L</td>
<td>VALIUM10</td>
<td>JF01963</td>
<td></td>
<td>attP2</td>
</tr>
<tr>
<td>25999</td>
<td>CG3302</td>
<td>FBgn0013767</td>
<td>Crz</td>
<td>y1 v1; P(TRiP)attP2</td>
<td>3R</td>
<td>VALIUM10</td>
<td>JF02023</td>
<td></td>
<td>attP2</td>
</tr>
</tbody>
</table>
Table B.2 (Continued)

<table>
<thead>
<tr>
<th>STOCK</th>
<th>CG#</th>
<th>FBGN</th>
<th>GENE SYMBOL</th>
<th>GENOTYPE</th>
<th>GENE LOC</th>
<th>VECTOR</th>
<th>TRIP#</th>
<th>HAIRPIN LOCATION</th>
<th>BALANCER</th>
</tr>
</thead>
<tbody>
<tr>
<td>29558</td>
<td>CG3533</td>
<td>FBgn0004055</td>
<td>zip</td>
<td>y1 v1; P{TRiP}attP2</td>
<td>2R</td>
<td>VALIUM10</td>
<td>JF03237</td>
<td>attP2</td>
<td></td>
</tr>
<tr>
<td>28032</td>
<td>CG3619</td>
<td>FBgn0000463</td>
<td>Dl</td>
<td>y1 v1; P{TRiP}attP2</td>
<td>3R</td>
<td>VALIUM10</td>
<td>JF02867</td>
<td>attP2</td>
<td></td>
</tr>
<tr>
<td>57320</td>
<td>CG4123</td>
<td>FBgn0026061</td>
<td>Mipp1</td>
<td>y1 sc v1; P{TRiP}attP40</td>
<td>3L</td>
<td>VALIUM20</td>
<td>HMS04521</td>
<td>attP40 CyO fl</td>
<td></td>
</tr>
<tr>
<td>38310</td>
<td>CG42343</td>
<td>FBgn0259245</td>
<td>CG42343</td>
<td>y1 sc v1; P{TRiP}attP40</td>
<td>X</td>
<td>VALIUM20</td>
<td>HMS01774</td>
<td>attP40 CyO fl</td>
<td></td>
</tr>
<tr>
<td>40889</td>
<td>CG42829</td>
<td>FBgn0262018</td>
<td>CadN2</td>
<td>y1 v1/y1 sc v1; P{TRiP}attP40/CyO</td>
<td>2L</td>
<td>VALIUM20</td>
<td>HMS02137</td>
<td>attP40 CyO</td>
<td></td>
</tr>
<tr>
<td>40908</td>
<td>CG43743</td>
<td>FBgn0264009</td>
<td>Glu-RIB</td>
<td>y1 sc v1; P{TRiP}attP40</td>
<td>3L</td>
<td>VALIUM20</td>
<td>HMS02156</td>
<td>attP40 CyO fl</td>
<td></td>
</tr>
<tr>
<td>54012</td>
<td>CG4750</td>
<td>FBgn0259795</td>
<td>loopin-1</td>
<td>y1 v1; P{TRiP}attP40/CyO</td>
<td>2R</td>
<td>VALIUM20</td>
<td>HMJ21436</td>
<td>attP40 CyO mix</td>
<td></td>
</tr>
<tr>
<td>25821</td>
<td>CG5076</td>
<td>FBgn0011589</td>
<td>elk</td>
<td>y1 v1; P{TRiP}attP2</td>
<td>2R</td>
<td>VALIUM10</td>
<td>JF01839</td>
<td>attP2</td>
<td></td>
</tr>
<tr>
<td>36697</td>
<td>CG5083</td>
<td>FBgn0038390</td>
<td>Rbf2</td>
<td>y1 sc v1; P{TRiP}attP2</td>
<td>3R</td>
<td>VALIUM20</td>
<td>HMS01586</td>
<td>attP2</td>
<td></td>
</tr>
<tr>
<td>55692</td>
<td>CG5649</td>
<td>FBgn0024887</td>
<td>kin17</td>
<td>y1 sc v1; P{TRiP}attP40</td>
<td>3L</td>
<td>VALIUM20</td>
<td>HMC03906</td>
<td>attP40</td>
<td></td>
</tr>
<tr>
<td>53315</td>
<td>CG5996</td>
<td>FBgn0032593</td>
<td>trpgamma</td>
<td>y1 sc v1; P{TRiP}attP2</td>
<td>2L</td>
<td>VALIUM20</td>
<td>HMC03542</td>
<td>attP2</td>
<td></td>
</tr>
<tr>
<td>27485</td>
<td>CG6445</td>
<td>FBgn0036715</td>
<td>Cad74A</td>
<td>y1 v1; P{TRiP}attP2</td>
<td>3L</td>
<td>VALIUM10</td>
<td>JF02635</td>
<td>attP2</td>
<td></td>
</tr>
<tr>
<td>57985</td>
<td>CG6472</td>
<td>FBgn0034166</td>
<td>CG6472</td>
<td>y1 sc v1; P{TRiP}attP40</td>
<td>2R</td>
<td>VALIUM20</td>
<td>HMC04689</td>
<td>attP40</td>
<td></td>
</tr>
<tr>
<td>61830</td>
<td>CG6744</td>
<td>FBgn0037901</td>
<td>CG6744</td>
<td>y1 v1; P{TRiP}attP40/CyO</td>
<td>2L</td>
<td>VALIUM20</td>
<td>HMJ23313</td>
<td>attP40 CyO</td>
<td></td>
</tr>
<tr>
<td>34027</td>
<td>CG7851</td>
<td>FBgn0032013</td>
<td>Scalpha</td>
<td>y1 sc v1; P{TRiP}attP2</td>
<td>2L</td>
<td>VALIUM20</td>
<td>HMS00997</td>
<td>attP2</td>
<td></td>
</tr>
<tr>
<td>33884</td>
<td>CG8048</td>
<td>FBgn0262511</td>
<td>Vha44</td>
<td>y1 sc v1; P{TRiP}attP2</td>
<td>2R</td>
<td>VALIUM20</td>
<td>HMS00821</td>
<td>attP2</td>
<td></td>
</tr>
<tr>
<td>56924</td>
<td>CG8093</td>
<td>FBgn0033999</td>
<td>CG8093</td>
<td>y1 sc v1; P{TRiP}attP40/CyO</td>
<td>2R</td>
<td>VALIUM20</td>
<td>HMC04363</td>
<td>attP40 CyO</td>
<td></td>
</tr>
<tr>
<td>44661</td>
<td>CG9023</td>
<td>FBgn0015872</td>
<td>Drirp</td>
<td>y1 sc v1; P{TRiP}attP2</td>
<td>2R</td>
<td>VALIUM20</td>
<td>HMC02945</td>
<td>attP2</td>
<td></td>
</tr>
<tr>
<td>41689</td>
<td>CG9155</td>
<td>FBgn0010246</td>
<td>Myo61F</td>
<td>y1 v1; P{TRiP}attP2</td>
<td>3L</td>
<td>VALIUM20</td>
<td>HMS02253</td>
<td>attP2 TM3,Sb fl</td>
<td></td>
</tr>
</tbody>
</table>